

How to Go Nonmonotonic

David Makinson

**Department of Computer Science
King's College London**

Contents

Purpose

1. Introduction

- 1.1. We are all nonmonotonic
- 1.2. Recalling some features of classical consequence
- 1.3. Some misunderstandings and a habit to suspend
- 1.4. Three ways of getting more conclusions out of your premises

2. Using background assumptions

- 2.1. From classical consequence to pivotal assumptions
- 2.2. From pivotal assumptions to default assumptions
- 2.3. Specializations and generalizations

3. Restricting the set of valuations

- 3.1. From classical consequence to pivotal valuations
- 3.2. From pivotal valuations to default valuations
- 3.3. Specializations and generalizations

4. Using additional rules

- 4.1. From classical consequence to pivotal rules
- 4.2. From pivotal rules to default rules
- 4.3. Generalizations and variants

5. Connections with probabilistic inference

- 5.1. Basic concepts and axioms for probability
- 5.2. Probabilistic characterizations of classical consequence
- 5.3. Supraclassical probabilistic consequence relations
- 5.4. Bringing probabilistic and qualitative inference closer together

Appendix: Proof of Theorem 4.3-1.

Acknowledgements

References

Purpose

This overview is directed to those who have heard of nonmonotonic reasoning and would like to get a better idea of what it is all about. What are its driving ideas? In what ways does it differ from classical logic? How does it relate to probability? More generally, how does it behave and how can it be used? We will try to answer these questions as clearly as possible, without undue technicality on the one hand, nor vagueness or hand waving on the other.

For the outsider, nonmonotonic logic can be a rather mysterious affair. Even from the inside, it can appear to lack unity, with multiple systems proposed by as many authors going in different directions. The few available textbooks tend to perpetuate this impression. Our main purpose is to take some of the mystery out of the subject and show that it is not as unfamiliar as may at first sight seem. In fact, it is easily accessible to anybody with a minimal background in classical propositional logic and a few basic mathematical tools - provided certain misunderstandings and a tenacious habit, which we will signal in the first sections, are avoided and put aside.

As we shall show, there are monotonic logics that act as natural bridges between classical consequence and the principal kinds of nonmonotonic logic to be found in the literature. These logics, which we call *paraclassical*, are very simple to define and easy to study. They provide three main ways of getting more out of your premises than is allowed by strict deduction, that is, good old classical consequence. In other words, they are principled ways of creeping, crawling or jumping to conclusions. Like classical logic, they are perfectly monotonic, but they already display some of the distinctive features of the nonmonotonic systems that they prefigure, as well as providing easy conceptual passage to them. They give us, in effect, three main bridges from the classical homeland to nonmonotonic shores.

We will examine the three bridges one by one. We begin with the simplest among them, whose moving idea is to use *additional background assumptions* along with current premises. Then we consider ways of getting essentially the same result by *excluding certain classical valuations*. And finally we examine a third means to the same end, by *adding rules* alongside the premises. In each case we obtain a monotonic bridge system that may be rendered nonmonotonic in a natural and meaningful way. Indeed, each of the bridge systems leads to quite a range of nonmonotonic ones, and this is where the diverse systems of the literature show themselves. But they are no longer a disorderly crowd. They fall into place as variants, natural and indeed to be expected, of three basic ideas. The overview ends with an examination of the subtle relations between logic and probability – more specifically, between classical logic, probabilistic inference, and nonmonotonic reasoning.

The text does not pretend to cover everything in the area of nonmonotonic reasoning. There are logical constructions exhibiting nonmonotonic behaviour, but which are based in different principles and which we leave aside. In particular, we will not be considering autoepistemic logics. They are formed by adding to the Boolean connectives a special kind of modal operator whose ‘introspective’ interpretation makes them behave nonmonotonically. Nor will we be considering systems based on the idea of iterated undercutting and defeat, as in the theories of defeasible inheritance nets and of argumentative defeat. Finally, no separate treatment is given of logic programming, although it can be seen as a special case of the use of additional rules but in a fragment of the usual Boolean language.

Thus the overview does not pretend to be anything near a comprehensive treatment of all the investigations that have gone under the name of nonmonotonic logic. But it does present some very central ideas, structuring them in a coherent perspective, and seeks to explain them as clearly as possible.

1. Introduction

1.1. We Are All Nonmonotonic

In the stories of Sherlock Holmes, his companion Watson often speaks of the master's amazing powers of deduction. Yet those who have some acquaintance with deductive logic, whether in the form it took at the end of the nineteenth century when Conan Doyle was writing, or as it is understood today, realize that the game is not the same. None of the conclusions drawn by Sherlock Holmes follow deductively, in the strict sense of the term, from the evidence. They involve presumption and conjecture, and the ever-present possibility of going wrong. According to Watson, Sherlock Holmes usually got it right. But in real life, the incidence of failure is very much greater than pleasant fiction allows.

Nevertheless, for all its fallibility, it is reasoning. Appeal is made not only to the observations explicitly mentioned but also, implicitly, to a reservoir of background knowledge, a supply of rules of thumb, a range of heuristic guides. Most important, conclusions may be withdrawn as more information comes to hand, and new ones may be advanced in their place. This does not necessarily mean that there was an error in the earlier reasoning. The previous inferences may still be recognized as the most reasonable ones to have made with the information then available.

This is the essence of nonmonotonic reasoning. It is not only performed by Sherlock Holmes, but also by medical practitioners, garage mechanics, computer systems engineers, and all those who are required to give a diagnosis of a problem in order to pass to action. Robots negotiating their way through obstacles have to make inferences of this kind, and to do it quickly 'in real time' so as to continue their work without pause. Archaeologists sifting through the debris of a site may see their early conclusions about the date, function and origin of an artefact change as more evidence comes to hand.

In some circumstances, we may decline to make such inferences, believing that the evidence is not yet sufficient to draw any conclusions at all. This can sometimes be an eminently rational attitude, and is preferable to deluding oneself by exaggerating the certainty of the conclusions that one might draw. But there are circumstances where urgent action is needed on the basis of conclusions whose uncertainty is all too apparent. For example, in some medical situations the patient may die if action is not taken swiftly. There may be no time to conduct further tests to enlarge the information base, and a diagnosis is needed immediately in order to decide what action to take. In such cases, we must reach conclusions in full recognition of their fallibility, even fragility. If we wait for deductive certainty before ever drawing a conclusion, we will in most contexts be waiting forever.

In more formal terms, we are said to be *reasoning nonmonotonically* when we allow that a conclusion that is well drawn from given information may need to be withdrawn when we come into possession of further information, even when none of the old premises are

abandoned. In brief, a consequence relation is *nonmonotonic* iff it can happen that a proposition x is a consequence of a set A of propositions, but not a consequence of some superset $A \cup B$ of A .

At this point the reader may interject “Of course, that’s the way in which all inference has to be; it can’t be otherwise. There is nothing new about all this. Surely standard systems of logic must already be prepared to deal with it”.

Indeed, although the word is recent, there is nothing new about the idea of nonmonotonicity. Epistemologists have recognized the phenomenon for hundreds of years; for example, both Locke and Hume insisted upon it in the seventeenth and eighteenth centuries. It has for long been familiar to writers on jurisprudence, as well as authors on the philosophy of the empirical sciences. But still today, mainstream systems of logic do not take uncertainty and nonmonotonicity into account. They focus on purely deductive inference, where the conclusion follows of necessity, without possible exception or doubt, from the premises. They are monotonic: when a proposition is a consequence of a set of premises, then it remains a consequence of any larger set formed by adding new premises, so long as none of the old ones are discarded.

This is fine when working in some contexts, notably that of pure mathematics. Indeed logic as we know it today was developed in order to obtain a deeper understanding of deduction in mathematics. The resulting system is known as classical logic, and is made up of two main parts: propositional (alias truth-functional) logic, and predicate (alias first-order, also quantificational) logic. Its quite extraordinary success in the analysis of mathematical reasoning has tended to hide its limitations in areas outside its intended domain.

This is not to say that there is anything wrong with deductive logic. Nor is it necessarily to be regretted that historically things developed in the way that they did, for a clear understanding of classically valid inference is needed before one can begin to make sense of any other kind of reasoning. Although our subject is nonmonotonic reasoning, we will constantly make use of classical logic. Indeed, we do this in two ways. As our enterprise is mathematical in nature, our own reasoning will be carried out in accord with the principles of classical deduction; and the nonmonotonic systems studied will be defined in ways that use classical consequence as a central component.

1.2. Recalling Classical Consequence

As mentioned in the preface, we assume some familiarity with classical propositional consequence. But we refresh the memory by recalling the definition of classical consequence, and some of its most salient characteristics that are not always covered in introductory texts. They will be indispensable for everything that follows.

Classical logic uses a formal language whose *formulae* are made up from an infinite list of elementary letters by means of a suitable selection of truth-functional connectives. For example, we may use the two-place connectives \wedge, \vee and the one-place connective \neg , understood in terms of their usual truth-tables, for conjunction, disjunction, and negation. These three connectives are together functionally complete, in the sense that all truth-functional connectives can be expressed with them. Formulae made up in this way will be called *Boolean*, and we write the set of all of them as L . Material implication \rightarrow and

equivalence \leftrightarrow , also understood in terms of their usual truth-tables, are introduced as abbreviations in the usual way.

An *assignment* is a function on the set of all elementary letters into the two-element set $\{1,0\}$. Each assignment may be extended in a unique way to a *Boolean valuation* (more briefly *valuation* when the context is clear), which is a function v on the set of all formulae into the two-element set $\{1,0\}$ that agrees with the assignment on elementary letters and behaves in accord with the standard truth-tables for the compound formulae made up using \wedge, \vee, \neg . When A is a *set* of formulae, we write $v(A) = 1$ as shorthand for: $v(a) = 1$ for all $a \in A$. To keep notation down, we will usually use the same letter v for both an assignment and the valuation that it determines.

Let A be any set of formulae, and let x be an individual formula. One says that x is a *classical consequence* of A iff there is no valuation v such that $v(A) = 1$ whilst $v(x) = 0$. The standard notation is $A \vdash x$, and the sign \vdash is called ‘gate’ or ‘turnstile’. When dealing with individual formulae on the left, the notation is simplified a little by dropping parentheses, writing $a \vdash x$ in place of $\{a\} \vdash x$.

Thus classical consequence is a *relation* between formulae, or more generally between *sets* of formulae on the left and individual ones on the right. It may also be seen as an *operation* acting on sets A of formulae to give larger sets $Cn(A)$. In effect, the operation gathers together all the formulae that are consequences of given premises. The two representations of classical consequence are trivially interchangeable. Given a relation \vdash , we may define the operation Cn by setting $Cn(A) = \{x: A \vdash x\}$. Conversely we may define \vdash from Cn by the rule: $A \vdash x$ iff $x \in Cn(A)$. Again we simplify notation by eliminating parentheses for singleton premises, writing $x \in Cn(a)$ for $x \in Cn(\{a\})$.

Both of the representations are useful. Sometimes one is more convenient than another. For example, it is often easier to visualize things in terms of the relation, but more concise to formulate and prove them using the operation. The same will be true when we come to non-classical consequence. For this reason, we will constantly be hopping from one notation to the other, and the reader should get into the habit of doing the same.

Consequence is intimately related with consistency. We say that a set A of formulae is *classically consistent* (or *satisfiable*) iff there is some Boolean valuation v with $v(A) = 1$, i.e. $v(a) = 1$ for all $a \in A$. Otherwise we say that it is *inconsistent* (or *unsatisfiable*). Clearly classical consequence and consistency are inter-definable: $A \vdash x$ iff $A \cup \{\neg x\}$ is inconsistent; in the other direction, A is consistent iff $A \not\vdash f$ where f is an arbitrary contradiction, e.g. $p \wedge \neg p$.

The relation of classical consequence has a number of useful properties. To begin with, it is a *closure relation*, in the sense that it satisfies the following three conditions for all formulae a, x and all sets A, B of formulae:

<i>Reflexivity</i> alias <i>Inclusion</i>	$A \vdash a$ whenever $a \in A$
<i>Cumulative Transitivity, CT</i> alias <i>Cut</i>	Whenever $A \vdash b$ for all $b \in B$ and $A \cup B \vdash x$ then $A \vdash x$
<i>Monotony</i>	Whenever $A \vdash x$ and $A \subseteq B$ then $B \vdash x$

Expressed in the language of Cn , this means that classical consequence is a *closure operation* in the sense that it satisfies the following conditions for all sets A, B of formulae.

<i>Inclusion</i> alias <i>Reflexivity</i>	$A \subseteq Cn(A)$
<i>Cumulative Transitivity, CT</i> alias <i>Cut</i>	$A \subseteq B \subseteq Cn(A)$ implies $Cn(B) \subseteq Cn(A)$
<i>Monotony</i>	$A \subseteq B$ implies $Cn(A) \subseteq Cn(B)$

The first condition has two names. It is usually called *reflexivity* when in its application to a relation, and *inclusion* when applied to an operation. But it is exactly the same rule in the two cases, under the translation from one notation to the other.

The second condition has various names in the literature, most commonly *cumulative transitivity*, abbreviated CT, or *cut*. Care should be taken with the latter term however. In work on Gentzen systems for classical logic, for example, it is used not only for this particular rule but also for a number of variants.

Some readers may be a little surprised to see cumulative transitivity for the relation \vdash where they are accustomed to seeing plain *transitivity*, either in its singleton version (whenever $a \vdash b$ and $b \vdash x$ then $a \vdash x$) or in its general version (whenever $A \vdash b$ for all $b \in B$ and $B \vdash x$ then $A \vdash x$). In fact, when we are given both reflexivity and monotony, then cumulative transitivity is equivalent to the general version of plain transitivity. But we will be focussing on relations that are *not* monotonic, and as first noticed by Gabbay (1985), in this context it turns out best to work with cumulative transitivity rather than its plain companion.

Other readers, already familiar with closure operations in logic or abstract algebra, may be surprised to see cumulative transitivity where they are accustomed to seeing *idempotence*, i.e. the condition that $Cn(A) = Cn(Cn(A))$. Once again, the two are equivalent, not as individual conditions, but given the other two conditions defining the notion of a closure operation. But here too, when considering relations that fail monotony it turns out best to focus on CT rather than idempotence.

The third condition, monotony, is the main character of our play, to be conspicuous by its absence after being murdered in the first act. As a special case, it implies *singleton monotony*: whenever $a \vdash x$ then $\{a, b\} \vdash x$. Classical consequence also satisfies a closely related property, known as *singleton conjunctive monotony*, alias the rule of *strengthening the premise*, alias $\wedge+(\text{left})$: whenever $a \vdash x$ then $a \wedge b \vdash x$. These two are not quite the same, as the latter concerns the connective of conjunction; but they are equivalent for any system in which the conjunction of two propositions behaves, as a premise, just like the set of its two conjuncts.

The three conditions defining the notion of a closure relation are examples of what are known as *Horn rules*, so called after Alfred Horn who drew attention to their importance. Roughly speaking, a Horn rule for a relation tells us that if such-and-such and so-and-so (any number of times) are all elements of the relation, then so too is something else. None of the

suppositions of a Horn rule may be negative – none of them can require that something is *not* an element of the relation. Nor is the conclusion allowed to be disjunctive – it cannot say that given the suppositions, *either* this *or* that is in the relation. Horn rules have very useful properties, most notably that whenever a Horn rule is satisfied by every relation in a family, then the relation formed by taking the intersection of the entire family also satisfies the rule.

The three conditions for consequence as a relation, taken together, are equivalent to their counterparts for their counterparts as an operation, under either of the two definitions. Indeed, they are equivalent taken separately, with the nuance that to pass from CT as here formulated for Cn to CT formulated for \vdash requires a little help from reflexivity.

Finally, we recall that classical consequence is *compact*, in the sense that whenever $A \vdash x$ then there is a finite subset $A' \subseteq A$ with $A' \vdash x$. In the language of operations, whenever $x \in Cn(A)$ then there is a finite $A' \subseteq A$ with $x \in Cn(A')$.

Intimately related to compactness is the following *maximalizability property*, which we will use as a tool many times: whenever $A \vdash x$ then there is a set $A^+ \supseteq A$ of formulae such that *maximally* $A^+ \vdash x$. This means that $A^+ \vdash x$ but for every set $A^{++} \supset A^+$ of formulae we have $A^{++} \not\vdash x$. The property can be stated more generally, in terms of consistency: whenever A is consistent with B , then there is a set $A^+ \supseteq A$ of formulae such that *maximally* A^+ is consistent with B . In other words, A^+ is consistent with B and there is no set $A^{++} \supset A^+$ of formulae consistent with B .

This property may be inferred from compactness (or obtained *en route* within certain proofs of compactness), or proven separately. Given the length of its name, we will loosely refer to it as ‘compactness’ as well.

All of these properties of classical consequence are *abstract*, in the sense that they make no reference to any of the connectives \wedge, \vee, \neg . Evidently, classical consequence also has a number of properties concerning each of these connectives, arising from their respective truth-tables – for example the two properties that $a \wedge b \vdash a \vdash a \vee b$. Many of these will be familiar to the reader already, and we will not enumerate them. But we recall one that plays a very important role in what follows.

This is the property of *disjunction in the premises*, alias OR, alias \vee +(left). It says that whenever $A \cup \{a\} \vdash x$ and $A \cup \{b\} \vdash x$ then $A \cup \{a \vee b\} \vdash x$. This too is a Horn rule. In the language of operations, it says $Cn(A \cup \{a\}) \cap Cn(A \cup \{b\}) \subseteq Cn(A \cup \{a \vee b\})$.

We end with a historical note. The notion of a closure operation goes back to work of Kuratowski in topology in the 1920s. However, as defined in logic it is more general than in topology; it does not require that $Cn(A \cup B) = Cn(A) \cup Cn(B)$ – which is false for classical and most other kinds of consequence. The notion was defined in essentially its present form by Tarski (1930), and gradually permeated work in abstract algebra, e.g. Cohn (1965), and logic, e.g. Brown and Suszko (1973), Wójcicki (1988).

1.3. Some Misunderstandings and a Habit to Suspend

For a person coming to nonmonotonic reasoning for the first time, it can be rather difficult to get a clear grip on what is going on. This is partly due to certain misunderstandings, which are very natural but distort understanding from the beginning. In this section we clear them away so that the reader can proceed without hidden impediments.

Weaker or Stronger?

The first thing that one hears about nonmonotonic logic is, evidently, that it is not monotonic. In other words, it fails the principle that whenever x follows from a set A of propositions then it also follows from every set B with $B \supseteq A$. By contrast, classical logic satisfies this principle, as we have seen in the preceding section.

Given this failure, it is natural to imagine that nonmonotonic logic is weaker than classical logic. And indeed, in one sense it is. Typically, the set of Horn rules satisfied by a nonmonotonic consequence relation is a proper subset of those holding for classical consequence. For example, with preferential consequence relations (to be explained in section 3) the rules of reflexivity and cumulative transitivity always hold while monotony may not. And since classical consequence is itself a limiting case of preferential consequence, any Horn rule that is satisfied by all preferential consequence relations is satisfied by classical consequence.

But in another and more basic sense, the nonmonotonic logics that we will be studying are *stronger* than their classical counterpart. Recall that classical consequence is a relation, i.e. under the usual understanding of relations in set-theoretical terms, it is a set of ordered pairs. Specifically, \vdash is a set of ordered pairs (A,x) , where A is a set of Boolean formulae and x is an individual Boolean formula. When written as an operation C_n , it is a set of ordered pairs (A,X) , where both components are sets of propositions. It is at this level that the most basic comparison of strength arises.

Suppose we take a nonmonotonic consequence relation $\vdash\sim$ (usually pronounced ‘snake’). It too is a set of ordered pairs (A,x) . Under the principal modes of generation that we will be studying, it is a *superset* of the classical consequence relation, over the same set of Boolean formulae. In other words, we have $\vdash \subseteq \vdash\sim \subseteq 2^L \times L$, where \subseteq is set inclusion.

Likewise, suppose that we take a nonmonotonic consequence operation, usually referred to by the letter C . Then we have $C_n \leq C$. Here \leq is not quite set inclusion between the two operations, but set inclusion between their values. That is, $C_n \leq C$ means: $C_n(A) \subseteq C(A)$ for all sets A of formulae.

In this sense, nonmonotonic consequence relations are typically stronger than classical consequence, and when they are so we will call them *supraclassical* relations.

To be fair, there is a lot of looseness in talk of ‘weaker’ and ‘stronger’ relations. A person might retort to the preceding paragraph that it spells out precisely a sense in which we should say that nonmonotonic consequence relations $\vdash\sim$ are weaker than classical \vdash . In common language, we say that the relation of sisterhood is stronger than that of sibling. It is the sub-relation that is described as stronger, and the super-relation that is described as weaker. Since \vdash is a subrelation of $\vdash\sim$ we should say that classical consequence is the stronger!

In fact, we seem to tolerate two contradictory ways of speaking. One is rooted in informal discourse and the other is anchored in more technical language. We do tend to say that the relation of being a sister is stronger than that of being a sibling: in this case the sub-relation is ordinarily seen as stronger. On the other hand, without exception, logicians describe classical consequence as stronger than, say, intuitionistic or relevantist consequence. In this case it is the super-relation that is described as stronger.

In the end, it does not really matter which way we speak, so long as we understand and agree on what we mean and do not allow the coexistence of two conflicting ways of speaking to engender confusion. Our terminology for consequence relations will be constant. The super-relation will be called stronger, and the sub-relation will be termed weaker. Moreover, any loss of Horn rules satisfied by the relation will be described as a loss of regularity, rather than as a loss of strength.

Classical or Non-Classical?

In so far as monotony fails, the logic of nonmonotonic consequence relations certainly differs from classical logic. But it would be quite misleading to refer to it as a kind of ‘non-classical logic’ as that term is usually employed, for example in reference to intuitionistic logic.

For in contrast to the intuitionistic case, our relations do not reject as incorrect any elements of classical consequence. As already remarked, each of them includes classical consequence. Nor do we suggest that there is anything incorrect about the Horn property of monotony. It is perfectly appropriate for purely deductive reasoning, but can fail for other kinds of reasoning.

As emphasized in the first section, there is nothing wrong with classical logic. Moreover, we need to understand and use it when trying to understand other kinds of reasoning. In effect, we will show how the ‘good old relation of classical consequence’ may be deployed in certain ways to define stronger relations that are of practical value, but happen to fail monotony. Rather than talk of non-classical logics, it is more illuminating to speak of more sophisticated ways of employing classical logic, which generate non-monotonic behaviour.

One Logic or Many?

There is a third common misunderstanding of what nonmonotonic logic is all about. From the classical context, we are familiar with the idea that there is just one core logic, apart from notational differences and features like the choice of primitive connectives. That core is classical logic, and it is also the logic that we use when reasoning ourselves in the metalanguage.

Even intuitionists and relevantists, who do not accept all of classical logic, feel the same way – but each about their own system, which is a subsystem of the classical one. They have some difficulties, one might maliciously add, in reconciling this view with their own practice in the metalanguage, where they usually use classical logic, but that is another matter...

Given the unicity of classical inference, it is natural for the student, puzzled by seeing several different kinds of nonmonotonic consequence, to ask: what is *real* nonmonotonic inference? Which is the *correct* nonmonotonic consequence relation? What is the one that we use in practice, even if we can study others?

The answer is that *there is none*. There is no unique nonmonotonic consequence relation, but indefinitely many of them. There are all the relations that can be generated from certain kinds of structure, whose ingredients are allowed to vary within the boundaries of suitable constraints (which we will be exploring). Likewise, there are all the relations satisfying certain syntactic conditions, (which we will also study). Moreover, if one tries to get away from non-uniqueness by intersecting all the many such relations, one ends up with just classical logic again.

Leaving aside technical details, the essential message is as follows. Don't expect to find *the* nonmonotonic consequence relation that will always, in all contexts, be the right one to use. Rather, expect to find several *families* of such relations, interesting *syntactic conditions* that they sometimes satisfy but sometimes fail, and principal *ways of generating* them mathematically from underlying structures.

A Habit to Suspend

In the following pages, we will be showing that there are systems that act as natural bridges between classical consequence and nonmonotonic logics. These bridge systems are also supraclassical, but they are perfectly monotonic, indeed they are closure operations.

“But if these bridge systems are supraclassical”, one may ask, “how are they possible? Surely classical consequence is *already* maximal, in the sense that there is no stronger closure operation in the same underlying language, other than the trivial one under which every proposition of the language implies every other one. So how can the ‘bridge logics’ be closure operations and at the same time supraclassical?”

Indeed, this maximality observation has been part of the folklore of logic since the early twentieth century. But the formulation above omits a vital element, which is not always made as explicit as it should be. That is the condition of being *closed under substitution*, which we now explain.

By *substitution* we mean what is often called, at greater length, *uniform substitution of arbitrary formulae for the elementary letters in a formula*. For example, when a is the formula $p \wedge (q \vee \neg r)$, where p, q, r are three distinct elementary letters, we may consider the substitution σ that replaces all occurrences of p by r , say; all occurrences of q by $\neg p$; and (simultaneously, not subsequently) all occurrences of r by $\neg(p \wedge s)$. That will give us $\sigma(a) = r \wedge (\neg p \vee \neg \neg(p \wedge s))$. Simplifications, such as the elimination of the double negation, are not part of the substitution, but possible later actions.

Substitution, in this sense, is a function, not an inference. For those accustomed to algebraic terminology, a substitution is an endomorphism on the absolutely free algebra of formulae into itself. In general, we do not have $a \vdash \sigma(a)$; one need only consider the counterexample where a is an elementary letter p and σ takes p to another elementary letter q . On the other hand it is true that when a Boolean formula a is a classical tautology, then so too is $\sigma(a)$ for any substitution σ . For example $p \vee \neg p$ is a tautology, and so too is $\sigma(p \vee \neg p) = \sigma(p) \vee \neg \sigma(p)$ for every substitution σ . In other words, writing T for the set of classical tautologies, and writing $\sigma(T)$ for $\{\sigma(a) : a \in T\}$, we have $T \subseteq \sigma(T)$ for every substitution σ . In this sense we say that the set of tautologies is *closed under substitution*. So too is the relation of classical consequence. In other words, whenever $A \vdash x$ then $\sigma(A) \vdash \sigma(x)$. In the language of operations, whenever $x \in Cn(A)$ then $\sigma(x) \in Cn(\sigma(A))$; more concisely $\sigma(Cn(A)) \subseteq Cn(\sigma(A))$.

The maximality of classical logic may now be expressed as follows.

Theorem 1.3-1

There is no supraclassical closure relation in the same language as classical \vdash that is closed under substitution, except for \vdash itself and the total relation.

Here, the total relation is the one that relates every possible premise (or set of premises) to every possible conclusion; as an operation it sends any set of formulae to the set of all formulae. The proof of the observation is straightforward, and we recall it here.

Proof

Let \vdash_+ be any closure relation over Boolean formulae that is closed under substitution and also properly supraclassical, i.e. $\vdash \subset \vdash_+$. We want to show that $B \vdash_+ y$ for arbitrary B, y .

By the second hypothesis, there are A, x with $A \vdash_+ x$ but $A \not\vdash x$. From the latter, there is a classical valuation v with $v(A) = 1, v(x) = 0$. Take the substitution σ that substitutes tautologies for elementary letters that are true under v , and contradictions for letters that are false under v . We see, by an easy induction on depth of formulae, that $\sigma(A)$ is a set of tautologies while $\sigma(x)$ is a contradiction. Since by the first hypothesis \vdash_+ is closed under substitution, $A \vdash_+ x$ implies $\sigma(A) \vdash_+ \sigma(x)$.

But since $\sigma(A)$ is a set of tautologies we have by classical logic that for arbitrary $B, B \vdash \sigma(a)$ for all $\sigma(a) \in \sigma(A)$. Likewise, since $\sigma(x)$ is a contradiction we have $\sigma(x) \vdash y$ for arbitrary formula y . Thus since $\vdash \subseteq \vdash_+$ we have $B \vdash_+ \sigma(a)$ for all $\sigma(a) \in \sigma(A)$, and also $\sigma(x) \vdash_+ y$. Putting these together with $\sigma(A) \vdash_+ \sigma(x)$, we can apply cumulative transitivity and monotony of \vdash_+ twice to get $B \vdash_+ y$, completing the proof.

The moral of this story is that the supraclassical closure relations that we will be offering as bridges between classical consequence and nonmonotonic consequence relations *are not closed under substitution*. Nor, for that matter, are the nonmonotonic relations that issue from them. This runs against deeply ingrained habit. Students of logic are brought up with the idea that any decent consequence relation should be purely formal or structural, and hence satisfy substitution. To understand nonmonotonic logic, it is a habit to suspend.

1.4. Three Ways of Getting More Conclusions out of your Premises

In this overview we will describe three different ways of getting more than is authorized by classical consequence out of a set of premises. Roughly speaking, the first method uses additional background *assumptions*. The second restricts the set of *valuations* that are considered possible. And the third uses additional background *rules*.

Each of these procedures gives rise to a corresponding kind of monotonic consequence operation. They are not entirely equivalent to each other. But they all give us closure operations in the sense defined in section 1.2, and they are all supraclassical. We will call consequence relations with these two properties *paraclassical*. In other words, a paraclassical consequence relation is defined to be any supraclassical closure relation.

The three kinds of paraclassical consequence serve as conceptual bridges to corresponding families of nonmonotonic consequence. They are formed essentially by allowing key elements of the respective constructions to vary in a principled way with the premises under consideration.

To change the metaphor, imagine a solar system. The sun of classical consequence illuminates the firmament from its centre. Three kinds of paraclassical consequence circle around it like planets: *pivotal-assumption*, *pivotal-valuation*, and *pivotal-rule* consequence. Their key ingredients are, respectively, a set of additional background assumptions, a reduced set of valuations, and an additional set of rules. By allowing these to vary in a principled manner with the premises of any given inference, we will obtain three satellite kinds of nonmonotonic consequence operation: *default-assumption*, *default-valuation*, and *default-rule* consequence.

All these concepts are developed over a purely propositional language using only Boolean connectives, without adding further connectives to the usual truth-functional ones.

2. Using Additional Background Assumptions

2.1. From Classical Consequence to Pivotal Assumptions

We begin by examining the simplest kind of paraclassical consequence and its transformation into a form of nonmonotonic reasoning, namely inference with additional background assumptions.

In daily life, the assumptions that we make when reasoning are not all equally conspicuous. Generally, there will be a few that we display explicitly, because they are special to the situation under consideration or in some other way deserve particular attention. There will usually be many others that we do not bother even to mention, because we take them to be common knowledge, or in some way trivial. There will be yet others of which we are only half aware. This kind of phenomenon was already known to the ancient Greeks, who used the term *enthymeme* to refer to an argument in which one or more premises are left implicit. That is the idea that we develop in this section.

We work with the same propositional language as in classical logic, with the set of all its formulae called L . Let $K \subseteq L$ be a fixed set of formulae. Intuitively K will be playing the role of a set of background assumptions or, as they are called in Gärdenfors and Makinson (1994), ‘expectations’. Let A be any set of formulae, and let x be an individual formula.

Definition: Pivotal-Assumption Consequence

- We say that x is a *consequence of A modulo the assumption set K* , and write $A \vdash_K x$ alias $x \in Cn_K(A)$ iff there is no valuation v such that $v(K \cup A) = 1$ while $v(x) = 0$. Equivalently: iff $K \cup A \vdash x$.
- We call a relation or operation a *pivotal-assumption consequence* iff it is identical with \vdash_K (resp. Cn_K) for some set K of formulae.

Thus there is not a unique pivotal-assumption consequence relation, but many – one for each value of K .

Since classical consequence is monotonic, pivotal-assumption consequence relations and operations are supraclassical in the sense defined earlier. That is, for every K we have $\vdash \subseteq \vdash_K$, in the operational notation $Cn \leq Cn_K$.

They also share a number of abstract properties with classical consequence. In particular, they satisfy inclusion, cumulative transitivity and monotony, and thus are closure operations. Being supraclassical closure operations they are by definition (section 1.4), paraclassical. They are also compact, and have the property of disjunction in the premises. These positive features may all be verified by straightforward applications of the definitions. For ease of reference we summarise them in a table.

Table 2.1. Some Positive Features of Pivotal-Assumption Consequence

Supraclassical		Closure Relation	Paraclassical
Reflexive			
Cumulatively Transitive (CT)			
Monotonic			
Disjunction in the Premises (OR)			
Compact			

On the other hand, since the relations \vdash_K are supraclassical closure relations, we would expect from Theorem 1.3-1 that they are not, in general, closed under substitution.

It is instructive to look at a simple example how this failure happens. Let $K = \{p\}$ where p is an elementary letter. Choose any other elementary letter q , and put $A = \{q\}$ and $x = p \wedge q$. Then $A \vdash_K x$ since $\{p, q\} \vdash p \wedge q$. Now let σ be the substitution that replaces every elementary letter by itself, except for the letter p , which is replaced by some elementary letter r distinct from itself and from q , so that $\sigma(p) = r$ and $\sigma(q) = q$. Then $\sigma(A) \not\vdash_K \sigma(x)$ since $K \cup \sigma(A) = K \cup \{q\} = \{p, q\} \not\vdash r \wedge q = \sigma(x)$. Analysing this example, we see what is going on: the substitution is applied to the explicit premises A and the conclusion x , but not to the background assumption set K , because K is held constant.

On the other hand, it is easy to verify that when the background assumption set K is itself closed under substitution then the corresponding consequence relation \vdash_K is also closed under substitution. But this special case is a degenerate one. For as we have just noticed, \vdash_K is always paraclassical (i.e. a supraclassical closure relation), so Theorem 1.3-1 tells us that when it is also closed under substitution, it must be either classical consequence itself (which happens when $K = Cn(\emptyset)$) or the total relation (which happens when $K \supset Cn(\emptyset)$). To reach the same point more directly: when K is closed under substitution, then either all of its elements are tautologies or one of its elements has a substitution instance which is a contradiction and is also in K , so that $K \cup A$ is inconsistent and thus $K \cup A \vdash x$ for any x .

A striking feature of pivotal-assumption consequence, distinguishing it from the other bridge systems that we will be describing, is that the positive properties that we have entered into the table above suffice to characterize it. In other words, we have the following representation theorem for pivotal-assumption consequence operations (and likewise for the corresponding relations).

Theorem 2.1-1

Let Cn^+ be any paraclassical consequence operation that is compact and satisfies the condition of disjunction in the premises. Then there is a set K of formulae such that $Cn^+ = Cn_K$.

This representation theorem is formulated in Rott (2001) (section 4.4 observation 5), but appears to have been part of the folklore for a long time. Before giving a proof, we draw attention to three important properties needed for it.

- We say that a consequence operation C satisfies *left classical equivalence* (LCE) iff whenever $Cn(A) = Cn(B)$ then $C(A) = C(B)$. Note carefully the pattern: the antecedent concerns classical consequence Cn ; the consequent concerns whatever consequence operation C we are interested in; thus the principle as a whole relates the two. This property is often known as ‘left logical equivalence’ (LLE), but this can be a little confusing unless one keeps in mind that ‘logical’ here means ‘classical’.
- We say that a consequence operation C satisfies *right weakening* (RW) iff whenever $x \in C(A)$ and $y \in Cn(x)$ then $y \in C(A)$. In relational notation: whenever $A \mid\sim x \mid\sim y$ then $A \mid\sim y$. Note again the interplay between classical Cn and the operation C under consideration.
- We say that a consequence operation C has the *free premise* property iff whenever $B \subseteq A$ then $C(A) = C(C(B) \cup A)$. This is not an interactive property, but concerns C alone. It says that whenever $B \subseteq A$ then all the propositions in $C(B)$ may be thrown in as free premises alongside A without affecting the consequences of A .

In the particular case that B is the empty set of formulae, the free premise property tells us that $C(A) = C(C(\emptyset) \cup A)$. In other words, the identity of $C(A)$ is not affected by adding into A the consequences (under the same consequence operation) of the empty set.

Classical consequence itself has all three of these properties, the first one trivially, the second as a form of transitivity, and the third not difficult to verify. To establish the representation theorem, we need to know that *every* paraclassical consequence operation has these properties.

Lemma for Theorem 2.1–1

Let Cn^+ be any paraclassical consequence operation. Then Cn^+ satisfies left classical equivalence, right weakening, and free premises.

The verification is straightforward. With it in hand, we can prove the representation theorem as follows.

Proof of Theorem 2.1–1

Let Cn^+ be any paraclassical consequence operation that is compact and satisfies the condition of disjunction in the premises. Put $K = Cn^+(\emptyset)$. We claim that $Cn^+ = Cn_K$. It will suffice to show both $Cn_K \leq Cn^+$ and $Cn^+ \leq Cn_K$.

For the inclusion $Cn_K \leq Cn^+$ we need to show that for any A , $Cn(Cn^+(\emptyset) \cup A) \subseteq Cn^+(A)$. Since by supraclassicality $Cn \leq Cn^+$, we have $Cn(Cn^+(\emptyset) \cup A) \subseteq Cn^+(Cn^+(\emptyset) \cup A) = Cn^+(A)$ by the free premise property of paraclassical operations.

For the converse inclusion $Cn^+ \leq Cn_K$ we need to show that for any A , $Cn^+(A) \subseteq Cn(Cn^+(\emptyset) \cup A)$. This is where we need compactness and disjunction in the premises.

Suppose $x \in Cn^+(A)$. Then by the compactness of Cn^+ there is a finite $B \subseteq A$ with $x \in Cn^+(B)$. Let b be the conjunction of all the finitely many elements of B . Using left classical equivalence we have $x \in Cn^+(b)$ and thus by right weakening $\neg b \vee x \in Cn^+(b)$. But also, $\neg b \vee x \in Cn(\neg b) \subseteq Cn^+(\neg b)$, using the supraclassicality of Cn^+ . Applying disjunction in the premises, we thus have $\neg b \vee x \in Cn^+(b \vee \neg b) = Cn^+(\emptyset)$ using left classical equivalence again.

Since $\neg b \vee x \in Cn^+(\emptyset)$, monotony tells us that to show that $x \in Cn(Cn^+(\emptyset) \cup A)$ as desired, it will suffice to show that $x \in Cn(\{\neg b \vee x\} \cup A)$. But by the construction of b we have $b \in Cn(A)$ and so using disjunctive syllogism for classical consequence, $x \in Cn(\{\neg b \vee x\} \cup A)$, thereby completing the proof.

Those familiar with abstract algebra will be accustomed to the use of the term ‘representation theorem’ for results like Theorem 2.1–1. Such theorems are legion there – for example, the representation of Boolean algebras as fields of sets, and of groups in terms of transformations. But those coming from classical logic (and its subsystems like intuitionistic logic) may be puzzled by the term. In logic they are accustomed to ‘completeness theorems’. They seem to be doing much the same sort of job, so why the different terminology?

The reason comes largely from the fact, underlined in section 1.3, that whereas classical consequence is a *unique* relation, the nonmonotonic consequences that we are studying are *multiple*: different consequence relations are produced by different choices of the parameters (e.g. the background assumption set K) in the generating apparatus.

For example, the completeness theorem for classical logic tells us that whenever x is a classical consequence of A , where this relation is understood as defined semantically in terms of valuations, then x may be obtained from A by the application of certain rules. In other words, whenever $A \vdash x$ then x is in the least superset of A that is closed under the syntactic rules. The rules entertained must be ones for which the notion of ‘the least superset’ makes sense. In other words, the intersection of any family of supersets of A satisfying the rules must also satisfy them. Usually they will be Horn rules.

On the other hand, a representation theorem tells us that *every* structure satisfying certain syntactic conditions, is identical with (or in some broader sense equivalent to) *some* structure of a semantically defined kind. For example, the representation theorem for Boolean algebras says that *every* Boolean algebra (defined by, say, equations) is isomorphic to *some* field of sets. Again, in the present context, our representation theorem for pivotal-assumption consequence (Theorem 2.1–1) tells us that *every* relation satisfying certain syntactic conditions (paraclassicality, compactness, and disjunction in the premises) is identical with *some* semantically defined relation (one of the kind \vdash_K).

For representation theorems, the syntactic conditions can take a wider variety of forms than for completeness theorems. As we are not closing any set under the conditions, they need not take the form of Horn rules. We will see some examples of this later with the conditions of rational and disjunctive monotony, which have negative premises.

What happens if we try to get a completeness theorem out of Theorem 2.1–1, say by intersecting *all* the relations satisfying the specified syntactic conditions, i.e. by looking at the

least such relation? We do get something: it is not difficult to show that the least paraclassical consequence operation that satisfies the condition of disjunction in the premises is also compact, and it coincides with the semantically defined relation \vdash_K where $K = \emptyset$. But this fact is quite uninteresting, for then both left and right hand sides of the theorem are just classical consequence all over again!

In general, when dealing with supraclassical inference relations, this pattern repeats itself: interesting representation theorems may be proven, but attempts to establish completeness theorems, by taking the least relation satisfying given Horn conditions, collapse into triviality.

2.2. From Pivotal Assumptions to Default Assumptions

What has all this to do with *nonmonotonic* inference? Pivotal-assumption consequence relations \vdash_K are, as we have seen, perfectly monotonic. But nonmonotonicity is created if we now allow the background assumption set K to *vary* with the premise set A . More precisely, when the part of K that we actually use is allowed to vary, in a principled way, with the current premise set A . This is done by imposing a consistency constraint, and diminishing the usable part of K when it is violated.

Specifically, we go nonmonotonic when we use only the maximal subsets K' of K that are consistent with A , and accept as output only what is common to their separate outputs. We call this relation *default-assumption consequence*, to bring out its close relation to the preceding pivotal-assumption consequence.

To give the definition more explicitly, let $K \subseteq L$ be a set of formulae, which again will play the role of a set of background assumptions. Let A be any set of formulae, and let x be an individual formula.

We say that a subset K' of K is *consistent with A* (more briefly: is *A -consistent*) iff there is a classical valuation v with $v(K' \cup A) = 1$.

A subset K' of K is called *maximally consistent with A* (more briefly: *maxiconsistent with A* , or *maximally A -consistent*) iff it is consistent with A but is not a proper subset of any $K'' \subseteq K$ that is consistent with A . We can now formulate the main definition.

Definition: Default-Assumption Consequence

- We define the relation $\vdash_{\sim K}$ of *consequence modulo the default assumptions K* by putting $A \vdash_{\sim K} x$ iff $K' \cup A \vdash x$ for every subset $K' \subseteq K$ that is maxiconsistent with A .

Writing C_K for the corresponding operation, this puts $C_K(A) = \bigcap \{Cn(K' \cup A) : K' \subseteq K \text{ and } K' \text{ maxiconsistent with } A\}$.

- We call a relation or operation a *default-assumption consequence* iff it is identical with $\vdash_{\sim K}$ (resp. C_K) for some set K of formulae.

The notation used here is designed to run in parallel with that for pivotal-assumption consequence. Gates become snakes, i.e. \vdash_K becomes $\vdash_{\sim K}$ and likewise Cn_K becomes C_K . Note

again that there is not a unique default-assumption consequence relation, but many – one for each value of K .

Example

To illustrate the definition, consider the following example, which may be called the “Möbius strip”. Put $K = \{p \rightarrow q, q \rightarrow r, r \rightarrow \neg p\}$, let the premise set $A = \{p\}$. Then clearly K is inconsistent with A , so if we were looking at pivotal-assumption consequence we would have $Cn_K(A) = L$. But for default-assumption consequence we need to find the maximal subsets $K' \subseteq K$ that are consistent with A .

- A little checking shows that these are the three two-element subsets of K , namely $K_1 = \{p \rightarrow q, q \rightarrow r\}$, $K_2 = \{p \rightarrow q, r \rightarrow \neg p\}$, $K_3 = \{q \rightarrow r, r \rightarrow \neg p\}$. So a formula x is a default-assumption consequence of A modulo the background assumptions K , i.e. $x \in C_K(A)$, iff $x \in Cn(K_i \cup \{p\})$ for $i = 1, 2, 3$.
- A bit more checking shows that $Cn(K_1 \cup \{p\}) = Cn(\{p, q, r\})$, while $Cn(K_2 \cup \{p\}) = Cn(\{p, q, \neg r\})$, and $Cn(K_3 \cup \{p\}) = Cn(\{p, \neg q, \neg r\})$. So $x \in C_K(A)$ iff x is in all three of these sets, i.e. iff x is a classical consequence of the disjunction $(p \wedge q \wedge r) \vee (p \wedge q \wedge \neg r) \vee (p \wedge \neg q \wedge \neg r)$.
- A final calculation shows that this disjunction is classically equivalent to $(p \wedge q) \vee (p \wedge \neg r)$, i.e. to $p \wedge (q \vee \neg r)$; so finally $C_K(A)$ is just $Cn(p \wedge (q \vee \neg r))$.
- Thus, neither q nor r is in $C_K(A)$. On the other hand, $C_K(A)$ is larger than classical consequence $Cn(A) = Cn(p)$, as it also contains $(q \vee \neg r)$.

Unlike their pivotal counterparts, default-assumption consequence operations/relations are nonmonotonic. That is, we may have $A \sim_K x$ but not $A \cup B \sim_K x$ where A, B are sets of propositions. Likewise, we may have $a \sim_K x$ without $a \wedge b \sim_K x$ where a, b are individual propositions. Since monotony can fail, default-assumption consequence operations are not in general closure operations.

To illustrate the failure of monotony, let $K = \{p \rightarrow q, q \rightarrow r\}$ where p, q, r are distinct elementary letters of the language and \rightarrow is the truth-functional (alias material) conditional connective. Then $p \sim_K r$ since the premise p is consistent with the whole of K and clearly $\{p\} \cup K \vdash r$. But $\{p, \neg q\} \not\sim_K r$, for the premise set $\{p, \neg q\}$ is no longer consistent with the whole of K . There is a unique maximal subset $K' \subseteq K$ that is consistent with $\{p, \neg q\}$, namely the singleton $K' = \{q \rightarrow r\}$; and clearly $\{p, \neg q\} \cup K' \not\vdash r$ – witness the valuation v with $v(p) = 1$ and $v(q) = v(r) = 0$. In general terms, by passing from premise p to premises p, q we *gained a premise*, but because of the consistency requirement we *lost a background assumption*.

Evidently, the same example with $p \wedge \neg q$ in place of $\{p, \neg q\}$ for the enlarged premise set illustrates the failure of singleton conjunctive monotony.

On the positive side, however, default-assumption consequence relations are supraclassical, as is immediate from the definition. They also satisfy cumulative transitivity and disjunction in the premises, although these take a bit more effort to verify.

There is a further important property of default-assumption consequence relations. Although they may fail monotony, they always satisfy a weakened version of it called *cautious monotony* (CM). This is a weakened or restricted form of monotony; it is implied by monotony but not conversely. In a singleton form, cautious monotony says: whenever $A \sim_K x$ and $A \sim_K y$ then $A \cup \{x\} \sim_K y$. More generally: whenever $A \sim_K x$ for all $x \in B$ and $A \sim_K y$ then $A \cup B \sim_K y$. In the succinct notation of operations: whenever $A \subseteq B \subseteq C_K(A)$ then $C_K(A) \subseteq C_K(B)$.

On the other hand, default-assumption consequence can fail compactness in quite a radical way. We illustrate this with an example of David Gabelaia (personal communication). Put $K = K_1 \cup K_2$ where $K_1 = \{p_i \wedge q : i < \omega\}$ and $K_2 = \{\neg p_i \wedge \neg q : i < \omega\}$. Put $A = \{p_i : i < \omega\}$. Then on the one hand $A \sim_K q$, while on the other hand $B \not\sim_K q$ when B is any proper subset of A (finite or infinite).

- The positive side $A \sim_K q$ holds because $K_1 \cup A \vdash q$ and K_1 is clearly the unique maximal A -consistent subset of K .
- The negative side $B \not\sim_K q$ can be checked as follows. Consider the set $K^* = \{\neg p_n \wedge \neg q : p_n \notin B\}$. Since $B \subset A$ there is an n with $p_n \notin B$ and so $\neg p_n \wedge \neg q \in K^*$. Also K^* is consistent with B and indeed $K^* \cup B \not\vdash q$ – witness the valuation v with $v(p_n) = 1$ iff $p_n \in B$ and $v(q) = 0$. It remains only to check that in fact, K^* is a maximal B -consistent subset of K . Let $K^* \subset J \subseteq K$. Then there is some i with either $p_i \wedge q \in J$, or $\neg p_i \wedge \neg q \in J$ and $p_i \in B$. In the former case, since J also contains some $\neg p_n \wedge \neg q$ it is inconsistent (look at q) and so inconsistent with B . And in the latter case, J is again inconsistent with B (look at p_i).

The failure here is radical in that there is no proper subset B of A , whether finite or infinite, with $B \sim_K q$. In other words, default-assumption consequence also fails the following weakened form of compactness, which we might call the *redundancy property*: whenever $A \sim x$ and A is infinite, then there is a proper subset $B \subset A$ such that $B \sim x$.

On a more general level, we note some interesting relations between classical logic, pivotal-assumption consequence, and default-assumption consequence.

- We constructed default-assumption consequences C_K by adding a consistency constraint to the definition of the monotonic pivotal-assumption consequences Cn_K . These in turn were obtained by adding a set of background assumptions to the definition of classical consequence. The *order of construction*, or conceptual order, is thus: Cn , then Cn_K , then C_K . The monotonic supraclassical notion can thus be seen as a natural halfway house or stepping-stone between the classical and nonmonotonic ones.
- However, this order of construction is not the same as the *order of inclusion* between the three operations. We have $Cn \leq C_K \leq Cn_K$, for the simple reason that by the monotony of classical consequence, $Cn(A) \subseteq Cn(K' \cup A) \subseteq Cn(K \cup A)$ whenever $K' \subseteq K$. In other words, as João Marcos has expressed it (personal communication), we can think of classical consequence Cn as giving a lower bound on all of our operations C_K , and for each choice of K the pivotal-assumption consequence Cn_K gives an upper bound on C_K . In other words, the non-monotonic default-assumption consequence operation C_K is interpolated between classical and pivotal-assumption consequence.

We will find this pattern of contrast between order of construction and order of inclusion repeated when we come to other ways of generating nonmonotonic inference operations.

- In the case that the current premise set A is consistent with the set K of background assumptions, the definitions clearly tell us that $C_K(A) = Cn_K(A) = Cn(K \cup A)$, i.e. the two supraclassical operations get the same value for argument A . From the point of view of pivotal-assumption consequence this case is the principal one. But from the point of view of default-assumption consequence it is a limiting case, for all the interest of the default operation arises when A is inconsistent with K .

A Dilemma

Despite its naturalness, the notion of default-assumption consequence faces a pragmatic dilemma. The dilemma arises when we ask the question: what kinds of assumption set K may usefully be used in generating operations?

- On the one hand, when K is *not* closed under classical consequence, i.e. when $K \neq Cn(K)$, then the identity of $C_K(A)$ can be sensitive to the manner of formulation of the elements of K ; in other words, it is syntax-dependent.
- On the other hand, when K is closed under classical consequence, i.e. when $K = Cn(K)$, then the operation $C_K(A)$ becomes devoid of interest, because in its principal case it collapses back into classical consequence.

We state these results more formally, prove them, and discuss their significance. First, the situation when it is allowed that $K \neq Cn(K)$.

Observation 2.2-1

Classically equivalent background assumption sets K, K' can give rise to quite different consequence operations C_K and $C_{K'}$.

Proof

We give a simple example. Put $K = \{p \rightarrow (q \wedge r), r \rightarrow \neg p\}$ where \rightarrow is material implication, while $K' = \{p \rightarrow q, p \rightarrow r, r \rightarrow \neg p\}$. These are classically equivalent, indeed quite trivially so, by the properties of conjunction. Now put $A = \{p\}$. Then A is inconsistent with K , and the subsets of K maximally consistent with A are its two singletons $\{p \rightarrow q \wedge r\}$ and $\{r \rightarrow \neg p\}$. Clearly $q \notin Cn(p, r \rightarrow \neg p)$ so $q \notin C_K(A)$. In contrast, the subsets of K' maximally consistent with p are its two pairs $\{p \rightarrow q, p \rightarrow r\}$ and $\{p \rightarrow q, r \rightarrow \neg p\}$. The remaining pair $\{p \rightarrow r, r \rightarrow \neg p\}$ is evidently inconsistent with p . Thus $q \in Cn(K' \cup \{p\})$ for both of these values of K' and so $q \in C_{K'}(A)$.

Roughly speaking, q gets lost when we process K but not when we process the trivially equivalent set K' , because K is more ‘chunky’ than K' and so forces us to discard more to ensure consistency with the current premises. Such syntax-dependence is often regarded as counter-intuitive and undesirable: surely the consequences of A given a background set K of assumptions should depend only on the content of K , not on the way in which it is formulated.

We now consider the other horn of the dilemma: the situation when K is closed under classical consequence, i.e. when $K = Cn(K)$. The unpleasant outcome may be stated as follows. Here, as always, ‘consistency’ means classical consistency.

Theorem 2.2–2

When $K = Cn(K)$, then $C_K(A) = Cn(A)$ whenever A is inconsistent with K .

Of course, as already noted, when A is consistent with K , then we have $C_K(A) = Cn(K \cup A)$ whether or not K is closed under classical consequence. There is nothing untoward about that. The theorem considers the principal case, that A is inconsistent with K and states that if K is closed under classical consequence then the nonmonotonic operation loses almost all of its power. The background assumption set K ceases to contribute anything, and we get nothing more than the classical consequences of A .

Proof of Theorem 2.2–2

Suppose $K = Cn(K)$ and A is inconsistent with K . We want to show that $C_K(A) = Cn(A)$. The right-in-left inclusion is just supraclassicality, so we need only show that $C_K(A) \subseteq Cn(A)$. Suppose $x \notin Cn(A)$; we need to show $x \notin C_K(A)$. By the definition of C_K it suffices to find a maximal A -consistent subset K' of K such that $x \notin Cn(K' \cup A)$.

The argument uses the compactness of classical consequence and Zorn's Lemma to construct the maximal set. The key step is where we appeal to the hypothesis that $K = Cn(K)$ towards the end. Our presentation hops back and forth as convenient between operational and relational notations.

From the supposition of the case we know by compactness that there is a finite subset $A_0 \subseteq A$ that is inconsistent with K . Let a be the conjunction of all its finitely many elements. Then a is also inconsistent with K . To complete the proof we find a maximal A -consistent subset K' of K such that $\neg a \vee \neg x \in K'$. This will suffice, because given $\neg a \vee \neg x \in K'$ we have $K' \cup A \vdash a \wedge (\neg a \vee \neg x) \vdash \neg x$ so since $K' \cup A$ is consistent we have $x \notin Cn(K' \cup A)$ as was desired.

To find such a K' , first note that $\neg a \vee \neg x$ is consistent with A ; for otherwise $A \vdash \neg(\neg a \vee \neg x) \vdash a \wedge x \vdash x$ contrary to the hypothesis $x \notin Cn(A)$. Next we note that $\neg a \vee \neg x \in K$; for we know that a is inconsistent with K , so $K \vdash \neg a \vdash \neg a \vee \neg x$, and thus by the crucial hypothesis that $K = Cn(K)$ we have $\neg a \vee \neg x \in K$. Putting these two points together, we have established that the singleton subset $\{\neg a \vee \neg x\}$ of K is consistent with A . Since classical consequence is compact, we may now apply Zorn's Lemma to conclude that there is a maximal A -consistent subset K' of K such that $\{\neg a \vee \neg x\} \subseteq K'$, which is what we needed to show.

Responses to the Dilemma

What are we to make of the dilemma? Is default-assumption consequence irrevocably flawed, or is there a way out?

Consider first the case that K is not closed under classical consequence. One reaction is to say that in this case the syntactical form of the assumptions in K *should indeed* have an impact on

the conclusions that can be drawn, even though the syntactical form of the explicit premises in A should not have any such influence. This is a brave reaction, but rather difficult to swallow.

Another reaction is to reduce the amount of syntactic dependence. For example, one might allow as elements of the background assumption set only formulae that are expressed in a certain canonical way. For example, we might accept in K only formulae that are disjunctions of *literals*, i.e. disjunctions of elementary letters and their negations. In general, a single formula (e.g. $p \wedge q$) will be expressed by a set containing several such disjunctions (e.g. $p \vee q$, $p \vee \neg q$, $\neg p \vee q$), which makes it less ‘chunky’. From classical logic we know that every formula may be expressed by a set of such disjunctions (remember conjunctive normal forms, and separate the conjunctions into separate items); so this restriction does not reduce expressive power. This regimentation will have the effect of reducing the level of syntax-dependence. Indeed, in the case of a finitely generated language, if one accepts in K only disjunctions that use all the elementary letters of the language then the problem of syntax-dependence disappears. That is, if K, K' are classically equivalent sets, all of whose elements are of this canonical form, then $C_K(A) = C_{K'}(A)$.

In the case that $K = Cn(K)$, a quite different kind of response is needed. We need to modify the definition of default-assumption consequence in such a way as to avoid the collapse into classical consequence. Indeed, there are many ways in which this can be done, and they will be the focus of the following section. Before leaving this section, however, we make some general remarks on terminology and on the notion of maximality.

Terminology: Consequence or Inference?

We are using the term ‘consequence operation’ to cover monotonic operations like Cn and Cn_K and also nonmonotonic ones like C_K . In this, we are following the usage of many investigators, including Kraus, Lehmann and Magidor (1990). But we are running against others, including Lindström (1991), Makinson (1994), Rott (2001) and Bochman (2001). They reserve the term ‘consequence’ operation for the monotonic ones and speak of ‘inference’ operations for the nonmonotonic ones.

Such a terminological separation between ‘consequence’ and ‘inference’ is attractive in principle. But in practice it turns out to be rather cumbersome. This is partly because the verbal contrast is not rooted in ordinary usage, and partly because we sometimes wish to talk about the operations generated in a certain manner *before* finding out whether they are monotonic, or about families of operations *some* of which are monotonic and others not. In this presentation, therefore, we will use the term ‘consequence’ for *both the monotonic and nonmonotonic cases*. Nevertheless, as already remarked, we mark the difference in our notation by using \vdash , Cn (with appropriate subscripts) in the monotonic case and $\vdash\sim$, C when they are not known to be monotonic.

A Warning about Maxtalk

For those who are not accustomed to working mathematically with the notion of maximality, a word of warning may be useful. To say of a subset K' of K that it is maximally (such-and-such) and is also so-and-so is not at all the same thing as saying that it is maximally (such-and-such and so-and-so). The two must be distinguished carefully.

The former does imply the latter. For suppose K' is maximally a ϕ , and is also a ψ . Consider any $K'' \supseteq K'$ that is both a ϕ and a ψ . Then it will be a ϕ , so by the supposition it will be identical with K' .

But the latter does not in general imply the former. For suppose K' is maximally ϕ -and- ψ . Then there is no $K'' \supset K'$ that is *both* a ϕ and a ψ , but there may well be a $K'' \supset K'$ that is a ϕ without being a ψ ; so that K' is not maximally a ϕ .

The point could be made in more set-theoretical language, writing $max_{\subseteq}(X)$ for the set of all items that are maximal, under the relation of set inclusion, among the elements of X , but the point should be clear enough without such formality. The important thing is to know what one intends to say in any context, and to express it unambiguously. This can require careful phrasing; some bracketing or hyphening can also help.

To illustrate the difference, suppose that K' is maximal among the proper subsets of K , and is consistent with A . This implies that it is maximal among the proper subsets of K that are consistent with A . But it is easy to construct an example where K' is maximal among the proper subsets of K that are consistent with A , but is not maximal among the proper subsets of K . For instance, put $K = \{a, b, (a \vee b) \rightarrow c\}$ and $A = \{\neg c\}$. Then K has two maximal (proper subsets consistent with A), namely $K_1 = \{a, b\}$ and $K_2 = \{(a \vee b) \rightarrow c\}$. Of these, K_1 is indeed a maximal (proper subset of K), but K_2 is not.

Nevertheless, in the case that $K = Cn(K)$ we have a special situation. Let A be a premise set inconsistent with K . It is easy to verify that since K is closed under classical consequence, so too is each maximal A -consistent subset of K . And we have the following important fact, closely related to Theorem 2.2–2 and also to Theorem 2.3–1 in the next section. Roughly speaking, it says that examples like that of the preceding paragraph are impossible when $K = Cn(K)$.

Observation 2.2–3

Suppose $K = Cn(K)$ and let A be a set of formulae that is inconsistent with K . Then the following three conditions are equivalent, for any $K' \subseteq K$:

- *K' is maximal among the subsets of K that are consistent with A*
- *K' is maximal among the classically closed proper subsets of K that are consistent with A*
- *K' is maximal among the classically closed proper subsets of K , and also K' is consistent with A*

Historical Remarks

Default-assumption consequence has a rather long and complicated history. Veltman (1976), (1985) and Kratzer (1981) used closely related constructions in the logic of counterfactual conditionals. Alchourrón and Makinson (1982) defined the operation in a study of the revision of a belief set K by a new belief a , i.e. the introduction of new information into a belief set while preserving its consistency. The collapse into classical consequence when $K = Cn(K)$ was noted (in various forms, for singleton premise sets), by all three authors. The family of all subsets of K maximally consistent with A was also studied by Poole (1988), but as part of a formal account of *abduction*, i.e. the formation of hypotheses to explain data.

2.3. Specializations, Variants and Generalizations

We open this section by describing a special case of the concept of default-assumption consequence that is of both historical and practical importance: the *closed world assumption* (CWA). We will then sketch some of the many variations and generalizations that may be found in the literature – about a dozen in total, depending on how you count them – beginning with those that most closely resemble our simple paradigm in the preceding section, and then passing on to others more distantly related.

A Special Case: The Closed World Assumption

Default-assumption consequence as we have defined it in the preceding section, is an abstraction upon a much more specific notion, known as *inference using the closed world assumption* (briefly, CWA), introduced by Reiter (1978) – one of the earliest formal studies of nonmonotonic reasoning in a qualitative context.

To explain this kind of inference, we need the notion of a Horn formula of the language of classical logic. In section 1.2 we already introduced the notion of a Horn rule, as a special kind of rule formulated in the metalanguage in which we talk about inference relations. Now we need the corresponding concept for formulae in the object language, to describe a special class of Boolean formulae. A *Horn formula* is simply one of the form $(p_1 \wedge \dots \wedge p_n) \rightarrow q$, where p_1, \dots, p_n, q are all elementary letters and $n \geq 0$. Equivalently, using only the primitive connectives \neg, \wedge, \vee : it is a formulae of the kind $\neg p_1 \vee \dots \vee \neg p_n \vee q$, with the same ingredients. That is, a disjunction of literals, of which just one is positive.

Suppose that our current premise set A consists solely of Horn formulae. Suppose also that our set K of background assumptions consists of the negations $\neg p$ of all the elementary letters of the language. In other words we wish to assume, as far as is compatible with our current premises, that all elementary letters are false. The set K is evidently not closed under classical consequence, but the problem of syntax-dependence discussed in the preceding section does not arise as only this one choice of K is allowed: it is held fixed in all applications.

In this special case it is not difficult to show that there is a *unique* maximal subset K' of K consistent with A . Consequently the definition of default-assumption consequence, which was $C_K(A) = \cap \{Cn(K' \cup A) : K' \subseteq K \text{ and } K' \text{ maxconsistent with } A\}$, may be simplified to $C_K(A) = Cn(K' \cup A)$ where K' is the unique maximal subset of K consistent with A .

Exactly the same situation holds if, instead of elementary letters p_i , our language uses elementary formulae of the form $P_{\underline{t}}$, where P is a predicate or relation symbol and $\underline{t} = t_1, \dots, t_n$ is an n -tuple of individual names, with the remaining formulae of the language generated from them by Boolean operations only (no individual variables, no quantifiers). Suppose that K consists solely of the negations $\neg P_{\underline{t}}$ of all elementary formulae, and our premise set A consists of formulae of the form $(P_{1\underline{t}_1} \wedge \dots \wedge P_{n\underline{t}_n}) \rightarrow Q_{\underline{s}}$, where $P_{1\underline{t}_1}, \dots, P_{n\underline{t}_n}, Q_{\underline{s}}$ are all elementary formulae. Then the definition of default-assumption consequence simplifies to $C_K(A) = Cn(K' \cup A)$ where K' is the unique maximal subset of K consistent with A .

The motivation for this construction arises from the fact that we may be dealing with a database about a certain domain. The database may reasonably be thought of as a set A of Horn formulae $(P_1 t_1 \wedge \dots \wedge P_n t_n) \rightarrow Q s$ (with $n \geq 0$) where the individual constants in the tuples t_1, \dots, t_n, s designate specific elements of the domain and the symbols P_1, \dots, P_n, Q designate specific predicates and relations over that domain. Closed world inference amounts to assuming, within the limits of consistency with the database, that the predicates and relations P_1, \dots, P_n, Q are *as small as possible*.

In section 3 we will see how this particular kind of default-assumption consequence may also be formulated in semantic terms, and generalized to become a form of inference known as circumscription.

Variations and Generalizations: Three Broad Kinds

There are many variations and generalizations of default-assumption consequence. In all of them the notion of consistency continues to play an important, and usually quite visible, role. To bring some order into the affair, we may conveniently sort them into three main groups:

1. The *partial meet operations*, which intersect the output of only *some* of the maximal subsets of K consistent with the current premises.
2. The *submaximal operations*, which intersect subsets of K that are consistent with the current premises, but *not necessarily maximal*.
3. The *intersection-free operations*, which specify a unique output without the need for any kind of intersection.

They all have the advantage, over the simple paradigm presented in the preceding section, of allowing the set K of background assumptions to be closed under classical consequence without thereby collapsing the nonmonotonic operation into the classical one in the manner described in Theorem 2.2–2. We illustrate each of these kinds with examples from the literature.

Partial Meet Approaches

These approaches continue to work with maximal subsets of the background assumption set K that are consistent with the current premise set A , but restrict attention to only some of them. In other words, they all put $C(A)$ to be the intersection of sets $Cn(K' \cup A)$ for certain $K' \subseteq K$ that are maxiconsistent with A , but they differ in how they pick them out.

We will consider four of them, in roughly increasing order of generality. They are: *screened consequence* (treating some elements of K as inviolable); *layered consequence* (treating some elements of K as preferred to others); *relational partial meet consequence* (taking some A -maxiconsistent subsets of K as preferable to others); and consequence via *selection functions* (the most general form covering, as special cases, the other three as well as the paradigm default-assumption consequence defined in the preceding section).

Screened Consequence

We might wish to regard some of our background assumptions as unquestionable, no matter what premises come along. To express this in mathematical terms, we fix our set K of background assumptions, and also fix a distinguished subset $K_0 \subseteq K$ of *fully protected* assumptions. We then define an operation $C_{KK_0}(A)$ of *screened consequence* by the equation $C_{KK_0}(A) = \bigcap \{Cn(K' \cup A) : K_0 \subseteq K' \subseteq K \text{ and } K' \text{ maxiconsistent with } A\}$. In this way, the protected assumptions in K_0 are maintained in the face of anything, while the remainder may be dropped when contradiction so requires.

This kind of operation was introduced by Makinson (1997) in the context of the logic of belief contraction and revision. In that context, it also carried additional twist for the limiting case that the set K_0 of protected background assumptions is inconsistent with the current premise set A . Clearly, in that case there are *no* maximal sets including K_0 consistent with A , so *every* formula in the language becomes a consequence of A under the definition above. This makes good sense when we are thinking of inference; but less so for revision, where we might refuse to revise a belief set K when the incoming information is in conflict with its protected core. For this reason, in the belief revision context, in Makinson (1997) the output in the limiting case was taken to be just K itself.

We note that the inclusion sequence $Cn \leq C_K \leq Cn_K$ that was observed for default-assumption consequence becomes for screened consequence $Cn \leq Cn_{K_0} \leq C_{KK_0} \leq Cn_K$. In other words, the pivotal-assumption operation Cn_K continues to serve as upper bound for the nonmonotonic screened consequence operation, while the lower bound C_{KK_0} moves up from classical Cn to the pivotal-assumption operation Cn_{K_0} .

As the lower bound is no longer classical Cn , we no longer get the collapsing phenomenon that we noted in the previous section in the case that $K = Cn(K)$ and is inconsistent with the current premise set A . But we do get a ‘lesser collapse’. Under those two hypotheses, we have $C_{KK_0}(A) = Cn_{K_0}(A)$, which is nearly as bad: the only conclusions that may be drawn are those that follow from the current premises taken together with the inviolable background assumptions. The proof is a straightforward re-run of the proof of Theorem 2.2-2. Thus we may say that this construction, without further refinements, is still not suitable for application to background assumption sets closed under classical consequence.

Layered Consequence

Another idea is to segment the set K of background assumptions into levels, K_1, \dots, K_n , and then maximize in a stepwise manner, giving early levels preference over later ones. In other words, given a premise set A we first take the maximal subsets of K_1 consistent with A , then extend each of them into a maximal subset of $K_1 \cup K_2$ consistent with A , and so on.

Formally, a *preferred subset of K* with respect to A (modulo the layering) is a set $J_1 \cup \dots \cup J_n \subseteq K$ such that for each $i \leq n$, $J_1 \cup \dots \cup J_i$ is a maximal A -consistent subset of $K_1 \cup \dots \cup K_i$. Clearly, every preferred subset of K is in fact a maximal A -consistent subset of K , but not conversely. We are thus restricting attention to a subfamily of the latter.

This construction was introduced by Brewka (1989), where it was described as ‘consequence via level default theories’. A variant, in which maximal cardinality replaces maximality under the subset relation, was introduced by Benferhat *et al* (1993).

This approach is designed for background assumption sets K that are not closed under classical consequence. When they are so closed, and the levels K_i are also closed under Cn , then we have a kind of ‘mini-collapse’. For in that case, when K is inconsistent with the current premise set A , the conclusions obtainable will be just the classical consequences of A taken together with the highest K_i with which A is consistent (or with the empty set if none of them are). The proof is again essentially the same as that for Theorem 2.2-2.

Relational Partial Meet Consequence

An alternative procedure for selecting a subfamily of the maximal A -consistent subsets of K is to fix a relation $<$ between subsets K' of K , and pick out from the maximal A -consistent subsets of K those that are minimal under that relation. Placing constraints (such as rankedness) on the relation $<$ can give rise to greater regularity of behaviour in the induced consequence relation.

When applied to sets closed under classical consequence, this approach avoids the collapse phenomenon of the ‘full meet’ construction of the previous section and its lesser forms noted above for screened and layered sequence.

The construction is essentially the same as one effected in the AGM account of belief change, as developed by Alchourrón, Gärdenfors and Makinson (1985), and there called ‘partial meet contraction and revision’. It is also closely related to a semantic approach to nonmonotonic reasoning called preferential inference, which we examine in detail in the next section.

Consequence via Selection Functions

A more general procedure is to use a *selection function*. This associates with each family \mathcal{K} of subsets of K a subfamily $\delta(\mathcal{K}) \subseteq \mathcal{K}$ of ‘distinguished’ elements of \mathcal{K} . The selection function may be taken as arbitrary, or it may be constrained in various ways. It is usually required that $\delta(\mathcal{K})$ is non-empty whenever \mathcal{K} is non-empty.

Using this concept, we define $C_{\mathcal{K}}(A) = \cap\{Cn(K' \cup A) : K' \in \delta(\mathcal{K}_A)\}$, where \mathcal{K}_A is the family of all subsets $K' \subseteq K$ maxiconsistent with A , and δ is a selection function. Like K , δ is fixed independently of the premise sets A . Evidently, we get different consequence operations for different choices of δ , as well as for different choices of K .

In the limiting case that the selection function makes no real choice, always putting $\delta(\mathcal{K}_A) = \mathcal{K}_A$ (in other words, when it is the identity function) we are back with the basic notion of default-assumption consequence defined in the previous section. We have seen how this leads to counterintuitive results in the case that K is already closed under classical consequence, i.e. $K = Cn(K)$.

At the opposite end of the spectrum is another limiting case, where the selection function always gives us a singleton, i.e. $\delta(\mathcal{K}_A) = \{K'\}$ for some $K' \in \mathcal{K}_A$ whenever the latter is non-empty. Such a selection function is usually called a *choice function*, and the consequence operation $C_{\mathcal{K}}$ that it engenders is called a *maxichoice* default assumption consequence. In this case, for every premise set A inconsistent with K , there is a subset $K' \subseteq K$ maxiconsistent with A such that $C_{\mathcal{K}}(A) = Cn(K' \cup A)$.

This limiting case also leads to a counterintuitive result for background assumption sets K closed under classical consequence. When $K = Cn(K)$, the maxichoice consequence operations always give consequence-sets that are *complete*, in the sense that they contain at least one of $x, \neg x$ for every formula x of the language. This is a strange and undesirable inflation. The result was first proven by Alchourrón and Makinson (1982), in the context of the logic of belief change.

Theorem 2.3–1

Let $K = Cn(K)$ and let A be a premise set A inconsistent with K . Then for any maximal A -consistent subset K' of K , $Cn(K' \cup A)$ is complete.

Proof

Let $K = Cn(K)$ and let A be any premise set inconsistent with K . Let K' be any maximal A -consistent subset of K . Let x be any formula. We want to show that either $x \in Cn(K' \cup A)$ or $\neg x \in Cn(K' \cup A)$. The argument has some elements in common with that for Theorem 2.2–2.

Since A is inconsistent with K , the compactness of classical logic tells us that there is a finite subset $A_0 \subseteq A$ that is inconsistent with K . Let a be the conjunction of the finitely many elements of A_0 . Then $a \in Cn(A)$. To complete the proof it will suffice to show that either $\neg a \vee x \in K'$ or $\neg a \vee \neg x \in K'$, for then $K' \cup A \vdash a \wedge (\neg a \vee x) \vdash x$ or $K' \cup A \vdash a \wedge (\neg a \vee \neg x) \vdash \neg x$.

Suppose for *reductio ad absurdum* that both $\neg a \vee x \notin K'$ and $\neg a \vee \neg x \notin K'$. Since A_0 is inconsistent with K we know that a is also inconsistent with K , so $\neg a \in Cn(K)$, so both of $\neg a \vee x, \neg a \vee \neg x \in Cn(K) = K$ using the (crucial) supposition that K is closed under classical consequence. On the other hand, by supposition K' is maximal among the subsets of K consistent with A . So $K' \cup \{\neg a \vee x\}$ and $K' \cup \{\neg a \vee \neg x\}$ are both inconsistent with A . By the monotony of classical consequence it follows that $K' \cup \{x\}$ and $K' \cup \{\neg x\}$ are both inconsistent with A . In turn it follows that K' is inconsistent with A , giving us the desired contradiction.

The moral of this story is that when our background assumption sets are closed under classical consequence, the selection function should not be extremal. We get one undesirable result when the selection function is just the identity function: this takes us back to the basic form of default assumption consequence – which as we saw in Theorem 2.2–2 collapses into classical logic in the case that $K = Cn(K)$. We get another undesirable result when we select a single maximal subset: this inflates the size of consequence sets, making them complete. The selection functions will have to be somewhere in the middle between these extremes, Goldilocks style.

The approach via selection functions is very general, and covers the three preceding ones as special cases. The connection with relational partial meet consequence has received particular attention. Given any relation $<$ between subsets K' of K , the function taking each K' to its $<$ -minimal elements is clearly a selection function. Conversely, if sufficiently powerful constraints are imposed on a selection function δ on subsets of K , then δ can be represented in this way, i.e. we can find a relation $<$ such that $\delta(K')$ consists of the minimal elements K' under $<$. The relation thus uncovered is sometimes called a ‘revealed preference’, a term borrowed from the theory of choice between goods, in economics, where the interplay between preference relations and selection functions was first studied.

There is a systematic study of the relationships between constraints on selection functions and the possibility of representing them by preference relations, in Rott (1993), (2001).

Sub-maximal Approaches

All of the variants considered above worked with *maximal* premise-consistent subsets of the background assumption set K . They differed only in which of those subsets they took into consideration. The following generalizations also allow us to take into account some of the less-than-maximal subsets of K .

Background Constraints

One idea, already considered by Poole (1988), is to add to the set K of background assumptions another set J of ‘constraints’. In the best of circumstances, it is consistent with the explicit premises of an inference. But it does not participate positively in the generation of conclusions. And when consistency fails, J is not diminished; in effect, it further reduces the background assumptions from K that may legitimately be used alongside A .

To be precise, the definition of default-assumption consequence is re-run as follows. Let $K, J \subseteq L$ be sets of formulae, which will play the role of background assumptions and constraints respectively. We define the relation \sim_{KJ} of *default-assumption consequence with constraints* by putting $A \sim_{KJ} x$ iff $K' \cup A \vdash x$ for every subset $K' \subseteq K$ with K' maximal among the subsets of K that are consistent with $A \cup J$. Writing C_{KJ} for the corresponding operation, this says $C_{KJ}(A) = \cap \{Cn(K' \cup A) : K' \subseteq K \text{ and } K' \text{ maxiconsistent with } A \cup J\}$.

If there are no such subsets K' , this definition puts $C_{KJ}(A) = L$. Clearly, this happens when J itself is inconsistent with A , for then K' is inconsistent with $A \cup J$ even for $K' = \emptyset$. It is also possible to show the converse, using the compactness of classical consequence. Thus $C_{KJ}(A) = L$ iff J is inconsistent with A .

As one might expect, this kind of operation is rather less regularly behaved than the one without constraints. For example, it fails disjunction in the premises. However, it does satisfy both cumulative transitivity and cautious monotony.

What, in this construction, is the difference between a background assumption and a constraint? Both are fixed independently of the premise sets. But constraints are more ‘protected’ than background assumptions. In case of conflict with a premise set, a background assumption may be withdrawn, but constraints are never dropped. Indeed, the presence of the constraints will tend to create inconsistencies where none existed between premises and background assumptions alone, and so force the retraction of more background assumptions than would otherwise be the case. Constraints thus play a powerful negative role.

On the other hand, they have no positive role at all: they are never conjoined with the current premises to obtain new conclusions. Thus the default consequence relation generated by a set of background assumptions plus a set of constraints is always weaker than (i.e. is a sub-relation of) the one determined by the background assumptions alone. If we give the constraints a positive role as well as a negative role, by putting $C_{KJ}(A) = \cap \{Cn(K' \cup J \cup A) : K' \subseteq K \text{ and } K' \text{ maxiconsistent with } A \cup J\}$, then the operation can be seen as just screened

consequence, as defined above, with $K' \cup J$ now serving as the utilized subset of the background assumption set $K \cup J$, and J as the fully protected ones among them.

Maximally Informative Subsets

There is another way in which we might wish to relax the requirement of maximality. We might wish to consider not the *biggest* subsets, but the most *informative* subsets, of K that are consistent with the premises. It is natural to assume that all of the maximally large subsets will also be maximally informative, but some among the non-maximal subsets may be just as informative as their maximal supersets. In this case, it is suggested, they should also be counted in the family scheduled for intersection, thus in turn diminishing the set obtained by that intersection.

Isaac Levi (1996) and elsewhere has argued for this approach, formulated in the context of belief revision but with an evident analogue for inference. Rott and Pagnucco (1999) have studied the formal properties of such consequence relations.

This construction may still be seen as considering the subsets of K that are maximally consistent with the current premises, but with maximality now understood with respect to a relation $<$ of degree of informativeness rather than with respect to set inclusion.

In Levi's account $<$ is a subrelation of \subset . This implies that every \subset -maximal A -consistent subset of K is a $<$ -maximal A -consistent subset of K , with the effect that in the case $K = Cn(K)$, intersection of all of the latter leads again to collapse into classical consequence. To avoid this, we would have to combine the idea with one of the partial meet approaches already considered, or else generalize by allowing the relation $<$ to be more loosely tied to set inclusion. Indeed, such a generalization would be appropriate if we reinterpret $<$ as representing the 'relative epistemic value' of subsets of K , calculated in some way that combines several competing factors, not all positively correlated with size. For example, it might balance size, informativeness, simplicity of formulation, simplicity of application, richness of connections, etc. Then, not all of the maximal subsets of K that are consistent with the premises need be of maximal value, nor conversely.

Best Epistemic States

The 'ultimate' level of abstraction among 'partial meet' approaches has been attained by Bochman (2001). Two ideas underlie his construction. One is to compare subsets of K by an *arbitrary* relation, abstracting entirely from set inclusion or relative epistemic value. The other is to allow some of the subsets of K to be omitted from consideration; in other words, as well as fixing K , we also fix some family $\mathcal{K} \subseteq 2^K$ of its subsets. We think of these as representing the range of all entertainable states of belief within the limits of K , and Bochman calls them *epistemic states*. We define $A \mid \sim_{\mathcal{K}} x$ to hold iff $K' \cup A \mid - x$ for every epistemic state (i.e. element of \mathcal{K}) that is maximally A -consistent under the relation $<$.

Strictly speaking, Bochman does place some constraints on this construction. In particular, the elements K' of \mathcal{K} are required to be closed under classical consequence, and the relation $<$ is required to be transitive and irreflexive. Moreover, for most purposes, a constraint like stoppering, alias smoothness (to be explained in the next section) is placed on $<$. But the

definition remains very abstract. Bochman notes that once one fixes the family $\mathcal{K} \subseteq 2^K$ we have no more need for K itself, and may dispense with it.

We put the word 'ultimate' in inverted commas when introducing this approach, because in fact there is no such thing as a most general level of abstraction in these matters. For example, as Bochman (2001) notes, the consequence relations defined by Makinson (1989) are in some respects more abstract than his, but in other respects less so.

Approaches that do not need Intersection

All of the variants considered so far consider certain subsets K' of K that are consistent with the current premise set A and, following the 'sceptical policy', intersect their respective outcomes $Cn(K' \cup A)$. We now consider certain approaches under which we can get a unique outcome without intersection.

Epistemic Chains

Suppose we take the very abstract construction of Bochman but, as in an Escher engraving, descend the stair of generalization in a new direction. Let us retain Bochman's idea of a family $\mathcal{K} \subseteq 2^K$ of subsets of K , but return to plain set-inclusion as the relation for comparing its elements. Thus $A \sim_{\mathcal{K}} x$ may be defined to hold iff $K' \cup A \vdash x$ for every epistemic state K' that is consistent with A and maximally so among elements of \mathcal{K} .

This brings us closer to our paradigm notion of default-assumption consequence in section 2.2. The only difference is that we maximize over the $K' \in \mathcal{K}$ rather than over the $K' \subseteq K$. But this small change can have considerable effect. A maximally A -consistent subset K' of K may be absent from \mathcal{K} . Conversely, an A -consistent subset K' of K may be maximally so among elements of \mathcal{K} without being maximally so among the entire collection of subsets of K , for there may be another K'' consistent with A with $K' \subset K'' \subseteq K$ but $K'' \notin \mathcal{K}$.

Now let us particularize further. Suppose that the set \mathcal{K} of epistemic states contains the empty set as smallest element, and is a *chain under set inclusion* – that is, for all K', K'' in \mathcal{K} either $K' \subseteq K''$ or $K'' \subseteq K'$. Suppose further that the language is finite. Since the language is finite, there are only finitely many non-equivalent formulae modulo classical equivalence. Without loss of generality, we may thus take \mathcal{K} to be a set of formulae k_1, \dots, k_n , linearly ordered by classical consequence, with the weakest one a tautology. These are called *epistemic chains*.

In this context, the definition of $A \sim_{\mathcal{K}} x$ reduces to the following: $A \sim_{\mathcal{K}} x$ iff either A is inconsistent, or else A is consistent and $\{k_i\} \cup A \vdash x$ where k_i is the strongest formula in \mathcal{K} that is consistent with A . Equivalently, iff either A is inconsistent, or else A is consistent and $\{k_i\} \cup A \vdash x$ for *some* formula in \mathcal{K} that is consistent with A .

This kind of consequence relation was introduced and studied intensively by Michael Freund (1998). As he has shown, the definition using epistemic chains is (in the finite case) the syntactic counterpart of a semantic one (namely ranked preferential consequence) that we will discuss in the following section.

It is interesting to compare this construction with the ‘layered consequence’ of Brewka described earlier in this section. They both make use of a collection K_1, \dots, K_n of subsets of K . Freund requires this to be a chain under set inclusion, whereas Brewka requires it to form a partition of K , i.e. a collection of disjoint non-empty subsets whose union is K . However, on closer inspection it is apparent that this is largely a difference of presentation. More significant is the fact that Brewka looks for maximal A -consistent *subsets* of each K_i while Freund treats each K_i as *indivisible* and looks for the largest A -consistent one. When K is not closed under classical consequence, this difference typically changes the output. But when K and all the K_i are closed under classical C_n then, as we have seen, the K_i also become indivisible under the layered approach of Brewka, which in that case essentially reduces to Freund’s epistemic chain system.

Safe Consequence

An interesting feature of Freund’s construction is that it permits us to designate a special A -consistent subset of K without having to intersect maximals. The cost is that we must require the family of possible epistemic states to be linearly ordered under set inclusion. We now describe an approach that also obtains its output without intersection, but needs only an *acyclic* relation $<$ over K that is, in the sense that there are no propositions a_1, \dots, a_n in K with $a_1 < a_2 < \dots < a_n < a_1$.

The essential idea is to keep only those elements of K that cannot reasonably be ‘blamed’ for the inconsistency of K with A . Suppose we are given a fixed relation $<$ over the set K of background assumptions representing some notion of degree of vulnerability, so that $k < k'$ means that k is more vulnerable than k' . Let A be the current premise set. We say that a proposition $a \in K$ is *safe with respect to A* (modulo $<$) iff a is not a minimal element of any minimal subset K' of K that is inconsistent with A . Note carefully that there are *two dimensions of minimality* here. The first one is under the relation $<$ between elements of K , while the second one is under set-inclusion between subsets of K . It is not difficult to show that the set of all safe elements deserves its name:

Observation 2.3–2

Let K be a set of background assumptions, and $<$ an acyclic relation over K . Let A be a consistent set of premises. Then the set S_A of all safe elements of K with respect to A (modulo $<$) is consistent with A .

Proof

The proof is a very elegant one, reminiscent of diagonal arguments in set theory. Suppose for *reductio ad absurdum* that A is consistent but S_A is inconsistent with A . Then by the compactness of classical consequence, there is a finite $S_0 \subseteq S_A$ that is inconsistent with A . Since A is consistent, S_0 is not empty. Since S_0 is non-empty and finite, and the relation $<$ is acyclic, there must be at least one $<$ -minimal element s of S_0 . Thus by the construction, s is a $<$ -minimal element of a \subseteq -minimal A -inconsistent subset of K , and so by definition is not safe with respect to A . On the other hand, $s \in S_A$, and so by supposition it is safe with respect to A . This gives us a contradiction, completing the verification.

The relation of *safe consequence* is then defined by putting $A \sim x$ iff $A \cup S_A \vdash x$. This is evidently a kind of default-assumption consequence. From the set K of background assumptions it selects a special subset S_A of safe elements, whose identity depends in a motivated manner on the current premise set A (as well as on a background acyclic relation $<$ over K). It then joins these special background assumptions to the premises.

One attraction of safe consequence is that, like the others in this subsection, it does not require intersecting a family of output sets. A further attraction is the very weak requirement of acyclicity that is needed on the guiding relation $<$. Safe consequence has been studied in detail by Alchourrón and Makinson (1985), (1986), as a form of belief contraction and revision.

Comparative Expectation Inference

One may also elaborate the notion of a background assumption in a quite different way. Instead of representing the background assumptions as a *set*, one may say that to be a background assumption is at heart *a matter of degree*. Any consistent proposition of the language may to serve as an additional assumption, to a degree limited by its plausibility. Plausibility may be represented as a *relation* $<$ over all formulae of the language. Thus $z < y$ is read as saying that y is strictly more plausible (or: less unexpected) than z .

Given such a relation of plausibility, we can define an appropriate notion of consequence that allows sufficiently plausible propositions to be used as additional assumptions. One way of doing this, using a definition of Rott (1991), is to put $a \sim_{<} x$ iff $x \in Cn(\{a\} \cup \{y: \neg a < y\})$. In words: $a \sim_{<} x$ iff x is a classical consequence of premise a supplemented by all those propositions that are strictly more plausible than $\neg a$. In other words, when the premise is a , all propositions y with $\neg a < y$ are allowed to serve as background assumptions. This is called *comparative expectation inference*. If sufficient conditions are imposed on the relation $<$, it can be shown that the set $K_a = \{y: \neg a < y\}$ of all background assumptions (in the context of a) is classically consistent with a whenever a itself is consistent.

The construction has a number of attractions. One is that we do not need to intersect outputs. To the premise a we simply add the *unique* set $K_a = \{y: \neg a < y\}$ and take the closure under classical consequence. Another gratifying feature is that the nonmonotonic inference relation $\sim_{<}$ thus generated has very regular properties. It satisfies not only cumulative transitivity and cautious monotony but also a non-Horn condition of *rational monotony*: whenever $a \sim x$ and $a \sim \neg b$ then $a \wedge b \sim x$.

However, there is also a negative side. The requirements that we need to place on the relation $<$ are very demanding. In particular, to show that the additional assumptions are consistent with the premise, we need the radical requirement that $\neg a < y_1$ and $\neg a < y_2$ together imply $\neg a < y_1 \wedge y_2$. To make the induced consequence relation well behaved, we also need $<$ to be not only transitive and irreflexive but also *ranked* (alias *modular*): whenever $a < x$ and $y < x$ then $a < y$.

Comparative expectation inference may be defined in another way: put $a \sim_{<} x$ iff either $\neg a < \neg a \vee x$ or $\neg a \in Cn(\emptyset)$, where Cn is classical consequence. This definition is equivalent to the first one if sufficiently many constraints are placed on the relation $<$. With this definition, we no longer need any conditions on $<$ to show that the current premise is consistent with background assumptions – for we no longer use additional assumptions! But the economy is

an illusion. The same radical requirement on $<$ is still needed in order to ensure that $|\sim_<$ is reasonably well behaved. And conceptually, we lose the illuminating *gestalt* of default assumptions that is apparent under Rott's definition.

Comparative expectation inference is investigated in detail in Gärdenfors and Makinson (1994), with a brief account in Makinson (1994) section 4.2. These presentations take as primitive a relation \leq , defining $x < y$ as $y \leq x$. But the basic idea goes back much further, and appears in a wide variety of guises, sometimes in dual form, in earlier studies of belief revision, conditionals, and generalizations of probability theory. The history is outlined in appendix A of Gärdenfors and Makinson (1994). We will explain the connection with probability in section 5.4. For detailed investigations of comparative expectation inference relations, with generalization to infinite premise sets, see Rott (2001), (2003).

Final Remarks

We have deliberately left aside many issues of a technical nature. For example, we have mentioned only in passing the main Horn conditions (and the non-Horn condition of rational monotony) that are satisfied by the various kinds of consequence relation. This matter is studied systematically in Makinson (1994). On a more difficult level, we have not discussed the representation theorems that may be proven for some of them, nor the maps that can be constructed to embed structures of one kind in those of another. These questions have been investigated in the literature, and a number of results, some deep and difficult, have been obtained. Some of them are brought together in the books of Bochman (2001) and Rott (2001).

In addition to the basic form of default assumption consequence examined in the previous section, we have sketched nearly a dozen others from the literature, obtained by processes of specialization, variation, abstraction, and re-particularization. Evidently there are many other ways in which one might play with the notion of background assumptions, using consistency constraints, ordering principles and other devices to obtain yet other versions. Evidently too, one may combine the different variations into composite ones. There is no point in ringing out the changes systematically. It is enough, at least for this overview, to have indicated a simple paradigm and some of the main ideas for elaborating on it.

By now the reader may be rather dismayed by the multiplicity in face. Which if any of these kinds of default-assumption is *the right* one? Or, failing that, which is the best one to use?

We would suggest that none of them is *the right* one, and none of them is *always* the best to use. From a theoretical perspective, they are all interesting ways of generating nonmonotonic inference operations. They are all reasonably well behaved; in particular they are all supraclassical and satisfy left classical equivalence, right weakening and cumulative transitivity. When applied to background assumption sets already closed under classical consequence, they do not collapse into classical consequence as does the simple 'full meet' default-assumption consequence of section 2.2 – although two of them (screened consequence and layered consequence) do manifest a partial collapse in that context. From a practical perspective, they should be regarded as elements of a toolbox to be applied as suitable and convenient. For one purpose one may be more suitable than another; for another, the reverse. Sometimes the choice may not matter much; sometimes none may be what we are looking for.

3. Restricting the Set of Valuations

3.1. From Classical Consequence to Pivotal Valuations

So far, we have described one way of defining supraclassical consequence relations. The essential idea was to augment the current premises A by a set K of background assumptions. This set K may be independent of A , in which case the generated supraclassical relation is monotonic, or may vary in a principled manner with A , producing a nonmonotonic consequence relation.

Definition and Basic Properties

This idea of adding background assumptions may be seen as a ‘syntactic’ manoeuvre. We now do almost the same thing in a ‘semantic’ manner. The basic idea is to *restrict the set of valuations* that are considered. In other words, we take a subset W of the entire set V of all Boolean valuations, and simply redefine consequence modulo W instead of modulo V .

It turns out that the consequence relations that can be so generated are almost the same as can be obtained by using pivotal assumptions; indeed, in the finite case they are exactly the same. But this way of proceeding points to quite different ways of allowing the new ingredient W to vary with A , leading to substantially different ways of generating nonmonotonic relations. In this way an apparently small change on the monotonic level opens new perspectives for going nonmonotonic.

Definition: Pivotal-Valuation Consequence

Let $W \subseteq V$ be a set of Boolean valuations on the language L . Let A be any set of formulae, and let x be an individual formula.

- We say that x is a *consequence of A modulo the valuation set W* , and write $A \mid\!-\!_W x$ alias $x \in Cn_W(A)$, iff there is no valuation $v \in W$ such that $v(A) = 1$ whilst $v(x) = 0$.
- We call a relation or operation a *pivotal-valuation consequence* iff it coincides with $\mid\!-\!_W$ (resp. Cn_W) for some set W of valuations.

Note again that there is not a unique pivotal-valuation consequence relation, but many – one for each value of W .

Immediately from this definition, pivotal-valuation consequence relations/operations are supraclassical, i.e. $Cn \leq Cn_W$ for any choice of W . They also satisfy inclusion, cumulative transitivity and monotony, and thus are closure operations. We are thus still in the realm of paraclassical inference. Pivotal-valuation consequence relations also have the property of disjunction in the premises. All these points are easily checked.

On the negative side, as we would already expect, pivotal-valuation consequence relations are not closed under substitution. Moreover – and this is a new feature compared to pivotal-assumption consequence – they are not always compact.

The following very simple example of David Gabelaia (personal communication) illustrates the failure of compactness. Indeed, the example shows that pivotal-valuation consequence can even fail the weakened version of compactness that we called the redundancy property (section 2.2).

Consider a language based on a countable set P of elementary letters. Let v_1 be the unique valuation that makes every elementary letter true. Let W be the set of all valuations except v_1 . Then we have $P \vdash_W (p \wedge \neg p)$, since there is no valuation in W that satisfies all the infinitely many letters in P , thus *a fortiori* none that does that and at the same time falsifies $(p \wedge \neg p)$. On the one hand, for no proper subset $Q \subset P$ do we have $Q \vdash_W (p \wedge \neg p)$, since W contains a valuation that makes all letters in Q true, and this valuation evidently makes $(p \wedge \neg p)$ false.

These observations may be summarised in the following table.

Table 3.1. Some Positive and Negative Features of Pivotal-Valuation Consequence

Supraclassical		Closure Relation	Paraclassical
Reflexive			
Cumulatively Transitive (CT)			
Monotonic			
Disjunction in the Premises (OR)			
(Not always compact)			

The failure of compactness shows that not every pivotal-valuation consequence operation is a pivotal-assumption one. For as we have seen, all of the latter are compact. On the other hand, the converse inclusion does hold. This follows easily from the following simple lemma, where V is the set of all Boolean valuations.

Lemma 3.1-1

Let K be any set of (Boolean) formula. Then $Cn_K = Cn_W$ where $W = \{v \in V: v(K) = 1\}$.

Proof

Clearly, it suffices to show that for every valuation v and every premise set A , $v(K \cup A) = 1$ iff $v \in W$ and $v(A) = 1$. But $v(K \cup A) = 1$ iff $v(K) = 1$ and $v(A) = 1$, and by the definition of W this holds iff $v \in W$ and $v(A) = 1$.

Call a subset W of V *definable* iff there is a set K of formulae such that $W = \{v \in V: v(K) = 1\}$. Then:

Theorem 3.1–2

Every pivotal-assumption consequence operation is a pivotal-valuation one. Indeed, the pivotal-assumption consequence operations are precisely the pivotal-valuation operations determined by a definable subset W of V .

Proof

We already have the left-in-right inclusion by the Lemma. For the right-in-left inclusion, let Cn_W be any pivotal-valuation consequence operation where W is a definable subset of V . This means that there is a set K of formulae such that $W = \{v \in V: v(K) = 1\}$, so we may again apply the Lemma.

Thus the family of pivotal-assumption consequence operations is strictly narrower than the family of pivotal-valuation ones. Can we say any more about the links between the two families? Is satisfaction of compactness enough to ensure that a pivotal-valuation consequence operation is a pivotal-assumption one? The answer is positive.

Theorem 3.1–3

The pivotal-assumption consequence operations are precisely the pivotal-valuation ones that are compact.

Proof

For the left-to-right implication, we have just shown in Theorem 3.1–2 that every pivotal-assumption consequence operation is a pivotal-valuation one. Also, we have already noticed in section 2.1 that pivotal-assumption consequence operations are always compact.

For the right-to-left implication, we have noted that any pivotal-valuation operation is a supraclassical consequence operation satisfying disjunction in the premises. So the representation theorem 2.1–1 tells us that if it is also compact then it is a pivotal-assumption relation.

By a *finitely generated Boolean language* (or more briefly, a *finite language*), we mean one generated by Boolean connectives from a finite set of elementary letters. It is well known that in such a language there are only finitely many formulae that are mutually non-equivalent under classical consequence. The same is therefore also true under any supraclassical consequence operation. It follows that any paraclassical operation on a finite language is compact. It is also easy to show that any set of valuations over such a language is definable, indeed definable by a single formula – just take an appropriate formula in disjunctive normal form. By either route, we have the following as a corollary of Theorem 3.1–2.

Corollary to Theorem 3.1–3

For finite Boolean languages the pivotal-assumption consequence operations are precisely the pivotal-valuation ones.

For the computer scientist, who always works in a finitely generated language, pivotal-assumption and pivotal-valuation consequence are thus equivalent. For the logician, who

takes perverse pleasure in the subtleties of the infinite, they are not. We persist with the infinite case from here to the end of the section; hard-line finitists will have to grin and bear it, or skip it.

The Representation Problem for Pivotal-Valuation Consequence

The question arises whether we can characterize the family of pivotal-valuation consequence operations over a countable Boolean language (i.e. where we have countably many elementary letters) in terms of the properties that these operations satisfy. This would give us a representation theorem for the family, paralleling Theorem 2.1–1 for pivotal-assumption consequence operations.

It would be pleasant to be able to report that the family is fully characterized by the properties of being a closure operation, supraclassical, and satisfying disjunction in the premises. However, inspection of the proof that we gave for Theorem 2.1–1 should make us rather suspicious about such a conjecture. For half of that proof depends on applying the compactness property, which is no longer available. It was needed in order to make a single formula do the work of an infinite set A of premises in so far as any given conclusion x is concerned.

In fact, an elegant counterexample has been devised by Karl Schlechta. He has constructed a supraclassical closure operation satisfying disjunction in the premises (indeed, satisfying a more infinitary version of it than we have considered here) that is not a pivotal-valuation consequence operation. Schlechta's subtle construction shows even more: his example is not even a preferential consequence operation (to be defined in the next section); indeed, it is not even the intersection of any family of preferential consequence operations (see Schlechta (1992) or the alternative presentation in Makinson (1994) Observation 3.4.10). But these features need not concern us for the present. All we need to know here is that his example is a supraclassical closure operation, satisfies OR, but is not equal to Cn_W for any $W \subseteq V$.

On the other hand, the failed representation theorem does hold in a finitary form, as follows.

Theorem 3.1–4

Let Cn^+ be any supraclassical closure operation satisfying OR. Then there is a pivotal-valuation operation Cn_W that agrees with Cn^+ on finite sets, i.e. such that $Cn_W(A) = Cn^+(A)$ for all finite A .

Proof

The proof is easy, given what we know already. Define an operation $Cn^{+'}$ from Cn^+ by the rule: $x \in Cn^{+'}(A)$ iff $x \in Cn^+(B)$ for some finite $B \subseteq A$. Clearly, using monotony, $Cn^{+'} \leq Cn^+$. Also, $Cn^{+'}$ agrees with Cn^+ on finite sets, and is compact. For this reason, we might call it the *compactification* of Cn^+ . It is also easy to check that $Cn^{+'}$ inherits from Cn^+ the properties of supraclassicality, being a closure operation, and satisfying OR. Hence by the representation Theorem 2.1–1, $Cn^{+'}$ is in fact a pivotal-assumption operation Cn_K for some $K \subseteq L$, and so by Theorem 3.1–2 is also a pivotal-valuation operation Cn_W for some $W \subseteq V$.

The question remains: Is there any 'strong' representation theorem for pivotal-valuation consequence, i.e. also covering infinite premise sets? In other words, can we add further

syntactic conditions, still satisfied by all pivotal-valuation consequence operations, such that every operation satisfying the enlarged collection of conditions is identical to some pivotal-valuation consequence? As far as the author knows, this question has not been answered, positively or negatively. Given the finitary result in Theorem 3.1–4, any such conditions would presumably need to be infinitary. Given the subtlety of Schlechta’s example, they might also need to be rather complex

Notice that there is an epistemological asymmetry between the two sides of this open question. For a positive answer, we don’t need a precise formal definition of what a representation theorem is. We just need to understand the notion well enough to recognize the one that we find. But to obtain a negative result we do need a precise formal definition of the concept.

Generalizing to Ideals

We now describe another, more abstract, way of restricting the set of valuations without losing monotony in the consequence operation. For finitely generated languages it is equivalent to pivotal-valuation consequence (and thus in turn to its pivotal-assumption counterpart). But it is rather more general in the infinite case. It will appeal particularly to those readers who have studied some abstract algebra or lattice theory, where the notion of an ideal is well established. Others may prefer to skip it. The notion of an ideal is dual to that of a filter, which is rather more familiar in logical studies. The definitions below could be formulated equivalently, but rather less intuitively, in terms of filters.

Recall that we defined $A \vdash_W x$ to hold iff:

- There is no valuation $v \in W$ such that $v(A) = 1$ whilst $v(x) = 0$.

Clearly this is equivalent to each of the following:

- Every valuation $v \in V$ such that $v(A) = 1, v(x) = 0$ is in $V-W$
- $\{v \in V: v(A) = 1, v(x) = 0\} \subseteq V-W$
- $\{v \in V: v(A) = 1, v(x) = 0\} \in 2^{V-W}$.

Now for any set $U \subseteq V$, its powerset 2^U has special properties. In particular, it contains the empty set, contains all subsets of each of its elements, and is closed under the union of any two of its elements (in fact, of any collection of them). We may generalize by introducing the abstract notion of an *ideal* over V . This is any family Δ of subsets of V such that $\emptyset \in \Delta, S \in \Delta$ whenever $S \subseteq T \in \Delta$, and $S \cup T \in \Delta$ whenever $S, T \in \Delta$.

An ideal is thus like a family of ‘small’ subsets, in so far as it contains the empty set and all of the subsets of any of its elements. But it is dissimilar from any numerical notion of smallness in two ways. First, an ideal over V need not contain all the subsets of V that are numerically as small as one of its elements; for example it may contain some singletons but not others. Second, by iterating indefinitely the last clause in the definition, an ideal contains the union of any finite number of its elements, and so may have elements of any finite size so long as there are enough smaller ones to join together.

Let Δ be any ideal over the set V of all valuations on the language L . Let A be any set of formulae, and let x be an individual formula. We say that x is a *consequence of A modulo the ideal Δ* , and write $A \vdash_{\Delta} x$ alias $x \in Cn_{\Delta}(A)$ iff the set of valuations that satisfy A but not x is Δ -

small'. In other words, iff $\{v \in V: v(A) = 1 \text{ and } v(x) = 0\} \in \Delta$. We call a relation or operation a *pivotal-exception consequence* iff it coincides with $|-_{\Delta}$ (resp. Cn_{Δ}) for some ideal Δ over V . Note again that there is not a unique pivotal-exception consequence relation, but many – one for each value of Δ .

Any pivotal-valuation consequence operation is evidently a pivotal-exception one – given Cn_W , simply put $\Delta = 2^{V-W}$, and Δ will be an ideal over V with $Cn_{\Delta} = Cn_W$. Evidently, the converse also holds whenever the ideal Δ is *principal*, i.e. whenever there is a $U \subseteq V$ such that $\Delta = 2^U$. We can thus say:

Theorem 3.1–5

The pivotal-valuation consequence operations are precisely the pivotal-exception ones generated by a principal ideal over V .

Pivotal-exception consequences behave very much like pivotal-valuation ones. On the positive side, they are always closure operations and satisfies disjunction in the premises. On the negative side, they are not always compact, as can be seen by the same counterexample as in section 3.1. Indeed, when a pivotal-exception operation is compact, then it is a pivotal-valuation one. For by the representation Theorem 2.1–1 such an operation will be a pivotal-assumption consequence and thus by Theorem 3.1–2 a pivotal-valuation one.

Here too the question of a representation theorem arises. As for the less general class of all pivotal-valuation consequences, the question appears to open.

One could in principle generalize further by using *downsets* rather than ideals. These are families Δ of subsets of V such that $\emptyset \in \Delta$ and $S \in \Delta$ whenever $S \subseteq T \in \Delta$. In other words, the requirement of closure under union of pairs is dropped. However, the pivotal-exception operations generated by downsets will be less regularly behaved than those generated by ideals. In particular they will not in general be closure operations (cumulative transitivity can fail) and may fail the rule of conjunction in the conclusion.

3.2. From Pivotal Valuations to Default Valuations

How can we go from pivotal-valuation consequence to its default-valuation counterpart? The essential idea is to allow the restricted set W of valuations, or more precisely the part of it that we actually use, to vary with the premise set A .

In principle, this could be done by taking the various maxiconsistent constructions and translating them into the language of Boolean valuations. In this translation, formulae and sets of formulae become sets of valuations (specifically, the set of all valuations making them true), intersections of sets of formulae become unions of the corresponding sets of valuations, and so on. But this is not the way in which the material actually developed, nor is a good way of reviewing it. The ‘basic’ default-valuation system is not a translation of the simple default-assumption system of section 2.2. It is a rather more sophisticated affair, involving not only a set of valuations but also a relation between valuations.

In the context of valuations, the basic default construction is due to Shoham (1988). His central idea was to focus on those valuations satisfying the premise set A , that are *minimal* under some background ordering over the set of all valuations.

Definitions: Preferential Models and Preferential Consequence

- A *preferential model* is understood to be a pair $(W, <)$ where W is, as in the monotonic case, a set of valuations on the language L – not necessarily the entire set V – and $<$ is an irreflexive, transitive relation over W .
- Given a preferential model $(W, <)$ we say that a formula x is a *preferential consequence* of a set A of formulae and write $A \mid\sim_{<} x$ iff $v(x) = 1$ for every valuation $v \in W$ that is minimal among those in W that satisfy A . Here minimality is understood in the usual way: when U is a set of valuations, u is a *minimal element of U* iff $u \in U$ and there is no $v \in U$ with $v < u$.
- We call a relation or operation a *pivotal-valuation consequence* iff it coincides with $\mid\sim_{<}$ for some preferential model $(W, <)$.

The definition of the consequence relation $\mid\sim_{<}$ makes it depend on both W and $<$, and strictly speaking the snake sign should also carry W as a subscript; but it is customary to omit it to simplify notation. When preferential consequence is read as an operation rather than a relation, we write it as $C_{<}(A)$. The relation $<$ over W should not be confused with any of the relations between formulae that were discussed in section 2.3, and which we also wrote as $<$ to keep notation simple. In the following section, we shall look at more general versions of the notion of a preferential model, notably ones that allow 'copies' of valuations, but for the present we remain with the basic version.

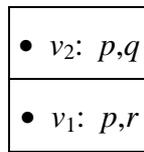
It is worth noting that the definition of minimality under $<$ makes perfect sense for an *arbitrary* relation over W . The notion of a preferential model is thus well defined even when we did not place any constraints on its relation, and in many presentations that is what is done. Nevertheless, the properties of preferential consequence become more regular when some such constraints are imposed, and to simplify life we do so here, requiring both irreflexivity and transitivity.

It is easy to see that pivotal-valuation consequence $A \mid\sim_W x$ is a special case of preferential consequence. It is the case where all elements of W are minimal under $<$, i.e. where no element of W is less than any other, i.e. where $<$ is the empty relation.

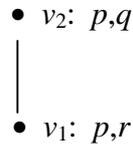
It is sometimes convenient to express the definition in a more concise manner. As is familiar, in classical logic one often writes $v \models A$ to mean that the valuation v satisfies A , i.e. that $v(A) = 1$, i.e. that $v(a) = 1$ for every $a \in A$. We now also write $v \models_{<} A$ and say that v *preferentially satisfies A* to mean that $v(A) = 1$ and there is no $v' \in W$ with $v' < v$ such that $v' \models A$. Again, strictly speaking the symbol $\models_{<}$ for preferential satisfaction should carry a further subscript, since it depends on W , but this is generally left as understood. In this notation the definition of preferential consequence says: $A \mid\sim_{<} x$ iff $v \models x$ whenever $v \models_{<} A$. This manner of expression tends to be handy when checking examples.

Another way of expressing the concept of preferential consequence more succinctly is to write $|A|_W$ for the set of all valuations in W that satisfy A , i.e. $|A|_W = \{v \in W : v(A) = 1\}$. Write $\min_{<}|A|_W$ for the set of all minimal elements of $|A|_W$. In this notation we have $A \sim_{<} x$ iff $v \in |x|$ whenever $v \in \min_{<}|A|_W$. Even more briefly: iff $\min_{<}|A|_W \subseteq |x|_W$. This notation leaves nothing implicit, and tends to be useful when proving general observations about preferential models.

Preferential consequence relations/operations are nonmonotonic. For a simple example, consider a language with three elementary letters p, q, r . Put $W = \{v_1, v_2\}$ where these two valuations are defined by setting $v_1(p) = v_2(p) = 1$, $v_1(q) = 0$, $v_2(q) = 1$, $v_1(r) = 1$, $v_2(r) = 0$. Finally, order W by putting $v_1 < v_2$. Informally, we can describe this preferential model by saying: let p be true in both valuations, q true in just the top one, and r true in just the bottom one. We can also represent it by a diagram with levels:



Here the convention is that all points on a lower level are less than all those on a higher level, but incomparable among themselves (i.e. none is less than any other). We can also use a diagram in which lines representing the relation between valuations. This is potentially more general in that we can also use it with less well-behaved relations:



In both kinds of diagram, the convention is that at each point, i.e. at each element $v \in W$, we list only the elementary letters that v satisfies, omitting those that it makes false. This convention reduces clutter, but readers who find it irksome can simply add at each point the negations of the letters omitted.

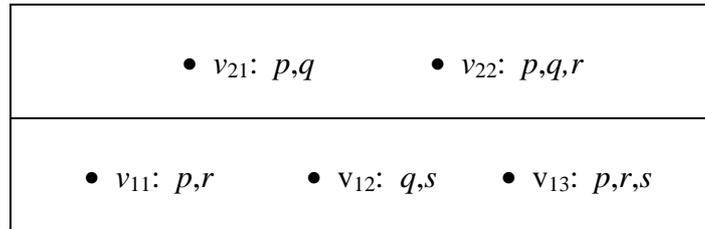
In this preferential model, $p \sim_{<} r$ since the least valuation in which p is true is v_1 , and r is also true there; but $p \wedge q \not\sim_{<} r$, since the least valuation in which $p \wedge q$ is true is v_2 , and r is false there.

This example also shows that preferential consequence $\sim_{<}$ can fail transitivity – even though the relation $<$ of the model is transitive. For we have $p \wedge q \sim_{<} p$ and $p \sim_{<} r$ but $p \wedge q \not\sim_{<} r$. In general, it is important not to confuse the properties of the relation $<$ of the model with the properties of the consequence relation $\sim_{<}$ that it generates. Evidently, imposing constraints on $<$ will tend to yield additional properties of $\sim_{<}$, but not simply the same properties.

This example is rather special in that the relation $<$ is not only transitive and irreflexive but also a *complete* in the sense that for all elements $v, v' \in W$, either $v < v'$ or $v' < v$ or $v = v'$.

Briefly, it is *linear*. As a result, any minimal element u of a subset $U \subseteq W$ is a *least* element of U , in the sense that $u < u'$ for every $u' \in U$ with $u \neq u'$. Indeed, as the relation is also asymmetric, any minimal element u of U is *the unique least* element of U .

But in general, when the relation is not complete, there may be more than one minimal element of U . Consider for example a language with four elementary letters p, q, r, s and the preferential model presented by the following diagram:



In the corresponding line diagram, each element in the bottom layer is linked with all elements in the top layer. Here there are two minimal valuations satisfying p (v_{11} and v_{13}) and they both satisfy r , so $p \sim_{<} r$ holds in the model. There are also two minimal valuations satisfying $p \wedge q$ (v_{21} and v_{22}) and one of them fails to satisfy r , so that $p \wedge q \not\sim_{<} r$. The relation in this example is still quite regular: it is *ranked* (or *modular*) in the sense that we defined in section 2.3: whenever $v < v'$ and not $v'' < v'$ then $v < v''$.

Apart from failing monotony, preferential consequence relations are remarkably well behaved. They are supraclassical and satisfy disjunction in the premises. They also satisfy cumulative transitivity and inclusion (the last implied anyway by supraclassicality) and so are also idempotent.

However, they lack some of the properties of their assumption-based analogues. For example, they do not always satisfy *cautious monotony*, which is in effect, a converse of cumulative transitivity. We recall the definition from section 2.2. In the language of operations it says: $A \subseteq B \subseteq C(A)$ implies $C(A) \subseteq C(B)$. In terms of relations: whenever $A \sim b$ for all $b \in B$ and $A \sim x$ then $A \cup B \sim x$. In the case that A, B are singletons it says: whenever $a \sim b$ and $a \sim x$ then $\{a, b\} \sim x$, which in its conjunctive counterpart says: whenever $a \sim b$ and $a \sim x$ then $a \wedge b \sim x$.

Although this property sometimes fails, it always holds for finite preferential models. More generally, it holds whenever there are no infinite descending chains in the preferential model. More generally still, it holds whenever the model satisfies a condition known as *stopping* (alias *smoothness*). This says that whenever $v \in |A|_W$ then either $v \in \min_{<} |A|_W$ or there is a $u < v$ with $u \in \min_{<} |A|_W$.

Another property that preferential consequence relations may lack is *consistency preservation*. This is the property that whenever $A \sim f$ then $A \vdash f$. Here, f is a classical contradiction and \vdash is of course classical consequence. Equivalently: whenever $C(A)$ is classically inconsistent then so is A itself. This holds for default-assumption consequences \sim_K . But it can fail for preferential consequences $\sim_{<}$ essentially because some classical valuations may be missing from the model. When valuations making a proposition a true are missing from W , we will

have $v(a) = 0$ for all $v \in W$ even though a is not a classical contradiction, so that also $\min_{<} a|_W = \emptyset$ and thus $a \not\sim_{<} f$.

Nevertheless, for models satisfying the stoppering condition, consistency is preserved with respect to a suitable paraclassical operation. Specifically, we may take this to be the pivotal-valuation consequence \vdash_W determined by the set W of all valuations in the preferential model. In other words, for stoppered preferential models, we have: whenever $A \sim_{<} f$ then $A \vdash_W f$.

Preferential consequence relations have been studied quite extensively. We do not wish to lose the reader in details; for them see Makinson (1994). The main point of this section is that conceptually, monotonic pivotal-valuation consequences Cn_W serve as a natural bridge between classical Cn and the nonmonotonic operations $C_{<}$ known as preferential consequence operations. The definitions of Cn_W and $C_{<}$ both allow restriction of the set of classical valuations; for $C_{<}$ we also allow this restriction to vary with the premises.

Looking at the inclusion relations between operations, we see that the pattern for preferential consequence resembles the one that we observed for screened consequence, described in section 2.3. For preferential consequence we have the inclusions $Cn \leq Cn_W \leq C_{<} \leq Cn_{\min(W)}$. In other words, the lower bound for the nonmonotonic consequence operation $C_{<}$ moves up from classical Cn to the pivotal-valuation operation Cn_W , while another pivotal-valuation operation $Cn_{\min(W)}$ serves as upper bound.

In effect the small set $\min(W)$ of valuations is the semantic counterpart of the large set K of background assumptions, while the large set W of valuations corresponds to the small set K_0 of protected background assumptions. If we are to construct formal maps, we will put $K = \{x: v(x) = 1 \text{ for all } v \in \min(W)\}$, i.e. we take K to consist of the formulae that are true when all is normal, while we put $K_0 = \{x: v(x) = 1 \text{ for all } v \in W\}$, i.e. we never entertain formulae that are false throughout W . This kind of opposition larger/smaller is typical of the duality between syntactic approaches with sets of formulae and semantic approaches with sets of valuations. Until one gets used to it, the reversal can be disorienting.

If one is to compare preferential consequence with any of the default assumption relations of the previous section, it is not with the basic one of section 2.2 but rather with a combination of screened and relational partial meet consequence from section 2.3. The sets W , $\min(W)$ provide a lower and upper bound respectively, as did K_0 and K for screened consequence, while the consequence relations in between those bounds is determined using the minimality rule as did partial meet consequence.

3.3. Specializations and Generalizations

We continue this account of default-valuation consequence operations by indicating briefly some of the better-known generalizations of our paradigm version of preferential consequence, including one that has acquired the position of ‘industry standard’. But first we describe a special case of considerable importance.

Circumscription

The operation of circumscription was introduced by John McCarthy (1980). It may be seen as a transposition to the semantic level, with considerable broadening, of the closed world assumption, which we presented in section 2.3 as a special kind of default-assumption reasoning.

Circumscription is typically applied in the context of classical first-order logic, where we have predicates, relation symbols, individual constants, individual variables, and quantifiers binding those variables, as well as the Boolean connectives that we have been considering. The essential idea is that when we are presented with a set A of first-order premises, we do not consider all the models of A but only those in which the extension of certain specified predicates are minimized, while the extensions of certain others are held constant and the remainder are allowed to vary freely. The consequences of A are just those formulae that hold in all the models of A that are in this sense minimal. No particular constraints are imposed on the form of the premises in A ; in particular, they need not be Horn formulae.

This may be given a precise definition as a special kind of default-valuation consequence, defined however on a classical language richer than propositional logic, so that the 'valuations' correspond not just to assignments of truth-values to elementary letters but to first-order models. Some forms of circumscription further refine the above definition of minimality – for example, by introducing an ordering on the predicates that are to be minimized, and performing the minimization lexicographically. Indeed, much of the literature on circumscription considers the question of what is the best definition of minimality to facilitate implementation and avoid unwanted consequences. Another part of the literature investigates the extent to which the semantic definitions in terms of minimality may be recast as syntactic ones – in our terminology, as default-assumption consequence relations. In order to do this without loss of generality it is necessary to move up one more linguistic level, formulating the additional background assumptions in the language of second-order logic. The resulting theory is highly intricate.

We will not enter into the labyrinth of these investigations, referring the reader to the summary overviews in Brewka, Dix and Konolige (1997) section 2.2, Antoniou (1997) chapter 12, and the extended survey in Lifschitz (1994). But to help bring out the relationship of circumscription to the closed world assumption on the one hand, and to preferential entailment on the other, we briefly extract its Boolean content.

Suppose that in the semantics of our first-order language we restrict attention a specific domain, perhaps infinite, and suppose that the language is equipped with an individual constant to serve as name for each element of the domain. Then, as is well known, a first-order model may be identified as an assignment of truth-values to the elementary formulae of the form $P\underline{t}$ where P is a predicate or relation symbol and $\underline{t} = t_1, \dots, t_n$ is an n -tuple of individual names. These formulae may in turn be associated with elementary propositional letters $p_{\underline{t}}$, which we may as well write without the subscripts.

On the Boolean level, a *circumscriptive model* may thus be understood as a preferential model $(V, <)$ in which V is the set of all Boolean valuations and $<$ is a relation between valuations defined in a particular way. The set S of all elementary letters is partitioned into three subsets S_1, S_2, S_3 . For valuations v, v' we put $v \leq v'$ iff $v(p) \leq v'(p)$ for all $p \in S_1$ while $v(p) = v'(p)$ for all $p \in S_2$. No constraints are placed on the values for letters p in S_3 . We put $v < v'$ iff $v \leq v'$ but not conversely (so that $v(p) < v'(p)$ for some $p \in S_1$).

Each choice of the sets S_1, S_2 determines a circumscriptive relation $<$ defined in this way, and thus a circumscriptive model. Each circumscriptive model determines a consequence operation $|\sim_<$ by the same rule as for preferential inference in general: $A |\sim_< x$ iff $v(x) = 1$ for every valuation v that is $<$ -minimal among those satisfying A . As the premise set A is allowed to be arbitrary (not just a set of Horn formulae), in general there need not exist a unique such valuation.

In the special case that $S_1 = S$ (so that $S_2 = S_3 = \emptyset$) and the premise set A consists only of Horn formulae, the $<$ -minimal valuation will be unique, and the construction clearly generates the same inference operation as does the closed world assumption. In both cases we are limiting attention to the valuations that satisfy as few elementary letters as possible. Circumscriptive inference, in a purely Boolean context, is thus a special case of preferential inference, but a generalization of closed world inference.

Multiple Copies of Valuations

We pass now to generalizations. It is evident that not all situations can be fully described in a meagre Boolean language. It is therefore natural to wish to allow distinct possible worlds to satisfy exactly the same Boolean formulae. In other words, we may wish to permit as many ‘copies’ v_i of a valuation v as we like in a preferential model.

Technically, this can be done by indexing valuations v by elements s of an arbitrary index set S . Equivalently, one may take valuations to be functions $v(a,s)$ of two arguments instead of one, where a ranges over formulae of the propositional language L as before, and the new argument s ranges over the index set S .

To be sure, such a step could already have been taken in the monotonic case for pivotal-valuation semantics. But, as is almost immediate from the definitions, it would have made no difference there, generating exactly the same consequence relations with a more complex apparatus. In the nonmonotonic case, however, it makes a significant difference, even for finitely generated languages. This is because two copies of a single valuation may have quite different positions in the ordering. A simple example of this will be given shortly, after formulating the precise definition.

For the definition, consider a set S of arbitrary items, whose elements we will call *states*. Consider also an arbitrary relation over S , written as $<$. With each state $s \in S$ we associate a valuation v_s to the formulae of the language. The function that takes a state s to its associated valuation v_s is often called the *labelling function*. It need not be onto the entire set V of valuations (i.e. there may be some $v \in V$ such that $v \neq v_s$ for all $s \in S$), thus allowing us to work with a proper subset of the classical valuations. Nor need it be *injective* (i.e. we may have $v_s = v_t$ although $s \neq t$), thus offering the desired possibility of multiple copies of a valuation.

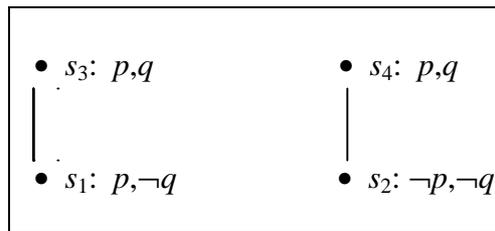
We remark in passing that terminology is fickle here, as it varies from author to author, and even between the papers of a single author. Usually, as here, the elements s of S are called ‘states’, but sometimes they are called ‘worlds’. Other authors reserve the term ‘worlds’ for the associated valuations v_s (or for the sets of formulae that they make true). This ambiguity makes no difference when the labelling function is injective, since the elements of S may then be identified with their images under the labelling. But if care is not taken, confusion can arise in the non-injective case, where several elements of S may correspond to a single valuation.

When multiple copies are allowed, a *preferential model* is defined as a structure made up of the above ingredients: a set S , a relation $<$ over S , and a labelling function associating a classical truth-value assignment v_s to each $s \in S$. The consequence operation is defined by the rule: $A \sim_{<} x$ iff $v_s(x) = 1$ for every state s that is minimal among those making A true, i.e. among those with $v_s(A) = 1$. In another notation introduced in section 3.2: iff $s \models x$ whenever $s \models A$. In a third notation introduced there: iff $\min_{<} |A|_S \subseteq |x|_S$, where this time $|A|$ means $\{s \in S: v_s(A) = 1\}$.

How can this introduction of 'copies' change anything? An elegant example due to Kraus, Lehmann and Magidor (1990) brings out the difference. It is a preferential model with just four states, two of which are labelled by the same valuation. We show that there is no preferential model without copies that determines the same consequence relation.

Example 3.3-1

Consider the Boolean language with just two elementary letters p, q (and thus just four possible valuations). Take the preferential model defined by the following diagram, where we use the same conventions as in section 3.2 except that it will be clearer in this example to list the negative part of each valuation as well as the positive part.



Note that this preferential model is *not ranked*. For instance, $s_1 < s_3$ and s_3 is incomparable with s_4 but $s_1 < / s_4$. Note also that it is *not injective*: states s_3 and s_4 are labelled with the same valuation. We claim that for any preferential model defined for the same language with elementary letters p, q , if it generates the same preferential consequence relation as this one does, then it must contain two distinct states labelled by the same Boolean valuation.

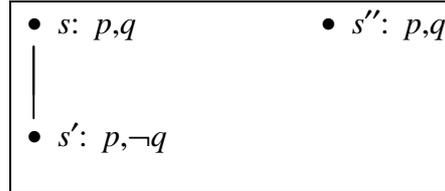
Verification

For simplicity of notation, we leave the subscript $<$ off the snake sign in what follows. First observe that in this preferential model we have $p \wedge q \sim / f$ (where f is any self-contradiction), $t \sim \neg q$ (where t is any tautology), $p \sim / \neg q$ and $p \leftrightarrow q \sim / \neg p \wedge \neg q$ (where \leftrightarrow is truth-functional co-implication). Now consider any preferential model generating the same consequence relation as this one. We want to show that it contains two distinct states labelled by the same Boolean valuation. We break the argument into cases.

Case 1. Suppose that the preferential model has more than four states. As there are only four valuations possible on our meagre language with just two elementary letters, there must be at least two states labelled by the same valuation.

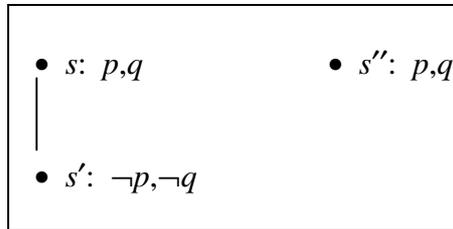
Case 2. Suppose that the preferential model has at most four states. Then it has finitely many states, and so is stoppered. Since $p \wedge q \sim / f$ there is a state s with $s \models_{<} p \wedge q$, and thus $s \models p \wedge q$. Evidently too, $s \models t$. But since $t \sim \neg q$ we know that s is not a minimal t -state. So by stoppering, there must be a minimal t -state s' with $s' < s$. Since $t \sim_{<} \neg q$ we have $s' \models \neg q$. Now either $s' \models p$ or $s' \models \neg p$. We consider these sub-cases separately.

Sub-case 2.1. Suppose $s' \models p$. Then $s' \models p, \neg q$. But we know that $p \sim / \neg q$ so there is also a minimal p -state s'' with $s'' \models q$, so we have $s'' \models p, q$. We can picture this in the following diagram:



Thus s, s'' are labelled by the same valuation, for they satisfy the same elementary letters of our meagre language. But they must be distinct. For on the one hand s'' is a minimal p -state, while on the other hand s is not since $s' < s$ and by the supposition of the sub-case $s' \models p$.

Sub-case 2.2. Suppose $s' \models \neg p$. Then $s' \models \neg p, \neg q$ so also $s' \models p \leftrightarrow q$. But we know that $p \leftrightarrow q \sim / \neg p \wedge \neg q$ so there is a minimal $(p \leftrightarrow q)$ -state s'' with $s'' \models \neg(\neg p \wedge \neg q)$. It follows that $s'' \models p, q$. The situation may be pictured as in the following diagram:



Thus again s, s'' are labelled by the same valuation, for they satisfy the same elementary letters of our meagre language. But they must be distinct. For on the one hand $s' < s$. On the other hand we do not have $s' < s''$, since s'' is a minimal $(p \leftrightarrow q)$ -state and $s' \models p \leftrightarrow q$.

Given the existence of examples like this, the multiple copies version of preferential consequence has become the ‘industry standard’, and is often known as *KLM consequence* after the seminal paper of Kraus, Lehmann and Magidor (1990).

The reader may suspect that there is another way of handling such examples, using a syntactic rather than a semantic manoeuvre. Instead of allowing multiple copies of valuations, we could increase the supply of elementary letters in the language so as to make the same model injective for the enlarged language. Thus in Example 3.3-1 we might add a third letter r , label the two top nodes with the valuations p, q, r and $p, q, \neg r$ and assign r anything we like in the bottom nodes. Similarly for any other example. If there are at most n copies of a valuation in the preferential model, we will need at most $\lceil \log_2 n \rceil$ fresh letters to do the job. If there is no finite upper bound on the number of copies, we can still do it with countably many fresh letters.

Clearly, the consequence relation that the preferential model determines on the enlarged language agrees, over the original language, with the original consequence relation. However, in practice when working with preferential consequence operations it is easier to hold the language constant and allow copies of valuations, than to keep on expanding the language and the valuations labelling the model and restricting attention to the original language.

Suppose, now, that we wish to work only with the consequence operations generated by injective preferential models, but hold fixed the supply of elementary letters in the language. Do these consequence operations satisfy any special syntactic conditions? Can we prove a representation theorem for them? The answer is not entirely clear. The rather intricate story, up to the time of writing, is as follows.

- Lehmann and Magidor (1992) showed that when we restrict attention to ranked preferential models, the generated consequence operations can always be determined by an injective model.
- Freund (1993) strengthened this result. He showed that it holds for all ‘filtered’ preferential models, i.e. all those such that for any two states non-minimally satisfying a formula, there is a state that is less than both of them that satisfies the formula. Ranked models are always filtered, but not conversely. As we remarked earlier, ranked models always satisfy the non-Horn condition of rational monotony. Filtered models always satisfy the weaker condition of *disjunctive rationality*. This is the non-Horn condition: whenever $a \vee b \sim x$ then either $a \sim x$ or $b \sim x$, or in the language of operations: $C(a \vee b) \subseteq C(a) \cup C(b)$.
- In the same paper (1993), Freund showed that when the language is finite, injective preferential models always satisfy the non-Horn condition $C(a \vee b) \subseteq C_n(C(a) \cup C(b))$. This is clearly weaker than disjunctive rationality, and we call it *Freund's condition*. Conversely, he showed that for a finite language we have a representation theorem: every preferential consequence relation satisfying this condition is generated by some injective preferential model.
- The restriction to finite languages in Freund’s representation theorem was eliminated by Pino Pérez and Uzcátegui (2000), and also by Zhu *et al* (2002). Using two different constructions, leading to distinct representations, these papers show, for both finite and infinite languages, that any inference relation satisfying Freund's condition may be represented by an injective preferential model.
- A gap remains: the soundness theorem can fail for infinite languages. For as Pino Pérez and Uzcátegui also show with an example, for an infinite language there are injective preferential models whose consequence operations do not satisfy Freund's condition. The issue of injectivity is thus quite complex, with a number of technical questions still unanswered in the infinite case.

Standing back from these technical issues, can we say what is the best framework for presenting preferential models? It is generally felt that the one chosen by Kraus, Lehmann and Magidor is the most convenient to work with. It allows copies of valuations in the preferential models. At the same time, attention is usually restricted to individual formulae as premises,

rather than sets of them. This is because the failure of compactness renders difficult the proof of representation theorems when infinite premise sets are considered.

Selection Functions instead of Relations

Another line of generalization abstracts on the preference relation over W (or over S , if we are allowing copies), and works instead with a selection function γ over its power set, with $\gamma(U) \subseteq U$ for every $U \subseteq W$. This selection function may be constrained by suitable conditions as desired. If sufficiently powerful constraints are imposed, then a preference relation can be recovered from the selection function.

Evidently, this move runs parallel to one for partial meet default assumptions that we considered in section 2.3. In that case, we could work either with relations between subsets of the set K of background assumptions, or with selection functions on families of such subsets. Now, on the semantic level, we may work with relations between valuations (or states), or selection functions on families of them. On the abstract mathematical level, essentially the same questions arise.

Lindström (1991), Rott (2001) and Lehmann (2001) investigate in detail the approach using selection functions. As remarked in section 2.3, Rott (1993) and (2001) looks deeply into the abstract connections between selection functions and preference relations.

Non-Classical Valuations

In all of the constructions of this section, we have been dealing with classical valuations, i.e. those functions on the set of formulae into $\{0,1\}$ that behave in the usual Boolean fashion. It is possible to relax this condition, either partially or entirely. Kraus, Lehmann and Magidor (1990) studied a partial relaxation. A more radical relaxation, taking valuations to be arbitrary functions on the set of formulae into $\{1,0\}$, was investigated by Makinson (1989) and more recently Lehmann (2002).

Working with Ideals

In section 3.1 we looked at a variation on pivotal-valuation consequence, which we called 'pivotal-exception consequence', in which the set of all Boolean valuations is restricted by looking at an ideal. That was on the monotonic level. We may pass to the nonmonotonic level by allowing the identity of the ideal Δ to vary with the premise set. Instead of taking Δ to be an ideal over the set V , we may take it to be one over the set $|A|$ of those valuations that satisfy the premise set A .

In more detail, consider any function δ that associates with each set $|A|$ an ideal over $|A|$. Then we may say that x is a *default-exception consequence of A modulo δ* , and write $A \sim_{\delta} x$ alias $x \in Cn_{\delta}(A)$, iff the set of valuations that satisfy A but do not satisfy x is an element of the ideal $\delta(|A|)$ over $|A|$, i.e. iff $\{v \in V: v(A) = 1 \text{ and } v(x) = 0\} \in \delta(|A|)$.

So defined, the relations \sim_{δ} remain supraclassical but become nonmonotonic. However, because the ideals associated with different sets $|A|$ need have no relationship to each other, properties such as CT and OR may also fail. In order to recover some or all of them, it is customary to add constraints on the relationship between the ideals $\delta(|A|)$ for different values of A . We shall not enter into the details, referring the reader to the papers mentioned below.

This approach was developed Ben-David & Ben-Eliyahu-Zohary (2000), Schlehta (1997), and Friedman & Halpern (2001), in a number of versions and with various names. Lehmann (2001) has made a detailed comparison with the default-valuation approach (in a version using selection functions and multiple copies of valuations).

4. Using Additional Rules

4.1. From Classical Consequence to Pivotal Rules

We come now to a third way of going supraclassical, monotonically and nonmonotonically. The basic idea is similar to that of adding background assumptions, but instead of adding *propositions* we add *rules*. This apparently small twist brings with it considerable divergence, which reveals itself even in the monotonic context – indeed in its finite case, where the pivotal-assumption and pivotal-valuation approaches are equivalent. In brief, *rules for propositions do not behave quite like propositions*.

By a rule for propositions (briefly, a *rule*) we mean any ordered pair (a,x) of propositions of the language we are dealing with. A set of rules is thus no more nor less than a binary relation R over the language, i.e. a set $R \subseteq L^2$. It must be confessed that although this terminology is standard in the area, it is a bit odd: it would be more natural to refer to the pairs (a,x) simply as pairs, and reserve the name ‘rule’ for the sets R of such pairs; but we follow the choice of words that is more common in the literature. It makes no difference to the content.

Given a set X of propositions and a set R of rules, we define the *image* of X under R , written as $R(X)$, in the standard manner of elementary set theory: $y \in R(X)$ iff there is an $x \in X$ with $(x,y) \in R$. A set X is said to be closed under R iff $R(X) \subseteq X$, i.e. iff whenever $x \in X$ and $(x,y) \in R$ then $y \in X$.

Let $R \subseteq L^2$ be a set of rules. Intuitively, they will be playing the role of a set of background ‘inference tickets’ ready for travel from any set of premises. Let A be a potential premise set, and let x be a potential conclusion.

Definition: Pivotal-Rule Consequence

- We say that x is a *consequence of A modulo the rule set R* , and write $A \mid_{-R} x$ alias $x \in Cn_R(A)$ iff x is in every superset of A that is closed under both Cn and the rule set R . In other words, iff x is in every set $X \supseteq A$ such that both $Cn(X) \subseteq X$ and $R(X) \subseteq X$.
- We call an operation a *pivotal-rule consequence* iff it is identical with Cn_R for some set R of rules.

The definition of $A \mid_{-R} x$ requires x to be in *every* superset of A that is closed under both Cn and the rule set R . Trivially, at least one such set always exists, namely the entire language L . Moreover, the intersection of any non-empty family of sets, each of which is closed under Cn (resp. R), is itself closed under Cn (resp. R). It follows that there is always a unique least such set, namely the intersection of all of them. Thus the definition may be put as follows: $Cn_R(A)$ is *the least* superset X of A such that both $Cn(X) \subseteq X$ and $R(X) \subseteq X$.

It is immediate from the definition that $Cn \leq Cn_R$ so that every pivotal-rule consequence Cn_R is supraclassical. It is not difficult to verify that it is monotonic and also satisfies cumulative transitivity, so that it is a closure operation. We are thus still in the realm of paraclassical

inference. Pivotal-rule consequence can also be verified to be compact – like its assumption-based counterpart but unlike the valuation-based one.

But it lacks an important property possessed by both pivotal-assumption and pivotal-valuation operations, namely disjunction in the premises. For example, if $R = \{(a,x), (b,x)\}$ then $x \in Cn_R(a)$ and also $x \in Cn_R(b)$ but $x \notin Cn_R(a \vee b) = Cn(a \vee b)$. This is because the last-mentioned set is vacuously closed under R , i.e. we have $R(Cn(a \vee b)) = \emptyset \subseteq Cn(a \vee b)$ since $a, b \notin Cn(a \vee b)$. We thus have the following picture.

Table 4.1. Some Positive and Negative Features of Pivotal-Rule Consequence

Supraclassical		Closure Relation	Paraclassical
Reflexive			
Cumulatively Transitive (CT)			
Monotonic			
(Does not always satisfy Disjunction in the Premises (OR))			
Compact			

Pivotal-rule consequence also fails contraposition. This follows from the failure of disjunction in the premises, since any paraclassical operation satisfying the former must satisfy the latter. For a direct counterexample, put $A = \{a\}$ and $R = \{(a,x)\}$. Then $x \in Cn_R(a)$ but $\neg a \notin Cn_R(\neg x) = Cn(\neg x)$, because the last-mentioned set is vacuously closed under R , i.e. we have $R(Cn(\neg x)) = \emptyset \subseteq Cn(\neg x)$ since $a \notin Cn(\neg x)$. This contrasts with the situation for pivotal-assumption and pivotal-valuation consequence. They both satisfy contraposition – although, it should immediately be added to avoid confusion, their default versions do not.

What is the relationship between the pivotal-rule and pivotal-assumption consequence operations? As one might expect, it is as follows.

Theorem 4.1-1

The pivotal-assumption consequence operations are precisely the pivotal-rule ones that satisfy disjunction in the premises.

Proof

On the one hand, every pivotal-assumption consequence operation is itself a pivotal-rule one – given Cn_K , simply put $R = \{(t,k): k \in K\}$ where t is a tautology, and it is easily checked that $Cn_K = Cn_R$. Conversely, when Cn_R is a pivotal-rule consequence then as we have remarked it is a compact supraclassical closure operation. So when in addition it satisfies disjunction in the premises, then by the representation result Theorem 2.1-1 it is also a pivotal-assumption consequence operation.

Combining this with results from the preceding sections, we thus have the following curious relationship between the three kinds of pivotal inference.

Corollary to Theorem 4.1–1

The set of pivotal-assumption consequence operations is the intersection of the set of pivotal-valuation and the set of pivotal-rule consequence operations.

Proof

We have already shown that the left set is included in each of the two right ones (Theorems 3.1–2 and 4.1–1). For the converse, when an operation is in both of the two right sets then it is paraclassical and satisfies both disjunction in the premises and compactness (Theorems 3.1–3 and 4.1–1), so that we may apply representation Theorem 2.1–1 again.

What happens if we try to get disjunction in the premises back on board? What if we define $Cn_{R\vee}(A)$ to be the least superset of A that is closed under Cn , R and disjunction in the premises? As the reader may suspect, this leads us straight back to pivotal-assumption consequence again. More specifically, we have the following identity, noticed in a more general context by Makinson and van der Torre (2000) section 6.

Theorem 4.1–2

$Cn_{R\vee}(A) = Cn(A \cup m(R))$ where $m(R)$ is the set of materializations of pairs in R , i.e. $m(R) = \{a \rightarrow x : (a, x) \in R\}$.

Sketch of Proof

Clearly $Cn_{R\vee}(A) \subseteq Cn(A \cup m(R))$ since the right side includes A and is closed under Cn , R , and OR, while the left side is by definition the least such set. The interesting part is the converse. It can be proven in several different ways, each instructive.

One proof begins by noting that $Cn_{R\vee}$ is by definition supraclassical and satisfies OR, and then checks that it is also a compact closure operation. We may thus apply the construction used in the proof of Theorem 2.1–1 to obtain $Cn_{R\vee}(A) = Cn(A \cup Cn_{R\vee}(\emptyset))$. We then check that $Cn_{R\vee}(\emptyset) = Cn(m(R))$, so that finally $Cn_{R\vee}(A) = Cn(A \cup Cn(m(R))) = Cn(A \cup m(R))$.

Another method is to argue from first principles. Use Zorn's lemma to show that whenever $x \notin Cn_{R\vee}(A)$ then there is a maximal $B \supseteq A$ with $x \notin Cn_{R\vee}(B)$, and show that such a maximal B must be a maxiconsistent set, with $m(R) \subseteq B$, so that its characteristic function is a Boolean valuation satisfying $A \cup m(R)$ but not satisfying x . In effect, this argument mimics that for Theorem 2.1–1 itself.

Yet another argument can be constructed by supposing $x \in Cn(A \cup m(R))$, using the compactness of classical consequence to infer that $x \in Cn(A \cup m(R'))$ for some finite subset $R' \subseteq R$, and then arguing by induction on the number of elements of R' .

Theorems 4.1–1 and 4.1–2 together throw light on the subtle relations between rules and propositional assumptions.

- The two are not the same. In general, adding propositional assumptions gives us *more* than adding rules. In other words, $Cn_R(A) \subseteq Cn(A \cup m(R))$ and the inclusion will usually be proper. This is so, not only for classical Cn but any closure operation satisfying modus ponens.
- But when we are also given closure under OR, the two kinds of addition have exactly the *same* power. Rules collapse into their materializations, that is, $Cn_{R \vee}(A) = Cn(A \cup m(R))$.

We also have a representation theorem for pivotal-rule consequence.

Theorem 4.1–3

For every compact supraclassical closure operation Cn^+ , there is a relation R such that for all A , $Cn^+(A) = Cn_R(A)$.

Proof

The construction is trivial, the proof easy. For the construction, simply put R to be Cn^+ restricted to singletons, i.e. put $R = \{(a,x) : x \in Cn^+({a})\}$. The proof is in two parts. First we show that $Cn^+(A)$ is a superset of A and is closed under both classical Cn and R . Then we show that it is the least such superset.

For the first part, we have $A \subseteq Cn^+(A)$ since Cn^+ is a closure operation. Also $Cn(Cn^+(A)) \subseteq Cn^+(A)$ since by supraclassicality and idempotence of Cn^+ we have $Cn(Cn^+(A)) \subseteq Cn^+(Cn^+(A)) \subseteq Cn^+(A)$. Again, $R(Cn^+(A)) \subseteq Cn^+(A)$ by the definition of R together with the monotony and idempotence of Cn^+ . This shows that $Cn^+(A)$ is a superset of A that is closed under both classical Cn and R .

To show that it is the least such superset, let X be another one; we need to check that $Cn^+(A) \subseteq X$. But if $x \in Cn^+(A)$ then by compactness of Cn^+ there is a finite $B \subseteq A$ with $x \in Cn^+(B)$. Also $Cn^+(B) = Cn^+(b)$ where b is the conjunction of all the finitely many elements of B , using supraclassicality and the supposition that Cn^+ is a closure operation. Since $x \in Cn^+(b)$ we have by the definition of R that $(b,x) \in R$, so $x \in R(\{b\}) \subseteq R(Cn(\{b\})) = R(Cn(B)) \subseteq R(Cn(A)) \subseteq R(Cn(X)) = R(X) = X$ since by hypothesis X includes A and is closed under both R and classical Cn .

However, it would be mistaken to read much into this representation theorem. It is far less significant than Theorem 2.1–1 for pivotal-assumption consequences. The proof works only because the definition of a pivotal-rule consequence relation places no constraints on the relation R . This permits us to ‘cheat’ by taking R to be *the very consequence relation that we are seeking to represent*, restricted to singletons. Evidently, it would be more interesting to have representation results in terms of relations R that are constrained in some natural way. That is an area that does not seem to have been explored.

To end this section, we note a point that will be vital when we pass to default-rule consequence in the next section. It is possible to reformulate the definition of pivotal-rule

consequence in an inductive manner. We have the following identity, whose verification uses the compactness of classical Cn .

Observation 4.1–4

$$Cn_R(A) = \cup\{A_n : n < \omega\} \text{ where } A_0 = Cn(A) \text{ and } A_{n+1} = Cn(A_n \cup R(A_n)).$$

Indeed, assuming that our language has at most countably many formulae so that R is also finite or countable, we can go further and reformulate the inductive definition so that it has only singleton increments in the inductive step.

To do this, fix an ordering $\langle R \rangle$ of the set R of rules, indexing them with natural numbers $0, 1, 2, \dots$. More precisely speaking, fix an ordering $r_i = (a_i, x_i)_{i < \alpha}$ of all the rules in R , without repetitions, with α a natural number if R is finite, or ω if R is countable.

Definition: Singleton Increment given an Ordering

- Given such an ordering $\langle R \rangle$ of a rule-set R , we define the operation $Cn_{\langle R \rangle}$ by putting $Cn_{\langle R \rangle}(A) = \cup\{A_n : n < \omega\}$ with $A_0 = Cn(A)$ and $A_{n+1} = Cn(A_n \cup \{x\})$, where (a, x) is the *first* rule in $\langle R \rangle$ such that $a \in A_n$ but $x \notin A_n$. In the case that there is no such rule, we put $A_{n+1} = A_n$.

To avoid a proliferation of letters, we are here using the same notation A_n for the terms of this sequence as for the previous sequence defining $Cn_R(A)$, but evidently it is quite a different one, making only singleton increments. Several features of the definition deserve special note.

- The sequence A_0, A_1, \dots is increasing, i.e. $A_n \subseteq A_{n+1}$ for all n .
- The construction does not require that rules are applied in the order in which they occur in $\langle R \rangle$. When we form A_{n+1} , we may be applying a rule (a, x) that occurs earlier in the sequence $\langle R \rangle$ than some rule that we have already applied, since its body a may become available only at A_n .
- By running through the sequence $\langle R \rangle$ from its beginning again at each stage of the construction, we ensure that no rule (a, x) is overlooked when its body eventually becomes available for use. Using terminology from programming theory, the inductive definition is 'fair'.
- By the terms of the definition, once a rule (a, x) is applied, it is never applied again, since its head x is in all the subsequent sets A_{n+k} . This prevents the sequence getting stuck with eternal reapplications of the same rule, unable to get over it to other ones.

Although the terms A_n of the sequence defining $Cn_{\langle R \rangle}(A)$ are not the same as those of the sequence defining $Cn_R(A)$, their union is the same. In other words:

Observation 4.1–5

$$Cn_{\langle R \rangle}(A) = Cn_R(A).$$

This in turn implies also that the choice of ordering $\langle R \rangle$ of R makes no difference to the final result, since this is in every case equal to $Cn_R(A)$.

These inductive definitions of pivotal-rule consequence may seem like long ways of expressing something that the original definition says quite briefly. And so they are! But we will see in the next section that they – and especially the singleton-increment one – provide a key for passing transparently to the nonmonotonic default-rule consequence operations.

Historical Remarks

The idea of using the monotonic operations Cn_R as a stepping-stone towards default-rule consequence is implicit in many papers, but rarely brought to the surface. One paper that does do so is Sandewall (1985). Sandewall presents the bridge system as a four-valued logic, but it can be translated into the language we are using. However the manner in which he then passes from the bridge system to the default one is quite different from the one that we will use in the following section, and is much closer to Reiter's fixed-point account.

4.2. From Pivotal Rules to Default Rules

The consequence operations Cn_R defined in the preceding section by pivotal rules are, as we have seen, monotonic. Once again we may obtain nonmonotonicity by allowing the set R of rules, or more precisely, the ones among them that may be applied, to vary with A .

As we are working on a syntactic rather than a semantic level, it will not be surprising to learn that this may be done by means of consistency constraints. There are a number of different ways of going about it. Roughly speaking, they are of two main kinds: those that impose consistency constraints on the rules themselves, and those that place constraints on their step-by-step application.

Formulated in such general terms, this may seem a rather nebulous and inconsequential distinction to make. But in fact it marks a real difference, and leads to quite different consequence relations, with quite different behaviour. The literature focuses mainly on the latter kind, where the rules are left untouched but their application is constrained, and we will do likewise. We will begin by describing one of the best-known accounts, due to Reiter (1980), referred to as his system of ‘normal defaults’. In the following section we will explain how this may be generalized to cover also his ‘non-normal defaults’, and will also outline an approach that restricts the rules themselves.

We take as our starting point the inductive characterization of pivotal-rule consequence using singleton increments under the guidance of an ordering $\langle R \rangle$ of the rules in R , which we gave in Observation 4.1–5. Recall that this put $Cn_R(A) = \cup\{A_n : n < \omega\}$ where $A_0 = Cn(A)$ and $A_{n+1} = Cn(A_n \cup \{x\})$, where (a, x) is the *first* rule in $\langle R \rangle$ such that $a \in A_n$ but $x \notin A_n$. In the limiting case that there is no such rule, we put $A_{n+1} = A_n$.

The essential idea now is to monitor these singleton increments, and allow rules to be applied only when they do not generate inconsistency.

Definition: Ordered Default-Rule Consequence with Normal Rules

- As before, we fix a finite or ω -ordering $\langle R \rangle$ of the given set R of rules. Still as before, we put $C_{\langle R \rangle}(A) = \cup\{A_n : n < \omega\}$, and set $A_0 = Cn(A)$.
- As before, we break the definition of A_{n+1} into two cases. But now the cases invoke a consistency check.

Case 1. Suppose there is a rule (a,x) in R such that $a \in A_n$ but $x \notin A_n$ and x is consistent with A_n . Then choose the first such rule and put $A_{n+1} = Cn(A_n \cup \{x\})$ as before.

Case 2. Suppose there is no such rule. Then put $A_{n+1} = A_n$ as before.

Thus the only change in definition, compared to that for pivotal-rule consequence in the form given by Observation 4.1–5, is the introduction of the italicised consistency check as a requirement for entering into the first case of the induction step: the head of the rule must be consistent with the material constructed so far. But this small addition has many effects.

- One is that choice of ordering now makes a difference. In other words, the content of $C_{\langle R \rangle}(A)$ varies with the particular ordering $\langle R \rangle$ of R . Roughly speaking, this is because the application of an earlier rule may introduce material that makes a later rule fail its consistency check.

For example, if $A = \{a\}$ and $R = \{(a,x), (a,\neg x)\}$, then if we order R in the manner written, we get $C_{\langle R \rangle}(A) = Cn(a,x)$ while if we take the reverse order we get $C_{\langle R \rangle}(A) = Cn(a,\neg x)$. Reason: in the former case we have $A_1 = Cn(A_0 \cup \{x\})$ since x is consistent with A_0 , while $A_2 = A_1$ since $\neg x$ is inconsistent with A_1 . In the other case, it is the reverse.

- This means, in turn, that even when the set R of background rules is fixed, there will be not one but many consequence operations $C_{\langle R \rangle}(A)$. Roughly speaking, there will be one for each ordering $\langle R \rangle$, although of course different orderings may happen to give rise to the same operation.
- Another effect of the consistency constraint is that the operations $C_{\langle R \rangle}(A)$ are nonmonotonic. Roughly speaking, this is because an increase in the premises may introduce material that makes some rules fail their consistency checks.

For example, if $A = \{a\}$ and $R = \{(a,x)\}$ then $C_{\langle R \rangle}(A) = Cn(a,x)$, but when A is increased to $B = \{a,\neg x\}$ then $C_{\langle R \rangle}(B) = Cn(B) = Cn(a,\neg x)$. Reason: in the case of $A = \{a\}$ we again have $A_1 = Cn(A_0 \cup \{x\}) = Cn(a,x)$ since x is consistent with $A_0 = \{a\}$, while $A_2 = A_1$ since already $x \in A_1$. But in the case of $B = \{a,\neg x\}$ we have $B_1 = B_0 = Cn(B) = Cn(a,\neg x)$ since x is inconsistent with B and thus with B_0 .

This inductive definition of the sets $C_{\langle R \rangle}(A)$ was first formulated explicitly by Brewka (1994), although with hindsight one might say that it is extractable from Reiter's own discussion of derivations for normal default systems, in the later sections of his seminal (1980) paper. It is equivalent to Reiter's original definition, in that the sets $C_{\langle R \rangle}(A)$ for the separate orderings $\langle R \rangle$ of R coincide with what Reiter called the 'extensions' of A under the normal rules R .

More explicitly, a set E of formulae is an extension, in the sense of Reiter, of A under the set R of normal rules iff $E = C_{\langle R \rangle}(A)$ for *some* well-ordering $\langle R \rangle$ of R . This result is a particular case of a more general theorem, covering also non-normal rules, stated and proven as Theorem 4.3–1 in the following section.

The notation used here is a little simpler than Reiter's. Whereas he writes a normal default rule as $a: Mx/x$, we write it simply as an ordered pair (a, x) . On our presentation, therefore, normal default rules are just the same things as plain rules; the difference lies in the way that they are used. There is also some variation in terminology. In a normal default rule (a, x) alias $a: Mx/x$ Reiter calls a the *prerequisite* of the rule and x its *conclusion*; we will also use the same terms as we did for plain rules – *body* and *head* respectively.

Much more interesting than these details, however, is the fact that the definition above differs in its structure from that of Reiter. In Reiter's original paper (1980), a set E is called an extension of A under R iff it is a *fixpoint* under a certain operation Γ , meaning that it coincides with the value of that operation with itself as argument – briefly, iff $E = \Gamma(E)$. It is important to realize that fixpoint definitions do not in general guarantee the uniqueness or even the existence of the objects being defined. In the particular case of normal default rules, Reiter's fixpoint definition does ensure existence, but not uniqueness. That is, a given set A of premises always has at least one extension under a given set R of normal default rules, but may have many of them. For non-normal default rules, neither uniqueness nor existence is guaranteed.

Reiter (1980) also shows that his extensions may be defined using what is sometimes called a *quasi-induction* or, in the terminology of Makinson (1994), an *end-regulated induction*. This looks like an ordinary inductive definition except that the induction step makes reference to the object to be defined – the definition of A_{n+1} calls not only on the values of A_m for $m \leq n$, but also on the value of the final product $\cup\{A_n : n < \omega\}$. While end-regulated inductions are similar in their appearance to genuine inductions, logically they behave like fixpoint definitions.

We have proceeded inductively for several reasons. With such a definition, it is much easier to get an intuitive feeling for what is going on. In practice, whenever teachers try to convey the general ideas of default-rule reasoning in the classroom, they use some kind of vague, quasi-inductive language before switching over to their official fixpoint definitions. In effect, we are showing that such inductive language can be used accurately and with a clear conscience. At the same time, we are making clear the way in which default-rule operations may be seen as natural elaborations of the perfectly monotonic pivotal-rule operations of the preceding section.

The reader may be puzzled how the inductive definition manages to give the same result as a fixpoint definition if the former gives a unique output and the latter allows multiple outputs. The job is done by the orderings $\langle R \rangle$ of the rules. In effect, the inductive definition transforms the multiplicity of extensions into uniqueness, *modulo the choice of ordering* $\langle R \rangle$. Thus different orderings in general give different extensions. When we generalize to non-normal rules in the next section, and not only uniqueness but also existence may fail, the inductive definition will need to be given a further twist.

The operations $C_{\langle R \rangle}$ are themselves of interest. Indeed, if one is equipped with a preferred ordering $\langle R \rangle$ of R , one may wish to go no further than the operation $C_{\langle R \rangle}$ defined modulo that

particular choice of order. But one can also define a consequence C_R that is independent of any particular order, by formulating a policy for amalgamating (or ‘flattening’, as is sometimes said) all the operations $C_{\langle R \rangle}$ for different values of $\langle R \rangle$.

If we are to respect a principle of equal treatment for all orderings, then only two such policies seem possible: union and intersection of values. In the present context, the union of extensions is hardly acceptable for a notion of consequence, for it will give us an inconsistent output whenever there was more than one extension to begin with. This is because, as is well known, the union of any two distinct Reiter extensions of A under R will be inconsistent. The usual policy is therefore that of intersecting, just as it was when analogous questions arose for default-assumption consequence – and in effect also for default-valuation, where we considered the propositions true under *all* the minimal valuations satisfying the premises. Such a policy for amalgamating multiple outputs is usually called the *sceptical* one, and is expressed in the following definition.

Definition: Sceptical Default-Rule Consequence

- We define *default-rule consequence* (under the sceptical policy for amalgamating multiple outputs) by putting $C_R(A) = \bigcap \{C_{\langle R \rangle}(A) : \langle R \rangle \text{ a finite or } \omega\text{-ordering of } R\}$. In relational notation: $A \sim_R x$ iff $A \sim_{\langle R \rangle} x$ for every such ordering. As we have remarked, it coincides with what is usually known as ‘consequence using normal Reiter default rules with the sceptical policy on extensions’.

Since the default-rule operations $C_{\langle R \rangle}$ are in general nonmonotonic, it follows that the sceptical operations C_R are also in general nonmonotonic. Indeed, the simple example given above may be used again, for there R is a singleton and so $C_{\langle R \rangle} = C_R$.

Continuing on the negative side, we note that default-rule operations C_R may also fail cautious monotony. This was noticed by Makinson (1989) with a simple example, although with hindsight one can also use example 6.1 of Reiter (1980) to the same end. The following even simpler example is taken from Makinson (1994) section 3.2. Let R consist of three rules: (t,a) , (a,b) , $(b,-a)$ where t is a tautology. It is easy to check that no matter what the ordering of R , the premise set $\{t\}$ has the same set extension $Cn(a,b)$ under that ordering. On the other hand, $\{b\}$ has two extensions: $Cn(a,b)$ under e.g. the ordering as listed, and $Cn(-b)$ under e.g. the reverse ordering. Thus $t \sim_{\langle R \rangle} a$ and $t \sim_{\langle R \rangle} b$ but $b \not\sim_{\langle R \rangle} a$, where $\langle R \rangle$ is the reverse ordering. Likewise for the sceptical relation \sim_R we have the same pattern. So both of them fail cautious monotony.

Still on the negative side, as we would expect from the behaviour of the underlying pivotal-rule operations, default-rule operations C_R may also fail the rule of disjunction in the premises (OR). Indeed, the same trivial example may be used to illustrate the failure.

Finally, default-rule operations C_R may fail compactness. This can be shown by taking the same example used in section 2.2 to illustrate non-compactness for default-assumption consequence and translating it into the language of rules, replacing the set K of background assumptions by the corresponding set of rules with tautologous bodies, i.e. by putting $R = \{(t,k) : k \in K\}$.

On the positive side, the default-rule operations $C_{\langle R \rangle}$ and their ‘meet’ versions C_R are of course supraclassical, and it is easy to check that they satisfy conditions right weakening (RW) and

left classical equivalence (LCE). They also satisfy cumulative transitivity, as was in effect proven by Reiter (1980), and verified explicitly in Makinson (1994).

Looking now at the relationship between classical, pivotal-rule, and default-rule consequence, we see a familiar pattern reappearing. The order of construction was: first classical consequence Cn , then pivotal-rule consequence Cn_R , then default-rule consequence C_R . The monotonic Cn_R is thus a conceptual bridge between the classical and nonmonotonic operations. However, the order of inclusion between the operations is different: we have $Cn \leq C_R \leq C_{\langle R \rangle} \leq Cn_{\langle R \rangle} = Cn_R$. In brief, picking out the key items, $Cn \leq C_R \leq Cn_R$. In other words, classical consequence Cn gives a lower bound on our default operations C_R , while pivotal-rule consequence Cn_R gives an upper bound, with the non-monotonic operations intermediate between the two.

In the case that the set $heads(R)$, consisting of all heads of rules in R , is consistent with the premise set A , we have $C_R(A) = Cn_R(A)$, reaching the upper bound. For in that case the consistency check in the inductive definition of an extension is always satisfied, so that the unique extension of A under R is the least superset X of A such that both $Cn(X) \subseteq X$ and $R(X) \subseteq X$, i.e. $Cn_R(A)$. At the opposite extreme, when every head of every rule in R is inconsistent with A , we have $C_R(A) = A_0 = Cn(A)$, down at the lower bound.

4.3. Generalizations and Variants

We begin with the best-known generalization of the normal default-rule consequence relations that were defined in the preceding section – its broadening to cover what are called 'non-normal' default rules.

Non-normal Default Rules

In his paper of 1980, Reiter already considered default rules of a much more general kind, termed 'non-normal'. They are no longer pairs (a,x) but triples (a,P,x) , where a and x are Boolean formulae and P is a finite, possibly empty, set of formulae.

As already mentioned in the preceding section, Reiter calls a the *prerequisite* of the rule, and x its *conclusion*; we will also continue to call them *body* and *head* respectively. The new ingredient is the set P . Reiter calls its elements the *justifications* of the rule. We would prefer to call them *constraints*, because their only role is to figure in consistency checks; but in deference to tradition we continue to use Reiter's terminology.

Normal rules, used in the preceding section, may be identified with those triples (a,P,x) in which P consists of the head x alone. Plain rules without consistency checks, as studied in section 4.1, may be identified with triples (a,P,x) in which P is empty.

In allowing the set P of justifications to be empty we are following a convention that is by now consensual. Strictly speaking, Reiter himself required all rules to have a non-empty set of justifications; but allowing the empty set as limiting case facilitates our formulations.

The essential idea involved in applying a non-normal default rule $r = (a,P,x)$ is to pass from a to x whenever this can be done without leading to violation of any elements of P . Notoriously, this turns out to be an intricate notion, and admits many variations. We will focus on the

version of Reiter, and will continue to use an inductive approach rather than one in terms of fixpoints. In effect, we have two problems to face in constructing an inductive definition: how to take account of multiplicity now that we are in the broadened context of non-normal default rules, and how to deal with possible non-existence of extensions.

We continue to handle the first problem by means of well-orderings of the rule-set, as we did in the normal case, but with the assistance of auxiliary sets R_n whose job is to track the rules applied in the course of the induction. This device has been used for various purposes by a number of authors, e.g. Łukaszewicz (1984/8) and (1990), Brewka (1991), and Rimmel as reported in Marek and Truszczyński (1993), and was employed to obtain essentially the same result in Antoniou (1997).

We deal with the second problem by using an abort clause at an appropriate point in the definition. The idea is to insert a sub-clause into the construction of A_{n+1} , telling us that in certain situations A_{n+1} does not exist and the construction self-destructs. For the definition to be a genuine induction, the abort clause must not refer to what is still to be constructed, but only to what has already been built.

Definition: Ordered Default-Rule Consequence with Non-Normal Rules

- As before, we fix a well-ordering $\langle R \rangle$ of the set R of rules, of order type at most ω .
- As before, we put $C_{\langle R \rangle}(A) = \cup\{A_n : n < \omega\}$, and we set $A_0 = Cn(A)$.
- Simultaneously with the sets A_n we define a sequence of sets R_n to record the rules applied in the course of the induction. We set $R_0 = \emptyset$.
- For the the induction step defining A_{n+1} and R_{n+1} , we proceed as follows, where $just(R_n)$ stands for the set of all justifications of rules in R_n :
 - *Case 1.* Suppose there is a rule $r = (a, P, x)$ in R such that $a \in A_n$, $r \notin R_n$, and A_n is consistent with each separate $p \in P$. Take the first such rule, and consider two sub-cases.
 - *Sub-case 1.1:* Suppose that also $A_n \cup \{x\}$ is consistent with each separate $p \in just(R_n) \cup P$. Then we put $A_{n+1} = Cn(A_n \cup \{x\})$ and $R_{n+1} = R_n \cup \{r\}$.
 - *Sub-case 1.2:* Suppose otherwise. Then the construction aborts in the sense that A_{n+1} and R_{n+1} are left undefined and we continue no further, taking all subsequent A_m and the union $\cup\{A_n : n < \omega\}$ alias $C_{\langle R \rangle}(A)$ also to be undefined.
 - *Case 2.* Suppose there is no rule as required for case 1. Then we put $A_{n+1} = A_n$ and $R_{n+1} = R_n$.

Following the terminology of Reiter, we will often refer to $C_{\langle R \rangle}(A)$, when it exists, as an *extension*, more fully as the extension of A generated by the ordered set $\langle R \rangle$ of rules.

Roughly speaking, the entry condition for Case 1 tells us that if everything is in order when we compare the part of the extension so far constructed with the justifications of the current rule, then we are forced to *try* to apply the rule. We cannot back out. But, having got so far, if the further test for entry into Sub-case 1.2 is not satisfied, then we are *condemned* – we are shunted into Sub-case 2.2 and the construction aborts. This second test takes into account the

head of the current rule as well as the part of the extension constructed so far, and past as well as present justifications.

As mentioned in the previous section, the first inductive definition of normal Reiter extensions was given by Brewka (1994). For non-normal systems, is implicit in the 'operational semantics' for default logic devised by Antoniou (1997). Although Antoniou presents matters in a procedural perspective using other terms, his construction is in effect inductive and, as he shows, equivalent to the Reiter fixpoint definition. The present formulation was given in Makinson (2003a).

Examples

To reduce notation, in these examples when P is a singleton $\{p\}$ we write the rule $(a, \{p\}, x)$ as (a, p, x) .

1. Consider the simple example where $A = \{a\}$ and R consists of the single triple $(a, \neg x, x)$, so that there is just one well-ordering of R . As always, $A_0 = Cn(a)$. To identify A_1 , note that the premise a of the rule is in A_0 , the rule itself is not in $R_0 = \emptyset$, and its justification $\neg x$ is consistent with A_0 . Thus we are in Case 1. On the other hand, the conclusion x of the rule is inconsistent with the justification $\neg x$, and so we are in Sub-case 1.2. Thus A_1 is undefined and the construction aborts. As R consists of just one rule, with only one possible well-ordering, this means that the example has no extensions.

2. It is instructive to compare this with the example where $A = \{a\}$ and R consists of the single triple $r = (a, \neg a, x)$, so that the justification now contradicts the premise of the rule rather than its conclusion. Again, $A_0 = Cn(a)$. To identify A_1 , we observe that the premise a of r is in A_0 and $r \notin R_0 = \emptyset$, but the justification $\neg a$ of r is inconsistent with A_0 . Thus we are in Case 2 and hence $A_1 = A_0 = Cn(a)$.

3. Finally, we look at the example where $A = \emptyset$ and R contains two rules $r_1 = (t, q, x \wedge \neg p)$ and $r_2 = (t, p, x)$ where t is a tautology. If we order the rules as numbered, then we get the extension $Cn(x \wedge \neg p)$. If we reverse the order, the construction of an extension aborts.

If one is fanatic about parsimony, one can streamline the above definition a little. A straightforward induction tells us that the sets A_n are fully determined by the corresponding sets R_n ; specifically, $A_n = Cn(A \cup heads(R_n))$, where $heads(R_n)$ is the set of heads of rules in R_n . One can therefore define just the sequence R_n inductively, then use the equality to get the corresponding A_n .

As in the case of normal rules, the inductive definition gives us exactly the extensions in the sense of Reiter. In other words:

Theorem 4.3-1

Let R be any set of (normal or non-normal) default rules, and A any premise set. Then the Reiter fixpoint extensions of A under R are precisely the sets $C_{\langle R \rangle}(A)$ for well-orderings $\langle R \rangle$ of R of order-type at most ω . That is:

(a) For every well-ordering $\langle R \rangle$ of R of order-type $\leq \omega$, if the set $C_{\langle R \rangle}(A)$ is well-defined then it is a Reiter extension of A under R .

(b) Every Reiter extension of A under R is a well-defined set $C_{\langle R \rangle}(A)$ for some well-ordering $\langle R \rangle$ of R of order-type $\leq \omega$.

As the proof is quite intricate and takes us out of the main flow of our exposition, we place it in an Appendix.

Other Variants in the Reiter Tradition

Many variants of the Reiter default systems have been studied in the literature. We mention a few of the most important among them.

Some vary the *policy for amalgamating multiple extensions*. One idea is to use *partial meets* rather than full meets when managing multiple extensions. Recall that given a set R of rules, the 'sceptical' operation C_R was defined as the meet of the default-rule operations $C_{\langle R \rangle}$ for all possible well-orderings $\langle R \rangle$ of R . That is, we put $C_R(A) = \bigcap \{C_{\langle R \rangle}(A) : \langle R \rangle \text{ an ordering of } R\}$. As in the analogous cases of default-assumption and default-valuation consequence, one could replace this full meet by a partial meet, intersecting only a selected subfamily of the sets $C_{\langle R \rangle}$.

Indeed, there are several ways in which this may be done. One would be to allow only those orderings $\langle R \rangle$ that include some given transitive irreflexive relation $<$ over R . In this way we obtain the *prioritized default logic* (acronym PDL) defined by Brewka 1994. Another approach would be to introduce a preference relation between orderings, and intersect only those sets $C_{\langle R \rangle}(A)$ such that $\langle R \rangle$ is a *minimal* ordering of R under that preference relation. Other devices are also possible; see e.g. Tan and Treur (1992).

Some other variants are motivated by examples that, arguably, are mistreated by the Reiter systems. Two are particularly well known.

Poole (1989) suggested that when building an extension we should require *joint consistency* of the justifications, rather than their separate consistency, with other material. It is easy to see how the definitions of $C_{\langle R \rangle}(A)$ given above, for normal and for non-normal rules, may be modified in this way – just replace the words 'each separate' with 'the conjunction of all' in them. Alternatively, one can simply replace the finite justification set of each rule by its conjunction and then apply the definition without modification.

Łukaszewicz (1984/8) and (1990) expressed dissatisfaction with the way in which the Reiter construction can abort in the context of non-normal rules, leaving us without extensions. He proposed a revised formulation of Reiter's fixpoint definition of an extension, which *guarantees that extensions always exist* even when the default rules are non-normal. Reiter extensions are always Łukaszewicz extensions, but not conversely.

The idea may be seen most clearly in terms of our inductive approach. Whenever our inductive generation of a Reiter extension of A under $\langle R \rangle$ succeeds, then it is also a Łukaszewicz extension. If on the other hand it aborts, say in the passage from A_n to A_{n+1} , then we also get A_n as a Łukaszewicz extension. We can obtain the Łukaszewicz extensions by suitably simplifying the induction clause of our definition. We drop the Sub-cases in the induction step, and reformulate Case 1 as follows: "Suppose there is a rule $r = (a, P, x)$ in R such that $a \in A_n$, $r \notin R_n$ and $A_n \cup \{x\}$ is consistent with each $p \in \text{just}(R_n) \cup P$. Then we take the first such rule, and put $A_{n+1} = Cn(A_n \cup \{x\})$. We also set $R_{n+1} = R_n \cup \{r\}$."

Which of the definitions – of Reiter, Poole, Łukasiewicz or other – gives the right treatment of the controverted examples? In the author's view, it would be misleading to insist on this question, for it has no answer. There is no 'correct' account of Poole's 'broken-arm' example, Łukasiewicz' 'fishing' examples, or others like them, for intuitions are too variable and confused to provide clear verdicts. This is why we have not gone into the details of the examples themselves. It should also be said that the use of everyday language in examples can distort vision, for the language employed usually carries trails of unsuspected implicit assumptions influencing our intuitive judgement.

On the other hand, in specific contexts we may find it more convenient to use one variant than another, without erecting this preference into a doctrine of correctness.

We remarked in section 4.2 that although Reiter default-rule inference systems satisfy cumulative transitivity, they can fail the converse *cautious monotony*, even when all rules are normal. The same is true of the systems of Poole and Łukasiewicz, and indeed appears to be endemic to systems in which consistency constraints are imposed on the application of rules in derivations. Nevertheless, systems have been introduced that restore the validity of Cautious Monotony by the trick of 'upping the ontology'. The essential idea is that the premises and conclusions that occur in a derivation are no longer just Boolean formulae, but rather Boolean formulae indexed by traces of the justifications used in their derivations. Systems of this kind were developed by Brewka (1991) for default-rule reasoning, and by Brewka, Makinson and Schlechta (1991) for logic programming.

Computational aspects of fixpoint and inductive definitions

In the preceding section we explained why we prefer an inductive account to a fixpoint one: essentially, they are easier to grasp, work with, and communicate.

On the other hand, it may be argued that from a computational point of view the passage from fixpoints to induction is a step backwards rather than forwards. Every extension of A under default rules R is known to be of the form $Cn(A \cup heads(S))$ for some subset $S \subseteq R$. So when the set of default rules is finite with n elements, to determine the family of extensions of A under R , it suffices to take the 2^n subsets S of R and check them, one by one, to see whether $Cn(A \cup heads(S))$ is a fixpoint of the operator $\Gamma_{R,A}$ used to define extensions. If done haphazardly, this is a guess-and-test procedure; carried out systematically, it is a list-and-test one.

On the other hand there are $n!$ well-orderings of an n -element set R and to determine the extensions inductively it suffices to take these orderings one by one, apply the induction, see which abort before completion, and retain the surviving constructions. But as is notorious, we have $n! > 2^n$ for all $n \geq 4$. Indeed, the function $n!$ grows faster than 2^n in the sense that $\lim_{n \rightarrow \infty} (2^n/n!) = 0$; in the usual shorthand, $2^n = o(n!)$ (see e.g. Hein (2002) section 5.5.2). Thus, other things being equal, the inductive procedure requires much more time to calculate.

Prima facie, this suggests that using an inductive definition to calculate the set of all extensions of a given finite system (A,R) , is less efficient than using a fixpoint one. However, the situation is not quite so clear-cut, as there are some other relevant factors.

- The most important is that in practice we may not wish to find the set of *all* extensions. As pointed out by Brewka (1994), the specific problem in hand may naturally delimit a relatively small subset of the possible well-orderings – for example,

all those that include some nearly-complete ordering – and we may be interested only in the extensions determined by them.

- Even when we wish to determine the set of all possible extensions, there may be some mitigating factors. In particular, if the set of all well-orderings of R is itself suitably well-ordered, information about the construction using an earlier ordering may perhaps be recycled towards the construction based on a later ordering. We do not attempt to resolve these computational questions here.

Induction versus fixpoints

Our presentation of Reiter default inference and some of its variants suggests a general questions about the relationship between definition by fixpoints and using standard mathematical induction. When may the former be reduced to the latter?

We have seen how the two principal difficulties were overcome in the case of Reiter default logic. The multiplicity of extensions was handled by introducing an ordering $\langle R \rangle$ of R as an additional parameter (and, in the non-normal context, tracking rules used); their non-existence was dealt with by introducing an appropriate escape clause into the induction step. Similar steps work, as we remarked, for the variants of Poole and of Łukasiewicz. But what are the limits? Without attempting to answer this question in its generality, we mention a further case known to be positive, another that we conjecture to be positive, and a third that seems to be negative.

The construction of Theorem 4.3–1 may be *generalized* to cover also logic programs with negation as failure under the so-called 'answer-set semantics'. Such a program is a set of expressions of the form $x \leftarrow a_1, \dots, a_n, \text{not}(a_{n+1}), \dots, \text{not}(a_{n+m})$ where x and all the a_i are literals (i.e. elementary formulae or their negations). These expressions may be seen as rules about a perfectly classical though very restricted object language. More specifically, they may be read as non-normal default rules with the literal x as head (conclusion), the conjunction of the elementary letters a_1, \dots, a_n , as body (prerequisites, in Reiter's terminology), and the negated literals $\text{not}(a_{n+1}), \dots, \text{not}(a_{n+m})$ as corresponding to justifications $\neg a_{n+j}$ of the rule. Theorem 4.3–1 can be generalized also to cover such logic programs (Makinson 2004).

Logic programs have been generalised to allow disjunctions of literals in the head. These disjunctions are not understood in a simple truth-functional manner, but as presenting alternative possible conclusions open for choice. In Makinson (2004) it is conjectured that such programs may be reduced to families of programs without disjunctive heads, thereby permitting an inductive definition of their answer sets.

In the area of default logic, the 'weak extensions' of Marek and Truszczyński (1989, 1993) are of interest, both in their own right and because, as those authors have shown, they correspond to the expansions of autoepistemic logic. It appears difficult to give an inductive account of the fixpoints defining these extensions. Intuitively, they are not visibly grounded in their premise-sets. Mathematically, they can exceed their unconstrained counterparts. Recall that for Reiter systems, whatever can be obtained by applying rules *with* their consistency constraints can already be obtained by applying the same rules *without* any consistency constraint. In other words, when E is any Reiter default extension of the premise set A under a rule-set R , then $E \subseteq Cn_R(A)$, where $Cn_R(A)$ is the (unique) closure of A under both classical consequence and all of the plain rules obtained by dropping the constraint sets from rules in R . This inclusion fails for weak extensions. To take a trivial example, when $R = \{(p, p, p)\}$ and A

$= \emptyset$, then both $Cn(\emptyset)$ and $Cn(p)$ are weak extensions E of A under R . Thus for $E = Cn(p)$, the inclusion $E \subseteq Cn_R(A)$ fails.

However, neither of these two considerations is conclusive. The former is little more than a vague feeling; and while the latter is a mathematical fact, it does not constitute a proof of the point at issue. Indeed, there is a trivial sense in which weak extensions *can* be given an inductive definition. Instead of inducing modulo well-orderings of the rule set R , do so modulo its subsets. For a given $S \subseteq R$, it is trivial to give a two-step inductive definition that puts $A_0 = Cn(A)$; puts $A_1 = S$ if S is a fixpoint of the Marek and Truszczyński version of the Γ function and aborts otherwise; and if A_1 is well-defined, puts each following $A_n = A_1$. Formally speaking, this is an inductive definition, but evidently it is quite uninteresting. The problem therefore has to be reformulated: is there a non-trivial and interesting inductive definition of weak extensions? Evidently, the terms of the question are not very precise.

Maxfamilies of Rules

A quite different approach to default rules is to use consistency constraints to cut back on the set R itself, rather than restrict its step-by-step application. Following the pattern for default-assumption consequence in section 2, one can form a family of suitably maximal subsets of R .

Consider any set R of normal default rules (a,x) . For each premise set A we consider the family of all maximal subsets $S \subseteq R$ such that the pivotal-rule consequence $Cn_S(A)$ is consistent. This is called the *maxfamily* of A under R . We define $C_{\underline{R}}(A)$ (this time with subscript underlined) to be the intersection of the sets $Cn_S(A)$ for all S in the maxfamily of A .

These operations are not the same as the default-rule consequences C_R for normal rules given in section 4.2. To illustrate the difference, consider the Möbius strip example, where $R = \{(p,q), (q,r), (r,\neg p)\}$. On the one hand, the premise set $\{p\}$ gives rise to exactly one default-rule extension, namely $Cn(\{p,q,r\})$. On the other hand, we have three maxfamilies S , namely the three two-element subsets of R . These yield three sets $Cn_S(A)$ – namely $Cn(p)$, $Cn(\{p,q\})$, and $Cn(\{p,q,r\})$, which form a chain under set inclusion and have just $Cn(p)$ as their intersection. In general, there are more maxfamilies than extensions, and so the intersection of the maxfamilies will be smaller than the intersection of the extensions.

It should be noted, however, that the output obtained still differs from the one we saw when we used maximal subsets of a background assumption set K . If we translate the rules in R into material conditional formulae in K , the possibility of applying modus tollens as well as modus ponens to material conditional propositions gives us larger extensions, with a larger intersection. This can be seen in detail by comparing the Möbius strip example above with its counterpart in section 2.2.

While the distinction between the default-rule consequence operations C_R and the maxfamily consequence operations $C_{\underline{R}}$ is subtle, it leads to significantly different outcomes. In effect, we are looking at two strategies for guarding against contradictions when applying rules to a set of premises. One restricts the *generating process*, giving us the operations $C_{\langle R \rangle}$ for well-orderings $\langle R \rangle$ of R . The other constrains the *generating apparatus*, giving us the operations $C_S(A)$ for sets $S \subseteq R$ in the maxfamily of A .

Maxfamily operations are studied by Makinson and van der Torre (2000) in a broader context where inputs (premises) are not necessarily authorized to reappear as outputs (conclusions).

5. Connections between Nonmonotonic and Probabilistic Inference

Up to now, we have not said a single word about probability, concentrating on getting a clear overview of the three main qualitative bridges to nonmonotonic consequence operations. It is now time to do so, for it is clear that there are connections, of some kind, between them.

On the one hand, there is an element of uncertainty in nonmonotonic inference: being supraclassical, there possible situations in which a conclusion is false although all premises are true. On the other hand, there is something nonmonotonic about probabilistic reasoning. To begin with, there is the notorious fact that the conditional probability of a proposition x given another one a , may fall, just as it may rise, when we are given additional information b to conjoin with a . In other words, we may have $p_{a \wedge b}(x) < p_a(x)$ where p_a is the conditionalization of the probability function p on the proposition a .

What then is the relation between probability and nonmonotonic logic? That is our principal question. But behind it lie a number of other questions that need to be untangled.

- What, after all, is the relation between probability and *classical* consequence?
- In what ways do probabilistic considerations lead to supraclassical inference relations, and how do they compare with the qualitatively defined ones that we have considered so far?
- Are all the probabilistic consequence relations nonmonotonic, or are there any natural probabilistic bridge-systems, akin to the qualitative ones outlined earlier?

In the present section 5.1 we recall and comment on the basic concepts and principles of probability theory, focussing on the Kolmogorov axioms and the notion of conditional probability. In section 5.2 we examine the relationship between probability and classical logic, showing in particular that classical consequence may be characterized in probabilistic terms – indeed, in a number of different but equivalent ways. In section 5.3 we show how these equivalent characterizations come apart from each other when we rework them to generate supraclassical consequence operations, with the most interesting one among them becoming nonmonotonic. In all cases, however, the supraclassical operations are quite different from those that we have so far considered; in particular, they fail the principle of conjunction in the conclusion. In section 5.4 we discuss this failure, and look at ways in which the probabilistic approach may be reconstructed to make it behave more like the qualitative one.

5.1. Basic concepts and axioms for probability

The Kolmogorov Postulates

As with most branches of mathematics, probability was used and studied long before it was axiomatized. The standard axiomatization is that of the Russian mathematician Kolmogorov,

published in German in 1933 but most accessible in the translation Kolmogorov (1950). It defines the notion of a (finitely additive) *probability function* or, as it is often called, *probability measure*.

Like all functions, these have a domain and a range. The range is the real interval $[0,1]$: every probability function is into the set of all real numbers from 0 to 1 inclusive. The domain can be expressed in several different ways. Probability theorists usually take it to be an arbitrary *field of sets*, that is, as a non-empty family of subsets of some fixed set X , such that whenever A, B are in the family then so too are $A \cap B$, $A \cup B$ and $X - A$. Equivalently, the domain may be taken to be an arbitrary *Boolean algebra*, for every field of sets is a Boolean algebra and conversely, by Stone's representation theorem, every Boolean algebra is isomorphic to a field of sets. A less common alternative, but which we will follow here to make more transparent our comparison with logic, is to take the domain to be the set of all Boolean formulae of a propositional language. This may be done because the quotient structure of such a language under the relation of classical equivalence is a Boolean algebra, and conversely every Boolean algebra is a homomorphic image of such a quotient structure for some propositional language with sufficiently many elementary letters.

There are many minor variations in the formulation of the axioms, equivalent when the axioms are taken together but not when the axioms are considered separately. It will be convenient for us to use the following one.

Definition: Probability Function

A probability function is any function defined on the formulae of a language closed under the Boolean connectives, into the real numbers, such that:

- (K1) $0 \leq p(x) \leq 1$
- (K2) $p(x) = 1$ for some formula x
- (K3) $p(x) \leq p(y)$ whenever $x \vdash y$
- (K4) $p(x \vee y) = p(x) + p(y)$ whenever $x \vdash \neg y$.

Here Cn alias \vdash is, as always, classical consequence. Thus postulate (K1) tells us that the range of any probability function is a subset of the real interval $[0,1]$. (K2) tells us that 1 is in that range. (K3) says that $p(x) \leq p(y)$ whenever x classically implies y ; and (K4) tells us that $p(x \vee y) \leq p(x) + p(y)$ whenever x is classically inconsistent with y .

Strictly speaking, these postulates, known as the Kolmogorov axioms, define the notion of a *finitely additive probability function*. They can be extended by strengthening (K4) to tell us something about infinite sums of probabilities. To do this one needs to enrich the background domain, working either with a field of sets closed under countable (or more) unions, or with a Boolean algebra likewise closed under infinite joins, or with a propositional language admitting infinite disjunctions. Such an extension is indeed necessary for many developments of probability theory, but will not be needed for our purposes.

Some further comments on the Kolmogorov postulates will help in our task of understanding the relations between probability and logic.

- The formulation of the axioms assumes concepts of classical logic – indeed two out of the four make explicit use of the notion of classical consequence. Of course, this dependence disappears, or rather takes a different form, if probability theory is formulated in terms of Boolean algebras or fields of sets. The axioms then make use of the subset relation \subseteq between sets or the corresponding relation \leq in Boolean algebra. Even when one retains Boolean formulae as domain, one can eliminate explicit reference to classical consequence by using a rather more extensive list of axioms, devised by Stalnaker (1970). But we have no quarrel with classical logic, and make free use of it whenever convenient.
- The real interval $[0,1]$ is infinitely larger and much more finely structured than is the two-element set $\{0,1\}$. Thus a probability function $p: L \rightarrow [0,1]$ can be much richer in information than a Boolean valuation $v: L \rightarrow \{0,1\}$. Accordingly probability theory is much more complex than truth-functional logic.
- Nevertheless, the Boolean valuations of classical logic are themselves limiting cases of probability functions, as they satisfy all four of the axioms, as can easily be verified.
- Conversely, any probability function whose values are all in the two-element set $\{0,1\}$ is a Boolean valuation. This is also easy to verify.
- Finally, it should be noted that the last two postulates are both conditionals. They are not biconditionals. Nor do their converses follow from what is given: each converse admits exceptions for suitably constructed probability functions.
- From the first three axioms it follows that $p(x) = 1$ whenever $x \in Cn(\emptyset)$. Here too the converse does not hold. The converse is sometimes added as a further condition defining a special class of probability functions, but that will not concern us here.

Unique Extension Properties for Logic and for Probability

Despite the close connections, there are also important differences. For those whose background is in logic, there is a habit of mind that must be broken from the beginning.

We know that in classical propositional logic, any assignment of truth-values to elementary letters can be extended in a unique way to a Boolean valuation over all formulae. But probability is not like that. There are assignments to elementary letters of real numbers in the interval $[0,1]$ that can be extended in more than one way to a function p on the entire Boolean language into $[0,1]$ that satisfies the Kolmogorov axioms.

For example, consider the language with just two elementary letters. To avoid possible confusion, as we are using the letter p in this section to stand for an arbitrary probability function, we will no longer use it for an elementary letter of a Boolean language, so we call the two letters of the example q, r . Consider the assignment f that puts $f(q) = f(r) = 0.5$. There are many ways in which this can be extended to a probability function $p: L \rightarrow [0,1]$. In one of them, call it p_1 , we have $p_1(q \wedge r) = p_1(\neg q \wedge \neg r) = 0.5$ and $p_1(q \wedge \neg r) = p_1(\neg q \wedge r) = 0$; while in another, call it p_2 , we have $p_2(q \wedge r) = p_2(q \wedge \neg r) = p_2(\neg q \wedge r) = p_2(\neg q \wedge \neg r) = 0.25$.

The same point can be put in another way. The example shows that probability functions are not compositional: there is no function $*$: $[0,1]^2 \rightarrow [0,1]$ such that for all probability functions p and all formulae a,b , $p(a \wedge b) = p(a)*p(b)$. For in the example we have $p_1(q) = p_2(q)$ and also $p_1(r) = p_2(r) = 0.5$, so for any function $*$ whatsoever, $p_1(q)*p_1(r) = p_2(q)*p_2(r)$. If compositionality held, we would thus have $p_1(q \wedge r) = p_1(q)*p_1(r) = p_2(q)*p_2(r) = p_2(q \wedge r)$, whereas in fact $p_1(q \wedge r) = 0.5 \neq 0.25 = p_2(q \wedge r)$.

Yet another way of putting this is that probability functions are not many-valued truth-functions, created by expanding the set of truth-values from $\{0,1\}$ with more elements – even if we take as our truth-values the uncountably many elements of the real interval $[0,1]$. Whatever the number of values, and whatever the rules of truth for connectives, many-valued truth-functions give compound formulae values that are uniquely determined by the values of their constituents. Failure to see this clearly has led to many quixotic enterprises

Nevertheless, when a probability function is defined over a finitely generated Boolean language, it does have a quite different 'unique extension property'. The simplest way of formulating this property is in terms of finite Boolean algebras; but to maintain the uniformity of our presentation we will explain it within our logical framework.

Consider classical propositional logic formulated with just n elementary letters q_1, \dots, q_n . A *state-description* for that language is understood to be any conjunction of n terms in a fixed order, the i -th term being either q_i or its negation $\neg q_i$. For example, in the language with just two elementary letters q,r , the four formulae $q \wedge r$, $q \wedge \neg r$, $\neg q \wedge r$, $\neg q \wedge \neg r$ mentioned above are its state-descriptions. In general, when there n elementary letters there will be 2^n state-descriptions. Further, every non-contradictory formula x is classically equivalent to a disjunction of a non-empty subset of the state-descriptions. Finally, the conjunction of any two distinct state-descriptions is a contradiction. State descriptions on the level of formulae correspond to atoms on the level of Boolean algebras.

Now consider any function f on the 2^n state-descriptions of L into the real interval $[0,1]$, such that their values sum to 1. Such a function is often called a *probability distribution*. Then we have the following 'unique extension property': *f can be extended in a unique way to a probability function $p: L \rightarrow [0,1]$ satisfying the Kolmogorov axioms.*

The proof is simple. For each non-contradictory formula x , take the disjunction $s_1 \vee \dots \vee s_k$ of all the state descriptions that classically imply it, and put $p(x) = f(s_1) + \dots + f(s_k)$. When x is contradictory, put $p(x) = 0$. Then it is easy to verify that p is a probability function, and that any probability function on L that agrees with f on the state-descriptions is identical with p .

This is an important feature of probability on finitely generated languages, and is used in many proofs and calculations. But it is also important to understand clearly the difference between this unique extension property and the homonymous property for Boolean valuations. One is based on state descriptions, the other on elementary letters.

Unfortunately, it is all too easy to be led into confusion by a terminological conflict between different communities. Algebraists use the term *atom* to stand for minimal non-zero elements of a Boolean algebra, and thus for state-descriptions (strictly, for the equivalence classes of state-descriptions under the relation of classical equivalence). From the algebraic point of view, elementary letters (strictly their equivalence classes under classical consequence) are the free generators of a free Boolean algebra. On the other hand, many computer scientists and

some logicians, particularly those working in logic programming, call the elementary letters of propositional logic 'atomic formulae' or even 'atoms', because they are syntactically indivisible units.

The two terminologies should not be confused. We will always use the term 'atom' in the algebraist's sense, and reserve the term 'elementary letter' for the syntactically simple units of a propositional language. Readers will end up speaking the language of the tribe they associate with, but should also be alert to the potential confusion.

Conditionalization

One of the most basic and useful concepts of probability theory is that of *conditional probability*, or *conditionalization*. It will also play a central role in our story. It is a limited kind of revision. If p is a probability function giving proposition a some value in the interval $[0,1]$, and we learn that a is in fact true, how should we modify p to take account of that information? The traditional answer is that, in the case $p(a) \neq 0$, we should pass from p to the function p_a defined by putting $p_a(x) = p(a \wedge x)/p(a)$. The function p_a is called the *conditionalization of p on a* .

Some comments may help appreciate the contours of this very important concept. First, we note three important mathematical properties of conditionalization.

- Given that p is a probability function and $p(a) \neq 0$, p_a is evidently also a function on the same Boolean language into the interval $[0,1]$. More significantly, it is itself a probability function, i.e. satisfies the Kolmogorov axioms, as is easily checked from the definition with the help of some elementary arithmetic.
- There is a sense in which the passage from p to p_a is the most conservative change possible that brings the probability of a up to 1 (and so by axioms (K3) and (K1) does the same for all b with $a \vdash b$): it preserves untouched the ratios of probabilities of all propositions b with $b \vdash a$. To be precise, whenever p is a probability function and $p(a) \neq 0$ then it is not difficult to show that p_a is the unique probability function satisfying the following two conditions: (1) $p_a(a) = 1$, (2) for all b, b' , if $b \vdash a$ and $b' \vdash a$ then $p_a(b)/p_a(b') = p(b)/p(b')$.
- Conditionalization is also conservative in another sense. It never diminishes the 'full belief' set $B_p = \{x: p(x) = 1\}$ determined by the probability function. In other words, when p_a is well defined then for any formula x , if $p(x) = 1$ then $p_a(x) = 1$. Using terminology from the logic of belief change, conditionalization serves as an *expansion* of the full belief set B_p , rather than a revision of B_p , even though it can be said to *revise* p itself.

On the other hand, if we define 'threshold' or 'partial' belief sets $B_{p,k}$ by setting $B_{p,k} = \{x: p(x) \geq k\}$ where k is a fixed real less than 1, then they can lose as well as gain elements when p is conditionalized. Some ideas for redefining conditionalization to allow B_p itself to lose as well as gain when p is conditionalized, are canvassed in chapter 5 of Gärdenfors (1988).

The limited range of the definition also deserves note. The conditionalization p_a of p on a is *defined only when $p(a) \neq 0$* , i.e. p_a is left undefined when $p(a) = 0$. One cannot simply drop

this restriction while retaining intact the defining equation, because notoriously division by zero is itself undefined in arithmetic. This can be bothersome when constructing proofs, because it obliges us to check that the condition is satisfied whenever we wish to define a conditional probability function, to be sure that it is well-defined. It is also rather disappointing intuitively. Surely we should also be able to revise a probability function on learning new facts, even when those facts contradict the function in a radical way. In other words, we would like to be able to revise p by a even when $p(a) = 0$.

For these reasons, there have been proposals to *extend the definition to ensure that p_a is defined in all cases* while agreeing with the usual definition in the principal case that $p(a) \neq 0$. A simple device sometimes used is to introduce the absurd probability function p^\perp , defined by putting $p^\perp(x) = 1$ for all propositions x . Once that is done, p_a may be set as p^\perp in the limiting case that $p(a) = 0$. The cost of this manoeuvre is that the postulate (K4) must then be weakened, placing it under the condition that $p \neq p^\perp$. In any case, the irritating check to determine whether $p(a) \neq 0$ does not really disappear. It still needs to be made when defining a conditional function, not to determine whether the function is well-defined but in order to determine whether it is an ordinary one, with one kind of behaviour, or the absurd one, with a quite different behaviour. More discriminating definitions have also been suggested, allowing different values for $p_a(x)$ when $p(a) = 0$ – see for example the discussion in chapter 5 of Gärdenfors (1988). However we will not need to use such extensions, and will remain with the traditional definition.

Another kind of generalization is also possible. Conditionalization is appropriate when we wish to give a proposition a probability 1 upon learning that it is true. Evidently, this is a special case of a more general question: how should we revise a probability function p to a new one $p_{a=t}$ that ascribes probability t to a given proposition a , where t is an arbitrary element of the interval $[0,1]$? A well-known answer to this problem was introduced by Jeffrey (1965) and is called *Jeffrey conditionalization*. It is defined by the equality $p_{a=t}(x) = t.p_a(x) + (1-t).p_{\neg a}(x)$, where p_a and $p_{\neg a}$ are well-defined Bayesian conditionalizations. To mark the contrast, the plain version is sometimes called *Bayesian conditionalization*. It can be seen as Jeffrey conditionalization in the limiting case that $t = 1$. However once again we will not need to use this kind of generalization, and we will remain with the initial Bayesian version.

Conditional Probability versus Probability of a Conditional

We end this section with a remark on notation, which at the same time brings to the surface a very deep conceptual issue. We have written $p_a(x)$ for the function defined by the ratio $p(a \wedge x)/p(a)$. It is common, indeed customary, to write $p(x|a)$ for the same function. However, such a notation can be highly misleading.

If we start from a probability function $p: L \rightarrow [0,1]$, then clearly its conditionalization as defined above is *a probability function on the same domain L* . It is not an extension of the initial probability function p to a wider domain consisting of the formulae in L plus new 'conditional objects' $x|a$ for formulae a, x in L . But that is just what the notation $p(x|a)$ suggests. For if we read the functional expression in the usual way, it tells us that $p(x|a)$ is of the form $p(\bullet)$. In other words, it presents the conditionalization of p as the same function p applied to a new argument, the 'conditional object' $x|a$.

Once this notation is used, it leads almost irresistibly to the feeling that these 'conditional objects' have some kind of existence, and are in effect, conditional propositions. In other

words, it is felt that there must be a conditional connective \Rightarrow such that for every probability function p on the language with $p(a) \neq 0$, the conditional probability $p_a(x)$ alias $p(x|a)$ of x given a is equal to the probability $p(a \Rightarrow x)$ of the conditional proposition $a \Rightarrow x$. This is sometimes known as the *CP=PC* property, where the acronym abbreviates ‘conditional probability = probability of a conditional’.

Such is the cognitive force of a bad notation that this feeling became a conviction among some investigators concerned with probability theory. The problem, as they saw it, was to find what, exactly, this conditional connective is. It is easy to see that it cannot be the ordinary truth-functional conditional $a \rightarrow x$, whose behaviour is incompatible with the properties desired. So, the line of thought went, it must be some kind of non-classical conditional.

This project received a severe setback from a celebrated result of David Lewis (1976). Lewis showed the impossibility, in a language that maintains the classical logic of the truth-functional connectives \neg, \wedge, \vee , of defining or even introducing a conditional connective with the *CP=PC* property. Lewis’ paper was followed by others, establishing the same point in different ways. Some of them have slightly stronger formulations – see e.g. the review Hájek (2001).

For most investigators, Lewis’ ‘impossibility theorem’ put an end to the dream of a conditional with the *CP=PC* property. But not all have accepted defeat. Following an idea going back to de Finetti (1936), it has been proposed to replace classical two-valued logic by a three-valued one, within which one may seek to define a conditional with the *CP=PC* property. In such a three-valued logic, certain of the familiar principles of the classical logic of \neg, \wedge, \vee will necessarily be lost. This program has been pursued by Dubois, Prade and colleagues in Europe, and by Calabrese in the USA. See e.g. Dubois and Prade (1994) and Calabrese (2003).

In the view of the present author, even if this program can be carried through without technical snags, to give up the simplicity of two-valued logic for the connectives \neg, \wedge, \vee in favour of a weaker three-valued logic is an unpalatable price to pay for a dubious gain. At best, we become able to identify the probability of a conditional with a conditional probability, but we give up the simplicity of classical propositional logic, and lose some of its basic principles. Few logicians would be willing to follow this expensive path.

In the author’s opinion, this issue illustrates a very general methodological point. It can be misleading to pose questions like the above as matters of doctrine, arguing that one point of view is correct and another incorrect. It can be a matter of cognitive convenience: we may think and talk coherently in various different ways, but some of them are ever so much more complicated and unmanageable than others.

5.2. Probabilistic characterizations of classical consequence

What does classical consequence look like from a probabilistic perspective? Can it be characterized in terms of probability functions? We say ‘characterized’ rather than ‘defined’ because, as noticed in the preceding section, the definition of a probability function, as given by the Kolmogorov axioms, already makes essential use of the relation of classical consequence, so that any reduction of the latter to the former would be circular. While reduction is not on the agenda, it nevertheless remains very interesting to know whether

characterization is possible, and more broadly how classical consequence appears in probabilistic terms. As well as giving new ways of looking at classical logic, it will help prepare the way for our analysis, in the following sections, of the relations between probabilistic and nonmonotonic inference.

As we have already noticed, the real interval $[0,1]$ is infinitely larger and much more finely structured than is the two-element set $\{0,1\}$. Thus a probability function $p: L \rightarrow [0,1]$ is much richer in content than a Boolean valuation $v: L \rightarrow \{0,1\}$, and accordingly probability theory is much more complex than truth-functional logic. One effect of this is that there is not just one way of characterizing classical consequence in probabilistic terms; there is a multiplicity of them. In this section we present some of the best known.

The first characterization is almost immediate from the Kolmogorov axioms, yet involves a subtlety that should not be overlooked.

Observation 5.2–1

Let a, x be any Boolean formulae. Then $a \dashv\vdash x$ holds iff for every probability function $p: L \rightarrow [0,1]$, $p(a) \leq p(x)$.

Proof

The implication from left to right is given by the Kolmogorov axiom (K3). Suppose that $a \dashv\vdash x$. Then there is a Boolean valuation $v: L \rightarrow \{0,1\}$ with $v(a) = 1$, $v(x) = 0$. As noted in section 5.1, v is itself a probability function, and clearly $v(a) > v(x)$.

Observation 5.2–1 should not be surprising, for it already has an analogue in the qualitative sphere. Quantifying over all probability functions is like quantifying over, say, all preferential models. Neither gives more than classical logic, since classical evaluation can be seen as a limiting case of evaluation in a preferential model, and likewise as a limiting case of probabilistic valuation.

It is important to note the positioning of the universal quantifier ‘for every probability function p ’ in Observation 5.2–1. We have not proven that for every probability function $p: L \rightarrow [0,1]$, $a \dashv\vdash x$ holds iff $p(a) \leq p(x)$, where the universal quantifier takes the whole biconditional as its scope. Indeed, the latter claim is wildly false. It does hold from left to right by (K3) again; but not from right to left. For example, let q, r be distinct elementary letters, and consider the Boolean valuation that puts $v(q) = 1$, $v(r) = 0$. Once again, v is itself a probability function and $v(q) > v(r)$. But clearly since q, r are distinct elementary letters we do not have $a \dashv\vdash x$.

This kind of pattern will repeat itself many times in what is to follow. In general, we must distinguish carefully between two different relationships, where the quantification $\forall p$ is understood to range over the domain of all probability functions.

- If every probability function $p: L \rightarrow [0,1]$ satisfies a certain condition, then so-and-so holds. Schematically: $(\forall p \varphi(p)) \rightarrow \psi$, where p does not occur free in ψ .

- For every probability function $p: L \rightarrow [0,1]$, if p satisfies a certain condition then so-and-so holds. Schematically: $\forall p (\varphi(p) \rightarrow \psi)$, where p does not occur free in ψ .

The difference is in the scope of the universal quantifier. The latter implies the former but not conversely. The right-to-left half of Observation 5.2-1 is of the form $(\forall p \varphi(p)) \rightarrow \psi$ and not of the form $\forall p (\varphi(p) \rightarrow \psi)$. In fact, the latter statement is false.

In addition to the probabilistic characterization of classical consequence given in Observation 5.2-1, several others are of interest. The following table adds to our first condition four more that do the same job. There is nothing sacred about the number – others may easily be formulated. But these five are particularly simple and well known.

Number	Condition
1	$p(a) \leq p(x)$
$2t$	if $p(a) \geq t$ then $p(x) \geq t$
$3t$	$p(a \rightarrow x) \geq t$
$4t$	$p_a(x) \geq t$ or $p(a) = 0$
5	$p(\neg x) \leq \Sigma(p(\neg a_i): i \leq n)$

In conditions (2t) through (4t), t is any real number $t \in [0,1]$ with $t \neq 0$, called the *threshold* of the condition. Thus to be strict, these are not single conditions, but families of conditions, each family having infinitely many members, one for each non-zero value of t . An important instance is the one where $t = 1$. In that case, condition (2t) for example says: if $p(a) = t$ then $p(x) = t$.

Conditions (1) and (2t) are straightforward. Condition (3t) could evidently be formulated equivalently as $p(\neg a \vee x) \geq t$, or as $p(a \wedge \neg x) \leq 1-t$, and in the limiting case that $t = 1$, as $p(\neg a \vee x) = 1$ or $p(a \wedge \neg x) = 0$. Condition (4t) makes use of the notion of (Bayesian) conditionalization, defined and discussed in the preceding section.

Condition (5) needs some explanation. It is understood for the situation where we may have finitely many premises a_1, \dots, a_n . Evidently, we could simply take the conjunction $a = \wedge a_i$ of all of them to do the work of the whole set. But when n is large, we may have information about the separate probabilities $p(a_n)$, but no direct knowledge of the probability $p(a)$ of their conjunction. In such a situation, condition (5) can be useful. Informally, it says that the improbability of the conclusion x is no greater than the sum of the improbabilities of the various separate premises.

This condition has been highlighted by Ernest Adams (1975, 1998). However, the reader should be careful with terminology. Adams paraphrases the condition as saying that the uncertainty of the conclusion x should be no greater than the sum of the uncertainties of the separate premises. This can be misleading since the term 'uncertainty' is already employed

with another meaning in probability theory. More recent expositions of his ideas, such as that in Edgington (2001), therefore use the term 'improbability'.

Theorem 5.2-2

Each of the quantified conditions $\forall p(1)$ through $\forall p(5)$, with p ranging over the set of all probability functions, characterizes classical consequence. They are thus also equivalent to each other. In full detail: let t be any real number in $[0,1]$ with $t \neq 0$. Let a, x be any formulae, and suppose for condition 5 that $a = a_1 \wedge \dots \wedge a_n$. Then $a \dashv\vdash x$ holds iff $\forall p(\phi)$ holds, where ϕ is any one of conditions 1, 2t, 3t, 4t, 5.

Proof

Two strategies of proof suggest themselves. One is to prove separately the equivalence of each of the probabilistic items to the logical one $a \dashv\vdash x$. Another is to construct a cycle of implications, with the logical condition $a \dashv\vdash x$ one of those in the cycle. The former uses more logic, the latter more probability theory. The former can be longer, for if there are n probabilistic conditions then there are n equivalences to verify, and if (as usual) most have to be proven as co-implications, then we have up to $2n$ implications to check. Using a cycle, there are only $n+1$ implications to prove.

We will follow a slightly mixed strategy, making use of the fact that we already know by Observation 5.2-1 that the logical condition $a \dashv\vdash x$ is equivalent to condition $\forall p(1)$. So we need only cycle around the probabilistic conditions in the table. As we do so, we take the opportunity to prove stronger implications whenever they hold, as they will be useful in the following section.

(i) We want to show $\forall p(1) \Rightarrow \forall p(2t)$. In fact we have the stronger implication $\forall p \in P (1 \Rightarrow 2t)$. Suppose that $p(a) \leq p(x)$. Then immediately if $t \leq p(a)$ then $t \leq p(x)$ and we are done.

(ii) We want to show $\forall p(2t) \Rightarrow \forall p(3t)$. It will be useful for later work to show a little more generally that $\forall p \in Q(2t) \Rightarrow \forall p \in Q(3t)$, where Q is any set of probability functions that is closed under conditionalization (i.e. whenever $p \in Q$ and $p(a) \neq 0$ then $p_a \in Q$). We argue contrapositively. Let Q be such a set and suppose that $p(a \rightarrow x) < t$ for some $p \in Q$. Then $p(a \wedge \neg x) = p(\neg(a \rightarrow x)) = 1 - p(a \rightarrow x) > 0$ so $p_{a \wedge \neg x}$ is well-defined. But clearly $p_{a \wedge \neg x}(a) = 1 \geq t$ while $p_{a \wedge \neg x}(x) = 0 < t$ as required.

(iii) We want to show $\forall p(3t) \Rightarrow \forall p(4t)$. Again we show that $\forall p \in Q(3t) \Rightarrow \forall p \in Q(4t)$, where Q is any set of probability functions closed under conditionalization. We argue contrapositively. Let Q be such a set and suppose that $p(a) \neq 0$ and $p_a(x) < t$ for some $p \in Q$. Clearly $p_a(\neg a) = 0$. Hence $p_a(a \rightarrow x) = p(\neg a \vee x) \leq p_a(\neg a) + p_a(x) = p_a(x) < t$ as desired.

(iv) We want to show $\forall p(4t) \Rightarrow \forall p(5)$. Again we show that $\forall p \in Q(4t) \Rightarrow \forall p \in Q(5)$, where Q is any set of probability functions closed under conditionalization. Again we argue contrapositively. Let Q be such a set, let $p \in Q$, and suppose that $p(\neg x) > \Sigma(p(\neg a_i): i \leq n)$, where $a = a_1 \wedge \dots \wedge a_n$. Now $\Sigma(p(\neg a_i): i \leq n) \geq p(\neg a_1 \vee \dots \vee \neg a_n) = p(\neg(a_1 \wedge \dots \wedge a_n)) = p(\neg a)$. So $p(\neg x) > p(\neg a)$ and thus $p(a) > p(x) \geq p(a \wedge x)$. Since $p(a) = p(a \wedge x) + p(a \wedge \neg x)$, it follows that $p(a \wedge \neg x) > 0$. Hence $p_{a \wedge \neg x}$ is well-defined, and clearly $p_{a \wedge \neg x}(a) = 1 \neq 0$ while $p_{a \wedge \neg x}(x) = 0 < t$ as required.

(v) To complete the cycle, we need to show that $\forall p(5) \Rightarrow \forall p \in P(1)$. We show that $\forall p \in Q(5) \Rightarrow \forall p \in Q(1)$, where Q is as before. Once more we argue contrapositively. Let Q be such a set and suppose that for some $p \in Q$, $p(a) > p(x)$ where $a = a_1 \wedge \dots \wedge a_n$. Then $p(a) \neq 0$ so p_a is well-defined. We have $p_a(a) = 1$, so $p_a(a_i) = 1$ for all $i \leq n$, so each $p_a(\neg a_i) = 0$, so $\Sigma(p_a(\neg a_i): i \leq n) = 0$. It remains to check that $p_a(\neg x) > 0$. Now $p_a(\neg x) = p(a \wedge \neg x)/p(a)$ so it suffices to show $p(a \wedge \neg x) > 0$. But $p(a) = p(a \wedge x) + p(a \wedge \neg x)$ so since by the supposition $p(a) > p(x) \geq p(a \wedge x)$ we have $p(a \wedge \neg x) > 0$ as desired and the proof is complete.

5.3. Supraclassical probabilistic consequence relations

Given the probabilistic characterizations of classical consequence expressed in Theorem 5.2–2, it is natural to ask what twists on them would make us go supraclassical, and eventually nonmonotonic.

A simple idea for going supraclassical, reminiscent of what we did with Boolean valuations in section 3.1, is to restrict the set of probability functions. Instead of quantifying over the set P of *all* probability functions on the language, we may take *some non-empty proper subset* $Q \subset P$ and quantify over it.

This will certainly give us properly superclassical consequence relations for a wide range of choices of Q . For example, in the case of condition 1, if there is some Boolean formula a that is not a tautology but such that $p(a) = 1$ for every $p \in Q$, then we have $a \sim t$ even though $a \not\sim t$.

However, the superclassical relation will remain monotonic for each of the conditions (1), (2t), (3t) and (5). Consider for example condition 1. Suppose $p(a) \leq p(x)$ for all $p \in Q$. Then immediately by postulate (K3), $p(a \wedge b) \leq p(a) \leq p(x)$ for all $p \in Q$. Similarly for the other conditions mentioned, quantified over an arbitrary domain $Q \subset P$.

The position with condition (4t) is subtler. Fix a domain $Q \subset P$. Suppose on the one hand that Q is closed under conditionalization. Then as is apparent from the proof of Theorem 5.2–2, the condition $\forall p \in Q(4t)$ is equivalent to the others and so monotonic. But when Q is not closed under conditionalization, monotony can fail. In particular, it can fail when Q is a singleton $\{p\}$. To show this, we need only find appropriate Boolean formulae a, b, x and a suitable probability function p with $p(a \wedge b) \neq 0$ (so that both p_a and $p_{a \wedge b}$ are well-defined) such that $p_a(x) \geq t$ while $p_{a \wedge b}(x) < t$. For example, suppose that the language has just two elementary letters q, r . Consider the probability distribution that gives each of the four atoms $q \wedge r, \dots, \neg q \wedge \neg r$ equal values 0.25, and choose the threshold value $t = 0.5$. Put $a = q \vee \neg q$, $x = q \vee r$, $b = \neg q \wedge \neg r$. Then $p(a) = 1$ so $p(a \wedge x)/p(a) = p(x) = p(q \vee r) = 0.75 \geq t$ while $p(a \wedge b \wedge x)/p(a \wedge b) = 0 < t$.

Condition (4t) is often regarded as in the most interesting probabilistic condition for defining consequence. So far, we have observed three points about it.

- When quantified over all probability functions, it gives us exactly classical consequence.

- When quantified over a domain of probability functions closed under conditionalization, it gives a supraclassical but still monotonic consequence relation.
- When instantiated to a specific probability function p (or, more generally, when quantified over a domain of probability functions not closed under conditionalization), the consequence relation typically becomes nonmonotonic.

However in probabilistic contexts the question of nonmonotonicity tends to leave centre stage, which is occupied by another property – conjunction in the conclusion (alias AND). To simplify the exposition, we write $(4tp)$ for the condition $(4t)$ instantiated to a specific probability function p ; and likewise for the other conditions. The consequence relations determined by these instantiated conditions *all fail conjunction in the conclusion*, irrespective of whether or not they are monotonic. Thus they differ in a very important respect from their qualitative cousins studied in earlier sections.

As the nonmonotonic condition $(4tp)$ is the most interesting one for defining consequence, we verify explicitly the failure of AND in its case. For a trivial example, take again the language with just two elementary letters q, r , and probability function that gives each of the four atoms $q \wedge r, \dots, \neg q \wedge \neg r$ equal values 0.25. Put a to be any tautology, $x = q, y = r$ and set the threshold value t at 0.5. Then $p_a(x) = p(x) = 0.5 \geq t$ and likewise $p_a(y) = p(y) = 0.5 \geq t$ so by condition $(4tp)$, we have $a \sim x$ and $a \sim y$. But $p_a(x \wedge y) = p(x \wedge y) = 0.25$, so that $a \not\sim x \wedge y$.

The failure of AND gives rise to other failures. There are rules that imply AND as a limiting case, and so also fail. One of these is right weakening (RW) in the generalized form: whenever $A \sim b$ for all $b \in B$ and $B \vdash x$ then $A \sim x$. In terms of operations, it can also be expressed as: $C(A) = Cn(C(A))$. Right weakening is valid for all the qualitative consequence relations, whether monotonic or nonmonotonic, that we have considered. Another rule that implies AND as a limiting case, assuming supraclassicality, is cumulative transitivity (CT) in its general form: whenever $A \sim b$ for all $b \in B$ and $A \cup B \sim x$ then $A \sim x$.

It may be argued that this is to the honour of the probabilistic approach, and that the qualitative one is misguided in validating the rule of conjunction in the conclusion. For while the rule may appear innocent when applied to the case of just two conclusions, it can become quite counterintuitive when applied by iteration to a large number of them. This is the message of the 'lottery paradox' of Kyburg (1961), (1970), and of the 'paradox of the preface' of Makinson (1965).

The *lottery paradox* observes that if a fair lottery has a large number n of tickets then, for each ticket it is highly probable that it will not win, and thus rational to believe that it will not do so. At the same time, it is certain (again, given that the lottery is assumed to be fair) that some ticket among the n will win, and so again rational to believe so. But these $n+1$ propositions are inconsistent. We thus have a situation in which on the one hand, it is rational to believe in each element of a large finite set of $n+1$ propositions, but on the other hand is not rational to believe in the conjunction of all of them, since it is a logically inconsistent proposition. Yet this is what the rule of conjunction in the conclusion would authorize us to do.

The *paradox of the preface* is similar in structure and conclusion, except that it makes no reference to probabilities. As author of a book making a large number of assertions, I may have checked and rechecked each of them individually, and be confident of each that it is correct. But sad experience in these matters may also teach me that it is inevitable that there

will be some errors somewhere in the book; and in the preface I may acknowledge and take responsibility for this (without being able to put numbers on any of the claims). Then the totality of all assertions in the main text plus the preface is inconsistent, and it would thus be irrational to believe their conjunction; but it remains rational to believe each one of them individually. Thus again, the rule of conjunction in the conclusion appears to lead us astray.

This is a difficult philosophical question, and it cannot be said that a consensus has been reached on it. The author's tentative opinion is that there is a tension between two components of rationality – *coherence* and *practicality*. The criterion of coherence would lead us to abandon the rule of conjunction; that of practicality encourages us to manage our beliefs with minimal bookkeeping and calculation. Probabilistic inference favours the former criterion while qualitative nonmonotonic logics lean towards the latter.

We end this section with two further remarks on the condition (4tp). For clarity, we repeat the full definition. For any given real number $t \in [0,1]$ with $t \neq 0$, and any given probability function p , the relation $a \mid_{\sim tp} x$ is defined to hold iff either $p_a(x) \geq t$ or, in the limiting case, $p(a) = 0$.

It should be appreciated that this definition expresses a notion of *sufficient* probability of x given a . It should not be confused with a corresponding notion of *incremental* probability of x given a , which would be defined by putting: $a \mid_{\sim tp^+} x$ iff $p_a(x) \geq p(x)$ in the principal case that $p(a) \neq 0$. This expresses the very different idea that the condition a leads to a probability of x that is higher than it would have in the absence of that condition. Clearly, this incremental inference relation is also nonmonotonic.

In the qualitative case, our nonmonotonic consequence relations were defined by relativizing the underlying supraclassical monotonic relations to the current premise set A . The relativization was effected by various means – consistency constraints on background assumptions, minimality conditions, and consistency checks on the application of rules. But this is not how we proceeded with the definition of the relations $\mid_{\sim tp}$. It is therefore natural to ask whether there is some monotonic notion that can serve as a stepping-stone to the nonmonotonic one $\mid_{\sim tp}$ in much the same manner as in the qualitative arena?

The answer is positive. The monotonic condition (3pt), i.e. $p(a \rightarrow x) \geq t$ for a specified probability function p , may be viewed in this light.

This can be seen clearly if we reformulate the two probabilistic conditions (3pt) and (4tp) in slightly different but equivalent ways. On the one hand, condition (3pt) may equivalently be written as $p(a \wedge \neg x) \leq 1-t$. On the other hand, in the principal case that $p(a) \neq 0$, condition (4tp) may equivalently be written as $p(a \wedge \neg x)/p(a) \leq 1-t$. Comparing these, we see that the latter uses the arithmetic operation of division to make the left-hand side depend on the probability of the premise. In other words, division by the probability of the premise serves to relativize condition (3pt) into (4tp). In this way, in probabilistic contexts the consequence relation defined by (3pt) serves as a (supraclassical, monotonic) bridge between classical and nonmonotonic inference.

We summarize the section. By taking any of the five probabilistic conditions that we listed as characterizing classical consequence, and restricting the set of probability functions, we get supraclassical consequence relations. When the chosen set is closed under conditionalization, all five conditions are equivalent, and all of them are monotonic. But when we look at

instantiations of the conditions, one of them, namely $(4tp)$, is notoriously nonmonotonic. All of the instantiated conditions lead to failure of the rule of conjunction in the conclusion, and thus of several other rules implying it. There are arguments to suggest that this failure is just what is needed, but opinion is unsettled. Finally, although the order of our presentation did not put the fact into the limelight, there is a version of probabilistic consequence that may be seen as a natural stepping-stone to the nonmonotonic one, in much the same manner as for the qualitative relations.

5.4. Bringing probabilistic and qualitative inference closer together

The results of the previous section lead to a further question: is there any way of bringing the probabilistic and qualitative approaches closer together? In particular, is there any way of adjusting the former so that the consequence relations that it engenders are more regularly behaved – specifically, so that the rule of conjunction in the conclusion (AND) is satisfied? Several ideas for such a rapprochement may be found in the literature. We outline two of them, quite different from each other.

One of them takes the concept of probability unchanged in its standard form as defined by the Kolmogorov postulates, but applies it in a more sophisticated manner so as to ensure satisfaction of AND. This is the *limiting probability* approach. The other transforms the notion of a probability function itself, replacing it by another kind of function into the real interval $[0,1]$ in which, roughly speaking, the operation of maximality takes over the role played by addition in standard probability. These are often called *possibility functions*, also known in a dual form as *necessity* or *plausibility* functions.

In this section we will sketch both of these ideas and the kinds of nonmonotonic consequence relations that they give rise to.

Limiting Probabilities

This approach uses ideas of Adams (1966), (1975) and was set out systematically by Pearl (1988), (1989) and Lehmann and Magidor (1992). It exists in two versions: one using epsilon/delta constructions, and the other using non-standard analysis, i.e. infinitesimals. They differ in certain respects. In particular, the version using infinitesimals validates the non-Horn rule of rational monotony, while the epsilon/delta one does not. We describe the epsilon/delta version. For simplicity we continue to assume that the Boolean language is finitely generated.

The construction takes as input any relation R over Boolean formulae, whose elements (b,y) are thought of as forming *the base* for our construction. We wish to define as output a nonmonotonic supraclassical relation $|\sim^R$ with $R \subseteq |\sim^R$.

Fix such an initial relation R , and let ε be any non-zero real in the interval $[0,1]$. We put $P_{R\varepsilon}$ to be the set of all probability functions p such that for every pair $(b,y) \in R$, $p(b) \neq 0$ (so that p_b is well-defined) and $p_b(y) \geq 1-\varepsilon$. Roughly speaking, $P_{R\varepsilon}$ is the set of all those probability functions that bring the conditional probability of every element of R ε -close to unity. We could call it the set of all probability functions that are *epsilon-good for R*. We can now define the relation $|\sim^R$ of ‘limiting probability consequence’, by the following rule:

$a |\sim^R x$ iff for every $\delta > 0$ there is an $\varepsilon > 0$ such that for every probability function $p \in P_{R\varepsilon}$, if p_a is well-defined then $p_a(x) \geq 1-\delta$.

In effect, $a |\sim^R x$ iff the conditional probability $p_a(x)$ can be brought arbitrarily close to 1 by bringing all the conditional probabilities $p_b(y)$ of base pairs sufficiently close to 1. We have written the index as a superscript rather than a subscript so as to avoid any confusion with the consequence relations defined using default rules.

It is almost immediate from the definition that every such relation $|\sim^R$ is supraclassical. Moreover, for finite languages, it is not difficult to show, using suitable arithmetic manipulations, that $|\sim^R$ satisfies conjunction in the conclusion, cumulative transitivity, cautious monotony and disjunction in the premises (see e.g. the verifications in section 3.5 of Makinson (1994)). For infinite languages, however, the generalized versions of the last three rules may fail. Finally, for finite languages, Adams (1975) and Lehmann & Magidor (1992) have proven a converse serving as a representation theorem.

Thus this construction, while remaining probabilistic, leads to nonmonotonic consequence relations which, for finite languages, have essentially the same properties as those generated in a qualitative manner. To be precise, for such languages it leads to exactly the same consequence relations as did the stopped preferential models (with copies).

Possibility Functions

The other way of making the probabilistic approach satisfy conjunction in the conclusion, and generally behave like a qualitative approach, is to modify the Kolmogorov postulates defining probability functions, essentially by substituting maximality for addition. Such functions are often called possibility functions, as they are thought of as representing degrees of possibility in some sense of the term. This approach has been studied in a number of publications by Dubois, Prade and their co-authors, e.g. Dubois, Lang and Prade (1994), Dubois and Prade (2001). To avoid confusion with probability functions in the ordinary sense of the term, we will write possibility functions as π rather than p . Again, we work with finite languages.

A *possibility function* may be defined to be any function π on the Boolean language into the real numbers satisfying the following conditions:

- ($\pi 1$) $0 \leq \pi(x) \leq 1$
- ($\pi 2$) $\pi(x) = 1$ for some formula x
- ($\pi 3$) $\pi(x) \leq \pi(y)$ whenever $x \vdash y$
- ($\pi 4$) $\pi(x \vee y) = \max(\pi(x), \pi(y))$

This is not the most parsimonious axiomatization, but it brings out forcefully the relationship with ordinary probability. The first three postulates are just the same as (K1) to (K3) for probabilities. The difference lies in the last one. Whereas probability required that $p(x \vee y)$ is the sum of $p(x)$ and $p(y)$ whenever x is classically inconsistent with y , these functions require that $\pi(x \vee y)$ is the greater of $\pi(x)$ and $\pi(y)$, and does so without any proviso at all.

From the first three postulates we have that $\pi(x) = 1$ whenever $x \in Cn(\emptyset)$, just as for probability. But from this and postulate ($\pi 4$) it follows that for any formula x , either $\pi(x) = 1$ or $\pi(\neg x) = 1$, which is quite unlike the equality $p(\neg x) = 1 - p(x)$ for probability. Disjunction thus becomes homomorphic with taking *max*, but negation ceases to be homomorphic with difference from unity. We gain one homomorphism but lose another.

Just as with probability functions, we have an extension property based on state-descriptions. Let f be any function on the state-descriptions of L into the real interval $[0,1]$, such that their maximum is unity, i.e. such that $f(s) = 1$ for at least one state-description s . Such a function may be called a *possibility distribution*. Then f can be extended to a possibility function $\pi: L \rightarrow [0,1]$ satisfying the above axioms. The construction parallels that of the probabilistic case.

For each non-contradictory formula x , take the disjunction $s_1 \vee \dots \vee s_k$ of all the state descriptions that classically imply it, and put $\pi(x) = \max(f(s_1), \dots, f(s_k))$. When x is contradictory, put $\pi(x) = 0$. Then it is easy to verify that π is a possibility function, and that any possibility function on L that agrees with f on the state-descriptions is identical with π , except possibly in the value that it assigns to contradiction (which, due to the replacement of (K4) by (π 4), need not be zero). The extension is thus ‘almost unique’.

It remains to define the consequence relations determined by possibility functions. For any possibility function π , we define the relation $|\sim_\pi$ by the rule:

$$a |\sim_\pi x \text{ iff either } \pi(a \wedge \neg x) < \pi(a \wedge x) \text{ or, as limiting case, } \neg a \in \text{Cn}(\emptyset).$$

Equivalently, given the properties of the function π and the irreflexivity of $<$:

$$a |\sim_\pi x \text{ iff either } \pi(a \wedge \neg x) < \pi(a) \text{ or, as limiting case, } \neg a \in \text{Cn}(\emptyset).$$

Then it is not difficult to show that every such relation $|\sim_\pi$ is supraclassical and, for finite languages, satisfies conjunction in the conclusion, cumulative transitivity, cautious monotony, disjunction in the premises, and the non-Horn rule of rational monotony. Here again it is possible to prove a converse as a representation theorem.

Although the definition of a possibility function is quantitative in appearance, a little reflection indicates that it is essentially qualitative in nature. It makes no use of arithmetic operations such as addition or multiplication. Instead it uses the operation of choosing the *maximum* of two items, which makes perfectly good sense in any chain. Also, the presence of the number 1 is quite unnecessary: we need only require that the chain has a greatest element, and require that this element to be the image of any tautology. The resulting logic is the same: we get exactly the same consequence relations $|\sim_\pi$.

Once we have shed the numerical trappings, we can also see that the approach is the same, in content, as one that we have seen in a qualitative context, namely comparative expectation inference, a variant of default-assumption consequence which we described in section 2.3.

Suppose we take a possibility function π on the Boolean language into a chain and ‘upend’ it. That is, suppose we define a function σ also on the same language by putting $\sigma(a) = \pi(\neg a)$ and at the same time inverting the chain order. Then the homomorphism of disjunction with respect to *max* becomes homomorphism of conjunction with respect to *min*. In other words, $\sigma(x \wedge y) = \pi(\neg(x \wedge y)) = \pi(\neg x \vee \neg y) = \max(\pi(\neg x), \pi(\neg y)) = \min(\sigma(x), \sigma(y))$, where maximality is understood with respect to the initial chain $<$ and minimality with respect to its inverse, which we write here as $<'$.

This is already beginning to look familiar. To complete the picture, define a relation $<''$ between Boolean formulae by putting $x <'' y$ iff $\sigma(x) <' \sigma(y)$. Then $<''$ satisfies the condition that whenever $x <'' y$ and $x <'' z$ then $x <'' y \wedge z$, which is a salient property of comparative expectation, as outlined in section 2.3. For if $x <'' y$ and $x <'' z$ then by definition $\sigma(x) <' \sigma(y)$ and $\sigma(x) <' \sigma(z)$ so that $\sigma(x) <' \min(\sigma(y), \sigma(z)) = \sigma(y \wedge z)$ and thus by definition again, $x <'' y \wedge z$. Indeed, all the properties required of comparative expectation relations emerge from this transformation, and we can also go in the reverse direction, transforming an arbitrary expectation relation into a qualitative possibility function.

The definition of the possibilistic inference relation $|\sim_{\pi}$ by the rule $a |\sim_{\pi} x$ iff $\pi(a \wedge \neg x) < \pi(a)$ or $\neg a \in Cn(\emptyset)$ becomes, under this transformation, equivalent to the one for comparative expectation inference in section 2.3. For in the principal case, $a |\sim_{\pi} x$ iff $\pi(a \wedge \neg x) < \pi(a)$, iff $\sigma(\neg a \vee x) < \sigma(\neg a)$, iff $\sigma(\neg a) <' \sigma(\neg a \vee x)$, iff $\neg a <'' \neg a \vee x$. We thus have $a |\sim_{\pi} x$ iff $\neg a <'' \neg a \vee x$ or $\neg a \in Cn(\emptyset)$. As remarked *en passant* in section 2.3, it is not difficult to verify, using the conditions imposed on comparative expectation relations, that this last holds iff $x \in Cn(\{a\} \cup \{y: \neg a <'' y\})$, which is the definition of comparative expectation inference.

It is also possible to see how possibility functions relate to ranked preferential models. In effect, given a finite ranked preferential model, we can define a function π on formulae by putting $\pi(a)$ to be the least rank of any state in the preferential model that satisfies a (and to be the top rank if a is not satisfied by any of its states). Then π will be a possibility function, in the abstract sense indicated above, into the chain that is the converse of the ranking order. Conversely, given a possibilistic model, we may build an equivalent ranked preferential one.

It is easy to get lost in the details of these connections. The important point to remember is that when we pass from probability functions to possibility functions by making maximality do the work of addition, the numerical clothing becomes redundant and we are really working with functions into a chain with greatest element. These functions are homomorphic with disjunction, but we may equivalently replace them by functions into the converse chain that are homomorphic with respect to conjunction, giving us the qualitative notion of an expectation ordering and its associated inference relations.

For further reading on the use of possibility functions to generate nonmonotonic consequence relations, see e.g. section 4.2 of Makinson (1994), which is presented in terms of the ‘upended’ functions σ homomorphic for conjunction with respect to *min*, there called ‘plausibility functions’. For a presentation in terms of the functions π , see Benferhat, Dubois and Prade (1997). A general survey of ‘possibilistic logic’ is given in Dubois, Lang and Prade (1994).

The border between logic and probability is long, with many crossing-points. We have visited one of them, directly related to our general theme: how can we use the notion of probability to characterize classical and supraclassical consequence relations, both monotonic and nonmonotonic?

It should be said that this is not what is most often discussed in examinations of the interconnections between probability and logic. Three main issues currently occupy the centre of the stage.

- The most commonly discussed preoccupation is whether probability may, in some mathematical or philosophical sense, be *reduced* to logic. In particular, whether the Kolmogorov postulates may be given a purely logical justification of some kind, perhaps even providing a logical meaning for the notion of probability. Also, whether purely logical considerations may be used to help us find, in the space of all possible probability distributions, ones that play a special role. We have not discussed these questions and as the reader may guess, we are sceptical about the chances of success for such quest. For an introduction to this question, see e.g. Williamson (2002) and Howson (2003).

- Conversely, some authors have raised the question of whether the concept of logical consequence may be reduced to that of probability without creating circularity of the kind that was noted at the beginning of section 5.2. We have not discussed the issue, and are even more sceptical of success.
- Finally, there is a group of issues on the borderline between logic, probability and the philosophy of science, concerning concepts such as confirmation, induction and abduction. For an introduction to some of them, see e.g. Skyrms (1999).

Appendix: Proof of Theorem 4.3–1

Theorem 4.3–1

Let R be any set of (normal or non-normal) default rules, and A any premise set. Then the Reiter extensions of A under R are precisely the sets $C_{\langle R \rangle}(A)$ for well-orderings $\langle R \rangle$ of R of order-type at most ω . That is:

(\Rightarrow) For every well-ordering $\langle R \rangle$ of R of order-type $\leq \omega$, if the set $C_{\langle R \rangle}(A)$ is well-defined then it is a Reiter extension of A under R .

(\Leftarrow) Every Reiter extension of A under R is a well-defined set $C_{\langle R \rangle}(A)$ for some well-ordering $\langle R \rangle$ of R of order-type $\leq \omega$.

Proof of \Rightarrow

The argument follows a natural, indeed 'inevitable' pattern given the definitions involved, although care is needed to set it out rigorously.

The definition of an extension in Reiter (1980) tells us that a set E of formulae is an extension of the premise set A under the default rules in R iff E is the least set X satisfying the following three conditions: (1) $A \subseteq X$, (2) $X = Cn(X)$, (3) whenever $(a, P, x) \in R$ and $a \in X$ and $\neg p \notin E$ for all $p \in P$, then $x \in X$.

Let $\langle R \rangle$ be a well-ordering of R , of order-type $\leq \omega$, and suppose that $C_{\langle R \rangle}(A)$ is well-defined, i.e. that its inductive construction does not abort at any stage. It will suffice to show that $C_{\langle R \rangle}(A)$ is the least set X satisfying conditions (1)-(3) with $C_{\langle R \rangle}(A)$ substituted for E , i.e. the three conditions: (i) $A \subseteq X$, (ii) $X = Cn(X)$, (iii) whenever $(a, P, x) \in R$ and $a \in X$ and $\neg p \notin C_{\langle R \rangle}(A)$ for all $p \in P$, then $x \in X$.

In other words, we need to show: (a) $C_{\langle R \rangle}(A)$ satisfies conditions (i)-(iii), and (b) whenever X satisfies conditions (i)-(iii), then $C_{\langle R \rangle}(A) \subseteq X$.

To prove (a), we need to show (ai) $A \subseteq C_{\langle R \rangle}(A)$, (aii) $C_{\langle R \rangle}(A) = Cn(C_{\langle R \rangle}(A))$, (aiii) whenever $(a, P, x) \in R$ and $a \in C_{\langle R \rangle}(A)$ and $\neg p \notin C_{\langle R \rangle}(A)$ for all $p \in P$, then $x \in C_{\langle R \rangle}(A)$. Now (ai) is immediate from the definition of $C_{\langle R \rangle}(A)$, and (aii) is verified easily using the compactness of classical Cn . The key verification is for (aiii).

Suppose $r = (a, P, x) \in R$ and $a \in C_{\langle R \rangle}(A)$ and $\neg p \notin C_{\langle R \rangle}(A)$ for all $p \in P$. We want to show $x \in C_{\langle R \rangle}(A)$. Since $a \in C_{\langle R \rangle}(A)$ we have $a \in A_n$ for some $n < \omega$. Since $\neg p \notin C_{\langle R \rangle}(A)$ for each p

$\in P$ we know by (a_{ii}) that p is consistent with $C_{\langle R \rangle}(A)$ and so by the monotony of classical consequence, is consistent with A_n . If already $x \in A_n \subseteq C_{\langle R \rangle}(A)$ then we are done, so suppose $x \notin A_n$. Then, as can be shown by an easy induction, $r \notin R_n$ so r satisfies the condition for entry into Case 1 of the definition of A_{n+1} . Indeed, we may assume without loss of generality that r is the first rule satisfying the entry condition for Case 1 of the definition of A_{n+1} . Since by hypothesis $C_{\langle R \rangle}(A)$ is well-defined, we cannot be in Sub-case 1.2, so we must be in Sub-case 1.1 and thus $A_{n+1} = Cn(A_n \cup \{x\})$, so that $x \in A_{n+1} \subseteq C_{\langle R \rangle}(A)$, completing the proof of (a_{iii}) and thus of (a) as a whole.

To prove (b), suppose X satisfies conditions (i)-(iii). We want to show that $C_{\langle R \rangle}(A) \subseteq X$. Clearly it suffices to show that $A_n \subseteq X$ for every $n < \omega$, which we do by induction. The basis is immediate since $A_0 = A \subseteq X$ by condition (i). For the induction step suppose $A_n \subseteq X$; we need to show that $A_{n+1} \subseteq X$. If there is no rule satisfying the test for entry into Case 1 of the definition of A_{n+1} , then we are in Case 2 and $A_{n+1} = A_n \subseteq X$ using the induction hypothesis. So suppose there is some rule satisfying the test for entry into Case 1 of the definition of A_{n+1} . Let $r = (a, P, x)$ be the first such rule.

Since by hypothesis $C_{\langle R \rangle}(A)$ is well-defined, A_{n+1} must be well-defined, so we must be in Sub-case 1.1. It follows that $A_{n+1} = Cn(A_n \cup \{x\})$ and $R_{n+1} = R_n \cup \{r\}$, and each element of $just(R_n \cup P)$ is consistent with $A_n \cup \{x\}$. Since by the induction hypothesis we have $A_n \subseteq X$ it remains only to show that $x \in X$. Since X satisfies condition (iii), it suffices to show $a \in X$ and $\neg p \notin C_{\langle R \rangle}(A)$ for all $p \in P$.

We have $a \in X$ since by supposition r satisfies the condition for entry into Case 1 of the definition of A_{n+1} so $a \in A_n \subseteq X$ again by supposition.

Let $p \in P$. We need to show that $\neg p \notin C_{\langle R \rangle}(A)$. By the entry condition to Sub-case 1.1, $\neg p \notin Cn(A_n \cup \{x\}) = A_{n+1}$ by the definition of A_{n+1} in that case. Also, by the definition of R_{n+1} in that case, we have $r \in R_{n+1}$ so $p \in just(r) \subseteq just(R_{n+1})$. From this, an easy induction shows that $\neg p \notin A_m$ for all $m > n+1$. Hence $\neg p \notin C_{\langle R \rangle}(A)$ as desired and the proof of (b), and thus of part (\Rightarrow) of the theorem, is complete.

Proof of \Leftarrow

Let E be any Reiter extension of A under R . We want to show that $E = C_{\langle R \rangle}(A)$ for some well-ordering $\langle R \rangle$ of R of order-type $\leq \omega$. We partition R into three subsets, which we call R_1, R_2, R_3 , with the indices 1,2,3 written on the same level as R so as not to confuse them with the first three elements of a sequence R_0, R_1, \dots defined inductively.

R_1 : the set of all rules $r = (a, P, x) \in R$ such that $a \in E$ and p is consistent with E for all $p \in P$;

R_2 : the set of all rules $r = (a, P, x) \in R$ such that $a \notin E$;

R_3 : the set of all rules $r = (a, P, x) \in R$ with $a \in E$ but p inconsistent with E for some $p \in P$.

Clearly these are disjoint, and their union is R . We begin by taking the set R_1 and putting its elements in any order $\langle R_1 \rangle$ of type ω (or finite if R is finite). Our strategy will be to show first that $C_{\langle R_1 \rangle}(A)$ is well-defined and is equal to E , and then to show that the elements of R_2 and R_3 may be inserted into the ordering $\langle R_1 \rangle$ to give an ordering of R which is still of type at most ω without altering the value of the induced output.

To show that $C_{\langle R1 \rangle}(A)$ is well-defined and is included in E we first note from the definition of fixpoint extensions that the conclusion of every rule in $R1$ is in E . A straightforward induction then shows that for every n , the conditions for entry into case 1 and subcase 1.1 are satisfied, so that A_{n+1} is well-defined and is a subset of E . Thus $C_{\langle R1 \rangle}(A)$ is well-defined and $C_{\langle R1 \rangle}(A) \subseteq E$. For the converse inclusion $E \subseteq C_{\langle R1 \rangle}(A)$ we recall that by definition, E is the least set X satisfying conditions (1)–(3) enunciated at the beginning of the entire proof. So we need only check that $C_{\langle R1 \rangle}(A)$ satisfies those same three conditions, i.e. that they hold with $C_{\langle R1 \rangle}(A)$ substituted for X . The first two are immediate, i.e. we have $A \subseteq C_{\langle R1 \rangle}(A)$ and $C_{\langle R1 \rangle}(A) = Cn(C_{\langle R1 \rangle}(A))$. For the third, suppose $(a, P, x) \in R$ and $a \in C_{\langle R1 \rangle}(A)$ and $\neg p \notin E$ for all $p \in P$. We need to show that $x \in C_{\langle R1 \rangle}(A)$. Since $C_{\langle R1 \rangle}(A) \subseteq E$ as already noted we have $a \in E$. By the definition of $R1$ and the fact that $E = Cn(E)$, this means that $(a, P, x) \in R1$. Again, since $C_{\langle R1 \rangle}(A) \subseteq E$ as already noted and $\neg p \notin E$ by supposition, we have $\neg p \notin C_{\langle R1 \rangle}(A)$ for all $p \in P$. Putting this together we have: $(a, P, x) \in R1$, $a \in E$, and $\neg p \notin C_{\langle R1 \rangle}(A)$ for all $p \in P$. We are thus in a position to re-run exactly the same argument as for item (aiii) in part (\Rightarrow) of the proof, with $R1$ substituted for R throughout, giving us $x \in C_{\langle R1 \rangle}(A)$ as desired.

We now need to show that the elements of $R2$ and $R3$ may be sandwiched into the ordering $\langle R1 \rangle$ to give an ordering of R which is still of type at most ω without altering the value of the induced output. We do it in two steps, first with the elements of $R2$ and then with the elements of $R3$.

To insert the elements of $R2$, write the ordering $\langle R1 \rangle$ as r_0, r_1, \dots and give $R2$ any ordering r_{20}, r_{21}, \dots of type at most ω . Insert each r_{2i} in the ordering $\langle R1 \rangle$ in any manner that keeps the sequence of order-type at most ω , e.g. insert r_{2i} just after r_i . By the definition of $R2$, the prerequisite of each rule r_{2i} is outside E and thus outside $C_{\langle R1 \rangle}(A)$ and so outside each A_n used in constructing $C_{\langle R1 \rangle}(A)$, so it is clear that none of the rules r_{2i} satisfies the entry condition of Case 1 in the induction determined by the new ordering, putting us in Case 2 where the rule is not applied. Hence the old ordering $\langle R1 \rangle$ determines the same induction output as the new ordering, which we will write as $\langle R1 \cup R2 \rangle$. That is, $C_{\langle R1 \rangle}(A) = C_{\langle R1 \cup R2 \rangle}(A)$.

To insert the elements of $R3$, write the ordering $\langle R1 \cup R2 \rangle$ as r_0, r_1, \dots and give $R3$ any ordering r_{30}, r_{31}, \dots of type at most ω . Consider any rule $r_{3i} = (a, P, x) \in R3$. Then $a \in E$ and p is inconsistent with E for some $p \in P$. We already know that $E = C_{\langle R1 \rangle}(A) = C_{\langle R1 \cup R2 \rangle}(A)$. Hence p is inconsistent with $C_{\langle R1 \cup R2 \rangle}(A)$ and so $\neg p \in A_n$ for some A_n in the inductive definition of $C_{\langle R1 \cup R2 \rangle}(A)$, with the value of n depending on that of i . Let f be the function putting $f(i)$ to be the first such n .

Insert all the r_{3i} in the ordering $\langle R1 \cup R2 \rangle$ in any manner that puts them somewhere after their counterparts $r_{f(i)}$ and keeps the sequence of length at most ω . For example, insert r_{3i} just after $r_{f(i)*i}$. Write the new ordering as $\langle R1 \cup R2 \cup R3 \rangle$. Now each inserted rule has a justification p that is inconsistent with some earlier A_n in the inductive definition of $C_{\langle R1 \cup R2 \cup R3 \rangle}(A)$, so the rule fails the entry condition of Case 1, putting us into Case 2 where the rule is not applied. Thus the construction of $C_{\langle R1 \cup R2 \cup R3 \rangle}(A)$ never aborts, and moreover $C_{\langle R1 \cup R2 \cup R3 \rangle}(A) = C_{\langle R1 \cup R2 \rangle}(A)$, so that $E = C_{\langle R1 \rangle}(A) = C_{\langle R1 \cup R2 \rangle}(A) = C_{\langle R1 \cup R2 \cup R3 \rangle}(A)$. Since $R1 \cup R2 \cup R3 = R$ we may order R by putting $\langle R \rangle = \langle R1 \cup R2 \cup R3 \rangle$ and we have $C_{\langle R \rangle}(A) = E$ completing the proof of part \Leftarrow and thus of the entire theorem.

Acknowledgements

The essential ideas underlying this overview were sketched in the author's paper 'Bridges between classical and nonmonotonic logic'. *Logic Journal of the IGPL* 11 (2003) 69-96: http://www3.oup.co.uk/igpl/Volume_11/Issue_01/. A textbook version, complete with exercises, answers etc, is in preparation as Makinson (to appear).

As the basic ideas were developed, they benefited from exposure in seminar, workshop and conference presentations, as well as in the classroom. Many people made helpful comments on particular points during the gestation process. They include Gerd Brewka, Björn Bjurling, Marc Denecker, Michael Freund, Dov Gabbay, David Gabelaia, Donald Gillies, Daniel Lehmann, Joao Marcos, Wlodek Rabinowicz, Hans Rott, Karl Schlechta, and an anonymous referee for the *Logic Journal of the IGPL*.

The book is dedicated to the memory of my former co-author Carlos Alchourrón, who died in January 1996. Despite his fascination with the logic of belief revision he distrusted its neighbour nonmonotonic logic, and I could never convince him otherwise. While writing this text, I realized that it is a continuation of our conversations.

References

This reference list contains only those items actually called in the text, together with some additional overviews, collections, and textbooks. It does not seek to be comprehensive. The literature on nonmonotonic reasoning is already vast, and an exhaustive list would itself take up a volume. The reader can find further references in the following books, each of which has its own focus and preferences: Antoniou (1997), Besnard (1990), Bochman (2001), Brewka, Dix and Konolige (1997), Grégoire (1990), Łukasiewicz (1990), Marek and Truszczyński (1993), Rott (2001), Schlechta (1997) and Schlechta (to appear). For bibliographies on special topics see the overview papers of Dubois *et al*, Horty, Konolige, Lifschitz, Makinson, all in Gabbay *et al* eds (1994). For logic programming, see the bibliography in Baral (2003).

Adams, Ernest W. 1966. 'Probability and the logic of conditionals', in J. Hintikka & P. Suppes eds, *Aspects of Inductive Logic*. North-Holland: Amsterdam.

Adams, Ernest W. 1975. *The Logic of Conditionals*. Reidel: Dordrecht.

Adams, Ernest W. 1998. *A Primer of Probability Logic*. CSLI: Stanford.

Alchourrón, Carlos and David Makinson 1982. 'On the logic of theory change: contraction functions and their associated revision functions', *Theoria* 48: 14-37.

Alchourrón, Carlos and David Makinson 1985. 'On the logic of theory change: safe contraction', *Studia Logica* 44: 405-422.

Alchourrón, Carlos and David Makinson 1985. 'Maps between some different kinds of contraction function: the finite case', *Studia Logica* 45: 187-198.

Alchourrón, Carlos, Peter Gärdenfors and David Makinson 1985. 'On the logic of theory change: partial meet contraction and revision functions', *The Journal of Symbolic Logic* 50: 510-530.

Antoniou, Grigoris 1997. *Nonmonotonic Reasoning*. MIT Press: Cambridge Mass.

- Baral, Chitta 2003. *Knowledge Representation, Reasoning and Declarative Problem Solving*. Cambridge University Press: Cambridge UK.
- Ben-David, Shai and Rachel Ben-Eliyahu-Zohary 2000. 'A modal logic for subjective default reasoning', *Artificial Intelligence* 116: 217-236.
- Benferhat, Salem, C.Cayrol, D.Dubois, J.Lang, H.Prade 1993. 'Inconsistency management and prioritized syntax-based entailment', pp 640-645 of *Proceedings IJCAI-93*. Morgan Kaufmann: Los Altos.
- Benferhat, Salem, Didier Dubois and Henri Prade 1997. 'Nonmonotonic reasoning, conditional objects and possibility theory', *Artificial Intelligence* 92: 259-276.
- Besnard, Philippe 1990. *An Introduction to Default Logic*. Springer: Heidelberg.
- Bochman, Alexander 2001. *A Logical theory of Nonmonotonic Inference and Belief Change*. Springer-Verlag: Berlin.
- Brewka, Gerhard 1989. 'Preferred subtheories: an extended logical framework for default reasoning', pp 1043-1048 of *Proceedings IJCAI-'89*. Morgan Kaufmann: Los Altos.
- Brewka, Gerhard 1991. 'Cumulative default logic: in defense of nonmonotonic inference rules', *Artificial Intelligence* 50: 183-205.
- Brewka, Gerhard 1994. 'Adding priorities and specificity to default logic', pp 247-260 of C. MacNish *et al* eds, *Logics in Artificial Intelligence: Proceedings of the JELIA '94 Workshop*. Springer Verlag: Berlin.
- Brewka, Gerhard, Jürgen Dix and Kurt Konolige (1997). *Nonmonotonic Reasoning – An Overview*. CSLI Publications: Stanford.
- Brewka, Gerhard, David Makinson and Karl Schlechta (1991). 'Cumulative inference relations for JTMS and logic programming', pp 1-12 of Dix, Jantke and Schmitt eds, *Nonmonotonic and Inductive Logic*. Berlin: Springer (Lecture Notes in Artificial Intelligence, n° 543).
- Brown, D.J. and R.Suszko 1973. 'Abstract logics', *Dissertationes Mathematicae* 102: 9-41.
- Calabrese, Philip 2003. 'Operating on functions with variable domains', *Journal of Philosophical Logic* 32: 1-18.
- Cohn, P.M. 1965. *Universal Algebra*. Harper and Row: New York.
- de Finetti, Bruno (1936). 'La logique de la probabilité', pp IV1-IV9 of *Actes du Congrès International de Philosophie Scientifique*, Paris 1935. Hermann: Paris.
- Dubois, Didier and Henri Prade 1994. 'Conditional objects as nonmonotonic consequence relationships', *IEEE Transactions on Systems, Man and Cybernetics* 24: 1724-1740.
- Dubois, Didier Jerome Lang and Henri Prade 1994. 'Possibilistic logic', pp 439-513 in Gabbay *et al* eds (1994).
- Edgington, Dorothy 2001. 'Conditionals', pp. 385-414 in Lou Goble ed., *The Blackwell Guide to Philosophical Logic*. Blackwell: Oxford.
- Freund, Michael 1993. 'Injective models and disjunctive relations', *Journal of Logic and Computation* 3: 231-247.
- Freund, Michael 1998. 'On rational preferences', *Journal of Mathematical Economics* 30: 210-228.

- Friedman, Nir and Joseph Y. Halpern 2001. 'Plausibility measures and default reasoning', *Journal of the ACM* 48: 648-685.
- Gabbay, Dov 1985. 'Theoretical foundations for nonmonotonic reasoning in expert systems', in K.Apt ed., *Logics and Models of Concurrent Systems*. Springer-Verlag: Berlin.
- Gabbay, Dov 1994. *Labelled Deductive Systems. Volume I*. Clarendon Press: Oxford.
- Gabbay, Dov, C. Hogger and J. Robinson. eds (1994). *Handbook of Logic in Artificial Intelligence and Logic Programming. Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*. Clarendon Press: Oxford.
- Gärdenfors, Peter 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press: Cambridge Mass.
- Gärdenfors, Peter and David Makinson 1994. 'Nonmonotonic inference based on expectations', *Artificial Intelligence* 65: 197-245.
- Gelfond, Michael and Nicola Leone 2002. 'Logic programming and knowledge representation – the A-prolog perspective', *Artificial Intelligence* 138: 37-38.
- Ginsberg, Matthew, ed. 1987. *Readings in Nonmonotonic Logic*. Morgan Kaufmann: Los Altos.
- Grégoire, Eric 1990. *Logiques non-monotones et intelligence artificielle*. Hermès: Paris.
- Hájek, Alan 2001. 'Probability, logic and probability logic', pp. 362-384 in Lou Goble ed., *The Blackwell Guide to Philosophical Logic*. Blackwell: Oxford.
- Hein, James 2002. *Discrete Structures, Logic, and Computability* (second edition). Jones and Bartlett: Boston.
- Horty, John 2001. 'Nonmonotonic Logic', pp 336-361 in Lou Goble ed., *The Blackwell Guide to Philosophical Logic*. Blackwell: Oxford.
- Howson, Colin 2003. 'Probability and logic', *Journal of Applied Logic* 1: 151-165.
- Jeffrey, R.C. 1965. *The Logic of Decision*. McGraw-Hill: New York.
- Kolmogorov, N. 1950. *The Foundations of Probability*. Chelsea Publishing Co: New York.
- Konolige, Kurt (1994). 'Autoepistemic logic', pp 217-296 in Gabbay *et al* eds (1994).
- Kraus, Lehmann and Magidor 1990. 'Nonmonotonic reasoning, preferential models and cumulative logics', *Artificial Intelligence* 44: 167-207.
- Kratzer, A. 1981. 'Partition and revision: the semantics of counterfactuals', *Journal of Philosophical Logic* 10: 201-216.
- Kyburg, Henry E. 1961. *Probability and the Logic of Rational Belief*. Wesleyan University Press: Middletown.
- Kyburg, Henry E. 1970. 'Conjunctivitis', pp 55-82 of M.Swain ed., *Induction, Acceptance and Rational Belief*. Reidel: Dordrecht.
- Lehmann, Daniel 2001. 'Nonmonotonic logics and semantics', *Journal of Logic and Computation* 11: 229-256.
- Lehmann, Daniel 2002. 'Connectives in quantum and other cumulative logics', Technical Report TR-2002-28 of the Leibniz Center for Research in Computer Science, Hebrew University of Jerusalem.

- Lehmann, Daniel and Menahem Magidor 1992. 'What does a conditional knowledge base entail?' *Artificial Intelligence* 55: 1-60.
- Levi, Isaac 1996. *For the Sake of the Argument*. Cambridge University Press: Cambridge UK.
- Lewis, David 1986. 'Probabilities of conditionals and conditional probabilities', *Philosophical Review* 85: 297-315; reprinted with postscripts pp 133-56 in David Lewis *Philosophical Papers Volume II*. Oxford University Press: Oxford.
- Lifschitz, Vladimir 1994. 'Circumscription', pp. 297-352 in Gabbay *et al* eds (1994).
- Lindström, Sten 1991. 'A semantic approach to nonmonotonic reasoning: inference operations and choice', Uppsala Prints and Reprints in Philosophy n°6, Department of Philosophy, University of Uppsala.
- Łukaszewicz, W. 1984/8. 'Considerations on default logic: an alternative approach' pp 165-193 in *Proceedings of the AAAI Workshop on Non-Monotonic Reasoning*. New Palz: New York; also *Computational Intelligence* 4: 1-16.
- Łukaszewicz, W. 1990. *Non-Monotonic Reasoning—Formalization of Commonsense Reasoning*. Ellis Horwood.
- Makinson, David 1965. 'The paradox of the preface', *Analysis* 25: 205-207.
- Makinson, David 1989. 'General theory of cumulative inference', pp. 1-17 of M. Reinfrank *et al* eds, *Nonmonotonic Reasoning*. Springer-Verlag: Berlin, Lecture Notes on Artificial Intelligence n°346.
- Makinson, David 1994. 'General Patterns in Nonmonotonic Reasoning', pp 35-110 in Gabbay *et al* eds (1994).
- Makinson, David 1997. 'Screened revision', *Theoria* 63: 14-23.
- Makinson, David 2003a. 'Bridges between classical and nonmonotonic logic'. *Logic Journal of the IGPL* 11 (2003) 69-96: http://www3.oup.co.uk/igpl/Volume_11/Issue_01/.
- Makinson, David (2004). 'Inductive versus fixpoint definitions in default reasoning' (unpublished manuscript).
- Makinson, David (to appear). *Bridges from classical to nonmonotonic logic*. King's College Publications: London, Texts in Computing vol. 5.
- Makinson, David and Leendert van der Torre 2000. 'Input/output logics', *Journal of Philosophical Logic* 29: 383-408.
- Marek, V.W. and M. Truszczyński 1993. *Nonmonotonic Logic: Context Dependent Reasoning*. Springer: Berlin.
- McCarthy, John 1980. 'Circumscription – a form of non-monotonic reasoning',. *Artificial Intelligence* 13: 27-39. Reprinted pp 145-152 of Ginsberg (1987).
- Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann: Los Altos.
- Pearl, J. 1989. 'Probabilistic semantics for nonmonotonic reasoning: a survey', pp 505-516 of Brachman *et al* eds, KR'89: *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*. Morgan Kaufmann: San Mateo.
- Pino Pérez, Ramón and Carlos Uzcátegui (2000). 'On representation theorems for nonmonotonic inference relations', *The Journal of Symbolic Logic* 65: 1321-1337.

- Poole, David 1988. 'A logical framework for default reasoning', *Artificial Intelligence* 36: 27-47.
- Poole, David 1989. 'What the lottery paradox tells us about default reasoning', pp 333-340 of *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*. Morgan Kaufmann: Los Altos.
- Reiter, R. 1978. 'On closed world data bases', pp 55-76 of H. Gallaire and J. Minker eds, *Logic and Data Bases*. Plenum: New York. Reprinted pp 300-310 of Ginsberg (1987).
- Reiter, R. 1980. 'A logic for default reasoning', *Artificial Intelligence*, 13: 81-132. Reprinted pp 68-93 of Ginsberg (1987).
- Rott, Hans 1991. 'Two methods of constructing contractions and revisions of knowledge systems', *Journal of Philosophical Logic* 20: 149-173.
- Rott, Hans 1993. 'Belief contraction in the context of the general theory of rational choice' *Journal of Symbolic Logic* 58: 1426-1450.
- Rott, Hans 2001. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*, Clarendon Press: Oxford (Oxford Logic Guides n° 42).
- Rott, Hans 2003. 'Basic entrenchment', *Studia Logica* 73: 257-280.
- Rott, Hans and Maurice Pagnucco 1999. 'Severe withdrawal - and recovery', *Journal of Philosophical Logic* 28: 501-547 (reprinted with corrections of corrupted text in *ibid* 2000, 29:121 ff).
- Sandewall, Eriik 1985. 'A functional approach to non-monotonic logic', *Computational Intelligence* 1: 80-87. Also appeared as Research Report LITH-IDA-R-85-07 of the Department of Computer and Information Science, Linköping University, Sweden.
- Schlechta, Karl 1992. 'Some results on classical preferential models', *Journal of Logic and Computation* 2: 676-686.
- Schlechta, Karl 1997. *Nonmonotonic Logics: Basic Concepts, Results and Techniques*. Springer: Berlin (LNAI Series n° 1187).
- Schlechta, Karl (to appear). *Coherent Systems*.
- Shoham, Yoav 1988. *Reasoning About Change*. MIT Press: Cambridge USA.
- Skyrms, Brian 1999. *Choice and Chance* (fourth edition). Wadsworth: Belmont, Cal.
- Stalnaker, Robert 1970. 'Probability and conditionals', *Philosophy of Science* 37: 64-80.
- Tan, Yao Hua and Jan Treur. 1992. 'Constructive default logic and the control of defeasible reasoning' pp 299-303 of B. Neumann ed., *Proceedings of the Tenth European Conference on Artificial Intelligence (ECAI '92)*. John Wiley: New York.
- Tarski, Alfred 1930. 'Über einige fundamentale Begriffe der Metamathematik'. *Comptes Rendus des Seances de la Société des Sciences et les Lettres de Varsovie*, d.III,23 : 22-29. Translated as 'On some fundamental concepts of metamathematics' in *Logic, Semantics, Metamathematics*. Clarendon Press: Oxford, 1956.
- Veltman, Frank 1976. 'Prejudices, presuppositions and the theory of counterfactuals', J. Groenendijk & M. Stokhof eds, *Truth, Interpretation and Information: Selected Papers from the Third Amsterdam Colloquium*. Foris: Dordrecht.
- Veltman, Frank 1985. *Logics for Conditionals*. Doctoral thesis, University of Amsterdam.

Williamson, Jon 2002. 'Probability logic', pp 397-424 in *Handbook of the Logic of Argument and Inference*, ed. D.Gabbay et alia. Elsevier: Amsterdam.

Zhu, Zhaohui, Zhenghua Pan, Shifu Chen, and Wujia Zhu 2002. 'Valuation structure', *The Journal of Symbolic Logic* 67: 1-23.