

To appear in a volume provisionally entitled *Degrees of Belief*, ed Franz Huber.

Levels of Belief and Nonmonotonic Reasoning

David Makinson

Department of Computer Science, King's College London, London WC2R 2LS, UK.
Email: david.makinson@kcl.ac.uk

Introduction

Behind any technical construction in logic, there is an underlying idea. Or at least there should be. The trouble is, things are not always so simple.

Often, the professional logician tends to let the idea emerge of its own accord from the details of the construction rather than formulate it explicitly, waving it through with a passing phrase, secure in the conviction that the real message lies in the formal version.

On a more substantive level, there are further complications. As is well known, a single intuitive idea can be given formal expression in a variety of distinct ways, differing to greater or lesser degree in their behaviour. Less well known is the fact that, conversely, a single formal construction may sometimes be read in several ways, expressing different ideas.

Nonmonotonic logic is no exception. In this chapter, we will explain how the general idea of *levels of belief* manifests itself in a number of ways in the constructions that have been proposed for modelling nonmonotonic reasoning and how, in most cases, a single formal work may be read in terms of more than one idea.

What is Nonmonotonic Reasoning?

It should be remembered that while nonmonotonic *logic* may be perceived as something new, exotic and mysterious, nonmonotonic *reasoning* is something that all of us do, all the time. For nonmonotonicity is a property – or rather, the failure of a property – that arises whenever our reasoning carries us ever so little beyond the bounds of what is strictly implied by the available information.

Despite loose talk in classics of detective fiction of sleuths *deducing* their conclusions from information whose significance is not seen by others, none of their chains of inference are purely deductive. They involve presumption and conjecture, and the ever-present possibility of going wrong. The conclusions do not follow of necessity from the

premises; it is logically possible that the former be false even when the latter are true. The procedure is defeasible.

Nevertheless, for all its fallibility, it is reasoning. We appeal not only to the observations explicitly mentioned but also, implicitly, to a reservoir of background knowledge, a supply of rules of thumb, a wide range of heuristic guides. Conclusions may be withdrawn as more information comes to hand, with new ones advanced in their place. When this is done, it does not mean that there was necessarily an error in the reasoning leading to the old conclusions, which may still be recognized as the best to have drawn with the limited information previously available.

Such reasoning is performed not only by Sherlock Holmes, but also by medical practitioners, garage mechanics, computer systems engineers, and indeed all those who are required to give a diagnosis of a problem in order to pass to action. Archaeologists sifting through the debris of a site may see their early conclusions about the date, function and origin of an artefact modified as more evidence comes to hand. We do it when we try to anticipate the weather by looking at the sky. Nobody other than the mathematician, logician, or professional of some other highly abstract domain such as theoretical economics or physics, spends much time in chains of pure deduction.

But what exactly is meant by calling such reasoning *nonmonotonic*? We are reasoning nonmonotonically when we allow that a conclusion that is well drawn from given information may need to be suspended when we come into possession of further information, even when none of the old premises are abandoned. In other words, if it can happen that a proposition x is a legitimate conclusion of a set A of premises, but not of some larger set B formed by adding further propositions to 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, epistemologists have for hundreds of years recognized this as an important phenomenon. It has for long been familiar to writers on jurisprudence, and to authors on the philosophy of the empirical sciences. But still today, mainstream systems of logic do not take uncertainty of inference into account. They deal only with purely deductive argument, where the conclusion follows of necessity from the premises without the remotest possible doubt or exception.

This narrow focus did in fact help mainstream logic analyse the kind of reasoning that is carried out in pure mathematics. Logic as we know it today was developed in the late nineteenth and early twentieth centuries to obtain a deeper understanding of the powers and limits of deduction in mathematics. Its remarkable success in the analysis of mathematical reasoning has tended to hide its limitations in areas outside that domain.

This is not to say that there is anything wrong with classical logic. Nor is it necessarily to be regretted that historically things developed in the way that they did. For despite its

limitations, an understanding of such inference is needed before one can begin to make sense of other modes of reasoning. And in what follows, we will have to assume that the reader does know a little about classical propositional logic.

Three Sources of Nonmonotonicity

It would be rash to try and enumerate all the sources of nonmonotonicity in our reasoning. They are many and open-ended. But we can specify the three that have received serious study from logicians over the last few decades.

One source is the capacity of ordinary languages, and some formal languages, to refer to themselves, and in particular to talk about the current limitations of the knowledge that is expressed in them. In other words, they have some capacity for *self-reference*. Because of this, the addition of further information may not only undermine the legitimacy of drawing certain conclusions, it may even transform some of them from being true to being false; it may *change their truth-values*. For example, I may say honestly that the only thing I know about a certain issue is so-and-so, but this statement will become false when further information about that issue is made available to me. Further conclusions based on this one may, in cascade, have their truth-values compromised.

Because this arises from the capacity for self-reference, it is a very special phenomenon. It is studied in *epistemic* and so-called *auto-epistemic* logic. But it will not occupy us here, for it has little to do with degrees of belief.

Another source of the failure of monotonicity that has received attention from logicians is associated with the notion of *dialogue*, or debate, between two or more people. Suppose that two discussants have access to stocks of information, real or apparent, which may or may not be the same. The first participant may begin by advancing something that he takes to be reasonable. The second may agree, query, challenge, undermine, or attack its supposed supports, or the link between the two. The first may give up, reinforce, counterattack, etc. And so on. At each stage of the discussion we may keep track of what conclusions have emerged, whether **from** a participant's angle, consensually, or from third-party perspective; and we may also consider those emerging from the final outcome (or from an infinite progression) of the exchange.

Such procedures are evidently highly nonmonotonic, and have been studied by logicians under names such as *defeasible reasoning*, *defeasible nets*, *dialogue logic* etc, using resources not only from logic itself but also from graph and game theory. But this too has little to do with degrees of belief, and will not occupy us here.

The third source of nonmonotonicity that has been studied by logicians lies in our constant attempt to go in a principled way beyond the limits of our meagre information, independently of any elements of dialogue or self-reference. It is the attempt to provide supplementary machinery, and at the same time controls and safeguards, for jumping,

creeping, or crawling to conclusions beyond those that may validly be derived using only the resources of classical (deductive, pure, certain, and monotonic) logic.

This kind of reasoning, is intimately connected with levels of belief. We will describe several ways in which it has been analysed, bringing out in each case the part played by underlying comparisons of commitment. We will keep formal details to a minimum; the reader interested in pursuing them further may consult the much more extended treatment in Makinson (2005), which also contains further references for all of the issues touched on here.

Additional Background Assumptions

When reasoning in daily life, the assumptions that we make are not all of the same level. Usually, there will be a few that we display explicitly, because they are special to the situation under consideration, or deserve to be highlighted in some other way. There will be many others that we do not bother even to mention, because we take them to be part of shared common knowledge, or too obvious to be made explicit without tedium. They may not even be clearly present to our conscious attention.

This phenomenon was already well known to the ancient Greeks. They used the term *enthymeme* to refer to an argument in which one or more premises are left implicit. That is one way in which we may go supraclassical (i.e. beyond the limits of classical consequence) and, with an extra twist, nonmonotonic.

Let K be any set of propositions, which will play the role of a set of background assumptions. Let A be another set of propositions, representing some explicitly articulated current premises for an inference. Finally, let x be an individual proposition, serving as a candidate conclusion.

We write $A \vdash_K x$ and say that x is a consequence of A *modulo the assumption set* K , iff $K \cup A \vdash x$, where \vdash is classical propositional consequence. In other words, iff there is no Boolean valuation v such that $v(K \cup A) = 1$ whilst $v(x) = 0$. And we call the relations \vdash_K , for all possible choices of background assumption sets K , *pivotal-assumption consequence relations*.

Clearly, such consequence relations are supraclassical, in the sense that whenever x is a classical consequence of A , then it is a consequence of A modulo K , for any choice of background assumption set K – even when it is empty. Given the way in which they are defined from classical consequence, these relations also inherit many of its properties. In particular, pivotal-assumption consequence relations remain perfectly monotonic. Whenever x is a consequence of premises A modulo background assumptions K , then by definition it is a classical consequence of $K \cup A$, so by the monotony of classical consequence, x is a classical consequence of $(K \cup A) \cup B$ alias $K \cup (A \cup B)$ for any B and so, by definition again, x is also a consequence of premises $A \cup B$ modulo K .

So far, by adding background assumptions, we have gone supraclassical; but we are still monotonic. How does nonmonotonicity come in? It arises when, instead of holding the background assumption set K fixed while the current premise set A varies, we allow it also to vary. More precisely, when the part of K that we actually use is allowed to vary, in a principled way, according to the current premise set A .

This is done by imposing a consistency constraint, and diminishing the usable part of K when the constraint is violated. Specifically, we use only the maximal subsets K' of K that are classically consistent with the current premises in A , and for safety's sake accept as legitimate conclusions only those propositions that are classical consequences of $K' \cup A$ for all those maximal A -consistent subsets K' . This relation is called *default-assumption consequence*, the name bringing out its close relation to the preceding pivotal-assumption consequence. It may be written as $A \mid \sim_K x$.

Why would we want to do such a thing, diminishing the usable power of our background assumptions? For the simple reason that when the current premises are inconsistent with the background assumptions then, notoriously, taken together they classically imply every proposition in the entire language. That leaves us with three alternatives: infer everything (hardly rational), truncate classical logic (rather desperate, and more difficult to do in a satisfactory manner than one might imagine), or work with less than all the available information (in particular, privileging the premises and abandoning some of the background assumptions).

With the introduction of the consistency checks, monotony goes out the window. We will illustrate this with a simple example, but the underlying reason is already apparent. When we add further premises to A , we increase its power, but by the same token we also increase its potential for conflict with the background assumption set K . To eliminate the inconsistency we have to reduce K to its maximal A -consistent subsets; but these are all weaker than K . So we lose background assumptions and thus risk losing conclusions.

For example, let $K = \{p \rightarrow q, q \rightarrow r\}$ where p, q, r are distinct elementary letters of a formal language and \rightarrow is the well-known truth-functional (alias material) conditional connective. Then $p \mid \sim_K r$, since the premise p is consistent with the whole of K and clearly $\{p\} \cup K \mid r$. But $\{p, \neg q\} \not\mid \sim_K r$, 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 $\{p, \neg q\} \cup K'$ does not classically imply r – witness the valuation v with $v(p) = 1$ and $v(q) = v(r) = 0$. In brief: we gained the premise $\neg q$, but lost the background assumption $p \rightarrow q$, losing thereby the conclusion r .

Despite this promising start, default-assumption consequence in the simple form that we have defined above faces a serious dilemma. This arises when we ask the question: what kinds of assumption set K may usefully be used in generating operations? Consider any such assumption set. Then either it is already closed under classical consequence, i.e. $x \in K$ whenever $K \mid x$, or it fails to be so closed, i.e. there is an x with $K \mid x$ but $x \notin K$. Each option leads to an unpleasant surprise.

- In the latter case, the identity of the consequence relation \vdash_K turns out to be sensitive to the manner of formulation of the elements of K ; in other words, it is syntax-dependent. It may be debated whether this is a shortcoming or just a feature, and whether anything can, or should, be done about it; but we need not go into such questions here, for it is the other case that concerns us most.
- When K is already closed under classical consequence, the difficulty is much more serious. The consequence relation \vdash_K becomes totally devoid of interest, for it can be shown that in all interesting situations it collapses back into classical consequence. To be precise: whenever the current premises A are inconsistent with the background assumptions K (this is the interesting case), the set of propositions x such that $A \vdash_K x$ coincides with the set of all classical consequences of A . None of the background assumptions ends up contributing anything at all to the authorized conclusions! It is this that leads us – almost forces us – into integrating the notion of degrees of belief into our construction of a nonmonotonic consequence relation, as we will now see.

Enter Levels of Belief

The standard way of dealing with the disastrous collapse that we have described is to refine the definition of default-assumption consequence. Recall that, for $A \vdash_K x$ to hold, we required in the unrefined definition x to be a classical consequence of $K' \cup A$ for *all* the maximal A -consistent subsets K' of K . We can liberalize by requiring $K' \cup A \vdash x$ only for *certain selected* maximal A -consistent subsets K' of K .

But how are they to be selected? By introducing a further background component into the modelling apparatus – typically a relation $<$ that prioritizes among the subsets of K , treating some as preferred over others. We can then define $A \vdash_K x$ to hold under the refined definition, iff x is a classical consequences of $K' \cup A$ *for each of the best* of the maximal A -consistent subsets K' of K , from the perspective of the relation $<$. From a mathematical point of view, ‘best’ here can be understood as *maximal* under the relation (if we are reading $<$ with better on the right) or as *minimal* (if, conversely, we are reading it with better on the left. We will look right, although left is just as common.

From a philosophical or intuitive viewpoint, it is natural to treat this relation $<$ as representing our confidence in the subset K' , in other words, the level of our belief in the truth of its elements, or (nuance!) the degree to which it deserves our belief. Under this kind of reading, the refined definition of $A \vdash_K x$ requires x to be a classical consequence of $K' \cup A$ for those maximal A -consistent subsets K' of the background assumption set K *in which we place (or should place) highest confidence*. We have a double maximalization here: once under the relation of set-inclusion (to determine the maximal A -consistent subsets K' of the background assumption set K), and again under the relation $<$ of level of belief (to find those in which we have or should have the greatest faith, or if one prefers, the least suspicion).

The terms ‘degree’ and ‘level’ here should not be overplayed. For the maximalization procedure to be well-defined, the relation need not be a linear ordering, as the term ‘degree’ might insinuate, nor even ranked (alias modular) in the mathematical sense as the term ‘level’ could suggest. Indeed, we could in principle work with an *arbitrary* relation between subsets of K . Usually, however, a condition is imposed on the relation to ensure that non-empty sets always have at least one maximal element.

This step, of considering only certain of the maximal A -consistent subsets of the background assumption set, has the negative virtue of avoiding the collapse that we noted above. It has the positive virtue of doing so with very little additional formal machinery, which can moreover be read in a natural manner. But from two points of view, it is still rather awkward.

Computationally, it is quite horrendous, because of the multiple consistency checks involved to find the inclusion-maximal A -consistent subsets of K , and also because of the job of determining the $<$ -maximal ones among them. But that need not concern the philosopher. The other point is that conceptually, the construction seems to be putting the cart before the horse. The relation $<$ compares the credibility of different *subsets* K' of the background assumption set K . But surely, some would say, that should be seen as a reflection of the credibility of the various *propositions* that are elements of K . From a philosophical point of view, it would be better to work with a credibility relation $<$ *between the propositions themselves*, rather than between sets of them. Can this be done?

Variations

Indeed it can, and in various different ways that have been worked out in the technical literature. Two of the best known are called *safe consequence* and *consequence using epistemic entrenchment*. We will sketch the former; the latter (and many others) being described in detail in Makinson (2005).

To define a relation of safe consequence, the only equipment that we need is a set K of background assumptions (whether or not closed under classical consequence – both options are acceptable for this construction) and a background relation $<$ between propositions in K . This relation is again read as representing level of confidence or credence, or, from the other end, suspicion and vulnerability. Taking level of belief as improving to the right (as is usually done in this context), we thus read $a < b$ as saying that we have more confidence in b than in a . The only constraint that need to impose on the relation in order to make the ensuing construction behave well, is that it is *acyclic* over K , in the sense that there are no propositions a_1, \dots, a_n in K with $a_1 < a_2 < \dots < a_n < a_1$. Evidently, this condition implies irreflexivity (take $n = 1$) and asymmetry (put $n = 2$).

The essential idea is as follows. When the current premises A are inconsistent with the background assumption set K , we use only those elements of K that cannot reasonably be ‘blamed’ for the inconsistency of K with A . In more formal terms, we say that a proposition a in K is *safe with respect to* A iff a is not a minimal element (under $<$) of any

minimal subset (under set-inclusion) J of K that is inconsistent with A . Note that once again there are two dimensions of minimality. One is with respect to the relation $<$ between elements of K , while the other concerns set-inclusion between subsets of K .

It is not difficult to show that the set of all safe elements deserves its name: no matter how inconsistent the current premise-set A may be with the entire background assumptions K , A is always consistent with the set of K 's elements that are, in this sense, safe with respect to A .

There are a great many alternative ways of proceeding, but all of them involve, in one way or another, some kind of machinery for selecting from among the propositions in K , or from among subsets of K , those which are deemed in some sense most worthy of preservation in the face of conflict with current premises A . The most straightforward way of doing this is by manipulating a relation in one way or another, but more abstract procedures make use of selection functions. The logician does not care very much about how, exactly, we may wish to read this relation or selection function. For the philosopher, one natural kind of reading is in terms level of belief, with however elbow room for nuances – for example, between the level of belief we actually have, and that which we should have.

Those who have studied the theory of belief change will find much of this familiar. The consistency constraints and credibility relations described above correspond to ones that are used to construct operations of belief contraction and revision, as well as closely related operations of update. This is only to be expected, for there are very close connection between the two fields: it is possible to map belief revision operations into nonmonotonic consequence relations, and vice versa, via what is known as the Gärdenfors-Makinson transformation.

If K is a set of beliefs, and a is a new input belief, then the result $K*a$ of revising K to accommodate a may be identified with the set of all propositions that are nonmonotonically implied by a , under a suitable nonmonotonic consequence relation $|\sim$ determined by K . Determined, indeed, in the simplest possible way: just put $a |\sim x$ iff $x \in K*a$. Conversely, given a nonmonotonic inference relation $|\sim$ and a premise a , the consequences x of a under $|\sim$ may be identified with the result $K*a$ of revising K to accommodate a , where K is chosen to be the set of consequences under $|\sim$ of the empty set of propositions (or, if one prefers, the set of all tautologies).

Preferred States

The constructions that we have been describing can also be carried out on what is called the semantic level. Instead of working with background assumption sets, which are sets of propositions, one works with abstract items called *states* (the rather neutral term usually used by computer scientists and mathematicians) or *possible worlds* (the rather grandiose one familiar to philosophers). To each state is associated the set of all elementary propositions of the language that it satisfies, and satisfaction of compound propositions is defined using the usual Boolean or first-order rules. It will be no surprise,

given the discussion above, that alongside the states a further piece of machinery is introduced. Again it is a relation, or more abstractly, a selection function, but this time between states rather than between propositions or sets of propositions. It is usually called a *preference relation*.

Suppose we are given a set of states, each associated with the set of propositions that it satisfies, and a preference relation $<$ between the states themselves. If A is a set of premises, then the *preferential consequences* of A modulo this machinery are defined to be the propositions x that are satisfied by the most preferred states that satisfy A (mathematically, the maximal ones, or the minimal ones if we are reading the relation in the other direction).

Despite its superficial differences, this kind of construction is in fact very similar to that in terms of background assumptions. Indeed, they can be mapped into each other in a way that shows that they are exactly equivalent in the finite case (i.e. when the Boolean language has only finitely many elementary letters, and there are only finitely many states to work with). The constructions differ only in the infinite case, where the semantic one is more general.

Once again, the pure logician does not care very much how the relation $<$ between states is to be understood. All that is needed, mathematically speaking, is *some* relation (if desired, satisfying certain regularity conditions) to discriminate among the states that satisfy a given set A of premises. Moreover, the logician will not be very concerned about what states ‘really are’; any items will do.

The philosopher, however, will usually want to give some kind of intuitive meaning to this equipment. The states are commonly understood as representing possible worlds, or possible reconfigurations of the actual world. The relation that prioritizes among them is often understood as indicating their relative plausibility. Sometimes, under the influence of earlier work of David Lewis and others on counterfactual conditionals, the relation is read in a rather different way, as comparing their ‘distance’, or level of dissimilarity, from some fixed world. These readings need not be the same; indeed the latter need have no connection at all with level of belief.

But in addition to these divergences of interpretation, there is another much less tangible difference of mindset. Some are content to see these interpretational notions as heuristic guides, handles on which the imagination can take hold to get a feeling for what is going on. Others want to take the talk literally, giving it a metaphysical ring. Such philosophers ask questions like ‘what, really, *are* these possible worlds?’ and ‘what is the *correct* way of ordering worlds according to plausibility?’ in the same way as in the philosophy of mathematics some ask ‘what really, are the natural numbers?’ What for the logician is a convenient mathematical device with a homely swing to it becomes for the philosopher a question of meaning – and for the metaphysician a matter of doctrine.

Additional Background Rules

Instead of allowing *propositions* to work as background assumptions alongside our current premises, we can make use of background *rules*. Isn't this the same? Not at all. A rule involving propositions is not itself a proposition, not even a conditional one and, as we will see, it behaves rather differently.

In this section we will see briefly how supraclassical (but still monotonic) consequence relations may be constructed in this way and how, with a further twist they may go nonmonotonic. Again, this twist typically involving some kind of ordering, this time of the rules, which may be read in terms of levels of acceptance.

By a *rule* we mean any ordered pair (a,x) of propositions of the language we are dealing with. A set of rules is thus an arbitrary binary relation R over the language. It would perhaps be more natural to use the term 'rule' for the relation, calling the pairs in it something else, but such is standard terminology.

Given a set X of propositions and a set R of rules, we recall from elementary set theory the standard definition of the *image* $R(X)$ of X under R : it is the set of all y such that for some $x \in X$ we have $(x,y) \in R$. 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 $y \in X$.

With this apparatus, we are ready to define pivotal-rule consequence. Let R be a set of rules. Intuitively, they will be playing the role of a set of background 'inference tickets' ready for application to any set of premises. Let A be a set of formulae, and let x be an individual formula. We say that x is a *consequence of A modulo the rule set R* , and write $A \vdash_R x$, iff x is in every superset of A that is closed under both classical consequence and the rule set R . In other words, writing classical consequence as an operation Cn rather than as a relation, iff x is in every set $X \supseteq A$ such that both $Cn(X) \subseteq X$ and $R(X) \subseteq X$. A relation is called a *pivotal-rule consequence* iff it is identical with \vdash_R for some set R of rules.

How does this differ from adding a set of background assumptions? In particular, from adding the material conditionals $x \rightarrow y$ for all the rules $(x,y) \in R$? A rule is a relatively inanimate object. It can be fired or remain inactive, but cannot legitimately be contraposed or subject to other manipulations. As a result, the generated consequence relation behaves less regularly.

Consider for example, the singleton rule set $R = \{(x,y)\}$ and premise set $A = \{\neg y\}$. Clearly we have $\neg x$ as a classical consequence of A taken with the material conditional $x \rightarrow y$, by modus tollens. But $\neg x$ is not in the least superset of A that is closed under both Cn and the rule set R . That superset is in fact just $Cn(\neg y)$, i.e. the classical closure of A , which is vacuously closed under the rule (x,y) for the simple reason that it does not contain x , so that $R(Cn(\neg y)) = \emptyset$.

Likewise, if we put $R = \{(x,z), (y,z)\}$ then z is a classical consequence of $x \vee y$ together with the two material implications $x \rightarrow z$ and $y \rightarrow z$. But while $x \vdash_R z$ and $y \vdash_R z$, we do not

have $x \vee y \vdash_R z$. This again is because the last-mentioned set is vacuously closed under R : we have $R(Cn(a \vee b)) = \emptyset \subseteq Cn(a \vee b)$ since $a, b \notin Cn(a)$.

Thus far, we have gone supraclassical. The reasons for making a further twist to go nonmonotonic are exactly the same as in the case of pivotal-assumption consequence – there are occasions where the current premises clash with the background rules, in the sense that the least superset of the premises that is closed under both the rules and classical consequence contains a contradiction and so in turn is the entire set of all propositions in the language. In this case, the same options present themselves: either swallow the entire language as conclusion set, or try to cut back classical logic, or apply consistency checks.

In this context, there are two ways of applying consistency checks: to the generating *apparatus* or to the generating *process*. The former runs parallel to what we did for assumptions. One considers the maximal subsets S of the set R of rules such that the closure of the current set of premises under both those rules and classical consequence is consistent; and then one intersects all, or a selected subset of, the outputs thus generated. We will not focus on that path, but rather on the other and better-known one that looks at the generating process. For that option a relation of level of confidence, acceptance or priority between individual rules is brought into play from the very beginning. Technically speaking, what we will be looking at is often known as Reiter default logic with normal rules; we will also call it *default-rule consequence*. As for all other constructions in this brief exposition, further details and references to the literature may be found in Makinson (2005).

We begin by fixing an ordering of the given set R of rules by listing them in a sequence (finite or infinite) r_1, r_2, \dots . For each such sequence s we define a consequence relation \vdash_s by building it up inductively. Reiter's original definition was not inductive but fixpoint, but the equivalent inductive formulation is easier to appreciate.

Given any set A of propositions as premises we set $A_1 = A$. For each $n \geq 1$ we put $A_{n+1} = Cn(A_n \cup \{x\})$, where (a, x) is the *first* rule in the sequence such that (1) $a \in A_n$ but (2) $x \notin A_n$ and also (3) x is consistent with A_n . In the limiting case that there is no such rule, then we tread water by putting $A_{n+1} = Cn(A_n)$. Finally, we say that $A \vdash_s y$ iff $y \in A_n$ for some n .

Clearly, it is condition (3) that puts on the brakes when contradiction threatens. It is here in the middle of the generating process, the inductive passage from A_n to A_{n+1} , that we impose our consistency constraint. Roughly speaking, if the conclusion of the rule is inconsistent with what we have built so far, then we do not apply it.

The sequencing of the rules serves as a prioritization: the higher the priority, the earlier will the rule have an opportunity for application. This makes a big difference to the output, because the successful application of one rule can introduce material that prevents another rule from being applied, because the conclusion of the second rule is inconsistent with it.

If we have a single preferred sequence s for the rules in R , then we may take as our final output the consequence relation that it thus generates. On the other hand, if we allow a number of sequences of the rules in R , perhaps even all possible such sequences, then our final output is defined to consist of those propositions that are in the outputs of all of the corresponding consequence relations.

Once again, the logician does not care much how exactly we might interpret the orderings of the background rules that are created by putting them in sequence; the non-committal term ‘prioritization’ is enough. Philosophers, on the other hand, are more interested in knowing how we might read the priority, in other words, what might be the criteria for giving one rule a higher priority than another. A natural candidate for this is that one rule is accepted with more confidence than another, so that when a conflict arises, we prefer to use it rather than the other one.

It would not be quite right to speak here of degrees of *belief*; but this is for a reason of English grammar rather than of logic. In English, we can say that we accept a rule, but we can’t say that we believe it. We can believe only items that carry a truth-value, i.e. are either true or false; and rules do not carry truth-values. As we emphasized earlier, they can be made up out of propositions, but they are not themselves propositions.

Using Probability

At this point, a reader may express puzzlement. Surely, a long-accepted way of expressing levels of belief is in terms of probability, of which we have a very well developed mathematical theory. Yet, so far, we have not so much as mentioned probability as an ingredient of nonmonotonic reasoning.

In fact, there is quite a deep cultural divide between those who work in probability theory and those who work in logic; not many researchers are active in both areas. Probability theory is much older as a discipline, and is generally practised by mathematicians. Historically, nonmonotonic logic developed quite recently at the hands of computer scientists and logicians, who often made a point of expressing its basic concepts in qualitative rather than quantitative terms. But these are features of the sociology of the domain rather than about its subject.

In fact, it is perfectly possible to express classical consequence in probabilistic terms. It is equally possible to tweak that representation to create supraclassical and nonmonotonic consequence relations. It turns out, however, that their behaviour is very different from that of the qualitative relations that we have considered so far. We will explain these points briefly, assuming that the reader has just a little familiarity with the elements of probability theory.

There are several equivalent ways in which classical propositional consequence may be characterized in probabilistic terms. The simplest of all is the following: Let a, x be any formulae of propositional logic. Then $a \mid\!-\! x$ holds iff for every probability function $p: L$

$\rightarrow [0,1]$ we have $p(a) \leq p(x)$. The left-to-right implication here is in fact one of the postulates of Kolmogorov's well-known axiomatization of probability theory; the converse implication is easily verified, noting the universal quantification over all possible probability functions in the right hand condition, and recalling that boolean valuations are themselves limiting cases of probability functions.

However, the characterization of classical consequence that is most interesting for our story is in terms of conditional probabilities. Fix any real number t distinct from zero in the interval $[0,1]$; this is called a threshold parameter. Let a,x be any formulae of propositional logic. Then it can be shown that $a \vdash x$ holds iff for every probability function $p: L \rightarrow [0,1]$ such that $p(a) \neq 0$ we have $p_a(x) \geq t$. Here p_a is the *conditionalization* of the function p , and is defined by putting $p_a(x) = p(a \wedge x)/p(a)$ for all arguments a such that $p(a) \neq 0$. It is easily verified that whenever p is a probability function, in the sense that it satisfies the Kolmogorov postulates, then so too is p_a .

Thus, in so far as we may take probability as a measure of degree of belief (whether actually held or rational to hold) then we may express classical logical consequence in terms of those degrees, provided we quantify over all probability functions (or a sufficiently large subset of them).

How can we go supraclassical? All we need to do is restrict the set of probability functions over which we quantify in the above characterization, severely enough to exclude some of the Boolean valuations (which, as we have remarked, are limiting cases of probability functions). In other words, whenever we fix a set P of probability functions and a threshold parameter t , we can define a consequence relation \vdash_P as follows. For any formulae a,x of propositional logic, $a \vdash_P x$ iff for every probability function $p \in P$ such that $p(a) \neq 0$ we have $p_a(x) \geq t$. The only difference between this definition and its predecessor is that the quantification is now over the probability functions that are in P , rather than over the set of all possible probability functions.

Are these consequence relations monotonic or nonmonotonic? It all depends on the choice of the subset P . Suppose that the set P is closed under conditionalization, i.e. that whenever $p \in P$ and $p(a) \neq 0$ (so that p_a is well defined) then $p_a \in P$. Then the consequence relation \vdash_P remains monotonic. But when P is chosen in such a way that it is not closed under conditionalization – at the extreme limit, when it is chosen to be a singleton – then monotony fails.

The failure can be illustrated by a very simple example. Suppose that our propositional 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 $t = 0.5$. Choose $a = q \vee \neg q$, $x = q \vee r$, $b = \neg q \wedge \neg r$. Then $p(a) = 1$ so $p(a \wedge x)/p(a) = p(x)/1 = p(x) = p(q \vee r) = 0.75 \geq t$ while $p(a \wedge b \wedge x)/p(a \wedge b) = 0/p(a \wedge b) = 0 < t$.

However in probabilistic contexts the question of nonmonotonicity leaves centre stage, which is dominated by another property – conjunction in the conclusion (alias AND). This is the condition that whenever $a \vdash_P x$ and $a \vdash_P y$ then $a \vdash_P x \wedge y$. It is easy to show that

the probabilistically defined consequence relations $|\sim_P$ defined above fail this condition for almost any choice of P (other than the set of all probability functions which as we have seen gives us exactly classical consequence, which does satisfy AND).

Indeed, the success or failure of the AND condition is quite a good criterion for distinguishing qualitative from quantitative approaches to reasoning. All of the qualitative consequence relations that we have considered in this essay satisfy the condition, while the probabilistically defined consequence relation defined above and its most natural variants, all fail it. To be sure, it is possible to regain AND by various contortions. One of these involves taking the limits of sequences of probability functions; another cheats by changing the very axioms defining probability functions. But in general, satisfaction of the AND rule is a litmus for determining the true nature of the construction, even if its presentation is devious.

Do we *want* the rule for conjunction to be satisfied? Is its failure a defect or a virtue of the probabilistic approach? While the rule appears innocent when applied to just two conclusions, it can become quite counterintuitive when iterated to apply to a large number of them. That is the message of the paradoxes of the lottery **and** the preface.

The *lottery paradox*, due to Kyburg, 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, taking the threshold probability t suitably close to unity.

The *paradox of the preface*, due to Makinson (and sometimes known as the paradox of the proof-reader), is similar in structure, but makes no reference to probabilities. It thus militates against AND even in situations where we may be quite unwilling to measure our degree of belief by a probability function. 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 issue, and it cannot be said that a consensus has been reached. The author's tentative opinion is that there is a tension between two components of rationality, *coherence* and *practicality*. On the one hand, the desire for coherence urges

us to abandon the rule of conjunction, for it leads us from inferences whose conclusions command high levels of belief to others whose conclusions have little credibility. On the other hand, the demand for practicality detests numerical calculations that cannot be carried out in the head or on the fly (and, even when we have pencil, paper, and computer, may need numerical input that is not available). It encourages us to manage our beliefs with minimal book-keeping. Once a conclusion is accepted, it is treated with full honours until thrown out. If its degree of belief is superior to the threshold, that is all we need to remember; we treat it as believed with no reservations at all, and no longer worry about how far above that threshold its credibility might really be.

Conclusions

Drawing the threads together, we may say that nonmonotonic logics as they have developed over the last quarter century, and more generally theories of defeasible or uncertain reasoning including those based on probability, are closely connected with the notion of levels of belief.

The most manifest of these connections is the one that we discussed last. Probability can be seen as expressing a numerical measure of level of credence (or credibility), and can be utilized in developing criteria for drawing conclusions. These criteria lead to our familiar (monotonic) classical logic, when all possible probability functions are taken into account in the definition of a consequence relation. They lead to supraclassical (and in general nonmonotonic) ones when only a subset of them is admitted. In this case, the familiar rule of conjoining conclusions is lost.

In qualitative approaches to nonmonotonic reasoning, the appeal to levels of belief is a little less direct. All of these approaches make use of some kind of ordering of propositions (or of states, possible worlds etc), which may be read as representing different degrees or levels of commitment. But for the logician, as contrasted with the philosopher, such a reading is a secondary, indeed marginal matter. The orderings employed establish priorities for performing certain operations (consistency checks, maximalizations, etc). So long as **that task can be performed**, the logician is happy. So long as the acceptability of a candidate output can be calculated efficiently (which at present is rarely the case) the computer scientist will be happy. To make the philosopher happy requires more. But it is up to the philosopher, not the logician, to articulate how much more.

Reference

Makinson, David, 2005. *Bridges from Classical to Nonmonotonic Logic*. London: King's College Publications. Series: Texts in Computing, vol.5.