Conditional Probability in the Light of Qualitative Belief Change

David Makinson

Abstract

We explore ways in which purely qualitative belief change in the AGM tradition can throw light on options in the treatment of conditional probability. First, by helping see why we sometimes need to go beyond the ratio rule defining conditional from one-place probability. Second, by clarifying criteria for choosing between various non-equivalent accounts of the two-place functions. Third, by suggesting novel forms of conditional probability, notably screened and hyper-revisionary. Finally, we show how qualitative uncertain inference suggests another, very broad, class of 'proto-probability' functions.

Key words: conditional probability, belief revision, AGM, Hosiasson-Lindenbaum, Kolmogorov, Popper, Rényi, cores, screened revision, hyper-revisionary probability, proto-probability.

Contents and Organization

To keep the main text reader-friendly, most of the verifications and historical remarks are placed in an extended appendix, whose sections run parallel to the main text.

1. Why Go Beyond the Ratio Rule?

Kolmogorov's postulates for one-place probability functions are simple, natural and easy to work with, and the ratio definition of conditional probability is convenient to use (see appendix). They have become standard. So why go beyond them?

The reasons advanced in the literature are of two main kinds: a metaphysical complaint and a pragmatic appeal for greater expressiveness. We outline them in this section, suggesting that while the metaphysical grounds do not stand up to scrutiny, there is a real need for greater expressive capacity. In the following section, we show how a comparison with the situation in qualitative belief revision makes the need all the more evident.

1.1. Metaphysical *vs* Pragmatic Considerations

It is commonly felt (see appendix) that all probability is at bottom conditional anyway, and we should bring this out from the very beginning of our formal treatment. From a subjective perspective: a probability judgement is always made given a whole lot of background information, and so is conditional on that information. From a frequency perspective: probability is some sort of limiting frequency of a type of item in a set, and if we enlarge or diminish the set, the frequency will in general change.

However, the argument has its limitations. On a substantive level, it may involve an infinite regress. This is most easily seen in the field-of-sets mode. Suppose we do take probability as a two-place function p: $F^2 \rightarrow [0,1]$ where F is a field of subsets of a set S. This still depends on the choice of S. Making it into a three-place function p: $F^3 \rightarrow [0,1]$ will not help, as that still depends on S, taking us one step further in an infinite regress. The only way to eliminate all such dependence is take the domain to be the *universal* class. But practising probabilists never do this, and if done, it might as well be done from the beginning with one-place functions.

Historically, the argument is reminiscent of an early way of looking at classical first-order logic, according to which universal quantifications $\forall x\varphi(x)$ are at bottom always conditionals, since their range depends on the choice of domain of discourse. On this view, the dependency should be made explicit from the beginning by always quantifying over the entire universe, writing the restricted generalizations as $\forall x[Dx \rightarrow \varphi(x)]$ where D is the intended domain. Such a view had some currency for a while despite the difficulties of talking about a universal set (so the universe was thought of as a class rather than a set). But we have become accustomed to working with the simpler mode of representing universal quantification without running into difficulty, and the philosophical worries have simply withered away.

The historical precedent carries a methodological lesson. Even if all quantification or probability can be said to be in some sense conditional, this does not imply that the conditionality should be brought into the formalism of the theory itself. In particular, when certain conditions are held fixed throughout an investigation, it may be more convenient to take them into account only at the stage of applying the theory to specific problems.

Taking all this into consideration, it would seem that the metaphysical reasons for going beyond the ratio rule and taking conditional probability as primitive are less than compelling. Nevertheless, a pragmatic need remains. When conditional probability is defined by the ratio rule, it has limited expressive capacity. We would like to allow propositions that have been accorded zero probability to serve as conditions for the probability of other propositions. This is impossible when $p(x|a)$ is put as $p(a \wedge x)/p(a)$, for it is undefined when $p(a) = 0$.

The most famous example of this expressive gap is due to Borel. Suppose a point is selected at random from the surface of the earth. What is the probability that it lies in the Western hemisphere, given that it lies on the equator? The condition of lying on the equator has probability 0 under the random selection, but we would be inclined to regard the question as meaningful and even as having 1/2 for its answer. Other examples are given by e.g. van Frassen 1976.

This complaint is more modest than the metaphysical one, pointing to a gap rather than alleging a defect. But it is much more productive. There is no escaping its basic point: it would be helpful to have a more general conception that covers what we will be calling *the critical zone* – the case where the condition a is consistent but of zero probability – and we should try to formulate it.

There are, of course, quite trivial ways of making the ratio definition cover the critical zone. One, due to Carnap, is to declare that the zone is empty: whenever $p(x) = 0$ then x is inconsistent. This is known as the regularity condition. It has the immediate effect that the ratio definition of $p(x|a)$ as $p(a \wedge x)/p(a)$ covers all instances of the right argument a except when a is inconsistent. For inconsistent a, one can then either leave $p(x|a)$ undefined, or take it to have value 1 for all values of the left argument x.

However, as remarked e.g. by Spohn 1986, this is more like a way of avoiding the problem than solving it. It abolishes by fiat the distinction between logical impossibility and total improbability. Moreover, as noted by Harper 1975 (page 229), it leads to an internal inelegance. Let p be a proper one-place function satisfying Carnap's regularity condition, and consider the two-place function $p(\cdot|\cdot)$ determined by the ratio definition. Now take a contingent proposition a with $1 \neq p(a) \neq 0$, and form the left projection $p_a(\cdot)$ alias $p(\cdot|a)$ of the two-place function. By the definition of left projections (see appendix) we have $p_a(x) = p(x|a)$ so substituting $\neg a$ for x, we have $p_a(\neg a) = p(\neg a|a) = p(\neg a \wedge a)/p(a) = 0/p(a) = 0$ since $p(a) > 0$. Thus $p_a(\neg a) = 0$ even though $\neg a$ is consistent, violating the regularity condition as applied to p_a. In other words, even when p satisfies the regularity condition, the left projection of its conditionalization under the ratio definition will not do so – which is a discord, to say the least.

Another trivial way of covering the critical zone is to put $p(x|a) = 1$ for every value of x when $p(a) = 0$. This might be called the *ratio/unit definition*. But while this renders the function always-defined, and is very convenient in many contexts, it does not fill the expressive gap satisfactorily. For under the rule, when a is in the critical zone the left projection $p_a(\cdot)$ from right value a of the two-place function is the constant function with value 1, i.e. the improper one-place probability function. Although this cannot be described as wrong, it is not very helpful. When a is consistent but $p(a) = 0$, hopefully we should be able to conditionalize to something more informative than the unit function; the two-place function should in some sense be essentially conditional.

In mathematical practice one can sometimes 'work around' the problem. The idea is that when a is in the critical zone, we should take $p(x|a)$ to be the limit of the values of $p(x|a')$ for a suitable infinite sequence of non-critical approximations a' to a. But this procedure is possible only for suitable domains (notably, fields based directly or indirectly on the real numbers) satisfying appropriate conditions. So while it serves well for some examples (such as the hemisphere/equator one mentioned earlier), it is not a general solution.

1.2. Some Notational Niceties

In the following sections, we will compare various options for axiomatizing conditional probability in the light of qualitative belief revision. When doing so, we follow certain notational conventions for clarity. In particular, we distinguish $p(x|a)$ from $p(x,a)$, writing:

- $p(x|a)$ with a bar when it is understood as a two-place operation defined from a one-place one by the ratio rule, i.e. by putting $p(x|a) = p(a \wedge x)/p(a)$ when $p(a) > 0$, possibly with the extension that puts $p(x|a) = 1$ when $p(a) = 0$ (in which case we call it the ratio/unit rule).
- $p(x,a)$ with a comma when taking p as an undefined (arbitrary or primitive) two-place operation defined over all or part of L^2.

Care will always be taken to specify the arity (number of places) of a function under consideration, either by mentioning it explicitly, or by using place-markers as in $p(\cdot)$, $p(\cdot|\cdot)$, $p(\cdot,\cdot)$.

Throughout, Cn is the operation of classical consequence; we also write $\approx$ for the relation of classical equivalence.

2. Exploring the Critical Zone

In this section we weigh the significance of the critical zone. We begin by showing how an analogous zone already arises on the qualitative level for AGM belief change, and how this helps bring out the conceptual options underlying different systems for two-place conditional probability. We then review those systems, presenting them in a modular way that makes manifest the rationales for apparently technical choices.

2.1. A Leaf from the AGM Book

It is instructive to compare the situation for probability change with that for qualitative belief change in the AGM tradition initiated in Alchourrón, Gärdenfors and Makinson 1985.

There, expansion is one thing, revision another. Let K be any belief set, i.e. a set of propositions closed under the operation Cn of classical consequence, i.e. $K = Cn(K)$. The *expansion* of K by a is defined simply by putting $K+a = Cn(K \cup \{a\})$. However *revision* is defined by putting $K*a = Cn((K-\neg a) \cup \{a\})$, where $-$ is a suitable contraction operation forming from K a subset that is consistent with a (when a is itself consistent) and satisfying certain regularity conditions.

We thus have two different kinds of change side by side. Again, they differ in the *critical zone* which, in this qualitative context, is the case where we modify the belief set K by a proposition a that is *itself consistent but inconsistent with K*. In this critical zone, expansion creates blow-out to the set of all propositions of the language, while revision forces contraction of the belief set. Outside the critical zone, the two operations coincide. This basic difference should not be obscured by talk of expansion being a special case of revision, which is a sloppy way of saying that the values of the two operations are the same outside the critical zone. Neither operation is a special case of the other.

This basic conceptual difference reflects itself in the different formal properties of expansion and revision. There are principles that hold for expansion but not for revision, and conversely. In particular:

- Expansion never diminishes the initial belief set, i.e. $K \subseteq K{+}a$. This is sometimes known as the principle of *belief preservation*. In contrast, revision will eliminate material from the belief set whenever the input a is in the critical zone.

- When a is inconsistent with K, expansion gives us blow-out: both $a, \neg a \in K{+}a = Cn(K{+}a) = L$ (the whole language). In contrast for revision, whenever a is itself consistent, so is $K{*}a$. This property of revision is known as the principle of *(input) consistency preservation*.

The pattern is replicated in the probabilistic context. There too we are looking at two different kinds of operation coinciding outside but differing inside the critical zone – which in this context, we recall, is the case where a is consistent but $p(a) = 0$. One is *expansionary*, the other is *revisionary*.

- The expansionary operation is given by the ratio/unit definition. It satisfies a probabilistic analogue of qualitative belief preservation: $p(x|a) = 1$ whenever $p(x) = p(x|\top) = 1$. Expressed with left projections, $p_a(x) = 1$ whenever $p_\top(x) = 1$. In other words, conditionalizing never reduces the corresponding belief set: writing $B(p)$ for $\{x\colon p(x) = 1\}$ we always have $B(p) \subseteq B(p_a) = \{x\colon p_a(x) = 1\} = \{x\colon p(x|a) = 1\}$ (see the appendix for a verification). No juice is lost. In contrast, a revisionary operation would allow for diminution of the associated belief set.

- More specifically, when $p(a) = 0$, the expansionary operation blows-out to the unit function (irrespective of a's own consistency): in that case $p_a(x) = p(x|a) = 1$ for all x, so that $B(p_a) = L$ (see appendix or a full verification). In contrast, a fully revisionary conditional probability function would never give us the unit function when the condition a is itself consistent.

These two kinds of conditionalization should not be thought of as competing for the position of 'the right one'. Like expansion and revision in the qualitative context, they can work side by side, as different kinds of conditionalization. But how can the revisionary conception best be expressed?

There are two main approaches to the problem. One is to define a family of revision operations that take one-place functions to one-place functions. That is the path taken by Gärdenfors in a pioneering paper of 1986 (integrated into his book of 1988). The other approach is to define a family of two-place probability functions. That is the path taken in varying manners by Hosiasson-Lindenbaum 1940, Rényi 1955, 1970, 1970a, Popper 1959 and others in their wake.

Although different in appearance, the two approaches are intimately related, as hinted by Gärdenfors 1988 and observed explicitly by Lindström and Rabinowicz 1989. Here, we consider only the approach using two-place probability functions. Our central question is: of the differing axiom systems for two-place probability, is there one that is preferable to the others, and on what grounds?

2.2. Bird's-Eye View of Available Systems

The usual presentation of axiom systems for two-place probability functions can be quite confusing. The systems themselves are not always formulated in an intuitively evident manner. They can also be difficult to compare due to differing choices of right domain – sometimes the whole of L, sometimes the consistent propositions in L, sometimes an arbitrary subset of L lying between $\{x: p(x,\top) > 0\}$ and L itself. To facilitate comparison and focus on essentials, we formulate all systems as functions defined on *the whole* of L^2. We also present the systems in a *modular way*, that is, with a common basis and differing in what is added to it.

The leading idea is to exploit Rényi's insight that for 'most' values of the right argument of the two-place function, the left projections should be proper one-place Kolmogorov functions, while in the remaining cases it should be the unit function. We obtain modularity by making a different specification of what counts as 'most' for each system.

We begin with the basic *van Fraassen system*, which was formulated (in the field of sets mode) by van Fraassen 1976 and 1995. Its axioms concern *all* propositions, rather than special subsets of them. They are the axioms of *right extensionality*, *left projection*, and *product* respectively, for two-place functions $p: L^2 \to [0,1]$.

(vF1) $p(x,a) = p(x,a')$ whenever $a \approx a'$

(vF2) p_a is a one-place Kolmogorov probability function with $p_a(a) = 1$

(vF3) $p(x \wedge y,a) = p(x,a) \cdot p(y,a \wedge x)$ for all formulae a, x, y

Note that (vF2), as formulated here, says that p_a is a one-place Kolmogorov function, but it does *not say whether it is proper or improper* (the unit function). Indeed, the axioms are consistent with p_a being the unit function for every $a \in L$.

Despite their modesty, the van Fraassen axioms have surprisingly many useful consequences. The following were already noticed between van Fraassen 1976, 1995, Arló Costa 2001, Arló Costa and Parikh 2005. For the convenience of the reader, we recall in the appendix the brief verifications.

- Left extensionality: $p(x,a) = p(x',a)$ whenever $x \approx x'$.
- When $y \in Cn(x)$ then $p(x,a) \le p(y,a)$.
- When $p(\cdot)$ is defined as $p(\cdot,\top)$, then we have the ratio rule (though not its unit extension to the critical zone, i.e. the ratio/unit rule).

- When a is a contradiction, then p_a is the unit function.
- The set Δ of all $a \in L$ such that p_a is the unit function is an ideal. That is, it is closed downwards (whenever $a \in Cn(b)$ and $a \in \Delta$ then $b \in \Delta$) and also closed under disjunction (whenever $a,b \in \Delta$ then $a \vee b \in \Delta$).
- p_a is the unit function iff $p(a,b) = 0$ for all b such that p_b is a proper Kolmogorov function.

Van Fraassen 1976, 1995 called the $a \in L$ such that p_a is a proper Kolmogorov function *normal*, and the a such that p_a is the unit function *abnormal* – of course, modulo the function $p(\cdot,\cdot)$. In that terminology, the set of all abnormal formulae form a non-empty ideal containing the contradictions, and that a formula a is abnormal iff $p(a,b) = 0$ for all normal b. Apart from that, they do not tell us anything about which formulae are normal, which abnormal.

Popper's system goes some way to filling the gap. It may be obtained by adding a single axiom, stating that p_a is normal whenever $p(a,\top) > 0$.

(Positive): when $p(a,\top) > 0$ then p_a is a proper Kolmogorov function.

More economically but less transparently, Popper's system may also be obtained by instead adding the statement that $p(x,b) \neq 1$ for some x,b (see appendix), i.e. that $p_b(\cdot)$ is not the unit function for some formula b, i.e. that $p(\cdot,\cdot)$ is not the unit two-place function.

This still leaves unspecified the status of p_a when a is in the critical zone, i.e. consistent but with $p(a,\top) = 0$. The other systems fill this gap in three different ways. *Carnap's system* does so trivially, by declaring that the zone is empty:

(Carnap) When a is consistent then $p(a,\top) > 0$.

This is equivalent to what we would get by keeping one-place functions as primitive, declaring that only contradictions can get the value 0, and using the familiar ratio definition to generate two-place functions,

The *Unit system* fills the gap almost as trivially, by adding instead an axiom saying that any left projection from a point in the critical zone has constant value 1:

(Unit) When a is consistent but $p(a,\top) = 0$, then p_a is the unit function.

This is equivalent to what we would get by keeping one-place functions as primitive and using the ratio/unit definition to generate two-place ones.

Hosiasson-Lindenbaum's system (briefly HL) regulates the critical zone by treating its elements just like consistent propositions outside the zone. It adds to the Popper axioms:

(HL) When a is consistent but $p(a,\top) = 0$, then p_a is a proper Kolmogorov probability function.

Thus, in terms of the leading idea mentioned above, 'most propositions' means, respectively:

- The van Fraassen system: an unspecified subset (possibly empty) of the consistent propositions,

- The Popper system: all propositions that are above the critical zone or in an unspecified subset (possibly empty) of it,

- The Unit system: for all propositions above the critical zone but no others,

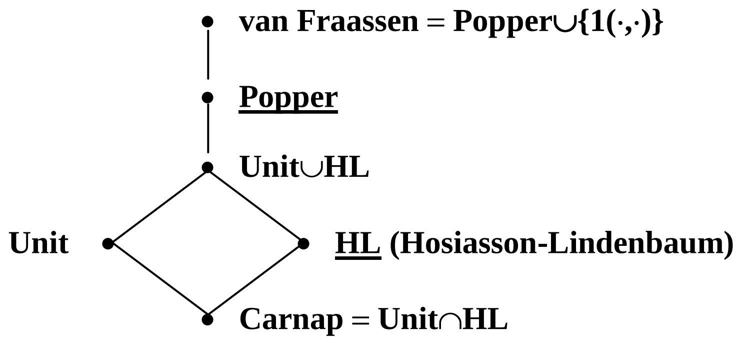
- The Hosiasson-Lindenbaum system: for all propositions above or in the critical zone,

- Carnap's system: we can say any of the last three, since the critical zone is declared empty.

It is easy to check that these axiom systems are equivalent to their usual presentations (see appendix), giving us the sets **Carnap**, **Unit**, **HL**, **Popper, van Fraassen** of functions. The modular arrangement makes it clear at a glance, from their very formulation, what the relations between the systems are. Specifically, we have **Carnap** = **Unit**$\cap$**HL** $\subset$ **Unit**, **HL** $\subset$ **Unit**$\cup$**HL** $\subset$ **Popper** $\subset$ **Popper**$\cup\{$**1(·,·)**$\}$ = **van Fraassen**.

The first four relations were established by Leblanc and Roeper (1989 theorems 4 and 15, table 5, figure 15; also 1999 chapter 3 section 2), with however rather laborious verifications from the usual formulations of the systems, and without mentioning the historical role of Hosiasson-Lindenbaum as a key contributor. With the present modular formulation, the inter-relations become trivial, except for the inclusion **van Fraassen** $\subseteq$ **Popper**$\cup\{$**1(·,·)**$\}$ and the proper part of the inclusion **Unit**$\cup$**HL** $\subset$ **Popper**. The former is checked in the appendix. For the latter, we need a 'mixed' function, failing axioms (Unit) and (HL) but satisfying the Popper axioms. Such a function was already supplied by Leblanc and Roeper 1989 in the form of a rather enigmatic 64-element table; in the appendix we provide the same example with an intuitive rule-based formulation. The relations between the classes are pictured in the Hasse diagram of Figure 1.

Figure 1. Hasse Diagram for Classes of Two-Place Probability Functions



The reader may be surprised that we have not mentioned the axiomatic system of Rényi 1955, also in his later books 1970, 1970a. This is not neglect: his work is indeed capital, providing the leading idea on which most subsequent presentations (including the present one) are based. Rather, Rényi's system takes a form rather different from those above. In effect, he presents a *scheme for a range of axiomatizations*, with the right domain of the function serving as a parameter. For a suitable choice of this parameter (and a little massage) we may obtain the axiomatization of Popper, and likewise of Hosiasson-Lindenbaum. Thus strictly speaking (and taking into account the chronology), Popper's axioms could well be called the Rényi/Popper postulates. These historical matters are reviewed more fully in the appendix.

3. Choosing between systems

Are there any good reasons for preferring one of these systems to another? From our discussion so far, there are only two serious contenders going beyond the ratio/unit account, namely the systems of Hosiasson-Lindenbaum and of Popper, underlined in the diagram. In this section we discuss possible criteria for preferring one to the other, coming to the conclusion that the choice is not a matter of correctness, but of how revisionary we want our revisionary conditional probability to be. We then give an example of how the difference between the two can sometimes make a difference to an enterprise using conditional probability.

3.1. Hosiasson-Lindenbaum *vs* Popper

The Hosiasson-Lindenbaum system is not just revisionary – it is radically so, satisfying without reserve the probabilistic counterpart of consistency preservation. That is, for *every* proposition a, if it is consistent then p_a is a proper Kolmogorov function. The only values of the right argument that project to the unit function are the inconsistent ones.

On the other hand Popper's system is 'variably revisionary': it leaves unspecified the extent to which a function satisfying the axioms is expansionary, and how far it is revisionary. As one extremal case it covers functions $p(\cdot,\cdot)$ that are purely expansionary, i.e. p_a blows out to the unit function for every a in the critical zone as well as for inconsistent a. These are the functions satisfying the Unit axiom (U) above. At the other extreme it covers the Hosiasson-Lindenbaum functions, where p_a never blows out in the critical zone. In between, it covers many 'mixed' functions, where for certain a,b in the critical zone p_b is the unit function but p_a is a proper Kolmogorov function. Intuitively, it is in the spirit of Leblanc 1989 who asked rhetorically: "Can't there be *some* statement of L that is 'utterly unbelievable', so unbelievable indeed that – should you believe it – you'd believe anything, and yet is not truth-functionally false?"

It would be unjustifiably doctrinal to regard one of these policies as right and the other wrong. They are two more options to be added to the traditional one of using the ratio or ratio/unit rule. One option may be appropriate in certain applications, another elsewhere. It is not a matter of choosing once and for all between candidates, but of knowing which candidate to call on for what employment.

It can also be said that at bottom the difference between the Popper and Hosiasson-Lindenbaum options does not stem from different conceptions of probability, but from different choices of the underlying consequence relation. This can be explained formally as follows.

We have already noticed that for any two-place function satisfying the Van Fraassen axioms, and thus *a fortiori* any Popper function $p(\cdot,\cdot)$, the set Δ of all $a \in L$ such that p_a is the unit function is a non-empty ideal. That is, it contains all contradictions, whenever $a \in Cn(b)$ and $a \in \Delta$ then $b \in \Delta$, and whenever $a,b \in \Delta$ then $a \vee b \in \Delta$. Hence the set $\nabla = \{a: \neg a \in \Delta\}$ is a filter, i.e. whenever $b \in Cn(a)$ and $a \in \nabla$ then $b \in \nabla$, and whenever $a,b \in \nabla$ then $a \wedge b \in \nabla$). From this in turn it follows that if we define a supraclassical consequence operation Cn' by putting $Cn'(A) = Cn(A \cup \nabla)$ we have: $\perp \in Cn'(a)$ iff $\perp \in Cn(\{a\} \cup \nabla)$ iff $\neg a \in Cn(\nabla) = \nabla$ iff $a \in \Delta$ iff p_a is the unit function. That is, p_a is the unit function iff a is inconsistent modulo Cn'. Moreover, it is easy to check that modulo Cn', the function $p(\cdot,\cdot)$ continues to satisfy all the van Fraassen axioms, given that it satisfied them modulo classical Cn, and so $p(\cdot,\cdot)$ is a Hosiasson-Lindenbaum function modulo Cn'.

In brief: any Popper function (modulo classical Cn) is a Hosiasson-Lindenbaum function modulo a suitably defined supraclassical consequence operation Cn', with the abnormal elements of the critical zone becoming Cn'-inconsistent.

Thus the gap between the broader and narrower classes of function is less impressive than might have been imagined. Which should we work with? Given that the tighter constraints of HL functions make them easier to handle, it would appear good practice to do so in applications that admit those constraints. Indeed, the following policy suggests itself: (1) When the application does not require any attention to the critical zone, stay with one-place functions as primitive, using the ratio or ratio/unit definition of conditional probability. Otherwise (2) take two-place probability as primitive with the Hosiasson-Lindenbaum axioms if the application admits doing so, otherwise (3) the Popper axioms

3.2 Does it Ever Make a Difference?

So, does it ever make a substantive difference which kind of essentially conditional probability we use? In some cases it it appears to do so. An example is the theory of 'cores' as set out by Arló Costa and Parikh 2005 using Popper functions, building on ideas of van Fraassen 1995 and Arló Costa 2001.

Cores were introduced to give a probabilistic account of the intuitive distinction between a broader class of 'plain' beliefs and a narrower one of 'full' beliefs, in such a way that *both* sets are closed under classical consequence (and hence under conjunction).

Translating from the field-of-sets mode used by the authors mentioned, a *core* for a Popper function p: $L^2 \rightarrow [0,1]$ is a formula c such that (1) c is normal, that is, the left projection p_c of p from the right value c is a proper Kolmogorov function, and (2) for any consistent formula a logically implying c and any formula b inconsistent with c, $p(b,a\vee b) = 0$.

Plain beliefs modulo p are then identified with those formulae logically implied by at least one core, while *full* beliefs are those implied by every core. The authors show that in the finite case for any Popper function p: $L^2 \rightarrow [0,1]$ there is a unique strongest core c_0 and a unique weakest one c_1; so that in that case plain beliefs are those formulae logically implied by c_0, while full beliefs are those implied by c_1. Indeed, in the field-of-sets mode we have the same whenever the underlying set is countable and we assume countable additivity.

However, for plain beliefs so defined, there is a difficulty. In the finite case they turn out to be just the formulae x with $p(x,\top) = 1$. In the field-of-sets mode, and assuming countable additivity, this also holds whenever the underlying set is countable. This is given as the 'coincidence lemma' of Arló Costa 2001 page 578, and is also an immediate consequence of Lemma 3.1 of Arló Costa and Parikh 2005. Thus in these contexts, the construal of *plain* belief in terms of cores gives us nothing new, no matter how we choose our Popper function. Nevertheless, as Parikh has urged (personal communication), when we are working in the uncountable case, or in the countable one but without countable additivity, we may not have the same collapse.

It does not seem to have been noticed in the literature that for full beliefs as defined via cores, there is another difficulty. If we work with Hosiasson-Lindenbaum functions rather than the broader class of all Popper functions, it turns out that in every case (finite, countable, uncountable) the full beliefs so construed are just the tautologies – which is hardly what we want. To show this, we need only verify that $\top$ is itself a core. Using the definition above, it suffices to check that $p(\neg\top,\top) = 0$ (which is easy) and that whenever a is consistent while b is inconsistent then $p(b,a\vee b) = 0$. But by the inconsistency of b we have $p(b,a\vee b) = p(b,a)$; and since p is a Hosiasson-Lindenbaum function, its left projection p_a fromconsistent a is a proper Kolmogorov function, so by the inconsistency of b again, $0 = p_a(b) = p(b,a)$.

Thus the usefulness of cores for defining a formal notion of full belief is not robust under these two axiomatizations of two-place probability. Some might take this as a point against the Hosiasson-Lindenbaum system; the author takes it as a point against the edibility of cores.

4. Reverse Direction: Belief Revision in the Light of Conditional Probability

We have been using AGM belief revision to explain why we should take seriously a revisionary reading of two-place probability functions, and to throw light on the options available for them. Insight can also be gained by looking in the other direction. There is a natural map from two-place probability functions satisfying the Hosiasson-Lindenbaum (HL) postulates into (in fact onto) the family of AGM belief revision operations modulo classical consequence. This section is rather more technical than the others; some readers may prefer to skip to the more exciting section perspectives of section 5.

4.1. A Map from Conditional Probability to AGM Belief Revision

Lindström and Rabinowicz 1989 constructed a map from the class of all Gärdenfors probability-revision operations into the class of AGM belief revision operations. The construction below essentially translates it, with some simplifications and an explicit verification of surjectivity, into a map from the class of Hosiasson-Lindenbaum probability functions onto the AGM operations.

We treat AGM belief revision functions as one-place operations $*$: $L \rightarrow 2^L$, with associated current belief sets K. Given any HL function p: $L^2 \rightarrow [0,1]$ as defined in section 2.2 or equivalently its appendix, we construct the associated function $*$: $L \rightarrow 2^L$ and the set $K = B(p)$, as follows.

- The operation $*_p$: $L \rightarrow 2^L$ is defined by putting $*_p(a) = \{x\colon p(x,a) = 1\}$.
- The set $B(p)$, also called the *top* of p, is defined by putting $B(p) = *_p(\top) = \{x\colon p(x,\top) = 1\}$.

Then we can show (see appendix) that for every HL function p: $L^2 \rightarrow [0,1]$:

- $B(p)$ is a consistent belief set.
- The operation $*_p$: $L \rightarrow 2^L$ satisfies the full set of AGM postulates (K$*$1) through (K$*$8) with respect to $B(p)$.

4.2. Properties of the Map: Surjective but not Injective

The passage from p to $*_p$ is not injective: a counterexample is given in the appendix. On the other hand, it is surjective for consistent belief sets and under the condition of finiteness (i.e. that the propositional language has only finitely many mutually non-equivalent formulae). That is: in such a language, for every consistent belief set K and every revision operation $*$: $L \rightarrow 2^L$ satisfying the AGM postulates with respect to K, there is a HL function p: $L^2 \rightarrow [0,1]$ with $* = *_p$ and $K = B(p)$.

The construction is quite straightforward. Given $*$ and consistent K for such a language, we define p: $L^2 \rightarrow [0,1]$ as follows:

- In the limiting case that a is inconsistent, put $p(x,a) = 1$ for all $x \in L$
- In the principal case that a is consistent, put $p(x,a)$ to be the proportion of ($K*a$)-worlds that are x-worlds.

Here, a *world* is a maximal consistent set of formulae, and an *X-world*, for $X \subseteq L$, is a world Y with $X \subseteq Y$. It is straightforward to verify (see appendix) that p satisfies the HL postulates, $* = *_p$, and $K = B(p)$ as desired.

Thus, we have a natural surjective though non-injective, map from the family of all HL conditional probability functions to the family of the AGM revision operations on consistent belief sets. This map helps us see the AGM postulates as reflections of the HL ones. To this extent, the AGM postulates may be said to go back to 1940!

The map and proof may be generalized to cover the Popper functions; we sketch what is involved. In the above construction, we have been taking AGM revision functions as formulated using classical consequence in the background. In fact, as was already made clear in Alchourrón, Gärdenfors and Makinson 1985, the same theory of belief revision carries through when formulated using arbitrary supraclassical consequence operations satisfying the Tarski closure conditions and disjunction in the premises. So, for a given function p: $L^2 \rightarrow [0,1]$ satisfying the Popper postulates modulo classical Cn, we first see it as a Hosiasson-Lindenbaum function modulo a suitable supraclassical Cn', in the manner described in section 3.1, checking also that this Cn' satisfies the above conditions. We then construct the same map as above, but with Cn' understood everywhere in place of Cn, and verify in the same manner as before.

5. Alternative forms of Conditionalization and Revision

Existing work on qualitative belief revision can suggest or provide perspective on novel forms of conditional probability. In this section we discuss two examples: screened and hyper-revisionary conditionalization. We then explain how recent work on qualitative but probabilistically supported inference leads to an interesting notion of proto-probability functions.

5.1. Screened Conditional Probability

Screened revision is a variant form of AGM belief revision. Its basic idea is to see the operation as made up of two steps: a pre-processing step possibly followed by application of an AGM revision. The pre-processor decides the question of *whether* to revise, and this is done by checking whether the proposed input is consistent with a central part of the belief set under consideration, i.e. a privileged subset. If the answer is negative, the belief set remains unchanged. If it is positive, we apply an AGM revision in a manner that protects the privileged material. Clearly, such a composite process will not satisfy all the postulates of AGM revision: to begin with, the postulate of success, $a \in K{*}a$, may fail.

What would a probabilistic analogue of this look like? Roughly speaking, using the language of Leblanc cited earlier, when a is too unbelievable to take seriously as a condition, we put the probability of x on condition a to be just the *unconditioned* probability of x. In other words, for values of a where for Hosiasson-Lindenbaum or Popper the left projection p_a would be the unit function, we now require that $p_a = p_{\top}$.

This forces modification of the Van Fraassen axioms. In particular, the axiom (vF2) of left projection must be weakened: we no longer always have $p_a(a) = 1$ since when a is inconsistent $p_a(a) = p_{\top}(a) = p(a,\top) = 0$. In another respect, (vF2) can be strengthened: we can require that the left projection from any point is always a proper Kolmogorov function, as we no longer have any use for the unit function. The product axiom (vF3) must also be weakened. To show this, consider any inconsistent a. Unrestricted use of the product axiom would give us that for all x: $p_{\top}(x) = p_a(x) = p(x,a) = p(x \wedge x,a) = p(x,a) \cdot p(x,a \wedge x) = p(x,a) \cdot p(x,a) = p_a(x) \cdot p_a(x) = p_{\top}(x) \cdot p_{\top}(x)$; so that for any x, $p_{\top}(x)$ is either 0 or 1 – which is quite undesirable behaviour.

The question of formulating adequate axiom systems for screened versions of the Popper and Hosiasson-Lindenbaum systems is open. These problems may well be worth investigating, for although screened conditional probability behaves in an unfamiliar way, it is a coherent, intuitively motivated, and possibly useful concept.

5.2. Hyper-revisionary Probability Functions

As is well known, for any van Fraassen function $p(\cdot,\cdot)$ and $a \in L$, if $p(a,\top) > 0$ then $p(x,a)$ is determined by a natural relativization of the ratio rule: $p(x,a) = p(a \wedge x,\top)/p(a,\top)$. Indeed, this equality is almost immediate: the product axiom gives us $p(a \wedge x,\top) = p(a,\top) \cdot p(x,a \wedge \top) = p(a,\top) \cdot p(x,a)$ by right extensionality, permitting division when $p(a,\top) > 0$.

As remarked by Jonny Blamey (personal communication), it may be suggested that this is too conservative, even when we give elements in the critical zone a radically revisionary treatment in the manner of Hosiasson-Lindenbaum. For if a has a very low positive probability – say, to fix ideas, $0 < p(a,\top) < 0.01$ – then a surprise occurrence of a might sometimes lead us to question whether the function $p(\cdot,\cdot)$ was really right to give $p(a,\top)$ such a small value. We should perhaps move to a function $q(\cdot,\cdot)$ which makes the truth of a less unexpected, i.e. puts $q(a,\top)$ well above $p(a,\top)$; and for such a q the value of $q(x,a)$ may be quite different from that of $p(a \wedge x,\top)/p(a,\top)$.

This interesting proposal has a number of repercussions, some of which may be seen as merits, others as drawbacks.

Philosophically, it drives a wedge between two different ways of 'adopting' a condition a. On the one hand, we may accept it because its truth has been revealed to us; on the other hand, we may entertain it to explore its consequences. The argument above suggests grounds for abandoning $p(\cdot,\cdot)$ when confronted with the truth of a proposition a for which p gives a very low value, but it does not suggest doing so when we merely entertain the truth of a to determine what effect it has on our probabilities. In this way, the proposal has the merit of putting the spotlight on a difference that tends to be hidden by the usual treatments of conditional probability.

Pragmatically, there can be no universally fixed cut-off point, like 0.01, at which we should revise rather than apply the relativized ratio rule. Where to draw the line would be a matter of context, purposes and subject matter, balanced in an informal judgement. This may be a source of frustration.

Formally, given the above short derivation of the relativized ratio rule from the product and left extensionality axioms, at least one of the two would have to be given up, or at least restricted in a suitable way. This is quite a loss.

Finally, it is not immediately clear how the new function $q(\cdot,\cdot)$ might be constrained by appropriate conditions. This could be seen as a disappointment, or as a challenge.

What would the qualitative analogue of such hyper-revisionary conditionalization look like? It would be to allow that even when input a is logically consistent with belief set K, we should not always take $K{*}a$ to be $Cn(K \cup \{a\})$. As well as adding in a, we should perhaps be contracting K, for despite the logical consistency of the two, a

latter may be so implausible in the eyes of K that the exposure of its truth may lead us to an 'agonizing reappraisal' of the latter.

This, of course, is counter to one of the basic postulates of AGM belief revision, which puts $K{*}a = Cn(K \cup \{a\})$ in *every case* that a is consistent with K. In brief, AGM does not admit any conflict less than consistency as forcing contraction, just as the standard forms of conditional probability do not allow any improbability other than zero to force us out of the ratio rule.

Perhaps there is room for systems of revision, and of probabilistic conditionalization, in which the background logic is not logical consequence but some form of uncertain inference. How to build on such an idea without circularity or obscurity is another question. On the probabilistic level, it would appear that such an investigation would have to make contact with the theory of 'error statistics' as developed by Fisher, Neyman and Pearson, which analyses grounds for preferring one statistical hypothesis to another when faced with evidence that is highly improbable given the latter but not the former. However, we do not attempt to explore this line of thought further in the present paper.

5.3. Proto-probability Functions for Qualitative Inference

In 1996, Hawthorne investigated rules of uncertain inference while, while qualitative, may be given a probabilistic justification, formulating an axiom system called Q. All of its axioms are in a natural sense probabilistically sound, although the converse has not yet been settled. The question arises: do we need the full force of the axioms of probability in order to justify the rules of Q, or can it be done with weaker constraints on the 'probability' functions? In this section we observe that considerable weakening is possible. We need only certain modest order-theoretic conditions from among those derivable in the system of van Fraassen, the weakest of those presented in section 2.2.

First, we recall Hawthorne's axioms. They concern consequence relations |~ (in words: snake) between formulae of classical propositional logic. There are six Horn rules O1–O6 defining a system O, and one 'almost Horn' rule NR whose addition gives Q. As always, Cn is classical consequence and ≈ is classical equivalence:

O1. a |~ a (reflexivity)

O2. When a |~ x and $y \in Cn(x)$, then a |~ y (RW: right weakening)

O3. When a |~ x and $a \approx b$, then b |~ x (LCE: left classical equivalence)

O4. When a |~ $x \wedge y$, then $a \wedge x$ |~ y (VCM: very cautious monotony)

O5. When a |~ x, b |~ x and $\neg b \in Cn(a)$, then $a \vee b$ |~ x (XOR: exclusive ∨+)

O6. When a |~ x and $a \wedge \neg y$ |~ y, then a |~ $x \wedge y$ (WAND: weak ∧+).

NR. When $a \vee b$ |~ x and $\neg b \in Cn(a)$, then either a |~ x or b |~ x.

As Hawthorne showed, these conditions are probabilistically sound in the sense that for any probability function $p(\cdot,\cdot)$ satisfying Popper's postulates and 'threshold' $t \in [0,1]$, if we define a relation by putting $a \mid\!\sim_{pt} x$ iff $p(x,a) \geq t$, then $\mid\!\sim_{pt}$ satisfies all the rules of Q.

Our question is: how much probability is really needed for the job? We show that it can be carried out with any function into an arbitrary linearly ordered set with greatest and least elements, satisfying certain very weak conditions in which no arithmetic operations appear.

Let D be any non-empty set equipped with a relation $\leq$ that is transitive and complete with a greatest element 1_D and a least element 0_D. A *proto-probability* function into D is any function p: $L^2 \to D$ satisfying the following six conditions:

P1. $p(a,a) = 1_D$

P2. $p(x,a) \leq p(y,a)$ whenever $y \in Cn(x)$

P3. $p(x,a) = p(x,b)$ whenever $a \approx b$

P4. $p(x \wedge y,a) \leq p(y,a \wedge x)$

P5. $p(x,a) \leq p(x,a \vee b) \leq p(x,b)$ whenever $p(x,a) \leq p(x,b)$ and $\neg b \in Cn(a)$

P6. $p(x,a) = p(x \wedge y,a)$ whenever $p(y,a \wedge \neg y) \neq 0_D$.

We call condition (P5) the principle of *disjunctive interpolation*. It is closely related to a principle of 'alternative presumption' of Koopman 1940, 1940a (details in the appendix).

Then, if we take any proto-probability function $p(\cdot,\cdot)$ and $t \in D$, and define a relation by putting $a \mid\!\sim_{pt} x$ iff $p(x,a) \geq t$, then $\mid\!\sim_{pt}$ satisfies all the rules of Q. Indeed, each condition (O*i*) follows directly from its counterpart (P*i*), with (NG) also following from (P5). The verifications are trivial, bit given the novelty of the notion of proto-probability, we give them in full in the appendix.

It is also easy to check that the axioms for proto-probability functions follow from those of van Fraassen, *a fortiori* from the stronger ones of Popper, Hosiasson-Lindenbaum, Carnap, and the Unit system. In fact, they are considerably weaker. Informally, it is clear that the left projection and product axioms of van Fraassen do not hold for all proto-probability functions, even when their top and bottom elements are chosen as the numbers 1,0, since our conditions for the latter make no use of either addition (which is implicit in the left projection axiom) or multiplication (explicit in the product axiom).

For a specific example of a proto-probability function that is not a van Fraassen one, take p: $L^2 \to [0,1]$ to be the characteristic function of the classical consequence relation, i.e. put $p(x,a) = 1$ when $x \in Cn(a)$, otherwise $p(x,a) = 0$. Clearly, this satisfies conditions P1 through P6, but it fails (vF2) since $p(x \vee \neg x, \top) = 1$ while $p(x,\top) = 0 = p(\neg x,\top)$ for contingent formulae x, so that $p_\top$ is not a Kolmogorov function. The example can be generalized (see appendix).

Thus, the proto-probability functions are defined by purely order-theoretic conditions that are strictly weaker than the axioms of any of the usual systems for conditional probability, but are strong enough to support the rules defining Hawthorne's system Q of probabilistic inference. In this way, the theory of qualitative uncertain inference, like that of qualitative belief change, provides new perspectives on conditional probability.

Appendix

This appendix runs parallel to the main text. It contains most of the formal work, verifications, references and historical remarks supporting the main text.

For Section 1: Why Go Beyond the Ratio Rule?

The Kolmogorov postulates

There are several modes for presenting the Kolmogorov postulates for one-place probability functions, according to what we take as their domain. It may be a field of sets (most common in mathematics and applications), or equivalently a Boolean algebra (the preferred way of algebraists), or the set of all formulae of a propositional language (whose quotient structure under classical equivalence will be a free Boolean algebra). In this paper we work in the propositional mode, with the following formulation (Makinson 2005) of the postulates.

A (one-place) *proper Kolmogorov function* p: $L \rightarrow [0,1]$ is any function defined on the set L of formulae of a language closed under the Boolean connectives, into the real numbers from 0 to 1, such that:

(K1) $p(x) = 1$ for some formula x

(K2) $p(x) \leq p(y)$ whenever $y \in Cn(x)$

(K3) $p(x \vee y) = p(x)+p(y)$ whenever $\neg y \in Cn(x)$.

Cn is classical consequence; we also write $\approx$ for classical equivalence. Thus postulate (K1) tells us that 1 is in that range of p; (K2) says that $p(x) \leq p(y)$ whenever x classically implies y; (K3), called the rule of finite additivity, tells us that $p(x \vee y) = p(x)+p(y)$ whenever x is inconsistent with y. It is sometimes extended so as to constrain the probability of countable unions (most easily expressed in the field of sets mode).

As remarked in the text (and observed by several authors, notably Harper 1975 and subsequently Gärdenfors 1988, Leblanc and Roeper 1989), in comparative contexts it is convenient to regard the *unit function* (i.e. the function p that puts $p(x) = 1$ for every $x \in L$) as also being a Kolmogorov function, and we will follow this convention. It can be formalized by the simple expedient of defining a *Kolmogorov function* as one that is *either* proper Kolmogorov function (i.e. satisfies the above postulates) *or* is the unit function.Equivalently, one could weaken axiom (K3) by putting it under the proviso that p is not the unit function. We refer to the unit function as the *improper* Kolmogorov probability function.

The ratio rule

The *ratio rule* for conditional probability uses an arbitrary Kolmogorov function p: $L \rightarrow [0,1]$ to define a two-place function, conventionally written as $p(x|a)$ and read as 'the probability of x given a', defined on $L\text{x}\{a \in L\text{: } p(a) > 0\}$ by the rule: $p(x|a) = p(a \wedge x)/p(a)$ when $p(a) > 0$ and otherwise undefined.

Left projections

We recall the standard concept of the left projection f_a: $X \rightarrow Y$ of a two-place function f: $X\text{x}A \rightarrow Y$ from point $a \in A$, defined by putting $f_a(x) = f(x,a)$ for all $x \in X$.

For Section 1.1. Metaphysical vs Pragmatic Considerations

Metaphysical considerations

Such metaphysical views have been expressed by a number of probabilists, notably Rényi 1955 and 1970, de Finetti 1974 and by some philosophers, e.g. Hájek 2003.

Rényi 1955 (page 286) puts it briefly: "In fact, the probability of an event depends essentially on the circumstances under which the event possibly occurs, and it is a commonplace to say that in reality every probability is conditional". The same idea recurs at greater length in his 1970 (page 35).

De Finetti 1974 (page 134) similarly remarks: "Every evaluation of probability is conditional; not only on the mentality or psychology of the individual involved, at the time in question, but also, and especially, on the state of information in which he finds himself at that moment."

More recently, Hájek 2003 writes: "...given an unconditional probability, there is always a corresponding conditional probability lurking in the background. Your assignment of 1/2 to the coin landing heads superficially seems unconditional; but really it is conditional on tacit assumptions about the coin, the toss, the immediate environment, and so on. In fact, it is conditional on your total evidence."

Carnap's regularity condition

Carnap's formulation of the additional 'regularity' condition may be found in his 1950 section 53 axiom C53-3 and also 1971 chapter 2.7 page 101.

We note in passing that the concept of a 'counterfactual probability function' discussed by Boutilier 1995 (building on Stalnaker 1970) also assumes that the critical zone is empty. That concept, defined in the finite case, is a curious mixture of quantitative and qualitative ingredients. It puts $p(x,a)$, called the counterfactual probability of x given a, to be the proportion of the best a-states of the model that are x-states. The emptiness of the critical zone is assumed to ensure that the denominator is non-zero for consistent formulae a.

Getting around the critical zone with limits

For a brief account of the approach by taking limits, see the Wikipedia entry under the heading 'Regular conditional probability'. This use of the term 'regular' is quite different from that of Carnap.

For Section 1.2. Some Notational Niceties

Two-place functions could alternatively be distinguished from one-place ones by different type-faces, e.g. lower case for one and upper case for the other. However that convention meshes poorly with the standard notation for left projection, which we also need to use extensively.

For Section 2.1. A Leaf from the AGM Book

How important is the critical zone?

Our view of the importance of the critical zone is in contrast with that of many writers who minimize it. For example McGee 1994: "The problem we have been examining, how to revise one's system of beliefs upon obtaining new evidence that had prior probability 0, is not a problem that has any great practical significance."

Conditional probability in the light of counterfactual conditionals

An argument for going beyond the ratio definition of two-place probability may also be made in terms of counterfactual conditionals rather than belief revision. Indeed, this is way in which it is usually done in philosophical literature going back to Stalnaker 1970. However, in the author's view, the comparison with belief revision affords a clearer view, and also lends itself to the construction of very simple formal maps, as shown in section 5 and the corresponding part of the appendix.

Verifications of properties of B(p)

We verify the claims made in bullet points about belief sets for probability functions. Let the belief set $B(p)$ corresponding to one-place function p be defined by putting $B(p) = \{x\colon p(x) = 1\}$. This is also sometimes called the *top* of the function. Write B+a for the qualitative expansion of B by a, i.e. $B+a = Cn(B\cup\{a\})$. With $p_a(\cdot)$ understood as the left projection from a of the conditionalization $p(\cdot|\cdot)$ defined from $p(\cdot)$ by the ratio/unit rule, we want show: (1) in all cases, $B(p) \subseteq B(p)+a \subseteq B(p_a)$ and (2) in the limiting case that $p(a) = 0$ we have belief explosion: $B(p)+a = L = B(p_a)$, where L is the set of all propositions of the language.

For (1), the first inclusion is immediate from the definition of expansion above. To check the second inclusion, note that since $B(p_a)$ is closed under consequence it suffices to show that $a \in B(p_a)$ and $B(p) \subseteq B(p_a)$. The former is immediate since when $p(a) > 0$ then $p_a(a) = 1$ by the ratio definition and the Kolmogorov postulates for one-place probability, and $p_a(a)$ is also 1 when $p(a) = 0$, by the unit part of the ratio/unit definition. For the latter, it suffices to show that whenever $p(x) = 1$ then $p_a(x) = 1$. This is immediate when $p(a) = 0$. When $p(a) > 0$ we have $p_a(x) = p(a\wedge x)/p(a) = p(a)/p(a) = 1$ since the hypothesis $p(x) = 1$ implies that $p(a\wedge x) = p(a)$. For (2), it suffices to show further that when $p(a) = 0$ we have $B(p)+a = L$. But when the hypothesis holds then $p(\neg a) = 1$, so $\neg a \in B(p)$ and thus $B(p)+a \supseteq Cn(\neg a,a) = L$.

For Section 2.2. Bird's-eye view of available systems

Verification of consequences of the van Fraassen axioms

Left extensionality: $p(x,a) = p(x',a)$ whenever $x \approx x'$. *Verification*: By left projection, p_a is either a proper Kolmogorov function or the unit function. In the former case, $p(x,a) = p_a(x) = p_a(x') = p(x',a)$ using the hypothesis. In the latter case, $p(x,a) = p_a(x) = 1 = p_a(x') = p(x',a)$ irrespective of the hypothesis.

When $y \in Cn(x)$ then $p(x,a) \leq p(y,a)$. *Verification*: If $y \in Cn(x)$ then $x \approx y \wedge x$ so by left extensionality and product, $p(x,a) = p(y \wedge x,a) = p(y,a) \cdot p(x,a \wedge y) \leq p(y,a)$.

When $p(\cdot)$ is defined as $p(\cdot,\top)$, then we have the ratio rule. *Verification*: Suitably instantiating the product axiom, $p(a \wedge x,\top) = p(a,\top) \cdot p(x,a \wedge \top) = p(a,\top) \cdot p(x,a)$ using right extensionality, so if $p(a,\top) > 0$ we have $p(x,a) = p(a \wedge x,\top)/p(a,\top) = p(a \wedge x)/p(a)$.

When a is a contradiction, then p_a is the unit function. *Verification*: $p_a(a) = p(a,a) = 1 \leq p(x,a) = p_a(x)$, using left projection and an inequality already established.

The set Δ of all $a \in L$ such that p_a is the unit function is an ideal. *Verification*: To show that Δ is closed downwards, suppose $a \in \Delta$ and $a \in Cn(b)$. Then $1 = p(b \wedge x,a) = p(b,a) \cdot p(x,a \wedge b) = 1 \cdot p(x,a \wedge b) = \cdot p(x,b) = p_b(x)$, using the first supposition, left projection, first supposition again, second supposition respectively. To show that Δ is closed under disjunction, suppose p_a, p_b are both the unit function. To show that $p_{a \vee b}$ is also the unit function it suffices, by the left projection axiom to show that it is not a proper Kolmogorov function. Suppose it is; we get a contradiction. From the van Fraassen axioms we have $p(\perp,a \vee b) = p(\perp \wedge a,a \vee b) = p(a,a \vee b) \cdot p(\perp,a \wedge (a \vee b)) = p(a,a \vee b) \cdot p(\perp,a) = p(a,a \vee b) \cdot 1 = p(a,a \vee b)$ using the supposition that p_a is the unit function. Likewise $p(\perp,a \vee b) = p(b,a \vee b)$. By the supposition that $p_{a \vee b}$ is a proper Kolmogorov function we have $p(\perp,a \vee b) = 0$ so $p(a,a \vee b) = 0 = p(b,a \vee b)$. By the same supposition, $p(a \vee b,a \vee b) \leq p(a,a \vee b) + p(b,a \vee b) = 0+0 = 0$, contradicting the left projection axiom.

Finally, we check that a is abnormal iff $p(a,b) = 0$ for all normal b. *Verification*: From right to left, suppose $p(a,b) = 0$ for all normal b, but a is not abnormal. Then a is normal, so $p(a,a) = 0$, contradicting the left projection axiom. From left to right, suppose a is abnormal and b is normal. Then $a \wedge b$ is abnormal as already established, so $0 = p(\perp,b) = p(\perp \wedge a,b) = p(a,b) \cdot p(\perp,a \wedge b) = p(a,b) \cdot 1 = p(a,b)$ as desired.

Verification of the alternative axiomatization of the Popper system

Assume first the van Fraassen axioms plus (Positive); we need to show that $p(x,b) \neq 1$ for some x,b. By left projection, $p(\top,\top) = 1 > 0$ so by (Positive) $p_\top$ is proper and thus $p(\perp,\top) = 0 \neq 1$ as desired. Now assume the van Fraassen axioms plus $p(x,b) \neq 1$ for some x,b. Suppose $p(a,\top) > 0$; we need to show that p_a is proper, for which it suffices to show that it is not the unit function. First note that $p(\perp,\top) = p(\perp \wedge b,\top) = p(b,\top) \cdot p(\perp,\top \wedge b) = p(b,\top) \cdot p(\perp,b)$; but since $p(x,b) \neq 1$ it follows that p_b is proper so $p(\perp,b) = 0$ and thus $p(\perp,\top) = 0$. But also $p(\perp,\top) = p(\perp \wedge a,\top) = p(a,\top) \cdot p(\perp,a)$, so since $p(a,\top) > 0$ we have $p(\perp,a) = 0$ so that p_a is not the unit function, as desired.

Example of a 'mixed' function

Leblanc and Roeper 1989 gave an example of a two-place function satisfying the Popper postulates, whose treatment of formulae with probability zero is a mix of the expansionary and revisionary policies. They presented it rather enigmatically as an 8·8 table (their Table 5). We provide it with a more transparent rule-based presentation, which for convenience we express as a field of sets.

Take the field F of all subsets of the three-element set $S = \{\alpha,\beta,\gamma\}$. For motivation, think of α,β,γ as of increasing levels of importance beginning from α, which has no importance at all. Put $p(x,a) = 1$ unless there is some item of positive importance in a and the item of greatest importance in a is not in x. Formally, we define p: $S^2 \to [0,1]$, in fact into $\{0,1\}$as follows:

1. If $\gamma \in a$ then $p(x,a) = 1$ if $\gamma \in x$, otherwise $p(x,a) = 0$
2. If $\gamma \notin a$ but $\beta \in a$ then $p(x,a) = 1$ if $\beta \in x$, otherwise $p(x,a) = 0$
3. If $\gamma \notin a$ and $\beta \notin a$ then $p(x,a) = 1$.

This function is a mix of the two kinds of conditional probability: $p(\{\beta\},S) = 0 = p(\{\alpha\},S)$ applying the first clause, but $p(\emptyset,\{\beta\}) = 0$ applying the second while $p(\emptyset,\{\alpha\}) = 1$ by the third. On the other hand, it is straightforward to check that it satisfies the Popper axioms.

Historical development of conditional probability

We review the historical steps in the construction of axioms for two-place probability functions, working backwards from Popper 1959. For ease of comparison, we consider them all in the propositional mode, and treat each as defined on the whole of L^2, but comment on particularities of the original formulations each as we go.

Popper's original postulates for two-place probability functions, contained in an appendix of Popper 1959 (recalled e.g. in Leblanc and Roeper 1989 and more accessibly Koons 2009) were in the propositional mode. They reflected a desire for the autonomy of probability theory from logic, abstract algebra and set theory and so avoided any use of concepts from those areas. But if we are happy to use concepts of classical logic in our presentation then, as shown by subsequent writers, they may be given more perspicuously, as in the following formulation of Hawthorne 1996, which require, for p: $L^2 \to [0,1]$, that:

(P0) $p(x,a) \neq 1$ for some formulae a, x
(P1) $p(x,a) = p(x,b)$ whenever $a \approx b$
(P2) $p(x,a) = 1$ whenever $x \in Cn(a)$
(P3) either $p(x \vee y,a) = p(x,a)+p(y,a)$ whenever $\neg(x \wedge y) \in Cn(a)$, or p_a is the unit function
(P4) $p(x \wedge y,a) = p(y,a) \cdot p(x,y \wedge a)$

Of course, if we are working in the context of fields of sets, (P1) becomes vacuous. Warning: The term 'Popper function' is sometimes used rather loosely, to refer to almost any primitive two-place probability function defined over the critical zone. For example, Lindström and Rabinowicz 1989 use the term to refer to the narrower class of Hosiasson-Lindenbaum functions, defined below.

Our modular presentation takes from Rényi 1955, 1970, 1970a his leading idea that for 'most' values of a, the left projection from a will be a proper Kolmogorov function giving a the value 1, and so is very similar in gestalt. But in its details, Rényi's system is rather different from any of those we have considered. Formulated in the field-of-sets mode, it treats the right domain as a *parameter*, allowing it to be chosen as any subset of the left domain that is consistent with the axioms. These axioms are just the product rule and the principle that p_a is a *proper* one-place Kolmogorov function with $p_a(a) = 1$, both formulated under the restriction that all values of the right argument takes a value in the restricted right domain. For values of the right argument outside that subset, the probability functions are left undefined. We are thus given a scheme for a *family* of axiom sets, one for each choice of right domain.

This yields the Popper axioms if we constrain the right domain to include $\{a$: $p(a,S) > 0$, where S is the set on which the field is based, and carry out the following editing: (a) put $p(x,a) = 1$ for all a with $p(a,S) = 0$, (b) ensure consistency by allowing in the left projection axiom that p_a may be improper (as in the axiom (vF2) of section 2.2), (c) for the one-place Kolmogorov functions mentioned in the left projection axiom, weaken Rényi's assumption of countable to finite additivity, and finally (d) translate from the field-of-sets mode to the propositional one.

The system of Hosiasson-Lindenbaum 1940 concerned for what she called 'confirmation' functions, writing them as $c(x,a)$ rather than $p(x,a)$ and working in the propositional mode. This ground-breaking work has been comparatively neglected, despite its accessible and respected place of publication. In particular, the paper is not mentioned in Rényi 1955, 1970, or 1970a, nor in the wide-ranging discussion of Harper 1975 or the comprehensive study of Roeper and Leblanc 1999. Popper 1959 does mention Hosiasson-Lindenbaum in passing, but with respect to other questions and without citing her 1940 paper. This contrasts with the explicit acknowledgement (note 12 in new appendix iv) of the influence of Rényi 1955 on his thinking.

Hosiasson-Lindenbaum excluded inconsistent propositions from the right domain. Restoring them, we get the following axioms:

(HL1) $p(x,a) = 1$ whenever $x \in Cn(a)$

(HL2) $p(x \vee y,a) = p(x,a)+p(y,a)$ whenever $\neg(x \wedge y) \in Cn(a)$, provided a is consistent

(HL3) $p(x \wedge y,a) = p(x,a) \cdot p(y,a \wedge x)$ for all formulae a, b, x, y

(HL4) $p(x,a) = p(x,b)$ whenever $a \approx b$.

Axiom (HL2) thus broadens the conditions under which the left projection of a two-place function satisfies additivity and is thus a proper Kolmogorov function, from the narrower condition that $p(a,\top) > 0$ to the wider one that a is consistent. The system may be obtained fron Rényi's scheme by setting the right domain at the set of all non-empty sets of S and editing by first putting $p(x,\emptyset) = 1$ and then as for Popper's system.

In what respect can it be said that Rényi's formulation was an advance on that of Hosiasson-Lindenbaum? For working mathematicians and statisticians, its use of the field-of-sets mode made application to practical problems more transparent. But at a deeper level, the step forward was *conceptual* – the realization that a rather arbitrary-looking axiom system becomes natural if we build it around the idea that for 'most' values of the right argument, the left projection will be a proper one-place probability function. As Rényi put it: "a conditional probability space is nothing else than a set of ordinary probability spaces which are connected with each other by [the product axiom]" (Rényi 1955 pp 289-290).

Mini-note: We reverse a correction made by Hailperin 1991 (page 75) to the effect that since Hosiasson-Lindenbaum's formulation is in the propositional mode, it needs a left companion to (HL4) stating that $p(x,a) = p(y,a)$ whenever $x \approx y$. In fact, this follows from the postulates as given. In the limiting case that a is inconsistent we have $p(x,a) = 1 = p(y,a)$ by (HL1), so suppose a is consistent and $x \approx y$. Then $\neg(x\wedge\neg y) \in Cn(a)$, so by the additivity axiom (HL2) we have $p(x\vee\neg y,a) = p(x,a)+p(\neg y,a)$. But the supposition also gives us LHS = 1 by (HL1), so $p(x,a)+p(\neg y,a) = 1$. Moreover, (HL1) and (HL2) imply that $p(\neg y,a) = 1 - p(y,a)$, and so by arithmetic $p(x,a) = p(y,a)$. Essentially this point was already made by Tarski with regard to the earlier axiomatization of Mazurkiewicz 1932 (discussed below), and was acknowledged in footnote 1 of that paper.

Hosiasson-Lindenbaum 1940 states that her axioms for two-place probability are 'analogous' to still earlier ones of Mazurkiewicz 1932. In fact, they constitute a major simplification and clarification of his quite complex system, which requires the left domain to contain *individual* propositions, while the right one contains consistent *sets* of propositions closed under classical consequence – the two kinds of proposition drawn, moreover, from intersecting and not very clearly defined languages. In his only example, Mazurkiewicz considers a game: the left argument of $p(x,A)$ can be filled by a proposition describing a state of play, while the right one can be occupied by a closed set of propositions containing the rules of the game, the current state of play, and any mathematical apparatus needed for deductions.

In turn, Mazurkiewicz states that he is taking as his starting point the axioms of Bohlmann 1909. However, Bohlmann's postulates are for *one*-place probability in a mode of unanalysed items called events and occurrences, which he supplements with an 'axiom' defining conditional probability by the ratio rule.

For some late nineteenth-century uses of conditional probability (without any attempt at axiomatization) see Hailperin 1988.

Thus our trail into the history of axiomatizations of two-place probability that cover the critical zone appears to end with Mazurkiewicz 1932 as first serious attempt, Hosiasson-Lindenbaum 1940 as the first really successful one, and Rényi 1955 for providing a clear gestalt.

For Section 3.1. Hosiasson-Lindenbaum *vs* Popper

We briefly review some interesting, but in the end inconclusive reasons that could be given for regarding one or other of the Hosiasson-Lindenbaum and Popper accounts as intrinsically preferable to the other.

An unconvincing argument for Hosiasson-Lindenbaum: fineness of grain

It might be said that since classical *Cn* makes finer discriminations than a supraclassical *Cn′*, probability theory with *Cn* as background consequence is more fine-grained than with *Cn′*, giving an advantage to Hosiasson-Lindenbaum over Popper. But this is not very convincing. In the propositional mode, we can use *Cn′* in the description of our probability functions but still maintain *Cn* for other purposes: we can have our cake and eat it. In the field-of-sets or algebraic mode, passage to the quotient structure determined by a suitable filter would indeed lose information, but we are not obliged to effect the passage. We can work with Popper functions in the original structure, simply knowing that we could, if desired, do so in the quotient structure.

A misdirected argument for Popper: strengthening conditions

As well as passing from an unconditional to a conditional function, we often need to strengthen the condition of an already conditional one. It could be useful to be able to express this an operation taking a two-place function $p(\cdot,\cdot)$ to another two-place function $p_{\wedge b}(\cdot,\cdot)$ by the rule $p_{\wedge b}(x,a) = p(x,a\wedge b)$. But this operation breaks the boundaries of Hosiasson-Lindenbaum functions. It may happen that while a is consistent, $a\wedge b$ is not, in which case for any function $p(\cdot,\cdot)$ satisfying the Hosiasson-Lindenbaum axioms, $(p_{\wedge b})_a$ is the unit function despite the consistency of a, so that $p_{\wedge b}$ does not satisfy axiom (HL).

However, Popper functions face a similar problem. Consider any Popper function $p(\cdot,\cdot)$, and let a be an inconsistent proposition. Then $p_{\wedge a}(a,\top) = p(a,a) = 1 > 0$, while for all values of x we have $(p_{\wedge a})_a(x) = p_{\wedge a}(x,a) = p(x,a) = 1$ since a is inconsistent, so that $(p_{\wedge a})_a$ is the unit function. These two facts contradict the distinctive Popper axiom (Positive).

The only way to keep our class of two-place functions closed under the ‘conjoined condition’ operation is to drop (Positive) and retreat to the van Fraassen system. Thus the convenience of being able to strengthen conditions could be seen as a point in favour of the usefulness of that very basic system.

Changing the underlying consequence operation

If one is working in the mode of fields-of-sets or Boolean algebras as carriers for the probability functions, then one can similarly express Popper functions as Hosiasson-Lindenbaum ones by passing to the quotient algebra determined by the same filter as in the propositional case. Essentially this construction was described, under a different light, by Harper 1975 (page 234) and more explicitly Harper 1976 (section 6).

For Section 4.1. A Map from Conditional Probability to AGM revision

We verify the claims made in the text about the map from HL conditional probability functions to AGM revision operations.

For HL functions, the left projection p_a of $p(\cdot,\cdot)$ from a is a proper Kolmogorov one-place probability function whenever a is consistent (section 2.5), so under that condition we can apply well-known properties of the one-place functions without detailed justification, as well as the HL axioms themselves.

To show that $K = B(p)$ is a belief set, suppose $y \in Cn(K)$; we need to check that $y \in K$. By compactness, $y \in Cn(\wedge x_i\text{: } i \leq n\}$ for some $x_1,.., x_n \in B(p)$, so each $p(x_i,\mathsf{T}) = 1$, $p(\wedge x_i,\mathsf{T}) = 1$ and thus $p(y,\mathsf{T}) = 1$ so that $y \in B(p)$. To show that $B(p)$ is consistent we need then only note that $p(\perp,\mathsf{T}) = 0$.

Let p: $L^2 \rightarrow [0,1]$ be any HL function. We need to check that the associated function $*$: $L \rightarrow 2^L$ satisfies each of the AGM postulates (K$*$1) through (K$*$8) with respect to $K = B(p)$. Two general remarks before the details:

- The AGM postulates for revision were first formulated in Gärdenfors 1984 and a convenient overview may be found in Peppas 2007, whose presentation we follow. We note in passing that the classic account in Alchourrón, Gärdenfors and Makinson 1985 focused on contraction, and its axiomatization of revision contains a confusion: it omits postulate (K$*$3) below, and treats the definition of revision from contraction via the Harper identity as if it were a postulate.
- We are not verifying satisfaction with respect to an *arbitrary* belief set K, but with respect to the *specific* belief set depending on the choice of p, namely $K = B(p) = \{x\text{: } p(x,\mathsf{T}) = 1\}$. This specification is needed for (K$*$3) and (K$*$4) – though not for the other postulates, in which K does not appear in unrevised form.

(K$*$1): $K*a = Cn(K*a)$. *Verification*: Same as the above for $B(p) = Cn(B(p))$, but replacing T by a.

(K$*$2): $a \in K*a$. *Verification*: We need $p(a,a) = 1$, immediate.from axiom (vF2).

(K$*$3): $K*a \subseteq Cn(K \cup \{a\})$. *Verification*: Suppose $y \in$ LHS, so that $p(y,a) = 1$. We need to show that $y \in Cn(K \cup \{a\}) = Cn(B(p) \cup \{a\}) = Cn(\{x\text{: } p(x,\mathsf{T}) = 1\} \cup \{a\})$, so it suffices to show that $\neg a \vee y \in \{x\text{: } p(x,\mathsf{T}) = 1\}$, i.e. that $p(\neg a \vee y,\mathsf{T}) = 1$. Now $p(\neg a \vee y,\mathsf{T}) = p(\neg a \vee (a \wedge y),\mathsf{T}) = p(\neg a,\mathsf{T}) + p(a \wedge y,\mathsf{T})$. But $p(a \wedge y,\mathsf{T}) = p(a,\mathsf{T}) \cdot p(y,a) = p(a,\mathsf{T})$ since $p(y,a) = 1$. Thus $p(\neg a \vee y,\mathsf{T}) = p(\neg a,\mathsf{T}) + p(a,\mathsf{T}) = p(\mathsf{T},\mathsf{T}) = 1$ as desired.

(K$*$4): $Cn(K \cup \{a\}) \subseteq K*a$ whenever a is consistent with K. *Verification*: Suppose $y \in Cn(K \cup \{a\})$ and a is consistent with K; we need to show $p(y,a) = 1$. By the first supposition, $\neg a \vee y \in Cn(\wedge x_i\text{: } i \leq n\}$ for some $x_1,.., x_n \in K = B(p)$ with each $p(x_i,\mathsf{T}) = 1$, so that $p(\wedge x_i,\mathsf{T}) = 1$ and thus $p(\neg a \vee y,\mathsf{T}) = 1$. Hence $p(a,\mathsf{T}) = p(a \wedge (\neg a \vee y),\mathsf{T}) = p(a \wedge y,\mathsf{T})$. But also we have $p(a \wedge y,\mathsf{T}) = p(a,\mathsf{T}) \cdot p(y,a)$. Putting these together, $p(a,\mathsf{T}) = p(a,\mathsf{T}) \cdot p(y,a)$. But by supposition, $\neg a \notin K = B(p)$ so $p(\neg a,\mathsf{T}) \neq 1$ so $p(a,\mathsf{T}) \neq 0$, so by arithmetic $p(y,a) = 1$ as desired.

(K$*$5): $K*a$ is consistent whenever a is consistent. *Verification*: Suppose a is consistent; we need $p(\bot,a) = 0$, which is immediate.

(K$*$6): If $a \approx b$ then $K*a \approx K*b$. *Verification*: Suppose $a \approx b$; we need $p(x,a) = 1$ iff $p(x,b) = 1$, again immediate.

(K$*$7): $K*(a\wedge b) \subseteq Cn((K*a)\cup\{b\})$. *Verification*: Suppose $x \in$ LHS, so that $p(x,a\wedge b) = 1$. It suffices to show that $\neg b\vee x \in K*a$, i.e. that $p(\neg b\vee x,a) = 1$. When a is inconsistent, this is immediate, so suppose that a is consistent. From the supposition, $p(\neg b\vee x, a\wedge b) = 1$. Now $p(b\wedge x,a) = p(b\wedge(\neg b\vee x),a) = p(b,a)\cdot p(\neg b\vee x,a\wedge b) = p(b,a)\cdot 1 = p(b,a)$, so since a is consistent we have $p(b\wedge\neg x, a) = 0$ so $p(\neg b\vee x, a) = 1$ as desired.

(K$*$8): $Cn((K*a)\cup\{b\}) \subseteq K*(a\wedge b)$ whenever b is consistent with $K*a$. *Verification*: Suppose that $y \in$ LHS and b is consistent with $K*a$; we need to show that $p(y,a\wedge b) = 1$. We have already verified $a \in K*a$, so the second supposition gives us the consistency of a. We now proceed along lines similar to the verification of (K$*$4). By the first supposition, $\neg b\vee y \in K*a$, i.e. $p(\neg b\vee y,a) = 1$. Hence since a is consistent, $p(b,a) = p(b\wedge(\neg b\vee y),a) = p(b\wedge y,a)$. But also $p(b\wedge y,a) = p(b,a)\cdot p(y,a\wedge b)$. Putting these together, $p(b,a) = p(b,a)\cdot p(y,a\wedge b)$. But by the second supposition again, $\neg b \notin K*a$ so $p(\neg b,a) \neq 1$ and thus $p(b,a) \neq 0$, so by arithmetic $p(y,a\wedge b) = 1$ as desired.

For Section 4.2. Properties of the Map: Surjective but not Injective

Failure of injectivity

For the failure of injectivity, it suffices to find two distinct HL functions $p \neq p'$ with $*_p = *_{p'}$, i.e. with $*_p(a) = \{x: p(x,a) = 1\} = \{x: p'(x,a) = 1\} = *_{p'}(a)$ for all $a \in L$, i.e. with $p(x,a) = 1$ iff $p'(x,a) = 1$ for all $a,x \in L$. For simplicity we do this with Boolean algebras rather than propositional languages. Take any finite Boolean algebra with $n \geq 2$ atoms, and two distinct probability distributions f,f' to these atoms with each atom getting a non-zero probability; extend them to one-place probability functions (for simplicity using the same names) on the entire algebra. Noting that every non-zero element of the algebra receives a non-zero probability under each of these functions, we can define two-place functions p,p': $L^2\to[0,1]$ by the ratio rule for non-zero right arguments and putting $p(x,0) = p'(x,0) = 1$. These are HL probability functions, in fact they are Carnap functions. Then for all a,x we have $p(x,a) = 1$ iff $p(a\wedge x) = p(a)$ i.e. iff $a \leq x$ and likewise for p', and so $p(x,a) = 1$ iff $p'(x,a) = 1$ as desired.

Surjectivity

Suppose that the language is finite, and let $*:L\to 2^L$ satisfy the AGM postulates with respect to a consistent set K. Define $p(\cdot,\cdot)$ by the rule: $p(x,a) = 1$ for all $x \in L$ in the limiting case that a is inconsistent, while in the principal case that a is consistent $p(x,a)$ is the proportion of $(K*a)$-worlds that are x-worlds. We need to show that (1) p satisfies the HL axioms, (2) $* = *_p$, and (3) $K = B(p)$.

For (1) it is convenient to check the HL axioms in the form given to them by Hosiasson-Lindenbaum 1940 (see appendix to section 2.2), as follows.

(HL1) $p(x,a) = 1$ whenever $x \in Cn(a)$. *Verification*: If a is inconsistent then we have $p(x,a) = 1$ by the definition for that case, so we may suppose that a is consistent. By AGM, $a \in K{*}a$ so if $x \in Cn(a)$ we have $x \in Cn(K{*}a) = K{*}a$. Thus all $(K{*}a)$-worlds are x-worlds, i.e. the proportion of $(K{*}a)$-worlds that are x-worlds is 1, so $p(x,a) = 1$ as required.

(HL2) $p(x \vee y,a) = p(x,a)+p(y,a)$ whenever a is consistent but $\neg(x \wedge y) \in Cn(a)$. *Verification*: Suppose a is consistent and $\neg(x \wedge y) \in Cn(a)$. By the first supposition, we need to consider proportions, and by the second the proportion of $(K{*}a)$-worlds that are $(x \vee y)$-worlds is the sum of the proportions of $(K{*}a)$-worlds that are, separately, x-worlds or y-worlds, and we are done.

(HL3) $p(x \wedge y,a) = p(x,a) \cdot p(y,a \wedge x)$. *Verification*: If a is inconsistent then so is $a \wedge x$ and LHS = 1 = RHS. Suppose a is consistent. If $a \wedge x$ is inconsistent then LHS = 0 while RHS = $0 \cdot 1 = 0$ and again we are done. If $a \wedge x$ is consistent then LHS is the proportion of $(K{*}a)$-worlds that are $(x \wedge y)$-worlds, while RHS is the proportion of $(K{*}a)$-worlds that are x-worlds multiplied by the proportion of $(K{*}a \wedge x)$-worlds that are y-worlds. If x is inconsistent with $K{*}a$ then both LHS and RHS equal 0, so we may suppose that x is consistent with $K{*}a$. Then by AGM axioms (K$*$7) and (K$*$8) the $(K{*}a \wedge x)$-worlds are just the $(K{*}a)$-worlds that are x-worlds. Hence RHS is the proportion of $(K{*}a)$-worlds that are x-worlds multiplied by the proportion of *those* that are y-worlds, which equals the proportion of $(K{*}a)$-worlds that are $(x \wedge y)$-worlds, equalling the LHS and we are done.

(HL4) $p(x,a) = p(x,b)$ whenever $a \approx b$. *Verification*: If a is inconsistent then so is b, so LHS = 1 = RHS. If a is consistent, then if $a \approx b$ the a-worlds are just the b-worlds, and the proportion of a-worlds that are x-worlds is the same as the proportion of b-worlds that are x-worlds.

To show that (2) $* = *_p$, consider first the principal case that a is consistent, where we need only note that by the definition of $*_p$ we have $x \in *_p a$ iff $p(x,a) = 1$ while, by the definition of p, also $p(x,a) = 1$ iff every $(K{*}a)$-world is an x-world, i.e. iff $x \in Cn(K{*}a) = K{*}a$. In the limiting case that a is inconsistent, $p(x,a) = 1$ for every x and by the AGM postulates, $x \in K{*}a$ for every x, so again we are done. Finally, to check (3) that $K = B(p)$ we need only show that $x \in K$ iff $p(x,\top) = 1$. But since K is consistent, the AGM postulates tell us that $K = K{*}\top$, and the equivalence $p(x,a) = 1$ iff $x \in K{*}a$ just established may be applied substituting $\top$ for a.

We conjecture that surjectivity fails in the infinite case. Evidently its present proof breaks down there, since one cannot meaningfully speak of proportions of infinite sets, thus blocking the definition of $p(\cdot,\cdot)$ above. Nor is it possible to repair the proof by replacing proportionality by some probability distribution that gives each world a non-zero value. For if the set of formulae is countable, there are continuum many worlds and as is well known, there is no probability distribution on a non-countable set that gives a non-zero value to each element.

For Section 5.3. Proto-probability Functions for Qualitative Inference

The system Q

For further information on systems O and Q see part I of Hawthorne and Makinson 2007.

Disjunctive interpolation

As remarked in the text, the principle of disjunctive interpolation is closely related to a rule discussed by Koopman 1940, 1940a. Called ‘alternative presumption’, it states that whenever both $p(x,a\wedge b)$, $p(x,a\wedge\neg b) \leq p(y,c)$ then $p(x,a) \leq p(y,c)$. To be precise, if we assume that the order $\leq$ is complete (as we do, although Koopman does not), alternative presumption is equivalent to the right half of disjunctive interpolation. *Verification*. To obtain Koopman: by completeness of $\leq$, either $p(x,a\wedge b) \leq p(x,a\wedge\neg b)$ or conversely; in e.g. the former case we have $p(x,(a\wedge b)\vee(a\wedge\neg b)) \leq p(x,a\wedge\neg b)$ and by right extensionality and transitivity of $\leq$ we are done. In the converse direction, suppose $p(x,a) \leq p(x,b)$ and $\neg b \in Cn(a)$, we want to show the right hand part of disjunctive interpolation, i.e. that $p(x,a\vee b) \leq p(x,b)$. We need only note that $p(x,b) \approx p(x,(a\vee b)\wedge b)$ and, given the last supposition, that $p(x,a) \approx p(x,(a\vee b)\wedge\neg b)$, then apply Koopman with a little help from left extensionality, taking (y,c) as (x,b).

Verification that Q is proto-probablistically sound

We check that when we take any proto-probability function $p(\cdot,\cdot)$ and $t \in D$, and define a relation by putting $a \mid\!\sim_{pt} x$ iff $p(x,a) \geq t$, then $\mid\!\sim_{pt}$ satisfies all the rules of Q. For (O1) we need $p(a,a) \geq t$, which is immediate from (P1). For (O2), we need that when $p(x,a) \geq t$ and $y \in Cn(x)$ then $p(y,a) \geq t$, which is immediate from (P2). For (O3), we need that when $p(x,a) \geq t$ and $a \approx b$ then $p(x,b) \geq t$, which is immediate from (P3). For (O4), we need that when $p(x\wedge y,a) \geq t$ then $p(y,a\wedge x) \geq t$, which is immediate from (P4). For (O5) alias XOR, we need that when $p(x,a) \geq t$, $p(x,b) \geq t$ and $\neg b \in Cn(a)$, then $p(x,a\vee b) \geq t$. Since the order on D is complete, either $p(x,a) \leq p(x,b)$ or conversely, consider e.g. the former. Then using the left half of (P5), $p(x,a) \leq p(x,a\vee b)$ and we are done by transitivity of $\leq$. For (O6) alias WAND, we need that when $p(x,a) \geq t$ and $p(y,a\wedge\neg y) \geq t$ then $p(x\wedge y,a) \geq t$. If $t = 0_D$ then this is immediate, and if $t \neq 0_D$ it is immediate from (P6). It remains to obtain the non-Horn rule (NR) of negation rationality. We need to show that when $\neg b \in Cn(a)$ and $p(x,a\vee b) \geq t$ then either $p(x,a) \geq t$ or $p(x,b) \geq t$. Since the order on D is complete, either $p(x,a) \leq p(x,b)$ or conversely, consider e.g. the former. Then using the right half of (P5), $p(x,a\vee b) \leq p(x,b)$ giving $p(x,b) \geq t$ as desired.

Generalizing an example

We can generalize the example given in the text of a proto-probability function that is not a van Fraassen function. Fix any consistent formula c and put $p(x,a) = 1$ when $x \in Cn(a\wedge c)$, else $p(x,a) = 0$. Then $p(\cdot,\cdot)$ satisfied (P1) through (P6), but choosing x as any formula independent of c it fails (vF2) as before.

Acknowledgements

Thanks to Horacio Arló Costa, Jonny Blamey, Richard Bradley, Franz Dietrich, Peter Gärdenfors, Jim Hawthorne, Rohit Parikh, Wlodek Rabinowicz and Miklos Redei for remarks on drafts and oral presentations. Also to Hykel Hosni, Jacek Malinowski, and Miklos again for assistance in obtaining copies of rather inaccessible texts of Bohlmann, Mazurkiewicz and Rényi.

References

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.

Arló Costa, Horacio 2001. 'Bayesian epistemology and epistemic conditionals: on the status of export-import laws' *The Journal of Philosophy* 98: 555-593.

Arló Costa, Horacio and Rohit Parikh 2005. 'Conditional probability and defeasible inference' *Journal of Philosophical Logic* 34: 97-119.

Bohlmann, G. 1909. 'Die Grundbegriffe der Wahrscheinlichkeitsrechnung in ihrer Anwendung auf die Lebensversicherung' pages 244-278 of *Atti del IV Congresso Internazionale dei Matematici, Roma 6-11 Aprile 1908* (Roma: Accademia dei Lincei), vol. III..

Boutilier, Craig 1995.'On the revision of probabilistic belief states' *Notre Dame Journal of Formal Logic* 36: 158-183.

Carnap, Rudolf 1950. *Logical Foundations of Probability* (Chicago University Press).

Carnap, Rudolf 1952. *The Continuum of Inductive Methods* (Chicago University Press).

Carnap, Rudolf 1971. 'A basic system of inductive logic' pages 33-165 of Carnap and Jeffries eds *Studies in Inductive Logic and Probability* (University of California Press).

de Finetti, Bruno 1974. *Theory of Probability* (New York: Wiley).

Gärdenfors, Peter 1986. 'The dynamics of belief: contractions and revisions of probability functions' *Topoi* 5: 29-37.

Gärdenfors, Peter 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States* (Cambridge Mass: MIT Press).

Goosens, William K. 1979. 'Alternative axioms of elementary probability theory' *Notre Dame Journal of Formal Logic* 20: 227-239.

Hájek, Alan 2003. 'What conditional probability could not be' *Synthese* 137: 273-323.

Hailperin, Theodore 1988. 'The development of probability logic from Leibniz to MacColl' *History and Philosophy of Logic* 9: 131-191.

Hailperin, Theodore 1991. 'Probability logic in the twentieth century' *History and Philosophy of Logic* 12: 71-110.

Harper, W.L. 1975. 'Rational belief change, Popper functions and counterfactuals' *Synthese* 30: 221-262. Reprinted with minor editorial changes in pages 73-115 of

W.L. Harper and C.A. Hooker eds *Foundations of Probability, Statistical Inference, and Statistical theories of Science, Vol I* (Dordrecht: Reidel).

Harper, W.L. 1976. 'Rational conceptual change' pages 462-494 of *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1976, Vol. II* (University of Chicago Press).

Hawthorne, James 1996. 'On the logic of nonmonotonic conditionals and conditional probabilities' *Journal of Philosophical Logic* 25: 185-218.

Hawthorne, James and David Makinson 2007. 'The quantitative/qualitative watershed for rules of uncertain inference' *Studia Logica* 86: 249-299.

Hosiasson-Lindenbaum 1940. 'On confirmation' *The Journal of Symbolic Logic* 4: 133-148.

Koons, Robert 2009. 'Supplement to Defeasible Reasoning' *Stanford Encyclopedia of Philosophy* http://plato.stanford.edu

Koopman, Bernard O. 1940. 'The axioms and algebra of intuitive probability' *The Annals of Mathematics* 41: 269-292.

Koopman, Bernard O. 1940a. 'The bases of probability' *Bulletin of the American Mathematical Society* 46: 763-774.

Leblanc, Hugues and Peter Roeper 1989. 'On relativizing Kolmogorov's absolute probability functions' *Notre Dame Journal of Formal Logic* 30: 485-512.

Lindström, Sten and Wlodzimierz Rabinowicz 1989. 'On probabilistic representation of non-probabilistic belief revision' *Journal of Philosophical Logic* 18: 69-101.

Makinson, David 1997. 'Screened revision' *Theoria* 63: 14-23.

Makinson, David 2005. *Bridges from Classical to Nonmonotonic Logic* (London: College Publications).

Makinson, David (to appear). 'Logical questions behind the lottery and preface paradoxes: lossy rules for uncertain inference' *Synthese*, issue commemorating Henry Kyburg.

Mazurkiewicz, S. 1932. 'Zur Axiomatic der Wahrscheinlichkeitsrechnung' *Comptes rendus des séance de la Société des Sciences et des Lettres de Varsovie* (Sprawozdania z posiedzeń Towarzystwa Naukowego Warszawskiego) 25 : 1-4.

McGee, Vann 1994. 'Learning the impossible' pages 179-199 of E. Eells and Brian Skyrms eds *Probability and Conditionals: Belief Revision and Rational Decision* (Cambridge University Press).

Peppas, Pavlos 2007. 'Belief revision' chapter 8 of F. van Harmelen, V. Lifschitz, B. Porter eds *Handbook of Knowledge Representation* (Amsterdam: Elsevier)

Popper, Karl 1959. *The Logic of Scientific Discovery* second edition (New York: Basic Books)

Rényi, Alfréd 1955. 'On a new axiomatic theory of probability' *Acta Mathematica Academiae Scientiae Hungaricae* 6: 268-335.

Renyi, Alfréd 1970. *Foundations of Probability Theory* (San Francisco: Holden-Day).

Rényi, Alfréd 1970a. *Probability Theory* (Amsterdam: North-Holland).

Roeper, Peter and Hugues Leblanc 1999. *Probability Theory and Probability Logic* (Toronto: University of Toronto Press).

Stalnaker, R.C. 1970. 'Probability and conditionals' *Philosophy of Science* 37: 64-80, reprinted in pages 107-128 of W.L. Harper et al eds *Ifs* (Dordrecht: Reidel, 1981).

Spohn, Wolfgang 1986. 'The representation of Popper measures' *Topoi* 5:69-74.

Spohn, Wolfgang 2009. 'A survey of ranking theory', pages 185-228 of Franz Huber et al eds *Degrees of Belief* (New York: Springer)

van Fraassen, Bas 1976. 'Representation of conditional probabilities' *Journal of Philosophical Logic* 5: 417-430.

van Fraassen, Bas 1995. 'Fine-grained opinion, probability, and the logic of full belief' *Journal of Philosophical Logic* 24: 349-377.

Wikipedia (http://en.wikipedia.org). Entry on 'Regular conditional probability', August 2009.

Department. of Philosophy, Logic and Scientific Method
London School of Economics
Houghton Street, London WC2A 2AE, United Kingdom
david.makinson@gmail.com

16 November 2009