A Logic of Comparative Support: Qualitative Conditional Probability Relations Representable by Popper Functions

James Hawthorne
University of Oklahoma

Alan Hájek and Chris Hitchcock, eds.,
Oxford Handbook of Probabilities and Philosophy, Chapter 13, Oxford University Press.

Keywords: Bayesian confirmation theory, comparative argument strength, comparative probability, conditional probability, qualitative probability, Popper functions

Abstract:
This article presents axioms for *comparative conditional probability relations*. The axioms presented here are more general than usual. Each comparative relation is a weak partial order on pairs of sentences (i.e. each relation is transitive and reflexive) but need not be a complete order relation. The axioms for these comparative relations are *probabilistically sound* for the broad class of conditional probability functions known as *Popper functions*. Furthermore, these axioms are *probabilistically complete*. Arguably, the notion of *comparative conditional probability* provides a foundation for Bayesian confirmation theory. Bayesian confirmation functions are overly precise *probabilistic representations* of the more fundamental logic of comparative support. The most important features of evidential support are captured by comparative relationships among argument strengths, realized by the *comparative support relations* and their logic.

1. Introduction

Underlying the usual numerical conception of probability is a more basic qualitative notion, that of *comparative probability*. This comparative notion is formally expressed by weak (partial) order relations among sentences or propositions of the form ‘$A \succcurlyeq B$’, read “A *is at least as probable as* B”. These relations may be employed to represent the *comparative confidence relations* for idealized agents. Interpreted this way, a relation of form ‘$A \succcurlyeq_\alpha B$’ says that agent α is at least as confident that A is true as that B is true.

Each comparative probability relation $\succcurlyeq$ that obeys certain reasonable constraints (expressed as axioms) can be represented by a corresponding probability function P – i.e. it can be proved that $A \succcurlyeq B$ holds just when $P[A] \geq P[B]$ holds, provided the relation $\succcurlyeq$ is a *complete* (rather than *partial*) order. (See B. de Finetti 1937, 1974, and L. J. Savage 1954.) Thus, one possible answer to the problem of where a Bayesian agent’s numerical degrees of belief come from is this: the agent is comparatively more confident in some claims than in others, and numerical probabilities merely provide a computationally convenient way of modeling these comparative confidence relations. Furthermore, when the comparative relation $\succcurlyeq$ is only a *partial order* it will instead be representable by a set of precise probability functions, each extending $\succcurlyeq$ to a complete order – where for each representing probability function P: (1) whenever $A \succcurlyeq B$ and *not* $B \succcurlyeq A$, $P[A] > P[B]$, and (2) whenever $A \succcurlyeq B$ and $B \succcurlyeq A$, $P[A] = P[B]$.

This comparative notion of probability cannot capture some important probabilistic concepts, such as the concept of probabilistic dependence and independence. This deficiency may be remedied by extending the comparative concept to a notion of *comparative conditional probability*. A *comparative conditional probability relation* is a weak (partial) order among pairs of sentences or propositions of form '$A|B \succcurlyeq_\alpha C|D$', read "A, given B, *is at least as probable as* C, given D".[1]

These relations may be employed to represent idealized agents' *comparative conditional confidence relations*. Interpreted this way, a relationship of form '$A|C \succcurlyeq_\alpha C|D$' says that agent α is at least as confident that A (is true), given that B, as she is that C, given that D. However, an alternative (arguably distinct) conception employs these relations to represent *comparative argument strengths*. Interpreted this way, a relationship of form '$A|B \succcurlyeq_\alpha C|D$' says that under an interpretation α of the inferential import (or inferential meanings) of statements of the language, conclusion A is supported by premise B at least as strongly as conclusion C is supported by premise D. J. M. Keynes suggested this kind of reading of conditional probabilities in his *Treatise on Probability* (1921). B. O. Koopman (1940) axiomatized this Keynesian conception in terms of a logic of *comparative conditional probability* of the kind developed below.

In this article I will present axioms for *comparative conditional probability relations* that are more general than usual. Each of these relations is a weak partial order on pairs of sentences – i.e. each relation will be transitive and reflexive, but need not be a complete order relation. The axioms presented here are *probabilistically sound* for the broad class of conditional probability functions known as *Popper functions* (which will be axiomatized in section 2).[2] That is, for each Popper function P, the corresponding *comparative conditional probability relation* $\succcurlyeq$ (defined by '$A|B \succcurlyeq C|D$' whenever $P[A \mid B] \geq P[C \mid D]$) will satisfy the axioms below for *comparative conditional probability relations*. Furthermore, these axioms are *probabilistically complete*: a representation theorem shows that for each relation $\succcurlyeq$ that satisfies these axioms, there is a corresponding Popper function P such that, for all sentences A, B, C, D: (i) whenever the relationship $A|B \succ C|D$ holds (i.e. whenever $A|B \succcurlyeq C|D$ but *not* $C|D \succcurlyeq A|B$), the corresponding probabilistic relationship $P[A \mid B] > P[C \mid D]$ holds; (ii) whenever $A|B \approx B|D$ holds (i.e. whenever $A|B \succcurlyeq C|D$ and $C|D \succcurlyeq A|B$), the corresponding probabilistic relationship $P[A \mid B] = P[C \mid D]$ holds.[3]

---

[1] Two examples: [the coin comes up *heads* in this case | the coin is fair and flipped properly in this case] *is at least as probable as* [the die lands *six* on this toss | the die is fair and tossed properly on this toss]; [it will rain here later today | the barometer fell rapidly earlier today] *is at least as probable as* [the Democrats win a senate seat in Arizona next election | there is no major change in party politics in the US before the next election].

[2] Koopman's (1940) axiomatization leaves relationships $A|B \succcurlyeq C|D$ undefined whenever B (or D) has "0-probability" – i.e. whenever , $(E\cdot\neg E)|(E\vee\neg E) \succcurlyeq B|(E\vee\neg E)$. However, Popper functions permit $P[A \mid B]$ to have well-defined values between 0 and 1 even in cases where $P[B \mid (E\vee\neg E)] = 0$. The axioms for the comparative relations provided below follow suit by permitting '$A|B \succcurlyeq C|D$' to remain defined for all pairs of sentence pairs.

[3] More generally, a comparative relation $\succcurlyeq$ may be representable by a set of distinct Popper functions that disagree on numerical values, but agree on the orderings among conditional

The axiomatic system I'll present is purely formal. So the *comparative conditional probability relations* the axioms characterize may be interpreted in terms of any of the usual probabilistic concepts. For example, one may interpret these relations in terms of some notion of *comparative conditional chance*. On this sort of interpretation a relationship of form '$A|B \succcurlyeq_\alpha C|D$' may be read to say, "for systems in state α, the chance of outcome A among those systems (or states of affairs) with attribute B *is at least as great as* the chance of outcome C among those with attribute D." On this reading the representation theorems will show that the usual numerical conditional chance functions provide a convenient way to represent a purely qualitative-comparative conception of conditional chance relations among states of affairs.

Although the abstractness of the formalism provides generality, the axioms for comparative conditional probability relations will be easier to motivate if we give the comparative relations $\succcurlyeq$ some uniform interpretive reading throughout. So, henceforth I'll read the each such comparative relation as expressing comparisons among arguments with respect to support-strength. Each relationship '$A|B \succcurlyeq C|D$' will be read to say, "conclusion A is supported by the conjunction of premises B *at least as strongly as* conclusion C is supported by the conjunction of premises D". Thus, henceforth we will be investigating *comparative conditional probability* as a logic of *comparative argument strength*, a qualitative logic which may provide a foundation for the Bayesian logic of evidential support. Readers interested in other conceptions of probability are invited to see how well those conceptions fit the axioms on offer here.

2. *Popper Functions*

Popper functions are a generalization of the usual classical notion of conditional probability. All classical conditional probability functions are (in effect) very restricted kinds of Popper functions – i.e. they satisfy the axioms for Popper functions, provided that in cases where classical conditional probabilities are left undefined, we define them as equal to 1.[4] However, among the Popper functions are conditional probability functions that make important use of conditionalization on statements that have probability 0. I'll say more about this below.

Various axiomatizations of the Popper functions are available. Karl Popper's original motivation was to develop a probabilistic logic that does not presuppose (and does not draw on) classical deductive logic, and to then show that classical deductive logical entailment arises as a special

---

probabilities. This provides an entry into theories of imprecise and indeterminate probabilities (Koopman, 1940). A detailed account is provided in the article by Fabio Cozman in this volume.

[4] A *classical probability function* on language *L* (for sentential or predicate logic) is any function p from sentences to real numbers between 0 and 1 that satisfies the following axioms: (1) if $\models A$, then $p[A] = 1$; (2) if $\models \neg(A \cdot B)$, then $p[(A \vee B)] = p[A] + p[B]$; (3) (definition) when $p[B] > 0$, $p[A \mid B] = p[(A \cdot B)] / p[B]$. When $p[B] = 0$, $p[A \mid B]$ is undefined (or, may be defined to equal 1). These axioms suffice. It follows from them that when $B \models A$, $p[B] \leq p[A]$ (given $B \models A$, $\models \neg(B \cdot \neg A)$, so $1 \geq p[(B \vee \neg A)] = p[B] + p[\neg A] \geq 0$, so $1 - p[\neg A] \geq p[B]$; and since $\models \neg(A \cdot \neg A)$ and $\models (A \vee \neg A)$, $1 = p[(A \vee \neg A)] = p[A] + p[\neg A]$; so $p[A] = 1 - p[\neg A] \geq p[B]$). From this it follows that logically equivalent sentence have the same probability.

case of a purely probabilistic notion of entailment.[5] I'll bypass this aspect of Popper's project here, and build the logic of the Popper functions atop classical deductive logic.

Popper functions turn out to have another important feature. They provide a significant way to generalize the classical notion of conditional probability. I'll say more about this in a bit. First, here is a fairly sparse way to axiomatize the Popper functions. These particular axioms are informative because, weak as they are, they provide close analogs of the axioms for *comparative conditional probability relations* introduced later. The following axioms only suppose that numerical values are real numbers – the restriction to values between 0 and 1 must be proved.

> *Sparse Axioms for Popper Functions*: Let *L* be a language having either the syntax of sentential logic, or alternatively, the syntax of predicate logic (including identity and functions).[6] Let '|=' represent the usual logical entailment relation for the logic (either sentential logic or predicate logic). Each Popper function is a function P from pairs of sentences of *L* to the real numbers such that for all sentences A, B, and C:
> 1. for some sentences E, F, G, H, $P[E \mid F] \neq P[G \mid H]$
> 2. $P[A \mid A] \geq P[B \mid C]$
> 3. if $B \models A$, then $P[A \mid C] \geq P[B \mid C]$
> 4. if $C \models B$ and $B \models C$, then $P[A \mid B] \geq P[A \mid C]$
> 5. $P[A \mid B] + P[\neg A \mid B] = P[B \mid B]$ or else $P[D \mid B] = P[B \mid B]$ for all D
> 6. $P[(A \cdot B) \mid C] = P[A \mid (B \cdot C)] \times P[B \mid C]$

This axiomatization is so weak that the usual probabilistic formulae are difficult to derive. It is useful for our purposes because of its close connection with the axioms for *comparative conditional probability relations* provided later. Here is an alternative, more usual axiomatization of the Popper functions.

> *Robust Axioms for Popper Functions*: Let *L* be a language having either the syntax of sentential logic or predicate logic (including identity and functions), where '|=' represents the usual logical entailment relation. Each Popper function is a function P from pairs of sentences of *L* to the real numbers such that for all sentences A, B, and C:
> (1) if $\models \neg A$ and $\models B$ (i.e. A is a contradiction and B is a tautology), then $P[A \mid B] = 0$
> (2) $1 \geq P[A \mid B] \geq 0$
> (3) if $B \models A$, then $P[A \mid B] = 1$
> (4) if $C \models B$ and $B \models C$, then $P[A \mid B] = P[A \mid C]$
> (5) if $C \models \neg(A \cdot B)$, then either $P[(A \vee B) \mid C] = P[A \mid C] + P[B \mid C]$ or $P[D \mid C] = 1$ for all D
> (6) $P[(A \cdot B) \mid C] = P[A \mid (B \cdot C)] \times P[B \mid C]$

---

[5] See the appendix to (Popper, 1959). Hartry Field (1977) shows how to extend Popper's project to probability functions for predicate logic. That is, Field shows how to construct a probabilistic semantics for predicate logic that takes the notion of probability assignments (rather than truth-value assignments) as basic. He proves that this semantics gives rise to a notion of logical entailment that is coextensive with the classical notion.

[6] Probabilistic logics are often restricted to a language for sentential logic, but everything here carries over to full predicate logic with identity and functions.

Clearly, the *sparse axioms*, 1-6, are derivable from the *robust axioms*, (1)-(6). The derivation of the robust axioms from the sparse axioms requires some effort (see the Appendix).

To understand the relationship between Popper functions and classical conditional probability functions, think of it like this. Given an *unconditional* classical probability function p, *conditional probability* is usually defined as follows: whenever p[B] > 0, p[A | B] = p[(A·B)]/P[B]; and when p[B] = 0, p[A | B] is left undefined. Let's make a minor modification to this usual approach, and require instead that classical conditional probability functions make p[A | B] = 1 by default whenever p[B] = 0. Thus, on this approach conditional probabilities are always defined. Specified in this way, each classical conditional probability function is a simple kind of Popper function (i.e. it satisfies the axioms for Popper functions).

More generally, a Popper function may consist of a ranked hierarchy of classical probability functions, where conditionalization on a probability 0 sentence induces a transition from one classical probability function to another classical function at a lower rank. The idea is that probability 0 need not mean "absolutely impossible". Rather, it means something like, "not a viable possibility unless (and until) the more plausible alternatives are refuted."

Here is how that works in more detail. For a given Popper function P, if we hold the condition statement B fixed, then the function P[| B] behaves precisely like a classical probability function – it always satisfies the classical axioms. However, when a statement C has 0 probability on B, P[C | B] = 0, the probability function P[| (C·B)] gotten by now holding the conjunction (C·B) fixed may remain well-defined, and may behave like an entirely different classical probability function. In general, a Popper function consists of a ranked hierarchy of classical probability functions, where the transition from a classical probability function at one level in the hierarchy (the statement B level) to a new classical probability function at a lower level (the statement (C·B) level) is induced by conditionalization on a statement (C·B) that has probability 0 at that higher (statement B) level. Finally, at the bottom level, below all other ranks associated with P, is the level of logical contradictions. This level may also include sentences that "behave like logical contradictions" – i.e. sentences E such that every sentence has probability 1 when conditionalized on E: P[A | E] = 1 for all A.

The fact that a Popper function may consist of this kind of ranked hierarchy of classical functions is not an additional assumption or stipulation. Rather, it follows from the above axioms (from 1-6, and also from (1)-(6)) without supplementation. (See Hawthorne, 2013, for a detailed account of the ranked structure of Popper functions, including proofs of these claims.)

Here is an illustration of a case where this kind of generalization of classical probability proves useful. Suppose that the probability that a randomly selected point will lie within the upper 3/4 of a specific spatial region, described by 'A', given that it lies somewhere within that whole three-dimensional region, described by 'B', is P[A | B] = 3/4. The probability that this same randomly selected point will lie precisely on a specific plane described by 'C' where it intersects the B-region should presumably be 0, so we have 0 = P[C | B] = P[(A·C) | B]. However, given that this random point does indeed lie within the C-plane within the B-region, the probability that it lies within the part of that region described by 'A' (which, say, contains half of the plane described by 'C·B') may again be perfectly well-defined: P[A | (C·B)] = 1/2. Furthermore, the probability

that this random point will lie on the part of a line segment described by 'D' within the C-plane should also presumably be 0, so we again have a situation where $0 = P[D \mid (C{\cdot}B)] = P[(A{\cdot}D) \mid (C{\cdot}B)]$. However, given that this random point does indeed lie within the part of the D-segment within the part of the C-plane within the B-region, the probability that it lies in the A-region (which, say, contains two-thirds of the D-line-segment within $(D{\cdot}(C{\cdot}B))$) may again be well-defined: $P[A \mid (D{\cdot}(C{\cdot}B))] = 2/3$. So, the general idea is that a specific Popper function may consist of a ranked hierarchy of classical probability functions, where conditionalizations on specific probability 0 statements at one level of the hierarchy can induce a transition to another perfectly good classical probability function defined at a lower level of the hierarchy.[7]

Bayesian confirmation theory employs conditional probability functions to represent the support of evidence for hypotheses, and Popper functions may serve in this role. However, the Bayesian approach to confirmation owes us an account of what the proposed numerical *degrees of support* come from, and what the probabilistic numbers mean or represent. Subjectivist Bayesians attempt to provide this account in terms of betting functions and Dutch book theorems – they maintain that confirmation functions are belief-strength functions, and that their numerical values represent ideally rational betting quotients, which must satisfy the usual probabilistic rules in order to avoid the endorsement of betting packages that would result in sure losses. However, for a logical account of confirmation functions (e.g. of the kind endorsed by Keynes), wherein confirmation functions represent argument strengths, another kind of answer to the "where do the numbers come from, and what do they mean?" question may be offered – an answer via a *representation theorem*. On this approach the idea is that confirmation theory derives from a deeper underlying *qualitative logic of comparative argument strength*. We now proceed to specify the rules that govern this deeper logic. We'll see that Popper functions merely provide a convenient way to calculate the comparative support relationships captured by this qualitative logic of comparative argument strength; the probabilities add nothing that the qualitative logic cannot already capture on its own.

3. Towards the Logic of *Comparative Argument strength*: the *Proto-Support Relations*

A *comparative support relation* $\succcurlyeq$ is a relation among pairs of sentence pairs. It should satisfy axioms that provide plausible restrictions on a reasonable conception of the notion of *comparative argument strength*. We will get to the axioms in a moment.

Associated with each relation $\succcurlyeq$ are several related relations, defined in terms of it. Here is a list of these, their formal definitions, and an appropriate informal reading for each.

A comparative support relation $\succcurlyeq$ is a relation of form $A|B \succcurlyeq C|D$,
read "A *is supported by* B *at least as strongly as* C *is supported by* D".
Define four associated relations as follows:
(1) $A|B \succ C|D$ abbreviates "$A|B \succcurlyeq C|D$ and *not* $C|D \succcurlyeq A|B$",
read "A *is supported by* B *more strongly than* C *is supported by* D";

[7] For a comparison of the Popper functions to other accounts of conditional probability functions, see the article by Kenny Easwaran in this volume.

(2) $A|B \approx C|D$ abbreviates "$A|B \succcurlyeq C|D$ and $C|D \succcurlyeq A|B$",
read "A *is supported by* B *to the same extent that* C *is supported by* D";
(3) $A|B \asymp C|D$ abbreviates "*not* $A|B \succcurlyeq C|D$ and *not* $C|D \succcurlyeq A|B$",
read "*the support for* A *by* B *is indistinctly comparable to that of* C *by* D";
(4) $B \Rightarrow A$ abbreviates "$A|B \succcurlyeq C|C$";
read "B *supportively entails* A".

The axioms I'll provide for $\succcurlyeq$ turn out to entail that for each such relation, the corresponding *supportive entailment relations* $\Rightarrow$ satisfies the rules for a well-known kind of non-monotonic conditional called a *rational consequence relation*. Indeed, the *rational consequence relations* turn out to be identical to the *supportive entailment relations*.[8]

With these definitions in place we are ready to specify axioms for the *comparative support relations*. Axioms 1-6 closely parallel the corresponding axioms for Popper functions. The axioms will only ensure that these comparative relations are partial orders on comparative argument strength: they are transitive and reflexive, but need not be complete orders – i.e. some argument pairs may fail to be distinctly comparable in strength.

Let *L* be a language having the syntax of sentential logic or predicate logic (including identity and functions). Each *proto-support relation* $\succcurlyeq$ is a binary relation between pairs of sentences that satisfies the following axioms:

0. If $A|B \succcurlyeq C|D$ and $C|D \succcurlyeq E|F$, then $A|B \succcurlyeq E|F$ (transitivity)

1. for some E, F, G, H, $E|F \succ G|H$ (non-triviality)
[Not all arguments are equally strong; at least one is stronger than at least one other.]

2. $A|A \succcurlyeq B|C$ (maximality)
[Self-support is maximal support -- at least as strong as any argument.]

3. If $B \models A$, then $A|C \succcurlyeq B|C$ (classical consequent entailment)
[Whenever B logically entails A, *the support for* A *by* C *is at least as strong as the support for* B *by* C. The reflexivity of $\succcurlyeq$ follows, since $A \models A$. Together with other axioms it yields: (i) If $(B \cdot C) \models A$, then $A|C \succcurlyeq B|C$; (ii) If $B \models A$, then $A|B \succcurlyeq C|D$.]

4. If $B \models C$ and $C \models B$, then $A|B \succcurlyeq A|C$ (classical antecedent equivalence)
[Logically equivalent statements support all statements equally well.]

5. If $A|B \succcurlyeq C|D$, then $\neg C|D \succcurlyeq \neg A|B$ or else $B \Rightarrow D$ for all D (negation-symmetry)

---

[8] The ranked structure of the *Popper functions* (structured as a hierarchy of classical probability functions) is just the ranked structure of the rational consequence relations. The comparative support relations turn out to share this ranked structure, captured by their associated supportive entailment relations. See (Hawthorne, 2013) for a detailed account of the rational consequence relations and their ranked structures.

[Whenever A *is supported by* B *at least as strongly as* C *is supported by* D, *the falsity of* C *is supported by* D *at least as strongly as the falsity of* A *is supported by* B; the only exception is in cases where premise B behaves like a contradiction, maximally supporting every statement D. This captures the essence of the additivity axiom for conditional probability, axiom 5 for the Popper functions.]

6.1 If $H_1|(A_1{\cdot}E_1) \succcurlyeq H_2|(A_2{\cdot}E_2)$ and $A_1|E_1 \succcurlyeq A_2|E_2$, then $(H_1{\cdot}A_1)|E_1 \succcurlyeq (H_2{\cdot}A_2)|E_2$

6.2 If $H_1|(A_1{\cdot}E_1) \succcurlyeq A_2|E_2$ and $A_1|E_1 \succcurlyeq H_2|(A_2{\cdot}E_2)$, then $(H_1{\cdot}A_1)|E_1 \succcurlyeq (H_2{\cdot}A_2)|E_2$
(6.1-6.2: composition)

6.3 If $(H_1{\cdot}A_1)|E_1 \succcurlyeq (H_2{\cdot}A_2)|E_2$ and $A_2|E_2 \succcurlyeq A_1|E_1$, then $H_1|(A_1{\cdot}E_1) \succcurlyeq H_2|(A_2{\cdot}E_2)$ or $E_2 \Rightarrow \neg A_2$

6.4 If $(H_1{\cdot}A_1)|E_1 \succcurlyeq (H_2{\cdot}A_2)|E_2$ and $A_2|E_2 \succcurlyeq H_1|(A_1{\cdot}E_1)$, then $A_1|E_1 \succcurlyeq H_2|(A_2{\cdot}E_2)$ or $E_2 \Rightarrow \neg A_2$

6.5 If $(H_1{\cdot}A_1)|E_1 \succcurlyeq (H_2{\cdot}A_2)|E_2$ and $H_2|(A_2{\cdot}E_2) \succcurlyeq H_1|(A_1{\cdot}E_1)$, then $A_1|E_1 \succcurlyeq A_2|E_2$ or $(A_2{\cdot}E_2) \Rightarrow \neg H_2$

6.6 If $(H_1{\cdot}A_1)|E_1 \succcurlyeq (H_2{\cdot}A_2)|E_2$ and $H_2|(A_2{\cdot}E_2) \succcurlyeq A_1|E_1$, then $H_1|(A_1{\cdot}E_1) \succcurlyeq A_2|E_2$ or $(A_2{\cdot}E_2) \Rightarrow \neg H_2$ (6.3-6.6: decomposition)

[Axioms 6.1-6.6 capture the comparative content of probabilistic conditionalization, axiom 6 for the Popper functions. Think of H_1 and H_2 as hypotheses, A_1 and A_2 as auxiliary hypotheses, and E_1 and E_2 as evidence statements. For k = 1,2, $P[H_k{\cdot}A_k \mid E_k] = P[H_k \mid A_k{\cdot}E_1] \times P[A_k \mid E_k]$. So when $P[H_1 \mid A_1{\cdot}E_1] \geq P[H_2 \mid A_2{\cdot}E]$ and $P[A_1 \mid E_1] \geq P[A_2 \mid E_2]$, we must have $P[H_1{\cdot}A_1 \mid E_1] \geq P[H_2{\cdot}A_2 \mid E_2]$, expressed by 6.1. When $P[H_1 \mid A_1{\cdot}E_1] \geq P[A_2 \mid A_2]$ and $P[A_1 \mid E_1] \geq P[H_2 \mid A_2{\cdot}E_2]$, we must have $P[H_1{\cdot}A_1 \mid E_1] \geq P[H_2{\cdot}A_2 \mid E_2]$, expressed by 6.2. When $P[H_1{\cdot}A_1 \mid E_1] \geq P[H_2{\cdot}A_2 \mid E_2]$ and $P[A_2 \mid E_2] \geq P[A_1 \mid E_1]$, we must have $P[H_1 \mid A_1{\cdot}E_1] \geq P[H_2 \mid A_2{\cdot}E]$ *unless* $P[A_2 \mid E_2] = 0$ (in which case $P[A_1 \mid E_1] = 0$ as well, and so $0 = P[H_1{\cdot}A_1 \mid E_1] = P[H_2{\cdot}A_2 \mid E_2]$, and the values of $P[H_1 \mid A_1{\cdot}E_1]$ and $P[H_2 \mid A_2{\cdot}E]$ are not determined by the values of $P[A_1 \mid E_1]$, $P[A_2 \mid E_2]$, $P[H_1{\cdot}A_1 \mid E_1]$, and $P[H_2{\cdot}A_2 \mid E_2]$), expressed by 6.3. axioms 6.4-6.6 are related Popper functions in a similar way.][9]

---

[9] From 6.3 and 6.1 we derive the following rule:
6.3* If $H_1|(A_1{\cdot}E_1) \succ H_2|(A_2{\cdot}E_2)$ and $A_1|E_1 \succcurlyeq A_2|E_2$, then $(H_1{\cdot}A_1)|E_1 \succ (H_2{\cdot}A_2)|E_2$ or $E_1 \Rightarrow \neg A_1$.
From 6.4 and 6.2 we derive:
6.4* If $H_1|(A_1{\cdot}E_1) \succ A_2|E_2$ and $A_1|E_1 \succcurlyeq H_2|(A_2{\cdot}E_2)$, then $(H_1{\cdot}A_1)|E_1 \succ (H_2{\cdot}A_2)|E_2$ or $E_1 \Rightarrow \neg A_1$.
From 6.5 and 6.1 we derive:
6.5* If $H_1|(A_1{\cdot}E_1) \succcurlyeq H_2|(A_2{\cdot}E_2)$ and $A_1|E_1 \succ A_2|E_2$, then $(H_1{\cdot}A_1)|E_1 \succ (H_2{\cdot}A_2)|E_2$ or $(A_1{\cdot}E_1) \Rightarrow \neg H_1$.
From 6.6 and 6.2 we derive:
6.6* If $H_1|(A_1{\cdot}E_1) \succcurlyeq A_2|E_2$ and $A_1|E_1 \succ H_2|(A_2{\cdot}E_2)$, then $(H_1{\cdot}A_1)|E_1 \succ (H_2{\cdot}A_2)|E_2$ or $(A_1{\cdot}E_1) \Rightarrow \neg H_1$.

7. If $A|(B\cdot C) \succcurlyeq E|F$ and $A|(B\cdot\neg C) \succcurlyeq E|F$, then $A|B \succcurlyeq E|F$ (alternate presumption)*
[Probabilistically this axiom follows from additivity together with conditionalization – i.e. since $P[A \mid B] = P[A \mid B\cdot C] \times P[C \mid B] + P[A \mid B\cdot\neg C] \times (1 - P[B \mid C])$, if both $P[A \mid B\cdot C] \geq r$ and $P[A \mid B\cdot\neg C] \geq r$, then $P[A \mid B] \geq r$. Axiom 7 is a qualitative version of this result. It can be proved from the other axioms if the relation $\succcurlyeq$ is assumed to be a *complete order* rather than merely a partial order relation – i.e. if $\succcurlyeq$ takes all argument pairs to be distinctly comparable in strength.]

All relations that satisfy these axioms are weak partial orders – i.e. they are transitive and reflexive. Transitivity is guaranteed by axiom 0; reflexivity, $A|C \succcurlyeq A|C$, follows from axiom 3. I call the relations that satisfy these axioms *proto-support relations* because the axioms still need a bit of strengthening to rule out some relations $\succcurlyeq$ that fail to behave properly. I'll say more about that later.

The asterisk on the name of axiom 7 is to indicate that it follows from the other axioms whenever the relation $\succcurlyeq$ is also *complete* – i.e. whenever, for all pairs of sentence pairs, the following *complete comparability rule* also holds (or is added as an additional axiom) for relation $\succcurlyeq$:

either $A|B \succcurlyeq C|D$ or $C|D \succcurlyeq A|B$ (complete comparability).

Adding the rule for *complete comparability* would require that all argument pairs are distinctly comparable in strength: for all A, B, C, D, $A|B \not\asymp C|D$. Any relation $\succcurlyeq$ that satisfies the above axioms together with *complete comparability* is a weak order relation rather than merely a weak partial order.[10]

The rules 6.3^*-6.6^* are a bit weaker than their 6.3-6.6 counterparts. The following additional rules are also derivable from 6.1-6.6. Furthermore, 6.3-6.6 are derivable from 6.3^*-6.6^* together with the following rules (with the aid of 6.1 and 2). So we could replace 6.3-6.6 with 6.3^*-6.6^* together with the following rules:

6.3^{**} If $H_1|(A_1\cdot E_1) \asymp H_2|(A_2\cdot E_2)$ and $A_1|E_1 \succcurlyeq A_2|E_2$, then $(H_1\cdot A_1)|E_1 \succ (H_2\cdot A_2)|E_2$ or $(H_1\cdot A_1)|E_1 \asymp (H_2\cdot A_2)|E_2$ or $E_1 \Rightarrow \neg A_1$.

6.4^{**} If $H_1|(A_1\cdot E_1) \asymp A_2|E_2$ and $A_1|E_1 \succcurlyeq H_2|(A_2\cdot E_2)$, then $(H_1\cdot A_1)|E_1 \succ (H_2\cdot A_2)|E_2$ or $(H_1\cdot A_1)|E_1 \asymp (H_2\cdot A_2)|E_2$ or $E_1 \Rightarrow \neg A_1$.

6.5^{**} If $H_1|(A_1\cdot E_1) \succcurlyeq H_2|(A_2\cdot E_2)$ and $A_1|E_1 \asymp A_2|E_2$, then $(H_1\cdot A_1)|E_1 \succ (H_2\cdot A_2)|E_2$ or $(H_1\cdot A_1)|E_1 \asymp (H_2\cdot A_2)|E_2$ or $(A_1\cdot E_1) \Rightarrow \neg H_1$.

6.6^{**} If $H_1|(A_1\cdot E_1) \succcurlyeq A_2|E_2$ and $A_1|E_1 \asymp H_2|(A_2\cdot E_2)$, then $(H_1\cdot A_1)|E_1 \succ (H_2\cdot A_2)|E_2$ or $(H_1\cdot A_1)|E_1 \asymp (H_2\cdot A_2)|E_2$ or $(A_1\cdot E_1) \Rightarrow \neg H_1$.

[10] Koopman also provides the following rule as an axiom:

For any integer $n > 1$, if $A_1, ..., A_n$, and $B_1, ..., B_n$ are collections of sentences such that

$C \not\Rightarrow \neg C$, $C \Rightarrow (A_1\vee...\vee A_n)$, $C \Rightarrow \neg(A_i\cdot A_j)$, $A_n|C \succcurlyeq ... \succcurlyeq A_2|C \succcurlyeq A_1|C$, and

$D \not\Rightarrow \neg D$, $D \Rightarrow (B_1\vee...\vee B_n)$, $D \Rightarrow \neg(B_i\cdot B_j)$, $B_n|D \succcurlyeq ... \succcurlyeq B_2|D \succcurlyeq B_1|D$,

then $A_n|C \succcurlyeq B_1|D$ (subdivision)*

This rule may not seem as intuitively compelling as the others, so I forego it here. Later we will see that comparative support relations should be *extendable* to *completely comparable* relations. *Subdivision* is derivable in the presence of *complete comparability*.

The above axioms for *proto-support relations* are probabilistically sound in the following sense: For each Popper function P, define the *corresponding comparative relation* to be the relation $\succcurlyeq$ such that, for all sentences A, B, C, D, $A|B \succcurlyeq C|D$ holds if and only if $P[A \mid B] \geq P[C \mid D]$. Then for each Popper function, the *corresponding comparative relation* can be shown to be a *proto-support relation* – it satisfies the above axioms.

The axioms for the *proto-support relations* are not *probabilistically complete*. Some *proto-support relations* are not probability-like enough to be representable by a Popper function. Below we add additional constraints (additional axioms) that suffice to characterize the "full" *comparative support relations*. These relations will turn out to be *probabilistically complete* in the sense that each such *comparative support relation* $\succcurlyeq$ is representable by a Popper function P.

The *proto-support relations* are sufficiently strong to provide as theorems some comparative forms of Bayes' theorem. Here is one example.

Bayes' Theorem 1: Suppose $B \not\Rightarrow \neg H_1$.
If $E|(H_1 \cdot B) \succ E|(H_2 \cdot B)$ and $H_1|B \succcurlyeq H_2|B$, then $H_1|(E \cdot B) \succ H_2|(E \cdot B)$.

Think of H_1 and H_2 as hypotheses, B as common background knowledge and auxiliary hypotheses, and E as the evidence. This is an analogue of the following version of Bayes' theorem:

Suppose $P[H_1 \mid B] > 0$. Then
$P[H_2 \mid E \cdot B] / P[H_1 \mid E \cdot B] = (P[E \mid H_2 \cdot B] / P[E \mid H_1 \cdot B]) \times (P[H_2 \mid B] / P[H_1 \mid B])$,
so, if $P[E \mid H_1 \cdot B] > P[E \mid H_2 \cdot B]$ and $P[H_1 \mid B] \geq P[H_2 \mid B]$, then $P[H_1 \mid (E \cdot B)] > P[H_2 \mid (E \cdot B)]$.

Here is a second form of Bayes' theorem satisfied by *proto-support relations*.

Bayes' Theorem 2: Suppose $B \not\Rightarrow \neg H_1$, $B \not\Rightarrow \neg H_2$, and $B \Rightarrow \neg(H_1 \cdot H_2)$.
If $E|(H_1 \cdot B) \succ E|(H_2 \cdot B)$, then
$H_1|(E \cdot B \cdot (H_1 \vee H_2)) \succ H_1|(B \cdot (H_1 \vee H_2))$ and $H_2|(B \cdot (H_1 \vee H_2)) \succ H_2|(E \cdot B \cdot (H_1 \vee H_2))$.

A straightforward probabilistic analogue goes like this:

Suppose $P[H_1 \mid B] > 0$, $P[H_2 \mid B] > 0$, and $P[H_1 \cdot H_2 \mid B] = 0$. If $P[E \mid H_1 \cdot B] > P[E \mid H_2 \cdot B]$, then the following posterior probabilities when comparing H_1 directly to H_2: $P[H_1 \mid E \cdot B \cdot (H_1 \vee H_2)] > P[H_1 \mid B \cdot (H_1 \vee H_2)]$ and $P[H_2 \mid E \cdot B \cdot (H_1 \vee H_2)] < P[H_2 \mid B \cdot (H_1 \vee H_2)]$.

This is a comparative expression of the relationship, $P[H_2 \mid E \cdot B] / P[H_1 \mid E \cdot B] < P[H_2 \mid B] / P[H_1 \mid B]$, since $P[H_2 \mid E \cdot B \cdot (H_1 \vee H_2)] / P[H_1 \mid E \cdot B \cdot (H_1 \vee H_2)] = P[H_2 \mid E \cdot B] / P[H_1 \mid E \cdot B]$ and $P[H_2 \mid B \cdot (H_1 \vee H_2)] / P[H_1 \mid B \cdot (H_1 \vee H_2)] = P[H_2 \mid B] / P[H_1 \mid B]$.

4. The *Comparative Support Relations* and their Probabilistic Representations

Consider the following additional rules.

8. $A|B \succcurlyeq C|D$ or $C|D \succcurlyeq A|B$ (complete comparability)

9. For each integer $m \geq 2$ there is an integer $n \geq m$ such that for n sentences $S_1, ..., S_n$ and some sentence G:
   (i) $G \not\Rightarrow \neg S_1$,
   and for all distinct i, j,
   (ii) $G \Rightarrow \neg(S_i \cdot S_j)$ and
   (iii) $S_i|G \approx S_j|G$. (existence of arbitrarily large equal-partitions)

An *equal-partition* (given G) is a collection of pair-wise mutually exclusive sentences (the sentences S_k) that are "equally likely" (given G). When sentences G and collection of n sentences, $S_1, ..., S_n$, satisfy rule 9 for relation $\succcurlyeq$, it can be shown that: (i) $(G\cdot(S_1\vee...\vee S_n)) \not\Rightarrow \neg S_1$, for distinct $i, j \leq n$, (ii) $(G\cdot(S_1\vee...\vee S_n)) \Rightarrow \neg(S_i\cdot S_j)$, (iii) $S_i|(G\cdot(S_1\vee...\vee S_n)) \approx S_j|(G\cdot(S_1\vee...\vee S_n))$, and (iv) $(G\cdot(S_1\vee...\vee S_n)) \Rightarrow (S_1\vee...\vee S_n)$. Thus, rule 9 guarantees that for arbitrarily large n, $\succcurlyeq$ has an exclusive and exhaustive equal-partition based on $(G\cdot(S_1\vee...\vee S_n))$. These partitions can be used to provide approximate probability values for the strengths of arguments:

when $(S_1\vee...\vee S_{k+1})|(G\cdot(S_1\vee...\vee S_n)) \succcurlyeq A|B \succcurlyeq (S_1\vee...\vee S_k)|(G\cdot(S_1\vee...\vee S_n))$ we effectively get a probability-like approximation for the strength of A|B, $(k+1)/n \geq P[A \mid B] \geq k/n$.

For arbitrarily large partitions (for arbitrarily large values of n) these partitions provide arbitrarily close probability-like bounds on the strength of each argument.

Rule 9 is not as strong as needed in most cases. Here is a stronger alternative:

9^+. If $A|B \succ C|D$, then for some $n \geq 2$ there are sentences $S_1, ..., S_n$ and a sentence F such that: $F \not\Rightarrow \neg S_1$, and for distinct i, j, $F \Rightarrow \neg(S_i\cdot S_j)$ and $S_i|F \approx S_j|F$, and $F \Rightarrow (S_1\vee...\vee S_n)$, and for some m of them, $A|B \succ (S_1\vee\ldots\vee S_m)|F \succ C|D$. (Archimedean equal-partitions)

Rule 9^+ implies 9, but adds to it a kind of "Archimedean condition": whenever $A|B \succ C|D$, there must be an equal-partition that, for sufficiently large n, squeezes a "strength comparison" between A|B and C|D. This forces these two arguments to exhibit distinct probabilistic values:

$$P[A \mid B] > m/n > P[C \mid D].$$

A comparative support relation that satisfies rules 8 and 9 but that fails to satisfy 9^+ must permit some argument pairs for which $A|B \succ C|D$, but where A|B and C|D infinitesimally close together in comparative strength – i.e. no segment of any equal-partition argument can fit between them. There are interesting cases where such non-Archimedean support relations are useful. So I'll treat the full range of relations that satisfy rule 9, as well as the better-behaved subclass of Archimedean relations, which satisfy the more restrictive rule 9^+.

Here is an intuitive example of the kind of partition required by rules 9 and 9^+. Let statement F (a.k.a. statement $G\cdot(S_1\vee...\vee S_n)$) describe a fair lottery consisting of exactly n tickets. Each of the sentences S_i says "ticket i will win". F *says* via *supportive entailment* (or via *logical entailment*, which is stronger):

(1) "at least one ticket will win": so $F \Rightarrow (S_1\vee\ldots\vee S_n)$;

(2) "no two tickets will win": so $F \Rightarrow \neg(S_i \cdot S_j)$, for each distinct pair of claims S_i and S_j;
(3) "each ticket has the same chance of winning" (and where the argument from F provides exactly the same support for the claim "ticket i will win" as for the claim "ticket j will win"): so $S_i|F \approx S_j|F$ for each distinct pair of claims S_i and S_j.
(4) furthermore, we suppose that F does not *supportively entail* "ticket 1 won't win": that is, $F \not\Rightarrow \neg S_1$ (formally this clause is equivalent to $F|F \succ \neg S_1|F$; it eliminates the possibility that F behaves like a contradiction, supportively entailing every statement).

We could require all comparative support relations to satisfy rule 9. This would not be too implausible – it would merely require that the language of each relation $\succcurlyeq$ have the ability to describe such lotteries for arbitrarily large finite numbers of tickets. Presumable our own natural language can do that. So this would be a fairly innocuous requirement.[11] Nevertheless, we won't require that comparative support relations employ languages this rich. Rather, it will suffice for our purposes to suppose that each support relation is (in principle) *extendable* to a relation that includes such lottery descriptions. I'll say more about *extendability* in a bit. Before doing so, let's consider rule 8 more closely.

In many cases a pair of arguments may fail to be distinctly comparable in strength; neither is distinctly stronger than the other, nor are they determinately equal in strength. Nevertheless, I will argue that each legitimate *comparative support relation* should be *syntactically extendable* to a *complete* relation, at least in principle. I'll provide that argument in a moment. Let's first define the relevant notion of *extendability*.

Definition: A *proto-support relation* $\succcurlyeq_\alpha$ is *extendable* to a *proto-support relation* $\succcurlyeq_\beta$ just when the language of $\succcurlyeq_\beta$ contains the language of $\succcurlyeq_\alpha$ (i.e. contains the same syntactic expressions, and perhaps additional expressions as well) and the following two conditions hold:
(1) whenever $A|B \succ_\alpha C|D$, then also $A|B \succ_\beta C|D$;
(2) whenever $A|B \approx_\alpha C|D$, then also $A|B \approx_\beta C|D$.

When $\succcurlyeq_\alpha$ is *extendable* to $\succcurlyeq_\beta$, all argument pairs that are *distinctly comparable* according to $\succcurlyeq_\alpha$ must *compare in the same way* according to $\succcurlyeq_\beta$. Each relation $\succcurlyeq_\alpha$ counts as an extension of itself. An extension $\succcurlyeq_\beta$ of $\succcurlyeq_\alpha$ may employ precisely the same language as $\succcurlyeq_\alpha$, and may merely extend $\succcurlyeq_\alpha$ by *distinctly comparing* some arguments that were *not distinctly comparable* according to $\succcurlyeq_\alpha$. More generally, $\succcurlyeq_\beta$ may include comparisons that involve new expressions, not already part of the syntax of the language for $\succcurlyeq_\alpha$. Furthermore, the relationship between $\succcurlyeq_\alpha$ and an *extension* of it, $\succcurlyeq_\beta$, need only be syntactic. There is no presumption that an *extension* of a relation must maintain the same meanings (the same semantic content) for sentences it shares with the relation it extends (although it certainly may do so).

The *proto-support relations* commonly permit a wide range of argument pairs to remain *incomparable in strength*. But only those relations among them that can be *extended* to *complete relations* (i.e. which satisfy rule 8) will be counted among the full-fledged *comparative support*

[11] Rule 9 does not presume that any such lotteries exist. It only supposes that we can construct arguments describing them and their implications for prospective outcomes.

relation. To see that *extendability* to a relation that satisfies the *completeness rule* is a plausible constraint, let's consider what a *proto-support relation* must be like if it cannot possibly be extended to a *complete relation*.

Extendability is a purely syntactic requirement. That is, an *extension* $\succcurlyeq_\beta$ of a relation $\succcurlyeq_\alpha$ need not take on any of the meanings that one might have associated with the sentences of $\succcurlyeq_\alpha$. Rather, an *extension* $\succcurlyeq_\beta$ of $\succcurlyeq_\alpha$ is only required to agree with the *definite comparisons* – those of form $A|B \succ_\alpha C|D$ and $E|F \approx_\alpha G|H$ already specified by $\succcurlyeq_\alpha$ (while $\succcurlyeq_\beta$ continues to satisfy axioms 0-7). Thus, a proto-support relation $\succcurlyeq_\alpha$ may only fail to be extendable to a complete relation when no complete extension of $\succcurlyeq_\alpha$ is consistent with the (purely syntactic) restrictions on orderings embodied by axioms 0-7. That is, for $\succcurlyeq_\alpha$ to be *unextendable* to a complete relation, the definite comparisons (of form $A|B \succ_\alpha C|D$ and $E|F \approx_\alpha G|H$) already specified by $\succcurlyeq_\alpha$ must, in conjunction with axioms 0-7, *require* that some argument pairs inevitably remain incomparable in strength *merely to avoid an explicit syntactic contradiction*. In other words, any proto-support relation for which there cannot possibly be a *complete extension* must already contain a kind of looming syntactic inconsistency, due to the forms of its *definite* argument strength comparisons. It only manages to stave off explicit formal inconsistency by forcing at least some argument pairs to remain incomparable in strength.

It makes good sense to declare specific argument pairs *incomparable in strength* when, given their meanings (their semantic content), there seems to be no appropriate basis on which to compare them. But let's disregard any comparative relation that *requires* argument forms to remain *incomparable* in order to avoid syntactic inconsistency. Thus, we disregard any relation that cannot possibly be extended to a *complete relation*, not even by radically changing the meanings of the sentences involved.

We now define the class of *comparative support relations* as those *proto-support relations* that can be extended to relations that satisfy rules 8 and 9. Those that satisfy rules 8 and 9^+ are a special subclass of *comparative support relations*, the *arch* (for "Archimedean") *comparative support relations*.

Definition: Classes of *Comparative Support Relations*:

1. A ***completely-extended*** *proto-support relation* is any *proto-support relation* that satisfies rules 8 and 9.
2. A ***completely-extendable*** *proto-support relation* is any *proto-support relation* that is *extendable* to a ***completely-extended*** *proto-support relation* (i.e. any *proto-support relation* is extendable to a *proto-support relation* that satisfies rules 8 and 9).

Define the *comparative support relations* to be the ***completely-extendable*** *proto-support relations*. They include the ***completely-extended*** relations as special cases.

3. An ***completely-arch-extended*** *proto-support relation* is any *proto-support relation* that satisfies rules 8 and 9^+.
4. A ***completely-arch-extendable*** *proto-support relation* is any *proto-support relation* that is *extendable* to a ***completely-extended*** *proto-support relation* (i.e. any *proto-support relation* is extendable to a *proto-support relation* that satisfies rules 8 and 9^+).

Define the *arch comparative support relations* to be the ***completely-arch-extendable*** *proto-support relations*. They include the ***completely-arch-extended*** relations as special cases.

When a *proto-support relation* $\succcurlyeq_\alpha$ is extendable to one that satisfies rule 9 (together with 8) but *not* to one that satisfies 9^+, for some argument pairs, although $A|B \succ_\alpha C|D$, the relation *requires* that A|B be *only infinitesimally stronger* than C|D. That is, the relationships among arguments already specified by $\succcurlyeq_\alpha$ must imply that no extension of it, $\succcurlyeq_\beta$, can permit an equal-partition argument, $(S_1 \vee \ldots \vee S_m)|F$ to fit between A|B and C|D (for any n-sized equal-partition, where $F \Rightarrow_\beta \neg(S_i \cdot S_j)$, $S_i|F \approx_\beta S_j|F$, $F \Rightarrow_\beta (S_1 \vee \ldots \vee S_n)$). Such *non-Archimedean support relations* turn out to have interesting features, so we won't entirely by-pass them here.

Each *completely-extendable relation* is representable by a Popper function – perhaps by more than one. A typically *completely-extendable* relation may be extend to a variety of distinct *completely-extended* relations. Each *completely-extended relation* is represented by a unique Popper function. The nature of this probabilistic representation is perfectly tight for the *arch* relations, and a bit looser for the other relations. Here are the precise details.

Representation Theorem for *comparative support relations*:

1. For each *completely-arch-extended comparative support relation* $\succcurlyeq$, there is a unique Popper function P such that for all sentence H_1, E_1, H_2, E_2 in the language of $\succcurlyeq$,
   $P[H_1 \mid E_1] \geq P[H_2 \mid E_2]$ if and only if $H_1|E_1 \succcurlyeq H_2|E_2$.
   Note 1: given that $\succcurlyeq$ is *completely-arch-extended*, this condition is equivalent to the conjunction of the following:
   (1) if $P[H_1 \mid E_1] > P[H_2 \mid E_2]$, then $H_1|E_1 \succ H_2|E_2$;
   (2) if $P[H_1 \mid E_1] = P[H_2 \mid E_2]$, then $H_1|E_1 \approx H_2|E_2$.
   Note 2: given that $\succcurlyeq$ is *completely-arch-extended*, this condition is also equivalent to the conjunction of the following:
   (1) if $H_1|E_1 \succ H_2|E_2$, then $P[H_1 \mid E_1] > P[H_2 \mid E_2]$;
   (2) if $H_1|E_1 \approx H_2|E_2$, then $P[H_1 \mid E_1] = P[H_2 \mid E_2]$.
2. For each *arch comparative support relation* $\succcurlyeq$ (which, by definition, must be a *completely-arch-extendable proto-support relation*), there is a (not necessarily unique) Popper function P such that for all sentence H_1, E_1, H_2, E_2 in the language of $\succcurlyeq$,
   if *not* $H_1|E_1 \asymp H_2|E_2$, then $P[H_1 \mid E_1] \geq P[H_2 \mid E_2]$ if and only if $H_1|E_1 \succcurlyeq H_2|E_2$.
   Note 1: this condition is equivalent to the conjunction of the following:
   (1) if $P[H_1 \mid E_1] > P[H_2 \mid E_2]$, then $H_1|E_1 \succ H_2|E_2$ or $H_1|E_1 \asymp H_2|E_2$;
   (2) if $P[H_1 \mid E_1] = P[H_2 \mid E_2]$, then $H_1|E_1 \approx H_2|E_2$ or $H_1|E_1 \asymp H_2|E_2$.
   Note 2: this condition is also equivalent to the conjunction of the:
   (1) if $H_1|E_1 \succ H_2|E_2$, then $P[H_1 \mid E_1] > P[H_2 \mid E_2]$;
   (2) if $H_1|E_1 \approx H_2|E_2$, then $P[H_1 \mid E_1] = P[H_2 \mid E_2]$.
3. For each *completely-extended comparative support relation* $\succcurlyeq$, there is a unique Popper function P such that for all sentence H_1, E_1, H_2, E_2 in the language of $\succcurlyeq$,
   if $P[H_1 \mid E_1] > P[H_2 \mid E_2]$, then $H_1|E_1 \succ H_2|E_2$.
   Note: given the *completeness* of $\succcurlyeq$, this condition is equivalent to the conjunction of the following conditions:
   (1) if $H_1|E_1 \succ H_2|E_2$, then $P[H_1 \mid E_1] \geq P[H_2 \mid E_2]$;
   (2) if $H_1|E_1 \approx H_2|E_2$, then $P[H_1 \mid E_1] = P[H_2 \mid E_2]$.

4. For each *comparative support relation* $\succcurlyeq$ (which, by definition, must be a *completely-extendable proto-support relation*), there is a (not necessarily unique) Popper function P such that for all sentence H_1, E_1, H_2, E_2 in the language of $\succcurlyeq$,
   if $P[H_1 \mid E_1] > P[H_2 \mid E_2]$, then $H_1|E_1 \succ H_2|E_2$ or $H_1|E_1 \asymp H_2|E_2$.
   Note: this condition is equivalent to the following pair of conditions:
   (1) if $H_1|E_1 \succ H_2|E_2$, then $P[H_1 \mid E_1] \geq P[H_2 \mid E_2]$;
   (2) if $H_1|E_1 \approx H_2|E_2$, then $P[H_1 \mid E_1] = P[H_2 \mid E_2]$.

The representation theorem shows that each *completely-arch-extended* relation is virtually identical to its uniquely representing Popper function. Whenever the representing *Popper function* P for ***complete-arch*** *relation* $\succcurlyeq$ assigns $P[A \mid B] = r$ for a *rational* number $r = m/n$, the relation $\succcurlyeq$ acts like the representing probability function via a rule 9^+ style satisfying partition for which $A|B \approx (S_1\vee...\vee S_m)|(G\cdot(S_1\vee...\vee S_n))$.[12] When the representing function for $\succcurlyeq$ assigns $P[A \mid B] = r$ for an *irrational* number r, the relation $\succcurlyeq$ supplies a sequence of increasingly large rule 9^+ style partitions (for ever larger n), which supply a sequence of relationships $(S_1\vee...\vee S_{m+1})|(G\cdot(S_1\vee...\vee S_n)) \succcurlyeq A|B \succcurlyeq (S_1\vee...\vee S_m)|(G\cdot(S_1\vee...\vee S_n))$, where the values of m/n and (m+1)/n converge to r as n increases. Associated with each *completely-arch-extendable relation* is the set of Popper functions that represent its various *complete-arch extensions*.[13]

More generally, each *completely-extended* relation (*arch* or not) is at least nearly identical to its uniquely representing Popper function. A completely-extended relation $\succcurlyeq$ that fails to be *completely-arch-extended* may exhibit a slight "looseness in fit" of the following kind: its representing Popper function P may assign $P[A \mid B] = P[C \mid D]$ in cases where although $A|B \succ C|D$ and where no equal-partition can be fitted between A|B and C|D (i.e. they are infinitesimally close together).

Associated with each *completely-extendable relation* is the collection of Popper functions that represent its various *complete extensions*. Each Popper function in the representing collection for $\succcurlyeq$ accurately preserves the precise orderings ($\succ$ and $\approx$) among the argument pairs it compares – except for infinitesimally close argument pairs, which are always represented as equal.[14]

5. Conclusion

Bayesian approaches to confirmation theory represent evidential support for hypotheses in terms of conditional probability functions, which assign precise numerical values to each argument

---

[12] I.e., for which $(S_1\vee...\vee S_m)|(G\cdot(S_1\vee...\vee S_n)) \succcurlyeq A|B \succcurlyeq (S_1\vee...\vee S_m)|(G\cdot(S_1\vee...\vee S_n))$.

[13] This is the essence of how the representation theorem is proved. The version of the proof provided by Koopman (1940) is easily adapted to the representation of *completely-extended* comparative support relations by Popper functions – clause 3 of the above theorem. The remaining clauses this theorem follow easily from clause 3.

[14] Classical probability functions are, in effect, just the one-level Popper functions, those where $P[A \mid B] = 1$ for all A whenever $P[B \mid C\vee\neg C] = 0$. The *comparative support relations* that are represented by strictly classical probability functions are just those that satisfy the following additional rule (which provides an additional restriction on the *proto-support relations*):

If $\neg B|(C\vee\neg C) \succcurlyeq D|D$, then $A|B \succcurlyeq D|D$ for all A (i.e. if $(C\vee\neg C) \Rightarrow \neg B$, then $B \Rightarrow A$).

A|B. In most cases these probability assignments are overly precise. For, in most real scientific contexts the strengths of plausibility arguments for various alternative hypotheses, as represented by prior probability assignments, are fairly indefinite in strength, and so not properly rendered by the kinds of precise numbers that conditional probability functions assign. Indeed, most plausibility arguments for hypotheses are better rendered in terms of their strength compared with arguments for alternative hypotheses, rather than in terms of precise probability values. Thus, Bayesian prior probabilities are best represented by ranges of numerical values that capture the imprecision in comparative assessments of the extent to which plausibility considerations support one hypothesis over another.

Furthermore, in many cases this problem of over-precision also plagues the assignment of Bayesian *likelihoods* for evidence claims on various hypotheses (which represent what hypotheses *say* about the evidence). In realistic cases these will often have rather vague values. Indeed, on Bayesian accounts, the import of evidence via likelihoods is completely captured by *ratios of likelihoods*, which represent how much more likely the evidence is according to one hypothesis than according to an alternative. In many realistic cases these likelihood comparisons will also be somewhat vague or imprecise – best represented by ranges of values.[15]

This situation places the Bayesian approach to confirmation theory in an embarrassing predicament. The Bayesian approach first proposes a probabilistic logic that assigns overly precise numerical values to all arguments A|B, and then backs off from this over-precision by acknowledging that in many realistic applications the proper representation of evidential support should employ whole collections of precise probability functions that cover ranges of reasonable values for the prior probabilities of hypotheses, and also ranges for likelihoods in many cases.

The qualitative logic of comparative support described here offers a rationale for this kind of Bayesian approach to confirmation, including the introduction of classes of confirmation functions to represent ranges of values. The overly precise *probabilistic confirmation functions* are mere representational stand-ins for a deeper qualitative logic of comparative argument strength, captured by the *comparative support relations*. Each representing probability function

---

[15] Example: Precisely how likely is the observed fit between Africa and South America if the Continental Drift hypothesis is true? The value of this likelihood is presumably rather vague. To assess how the evidence supports the Drift Hypothesis, H_D, over the alternative Contractionist Hypothesis (that the continents have remained in place since the early molten Earth cooled and contracted), H_C, we need only assess *how much more likely* the evidence is according to Drift as compared to Contraction, $P[E \mid H_D{\cdot}B] / P[E \mid H_C{\cdot}B]$ (where B consists of relevant background and auxiliaries). Even this likelihood comparison will be somewhat imprecise, best represented by some range of numbers that capture the vagueness of the comparison. For confirmational purposes the approximate size of $P[E \mid H_D{\cdot}B] / P[E \mid H_C{\cdot}B]$ (very large, or extremely small) is all that matters. The resulting Bayesian assessment compares posterior probabilities on the basis of comparisons of prior plausibility, $P[H_D \mid B] / P[H_C \mid B]$ (where B contains plausibility considerations) together with evidential likelihood ratios:

$$P[H_D \mid E{\cdot}B] / P[H_C \mid E{\cdot}B] = (P[E \mid H_D{\cdot}B] / P[E \mid H_C{\cdot}B]) \times (P[H_D \mid B] / P[H_C \mid B]).$$

Each of these ratios may best be represented by a range of values, characterized by the collection of probability functions that provide values within the appropriate ranges.

reiterates the qualitative comparisons of argument strength endorsed by the *comparative support relation* it represents. Furthermore, the underlying qualitative logic can directly express the *incomparability in strength* among argument pairs that is common among real arguments. Although each representing probability function for a given relation $\succcurlyeq$ fills-in with precise comparisons among all argument pairs, the whole collection of representing probability functions for $\succcurlyeq$ captures the incomparability in terms of the available range of ways in which comparisons could, in principle, be filled-in, given the definite comparisons provided by $\succcurlyeq$.[16]

Thus, a probabilistic Bayesian confirmation theory employs overly precise *probabilistic representations* because they are computationally easier to work with than the *comparative support relations* they represent. Nevertheless, the features of evidential support that we really care about are captured by the comparative relationships among argument strengths, realized by the *comparative support relations* and their logic. The probabilistic representation of this logic merely provides a felicitous way to represent the deeper *qualitative logic of comparative support*.

Acknowledgements: I presented an early version of this paper at the *Workshop on Conditionals, Counterfactuals, and Causes in Uncertain Environments* at the University of Dusseldorf (May 20-22, 2011). Thanks to Gerhard Schurz and Matthias Unterhuber for making this workshop work so well. Thanks to all of the workshop participants for their extremely helpful questions and feedback. I owe a special debt to John Cusvert, David Makinson, Alan Hájek, and Chris Hitchcock. Each provided comments on evolving drafts, resulting in major revisions and improvements.

References

Cozman, F. (2013), 'Imprecise and Indeterminate Probabilities' (this volume).
de Finetti, B. (1937), 'La Prévision: Ses Lois Logiques, Ses Sources Subjectives', in *Annales de l'Institut Henri Poincaré* 17: 1-68. English translation in H. E. Kyburg, and H. E. Smokler (eds.) (1964), *Studies in Subjective Probability* (NewYork: Wiley).
de Finetti, B. (1974), *Theory of Probability*, Vols. 1 & 2, English translation by A. Machí and A. Smith (New York: Wiley).
Easwaran, K. (2013), 'Conditional Probability' (this volume).
Field, H. (1977), 'Logic, Meaning, and Conceptual Role', in *Journal of Philosophy* 74: 379-409.
Hawthorne, James (2014), 'A Primer on *Rational Consequence Relations*, *Popper Functions*, and their Ranked Structures', *Studia Logica*, v. 102, no. 4, 2014, pp. 731-749.
Keynes, J. M. (1921), *Treatise on Probability* (London: Macmillan).
Koopman, B. O. (1940), 'The Axioms and Algebra of Intuitive Probability', in *Annals of Mathematics* 41: 269–292.
Popper, K. R. (1959), *The Logic of Scientific Discovery* (London: Hutchinson).

[16] Transitivity (i.e. axiom 0) yields the following constraint on incomparable argument pairs: Given any transitive relation $\succcurlyeq$, whenever $H_2|E_2 \asymp H_1|E_1$ (i.e. whenever *not* $H_2|E_2 \succcurlyeq H_1|E_1$ and *not* $H_1|E_1 \succcurlyeq H_2|E_2$), every argument $H|E$ that is distinctly comparable to both $H_1|E_1$ and $H_2|E_2$ (i.e. $H|E \nsim H_1|E_1$, $H|E \nsim H_2|E_2$) must either be distinctly stronger than both (i.e. $H|E \succ H_1|E_1$ and $H|E \succ H_2|E_2$) or distinctly weaker than both (i.e. $H_1|E_1 \succ H|E$ and $H_2|E_2 \succ H|E$).

Savage, L. J. (1954), *The Foundations of Statistics* (New York: John Wiley & Sons; 2nd ed., 1972, New York: Dover).

Appendix:

From the Sparse Axioms 1-6 for Popper functions, we derive the Robust Axioms (1)–(6), as follows.

Notice that (6) is identical to 6, and (4) follows easily from 4; so we only need derive (1)-(3) and (5). We derive useful intermediate results along the way.

(3^*) if B |= A and A |= B, then P[A | C] = P[B | C]: directly from 3.

(3') if B |= A, then P[A | B] = P[C | C]: suppose B |= A; in 3 replace 'C' with 'B' to get P[A | B] ≥ P[B | B]; use 2 twice to get P[C | C] ≥ P[A | B] ≥ P[B | B] ≥ P[C | C].

(2^*) P[C | C] = 1: from ax. 1, for some A and B, P[A | B] ≠ 0; from (3^*) and 6, P[A | B] = P[A·A | B] = P[A | A·B] × P[A | B], so P[A | A·B] = 1; since A·B |= A, from (3') we have 1 = P[A | A·B] = P[C | C].

(3) if B |= A, then P[A | B] = 1: by (3') and (2^*).

(2') 1 ≥ P[A | B]: from (2^*) and 2.

($2^\#$) P[A | B] ≥ 0: if P[A | C] < 0, then P[A | C] ≠ 1 = P[C | C], so 5 yields P[¬A |C] = P[C | C] – P[A | C] = 1 – P[A |C] > 1 contradicting (2').

(2) 1 ≥ P[A | B] ≥ 0: by (2') and ($2^\#$).

(5^*) P[A | C] + P[¬A | C] = 1 or else P[D | C] = 1 for all D: by 5 and (2^*).

(1) if |= ¬A and |= B, then P[A | B] = 0: suppose |= ¬A and |= B; from (5^*) and (3), P[A | B] = 1 – P[¬A | B] = 1 – 1 = 0 (*done*) *unless* P[D | B] = 1 for all D; but "P[D | B] = 1 for all D" when |= B contradicts axiom 1, since it would have P[E | F] = 1 for all E and F, as follows: 1 = P[E·F | B] = P[E | F·B] × P[F | B] = P[E | F·B] × 1 = P[E | F] (by 6 and (4), since B·F is logically equivalent to F when |= B).

(5) if C |= ¬(A·B), then either P[(A∨B) | C] = P[A | C] + P[B | C] or P[D | C] = 1 for all D: Suppose C |= ¬(A·B) but *not* "P[D | C] = 1 for all D".

First we derive the following useful intermediate results:

P[A·B | C] = 0, P[B·¬B | C] = 0, P[A·¬A | C] = 0, P[¬B | A·C] = 1, and P[¬A | B·C] = 1: Since C |= ¬(A·B), by (3) and (5^*), P[¬(A·B) | C] = 1 = P[(A·B) | C] + P[¬(A·B) | C], so P[A·B | C] = 0. Since C |= (B∨¬B), by (3) and (5^*), P[(B∨¬B) | C] = 1 = P[(B∨¬B) | C] + P[¬(B∨¬B) | C], so 0 = P[¬(B∨¬B) | C] = P[B·¬B | C], by (3^*). Similarly, P[A·¬A | C] = 0. Also, from C |= ¬(A·B) we have both (A·C) |= ¬B and (B·C) |= ¬A, so from (3) P[¬B | A·C] = 1 and P[¬A | B·C] = 1.

Now we consider three cases, showing in each we get P[(A∨B) | C] = P[A | C] + P[B | C].

Case 1: Suppose P[B | ¬B·C] = 1: 0 = P[B·¬B | C] = P[B | ¬B·C] × P[¬B | C] = P[¬B | C], so P[¬B |C] = 0, so P[B | C] = 1, by 6 and (5^*). Since ¬(A∨B) |= ¬B, 0 = P[¬B | C] ≥ P[¬(A∨B) | C] ≥ 0 by 3 and (2), so P[¬(A∨B) | C] = 0, then by (5^*), P[(A∨B) | C] = 1 – P[¬(A∨B) | C] = 1 = P[B | C]; thus, P[A∨B | C] = P[B | C]. Also, 1 = P[¬B | A·C], so from ¬B·A |= ¬B, 3 and 6, 0 = P[¬B | C] ≥ P[¬B·A | C] = P[¬B | A·C] × P[A | C] = P[A | C], so P[A | C] = 0. Thus P[(A∨B) | C] = P[A | C] + P[B | C].

Case 2: Suppose P[A | ¬A·C] = 1: Then (as in Case 1) P[(A∨B) | C] = P[A | C] + P[B | C].

Case 3: Suppose P[B | (¬B·C)] < 1 and P[A | (¬A·C)] < 1. Then by repeated instances of (5^*), (3^*) and 6: 1 – P[(A∨B) | C] = P[¬(A∨B) | C] = P[¬A·¬B | C] = P[¬A | ¬B·C] × P[¬B | C] = (1 – P[A | ¬B·C]) × P[¬B | C] = P[¬B | C] – P[A·¬B | C] = P[¬B | C] –

$P[\neg B \cdot A \mid C] = P[\neg B \mid C] - P[\neg B \mid A \cdot C] \times P[A \mid C] = P[\neg B \mid C] - (1 - P[B \mid A \cdot C]) \times P[A \mid C] = P[\neg B \mid C] - P[A \mid C] + P[B \cdot A \mid C] = 1 - P[B \mid C] - P[A \mid C] + 0$; so $1 - P[(A \vee B) \mid C] = 1 - P[B \mid C] - P[A \mid C]$; thus, $P[(A \vee B) \mid C] = P[B \mid C] + P[A \mid C]$.