

September 15, October 23, 2008; January 3, 2010.

Deductively Definable Logics of Induction

John D. Norton

Center for Philosophy of Science

Department of History and Philosophy of Science

University of Pittsburgh

Running head: Logics of Induction

Keywords: Bayesianism, confirmation, induction, non-probabilistic

A broad class of inductive logics that includes the probability calculus is defined by the conditions that the inductive strengths $[A|B]$ are defined fully in terms of deductive relations in preferred partitions and that they are asymptotically stable. Inductive independence is shown to be generic for propositions in such logics; a notion of a scale-free inductive logic is identified; and a limit theorem is derived. If the presence of preferred partitions is not presumed, no inductive logic is definable. This no-go result precludes many possible inductive logics, including versions of hypothetico-deductivism.

1. Introduction

What if one does not believe that the probability calculus of the Bayesians is the universally applicable logic of induction? What if one believes, as I do (Norton 2003, 2005), that there is no universally applicable logic of induction and that different domains of application may call for different inductive logics? Then one asks what different sorts of logics of induction are possible. The purpose of this paper is to map out a large class of inductive logics. Their central term will be the quantity “[$A|B$],” which is, informally, the strength of inductive support proposition B lends to proposition A. It is modeled on the probability calculus’ $P(A|B)$ but the whole point of what follows is that $[A|B]$ need not and generally will not be a probability.

The class of inductive logics under consideration will contain those logics whose rules conform to two conditions. The first condition derives from the familiar, loose idea that inductive inference is a kind of partial or even inverted form of deductive inference. That makes it natural to expect deductive relations to figure in the definition of the rules of the inductive logic. The strongest form (shortly to we weakened) of this requires:

1. (Deductive definability) The strengths $[A|B]$ are defined fully in terms of the deductive relations obtaining among all the propositions of the Boolean algebra containing A and B . The simplest example of satisfaction of this condition arises in the bare form of hypothetico-deductive confirmation. In it, $[H|E]$ has one, non-trivial value only: “supports,” which is assigned when the theory tells us that evidence E inductively supports hypothesis H . By definition, $[H|E]$ has that value just if H deductively entails E .

While expanding the algebra of propositions by splitting its propositions into disjunctive parts may provide new information that will lead to adjustments in the strengths $[A|B]$, the second condition derives from the presumption that eventually the amount of new information supplied diminishes to the point of irrelevance and that the strength $[A|B]$ will stabilize.

2. (Asymptotic stability under disjunctive refinement) The strength $[A|B]$ assigned to fixed pairs of propositions A and B eventually stabilizes asymptotically as the algebra of propositions in which A and B reside is enlarged arbitrarily by replacing atomic propositions with disjunctions.

Conditions 1. and 2. define a class of inductive logics that we may call “deductively definable, asymptotically stable.” These two notions are introduced in Section 5 below. The sections before it develop background notions needed in Section 5. Sections 2 and 3 review the deductive structure of finite Boolean algebras of propositions and the symmetries of this structure. Section 4 introduces a weak notion of inductive logic and shows how an inductive logic is restricted by the condition that it shares the symmetries of the deductive structure.

Section 6 develops a “no-go” result for all deductively definable, asymptotically stable inductive logics: there are none possible, non-trivially. This no-go result immediately precludes some familiar attempts at inductive theorizing. For example, it precludes hypothetico-deductive approaches, even if they are richly supplemented by further restrictions concerning simplicity, explanatory power and the like, as long as these supplements are also defined in terms of deductive relations.

The consequence of the no-go result is that at least one of conditions 1. and 2. must be weakened if any non-trivial inductive logic at all is to be recovered. In condition 1., strengths [A|B] are required to be deductively definable in every algebra of propositions. It is shown in Section 7 that asymptotically stable inductive logics are possible if deductive definability is weakened so that it holds only in a subset of all algebras:

1'. (Deductive definability in preferred partitions) The strengths [A|B] are defined fully in terms of the deductive relations obtaining between propositions in inductively adapted partitions; and it is assumed that there are inductively adapted partitions of arbitrarily large size.

The presumption of preferred, inductively adapted partitions adds information about the inductive logic that cannot be derived from the deductive structure. It may seem like an unusual maneuver. It is, however, one of the oldest in probability theory. Requiring the inductive logic to be definable solely in terms of the deductive structure of these preferred partitions imposes certain symmetries on the logic. Most notably, it forces each atom of the algebra to enter equally into the inductive logic. When the precursor of probability theory was founded by Pascal and Fermat in the 17th century, their basic theoretical notion was essentially this, the equally likely case, such as the equally likely outcomes of a die throw. The addition of preferred partitions introduces analogous inductively equivalent cases in the new logics. The old rule for computing the probability of an outcome is an example of the rules of 1': the probability of an outcome is just the ratio of the number of its logical atoms to the total number of atoms.

We shall say that the resulting inductive logics are “deductively definable in preferred partitions and asymptotically stable.” Three results will then be demonstrated for this general class of inductive logics:

- Inductive independence is generic: If the Boolean algebra of propositions is large, inductive independence obtains between most pairs of propositions. (Sections 8)
- A limit result for narrow¹ inductive logics: If an hypothesis H with non-minimal prior support entails evidence E_1, E_2, E_3, \dots , then the inductive support E_n accrues from $E_1 \& \dots \& E_{n-1}$ approaches certainty with large n. (Section 9)
- There is a sub-class of scale-free inductive logics whose inductive relations remain unchanged as the size of the algebra of propositions is changed. (Section 10, 11)

These ideas are familiar in so far as they are well known in probability theory. They are derived here for a much broader class of logics and, as a result, from much weaker assumptions. It will become apparent that these results rely heavily on the deductive relations in the Boolean algebra and less on specific assumptions about the inductive logic. The results concerning inductive independence and the limit theorem can be seen as facts about the deductive structure merely reflected through good adaptation in the inductive logic.

This last point is itself an important moral. It shows that distinctive features of inductive logics, known to us through the probability calculus, may have a different origin. The requirement that an inductive logic be adapted to the deductive relations obtaining between propositions places strong restrictions on the inductive logic. The derivations of the results cited proceed *without assuming* that the strengths are additive or that there is a Bayesian dynamics for conditionalization.

Section 12 contains a summary conclusion. Supporting technicalities are found in five appendices.

2. Finite Boolean Algebras of Propositions

We begin by reviewing the deductive relations between propositions. We shall consider only the simplest case of Boolean algebras of finitely many, logically distinct propositions; that is, of sets of propositions closed under the operations of \sim (negation), $\&$ (conjunction) and \vee (disjunction). The deductive structure of the algebra is the set of all deductive entailment relations obtaining among the proposition, where an entailment relation is just the fact of whether one proposition entails another. The size of the algebra is measured conveniently by the number of atoms; that is, the number of deductively strongest, non-contradictory propositions it contains. The algebra Ω_3 with three atoms a_1 , a_2 and a_3 can be pictured helpfully by the directed graph of Figure 1:

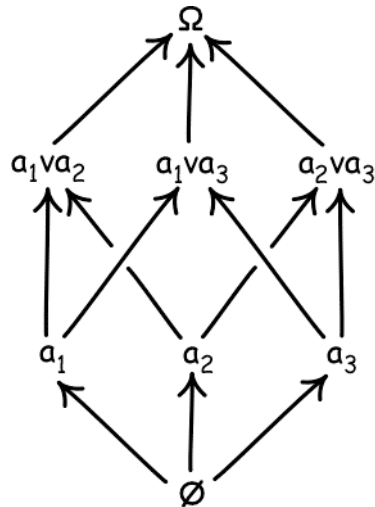


Figure 1. A Three-Atom Boolean Algebra

The nodes are the propositions of the algebra and the directed edges represent deductive entailment. (For more details, see *Appendix 1. Characterizing Finite Boolean Algebras.*)

The notion of asymptotic stability to be introduced below requires that we have a means of relating algebras of propositions of different sizes. This is achieved by the operation of “disjunctive refinement.” In it, each atom of the algebra is replaced by a disjunction of one or more new atoms. For example, we make take a three atom algebra with atoms a_1 , a_2 and a_3 and replace its first atom a_1 by two new mutually exclusive atoms $a_1 = b_1 \vee b_2$. Thereby, the algebra Ω_3 has been refined to the new algebra Ω_4 with atoms b_1 , b_2 , a_2 and a_3 , as shown in the Figure 2.

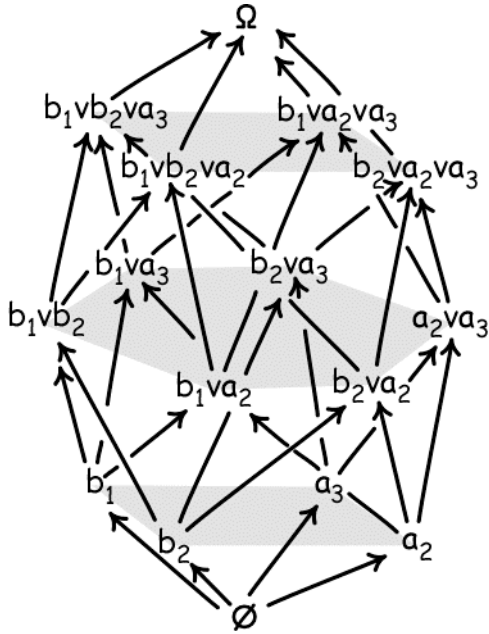


Figure 2. Refinement of a Three-Atom Boolean Algebra

The reverse operation is a disjunctive coarsening. We shall generally take the refinements and coarsening as connecting different partitions of the same space of possible outcomes. In general, each atom is refined to a disjunction of different numbers of atoms. In a *uniform* refinement, each atom is refined to a disjunction of the same number of atoms. In an *α -fold uniform refinement*, each atom is replaced by α atoms. It is presumed that, for any algebra, further disjunctive refinement of each atom is always possible. As a result, Boolean algebras of arbitrarily large, but still finite size are accessible.

3. Symmetries of a Boolean Algebra

Our present interest lies with the deductive relations among propositions in each of the Boolean algebras. The structure these relations form differs only according to the number of atoms. All three-atom Boolean algebras, for example, have the same deductive structure. Each presentation of a Boolean algebra will have surplus devices that are not part of the deductive structure. They are the labels used to distinguish the atoms and propositions from one another (in so far as they are distinguished other than through their deductive relations).

As far as the deductive structure is concerned, any system of labeling is as good as any other as long as it keeps distinct propositions distinct. In a three atom algebra, we may call the

atoms a_1, a_2 or a_3 ; or $\text{atom}_1, \text{atom}_2$ or atom_3 ; or Tom, Dick and Harry. The two algebras of Figure 3 have the same deductive structure:

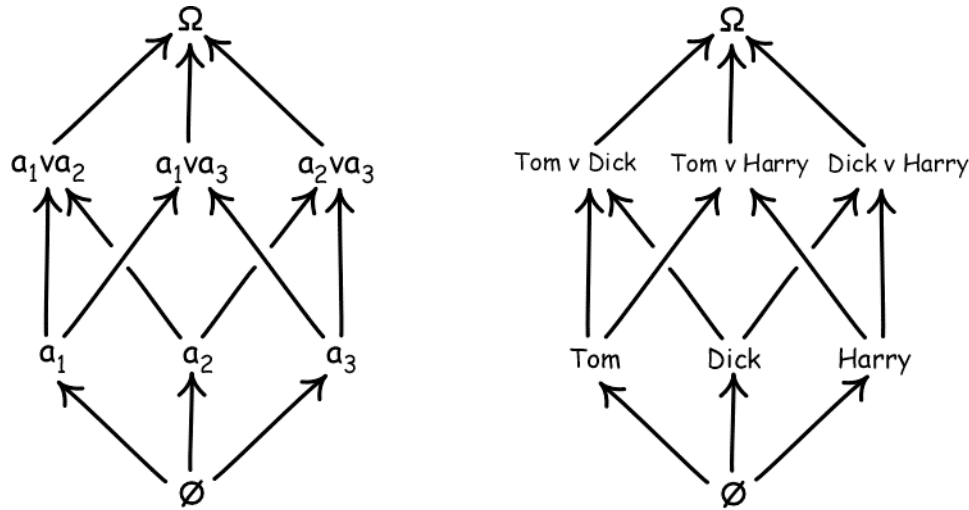


Figure 3. Renaming Atoms Yields Equivalent Three-Atom Boolean Algebras

The deductive structure is unchanged if we relabel the atoms; and it is unchanged if we merely switch the atoms around.

This last transformation of permuting atomic labels or, equivalently,² of permuting atoms is a symmetry of the deductive structure; that is, it leaves the deductive structure unchanged or invariant. For example, the three atom algebras of Figure 4, in which the three atoms have been permuted have the same deductive structure.

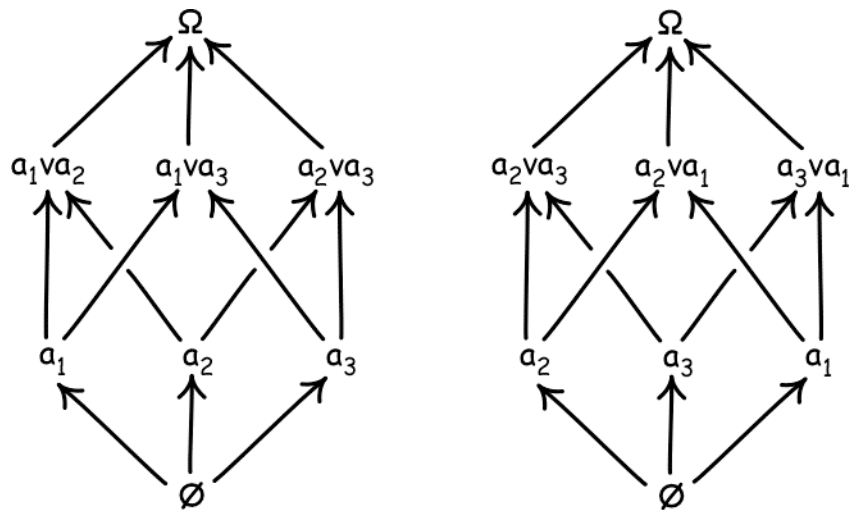


Figure 4. Permuting Atoms Yields Equivalent Three-Atom Boolean Algebras

These algebras are very rich in symmetries. It is easier to identify the symmetries if we visualize them geometrically. The simplest way of doing this for a three atom algebra is to arrange the atoms a_1 , a_2 and a_3 at the corners of an equilateral triangle as shown in Figure 5 below. A permutation of atoms in which a_1 is mapped to a_2 , a_2 to a_3 , and a_3 to a_1 is represented by the rotation of the entire structure by 120° around the dashed vertical axis.

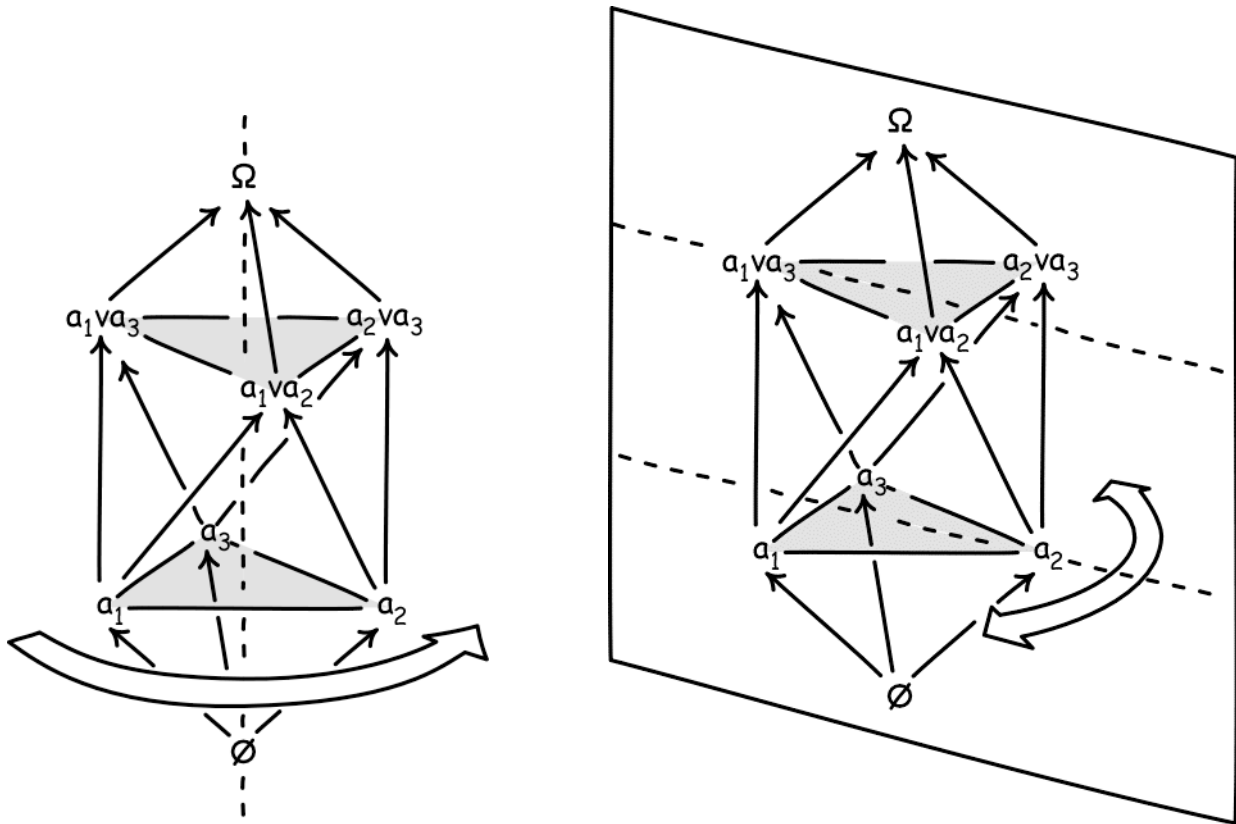


Figure 5. Symmetries of a Three Atom Boolean Algebra

The same figure represents another symmetry, the exchange of a_1 and a_3 , by a reflection through a plane. More generally, the deductive structure is given by the directed edges of the graphs shown; it is pictured best by a figure that represents each node as intrinsically the same and distinguished only by its place in the network of entailments, as shown in Figure 6:

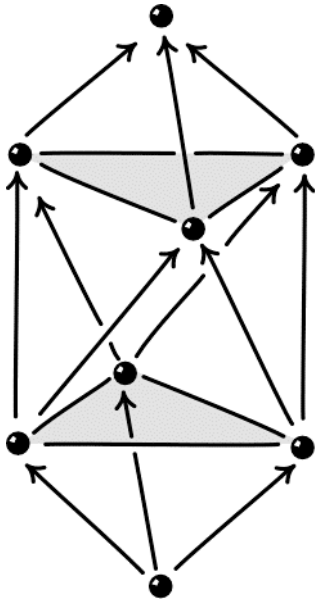


Figure 6. The Bare Picture of the Deductive Structure of a Three-Atom Boolean Algebra

Boolean Algebras with more atoms are also very rich in symmetries. Because they are more complicated structures, it is harder to visualize the full range of symmetries by the figures used so far. (Figure 2 of a four-atom algebra shows the problem. What are its symmetries?) Another geometric representation helps us to see this full range. In it, the propositions of Ω_3 (excluding the contradiction \emptyset) are pictured as an equilateral triangle. Each corner is one of the atoms; each side is a disjunction of two atoms; and the area is a disjunction of all three, Ω_3 , as shown in Figure 7.

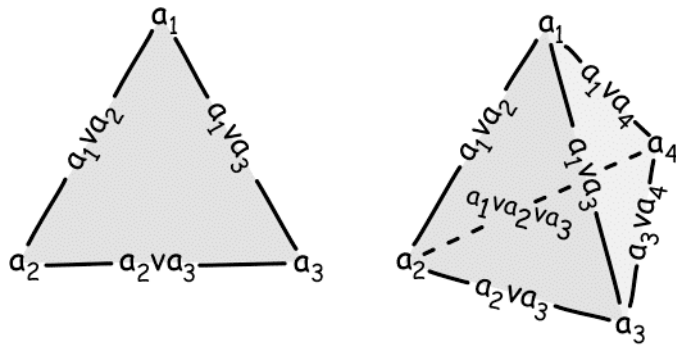


Figure 7. Three- and Four-Atom Boolean Algebras Represented as Simplices

The symmetries in which the atomic labels a_1 , a_2 or a_3 are rearranged correspond to the geometric symmetries of an equilateral triangle. A reflection about an axis of symmetry

corresponds to the switching two atomic labels. A rotation by 120 degrees corresponds to permuting all three.

The triangle of Figure 7 is a 2-simplex. Higher order simplices can represent Boolean algebras with more atoms. A tetrahedron, that is, a 3-simplex, represents a four-atom Boolean algebra as shown in Figure 7. Its symmetries are recovered from the geometric symmetries: a reflection through a plane switches two atoms and a rotation by 120° cyclically permutes three atoms. The various combinations enable access to all $24=4!$ permutations of the atoms.

4. Inherited Symmetries of an Inductive Logic: The Symmetry Theorem

We shall introduce an inductive logic through the quantity $[A|B]$, which is the strength of inductive support proposition B affords to proposition A. An inductive logic is the totality of the assignments and the logic is specified by the rules we choose for assigning these strengths to pairs of propositions A and B. These strengths or degrees $[A|B]$ could be real-valued, conditional probability measures. But they need not be. Very little is assumed at the outset for the set of values that the strength $[A|B]$ can take. In this value set, there is a minimal strength $[\emptyset|\Omega] = [\emptyset|A]$, for any non-contradictory proposition A, and a maximal $[\Omega|\Omega] = [A|A]$, corresponding to maximal support. Further properties will be induced in the discussion below by adapting the inductive logic to the deductive structure of the Boolean algebra.³ Indeed the major point of this paper is that we can arrive at strong and interesting results about inductive logics by assuming far less about the degrees of support than is normally done.

We have seen that a Boolean algebra of propositions is rich in symmetries. It can happen that all these symmetries are inherited by an inductive logic defined on the algebra. That circumstance arises most naturally if the inductive strengths are defined fully in terms of the deductive structure. Should this happen, since a Boolean algebra is rich in symmetries, the inductive logic is constrained powerfully. Those constraints are expressed by the Symmetry Theorem described below, as we now show.

Since the deductive relations in the algebra are unaffected by a permutation of the atoms a_1, a_2, \dots , then the inductive logic must be correspondingly unaffected, if its symmetries include

those of the deductive structure. To see the consequences of this symmetry, consider an algebra with 13 atoms, a_1, a_2, \dots, a_{13} , in which we form two degrees of support

$$[a_1 \vee a_2 \mid a_2 \vee a_3] \quad \text{and} \quad [a_{11} \vee a_{12} \mid a_{12} \vee a_{13}]$$

A simple permutation of the atoms will switch

$$a_1 \leftrightarrow a_{11} \quad a_2 \leftrightarrow a_{12} \quad a_3 \leftrightarrow a_{13}$$

and leave the rest unchanged. Since this is a symmetry of the Boolean algebra, the deductive relations among the atoms a_1, a_2 and a_3 and the remaining propositions will be the same as the deductive relations among the atoms a_{11}, a_{12} and a_{13} and the remaining, correspondingly permuted propositions in the algebra. By supposition, the inductive relations obtaining among these propositions must also be unaffected by the permutation. It follows that these two strengths must be the same. That is,

$$[a_1 \vee a_2 \mid a_2 \vee a_3] = [a_{11} \vee a_{12} \mid a_{12} \vee a_{13}]$$

The argument just given is easily generalized. Let us say that we wish to compare two degrees of support $[A \mid B]$ and $[A' \mid B']$ where the propositions are drawn from an algebra with N atoms. It is sufficient for the two degrees to be equal that the following equalities obtain:

$$\#A \& B = \#A' \& B' \quad \#A \& \sim B = \#A' \& \sim B' \quad \#\sim A \& B = \#\sim A' \& B' \quad \#\sim A \& \sim B = \#\sim A' \& \sim B'$$

where the notation “#Proposition” means “number of atoms in Proposition.” For then we can always find a permutation of the N atoms that will map just the atoms $A \& B$ onto those of $A' \& B'$; and conversely; and so on for the remaining three quantities.⁴ It now follows that all the deductive relations among the propositions A, B and the remaining propositions in the algebra will be the same as those among A', B' and the remaining, correspondingly permuted propositions in the algebra. Therefore the inductive relations among these propositions will also be the same. That is, we will have

$$[A \mid B] = [A' \mid B'].$$

All that is needed to arrive at this equality is agreement in the number of atoms in the disjunctive parts $A \& B, A \& \sim B, \sim A \& B, \sim A \& \sim B$ with those of the corresponding parts $A' \& B', A' \& \sim B', \sim A' \& B', \sim A' \& \sim B'$. Thus it follows that that the number of atoms in $A \& B, A \& \sim B, \sim A \& B, \sim A \& \sim B$ fixes the strength $[A \mid B]$. This important conclusion can be stated compactly as:

Symmetry theorem. If the deductive symmetries of a Boolean algebra with N atoms are all also symmetries of an inductive logic [.I.], then

$$[A|B] = f_N(\#A\&B, \#A\&\sim B, \#\sim A\&B) \quad (S)$$

for some function f_N from numbers of atoms to strength of support.⁵

5. Deductively Definable, Asymptotically Stable Inductive Logics

These inductive logics meet two conditions, which will be discussed separately:

5.1 Deductive Definability

We expect some sort of adaptation of an inductive logic to deductive relations. This is already suggested by the crude picture of induction as partial deductive entailment or as an inverse of deduction. In the older tradition, induction is the inference from instances to generalizations; it is just the reverse of the deductive inference from generalizations to their instances. In so far as we believe that the inferential relations among propositions are already fixed by the deductive relations between them, then we would expect these deductive relations to fix the inductive relations. This strongest adaptation of an inductive logic to the deductive structure arises if the first of the two conditions to be explored here are met:

Deductive definability. The rules for assigning degrees of inductive support are defined solely in terms of the deductive structure.⁶

It might seem that this strong adaptation could not arise in a logic of induction, since induction is, by its nature, non-deductive. That is not so. That a logic is defined solely by means of deductive structure does not mean that the resulting logic is itself deductive. It may well be inductive.

It turns out to be quite common for inductive relations to be defined fully in terms of the deductive structure. A familiar example is the formal theory of instance confirmation. The notion of an instance of a generalization is definable fully in terms of deductive relations. For example, in predicate logic, the formula $P(a)$ is an instance of the universally quantified proposition $(x)P(x)$. The instance is inferred from the generalization by the deductive rule of universal

instantiation. Analogously⁷ the formula $p(c_1) \vee \dots \vee p(c_n)$ is an instance of the existentially quantified proposition $(\exists x)P(x)$ in a domain with individuals c_1, \dots, c_n . This gives us a rule for specifying an inductive logic:

A instance-confirms B if A is an instance of B.

This relation “instance-confirms” is defined solely in terms of deductive relations. This example cannot be explored further here, however, since we are not considering predicate logic.

In the context of the Boolean algebras employed here, the important example is hypothetico-deductive confirmation. Its basic rule is just:

E HD confirms H if H deductively entails E.

so that once again the inductive relation “HD confirms” is defined solely in terms of deductive relations between the evidence E and the hypothesis confirmed H. Of course this is the very simplest form of HD confirmation. It must be strengthened in two ways if it is to approach the sort of inductive relations routinely employed in science.

First, the simple rule must be augmented in such a way that it admits degrees of confirmation. That it had been cold the first six days of the week HD confirms that it will be cold for the whole week; and it HD confirms that it will be cold now and forever. However we would expect that the first should be judged much more strongly confirmed than the second. This suggests a modification to the rule that somehow accords stronger HD confirmation to an hypothesis the shorter the deductive distance from the hypothesis to the evidence.

Second, there seems to be right and wrong ways for the hypothesis deductively to entail the evidence. Most of us find the evidence of lightning strikes to confirm better the existence of strong electric charges distributions in clouds than thunder gods. That is so, even though both hypotheses deductively entail the evidence. This and many other examples like it are routinely handled by requiring that the hypothesis not just deductively entail the evidence, but it must also explain it. Another familiar addition limits the hypotheses that may be confirmed to those that are simple; and another to those that have not been contrived ad hoc. Yet another requires that the entailing of the evidence arise within a context in which it mounts a severe test of the hypothesis; that is, it would be unlikely for the evidence to obtain were the hypothesis false. (For a short survey of the many ways that the simple HD scheme has been augmented, see Norton, 2005, pp. 14-17.)

These additional considerations will lead to an augmented rule for HD confirmation that will look something like

E HD confirms H with strength [H|E],
if H deductively entails E and condition D obtains.

The added condition D is intended to enable the augmented version of HD confirmation to meet the added demands just sketched. My concern lies not in the details of just what this possibly quite complicated condition D should say. My expectation is that it will be hard to find any formulation for it that attracts more than parochial assent.

Rather my concern lies just in the question of whether the condition D is formulated purely in terms of the deductive relations obtaining between the propositions in the Boolean algebra. That might happen if we measure deductive distance between propositions through the number of atoms added disjunctively in the deduction. We add one atom in the deduction from a_1 to $a_1 \vee a_2$; we add nine in the deduction from a_1 to $a_1 \vee \dots \vee a_{10}$. That measure of deductive distance might then be related to a varying strength of inductive support. Or, if the algebra is very large, we might find some characteristic of a particular type of deductive inference that enables us to designate it as explanatory; or, if the characteristic is lacking, non-explanatory. Again, we may find some simplicity measure for hypotheses that is defined solely in terms of the deductive relations obtaining between the propositions.

Successful rules, such as these, if they can be found, would have the very appealing character of being self-contained in the sense that no additional inductive structures need be posited to complete the formulation of condition D.

5.2 Asymptotic Stability

The second condition expresses an expectation about how an inductive logic will behave under repeated disjunctive refinement:

Asymptotic stability under disjunctive refinement. The inductive degrees of support approach fixed values arbitrarily closely under repeated disjunctive refinement.

To see what is at issue here, consider two proposition $A = a_1$ and $B = a_1 \vee a_2$ and form [A|B]. Does this strength [A|B] change if we disjunctively refine the algebra so that $a_1 = b_{1,1} \vee b_{1,2}$ and $a_2 = b_{2,1} \vee b_{2,2}$? Whether this strength changes will depend on the particular inductive logic implemented. The refinement reveals new disjunctive parts of A and B and the inductive logic

may judge that this new information requires it to adjust the strength $[A|B]$. What we do expect, however, is that the adjustments will diminish as we continue to refine further. That is, after massive refinement has been completed, we expect that further refinement will not appreciably affect the value of $[A|B]$. Further refinement eventually becomes immaterial to the strength $[A|B]$; it is just the repeated splitting of logical hairs.

This condition of stability has been defined vaguely in that a measure for the diminution of adjustment has not been specified. The expectation is that the sequence $[A|B]_M, [A|B]_N, [A|B]_P, \dots$ will eventually converge as we proceed through refinements $\Omega_M, \Omega_N, \Omega_P, \dots$, where $M < N < P < \dots$. This expectation requires a measure of closeness for two strengths $[A|B]$ and $[C|D]$. It is assumed that such a measure exists. If the strengths are real valued or real-vector valued, that measure could be just the norm $\|[A|B] - [C|D]\|$. For more discussion, see *Appendix 4. A Notion of Continuity for Strengths of Inductive Support*.⁸

It may be helpful to illustrate the reasonableness of asymptotic stability by considering a slightly more detailed example. Take the simple case of two hypotheses H and L that assert that some unknown parameter lies in a “high” or a “low” value set respectively. The evidence E narrows the possible values of the parameter to a range that contains both the high and low values sets. That is, each of H and L entail the evidence E . Hence we might say that each of H and L gain some inductive support from E . Knowing only that each of H and L entail E gives us no basis for assigning one greater support than the other. That can change if we learn more about H, L and E through a disjunctive refinement. Consider the refinement to $\Omega = a_1 \vee a_2 \vee \dots \vee a_{100}$, where a_n locates the parameter in the real interval $[n-1, n]$. We may then have $L = a_1, H = a_{10} \vee a_{11} \vee \dots \vee a_{50}$ and $E = a_1 \vee a_2 \vee \dots \vee a_{50}$. With this extra information in hand, we might reasonably expect our inductive logic to assign a greater strength of support to H than L . There will, however, be a limit to how much inductively useful information can be supplied by further refinement. The next refinement may merely divide each atom into ten disjunctive parts, which would likely have no effect on our assessment of H and L . We can certainly find disjunctive refinements that are manifestly irrelevant to our inductive assessments. Let X be any proposition that is irrelevant to the value of the parameter mentioned in H and L . Proposition X induces a disjunctive refinement that replaces each a_i by the disjunction $(a_i \& X) \vee (a_i \& \sim X)$. We would not expect our inductive assessments to be altered by this disjunctive refinement.

The requirement of asymptotic stability under disjunctive refinements asserts that this effect is universal. While disjunctive refinement may initially quite properly lead us to adjust our inductive assessments, eventually further refinement becomes immaterial and we expect our inductive assessments to stabilize. There is also a pragmatic virtue in an asymptotically stable logic: the assurance that its inductive assessments will stabilize means that it provides a more secure platform for further inference in the long run.

6. A “No-Go” Result for Them

An immediate consequence of the symmetry theorem of Section 4 is that there are no inductive logics that are both deductively definable and asymptotically stable in the sense of Section 5. The essence of the problem stems directly from the fact that the relation of inductive support must remain adapted to the Boolean algebra through relation (S) no matter how we refine it and this precludes asymptotic stability. Imagine, for example, that we have a 5 atom Boolean algebra and two mutually exclusive propositions A and B, each with two atoms. It follows from (S) that their support on the background Ω is the same:

$$[A|\Omega] = f_5(2, 0, 3) = [B|\Omega]$$

We can disjunctively refine just the atoms in A, so that its two atoms are replaced by 97 atoms, but those of B are unchanged. Then, under reasonable interpretations of the function f_{100} , A has gained vastly more support than B:

$$[A|\Omega] = f_{100}(97, 0, 3) \quad [B|\Omega] = f_{100}(2, 0, 3)$$

An alternative refinement of just the atoms in B gives exactly the reversed result:

$$[A|\Omega] = f_{100}(2, 0, 3) \quad [B|\Omega] = f_{100}(97, 0, 3)$$

This example indicates that the inductive relation is fatally unstable under disjunctive refinement. For any refinement that leads to a favoring of A, there will be a “clone” refinement that leads to an exactly equivalent favoring of B. It is not quite the “no-go” result sought since we have relied on expectations of reasonableness to conclude that $f_{100}(97, 0, 3)$ represents a far greater strength of support than $f_{100}(2, 0, 3)$.

To recover the no-go result, we proceed more cautiously. We could stipulate a small set of conditions that expresses what it is for a function f to be “reasonable.” It turns out that we need no new assumptions of this type. If the logic allows any inequalities among the limit values

to which the degrees of support stabilize, these inequalities give the symmetries of the logic sufficient purchase to force a contradiction. This sustains the following result, where the proof has been relegated to Appendix 2.

No-go result for an inductive logic defined fully in terms of deductive structure. The only such inductive logic that is asymptotically stable. Under arbitrary disjunctive refinement, it assigns the same limiting strength of support to all contingent propositions in the algebra and their disjunctive parts; and it affirms that in this limit an item of evidence confirms equally all contingent hypotheses that deductively entail it.

The result applies most importantly to any hypothetico-deductive logic of induction in which an hypothesis is confirmed by its deductive consequences; and to all elaborations of it that seek to distinguish stronger and weaker strengths of support through further considerations, such as the simplicity or explanatory power of the hypotheses, as long as these considerations are expressible in terms of deductive relations. It is also applicable elsewhere, whenever inductive relations are defined solely in terms of deductive structure. Objective Bayesians, for example, seek conditions under which a unique prior probability may be specified. The result applies to such efforts in so far as the rules that specify the unique prior probability depend entirely on deductive structure. The simplest such rule might assign the same prior probability to each of the atoms of a finite Boolean algebra of propositions; that rule is clearly unstable under disjunctive refinement.⁹

7. Inductive Suppositions: Deductive Definability in Preferred Partitions Only

The outcome of the analysis of the preceding sections is that the deductive structure of a Boolean algebra is insufficient by itself to enable the defining of an admissible inductive logic. Some further structure must be introduced. While that may be done in many ways, here it shall be done in a way that is closely related to the theoretical notions already introduced. We shall posit:

Inductively adapted partitions. Among all possible partitions of the Boolean algebra Ω into mutually exclusive disjunctive parts, there is a strict subset in which the symmetries of the deductive structure are included in those of the inductive logic.

These adapted partitions amount to a division into the inductively uniform cases. For probability theory, they correspond to identification of the equiprobable cases. A simple example pertains to the outcomes of a fair die toss. The partition into outcomes $a_1 = \text{odd}$ and $a_2 = \text{even}$ is adapted. The deductive relations are unaltered if we permute the labels a_1 and a_2 ; correspondingly, the inductive relations, in this case represented by the probabilities of the various outcomes, are also unaffected by the permutation. Analogously, partition into outcomes $b_1 = \text{one}$, $b_2 = \text{two}$, \dots , and $b_6 = \text{six}$ is adapted. However the partition $c_1 = \text{one}$, $c_2 = \text{two}$, $c_3 = \text{three or four or five or six}$ is not adapted. While permuting c_1 and c_3 is a symmetry of the deductive structure in the algebra Ω_3 , it is not a symmetry of the inductive structure. Outcomes c_1 and c_3 have different probabilities.

A related notion recalls that not all disjunctive refinements have the same status:

Inductively adapted disjunctive refinements are those that relate inductively adapted partitions.

We can now replace the earlier requirement of deductive definability of Section 5.1 by the weaker:

Deductive definability in preferred partitions. The rules for assigning inductive degrees of support are defined solely in terms of the deductive structure in these inductively adapted partitions; and it is assumed that there are inductively adapted partitions of arbitrarily large size.¹⁰

We shall henceforth consider inductive logics that meet this weaker requirement and also the requirement of *Asymptotic stability under disjunctive refinement* of Section 5.2 in these inductively adapted partitions.

The important observation is that the result (S) of the symmetry theorem still applies, but it applies only in the special case of inductively adapted partitions; for arbitrarily chosen partitions, it fails. This observation enables us to escape the no-go result of the last section. Its derivation depended essentially on our being free to apply any disjunctive refinement and still arrive at a partition in which the result (S) holds. With the added structure of adapted partitions, that freedom is no longer assured.

While the identification of equiprobable cases is an illustration of an adapted partition, the presumption that such partitions exist is weaker than the positing of an additive measure. The

inductive strength [A|B] in an adapted partition is merely constrained by (S) to some unspecified function $f_N(\#A \& B, \#A \& \sim B, \#\sim A \& B)$. The presumption that inductive strengths are probabilities is stronger in that it forces a particular choice for this function¹¹

$$f_N(\#A \& B, \#A \& \sim B, \#\sim A \& B) = \frac{\#A \& B}{\#A \& B + \#\sim A \& B} = \frac{\#A \& B}{\#B} \quad (P)$$

Note that this particular choice greatly restricts the logic. For example, it discards the values of $\#A \& \sim B$ thereby assuming what is defined as “Narrowness” in Section 9.2 below.

8. Inductive Independence is Generic

The logical structure of the Boolean algebra places powerful and interesting constraints on an inductive logic. One of the most interesting arises when the Boolean algebra is very large. Then it turns out that inductive independence between pairs of propositions in the algebra is generic; inductive dependence arises only rarely. Informally this means that learning the definite truth or falsity of one proposition in the algebra does not lend differential support to virtually all of the other propositions. In this sense, the inductive reach of most logics is weak.¹²

That independence is generic is a fact derived largely from the deductive structure of a large Boolean algebra and, more specifically, from the combinatorics of its atoms. We need very little from the inductive logic itself to arrive at this result. The precise argumentation needed to arrive at the result requires some involved accountancy. These arguments have been reserved for *Appendix 3. Inductive Independence in Large Boolean Algebras*. The essential ideas behind the combinatorial arguments are quite simple and will be sketched here.

We shall consider a Boolean algebra Ω_N , with N atoms a_1, \dots, a_N , where N is very large. As a preliminary, to show how strongly combinatorics constrain us, consider where most propositions are to be found. Virtually all the propositions will lie in the middle of the algebra and will be disjunctions of roughly $N/2$ atoms. This is not what happens in small Boolean algebras, such as pictured in Figures 1 and 2 above. There we see only a slight tendency of the propositions to concentrate around those with $N/2$ atoms. That tendency grows stronger without limit as N gets very large. So a more realistic picture of a very large Boolean algebra is shown in Figure 8 with virtually all propositions arising midway through the graph where propositions formed from $N/2$ atoms are found:

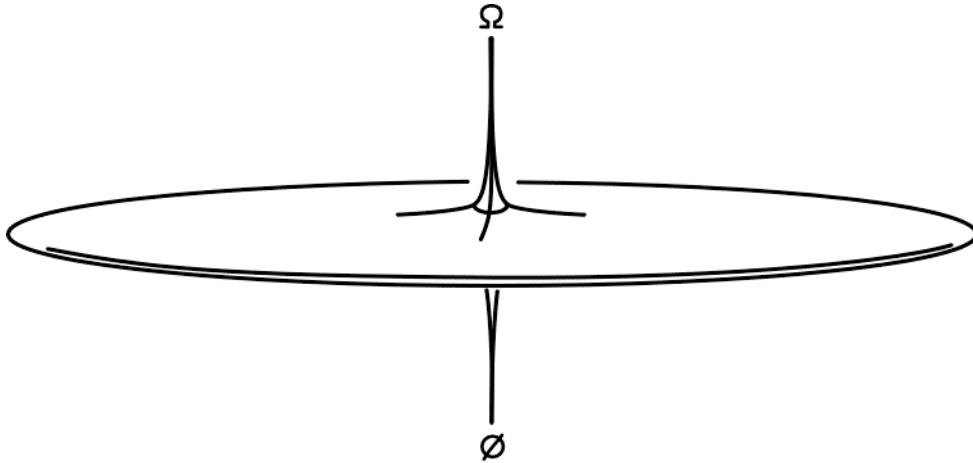


Figure 8. Very Large Boolean Algebra

This concentration is a direct result of combinatorics. Propositions are formed as disjunctions of atoms. Each distinct subset of the N atoms under disjunction forms a distinct proposition. We can form propositions with n disjuncts in $\frac{N!}{n!(N-n)!}$ ways. For large N , this combinatorial quantity has a very sharply peaked maximum at $n=N/2$, which is where most propositions are found.

This type of combinatorial effect is what underlies the generic character of inductive independence. To see this, consider any two propositions A and B in the algebra. Overwhelmingly, A will be a disjunction of roughly $N/2$ atoms; and B will also be a disjunction of roughly $N/2$ atoms. How will these two sets of atoms overlap? For most propositions B , roughly half of B 's atoms will be among A 's and the other half will not. The most frequent case is illustrated in the Figure 9:

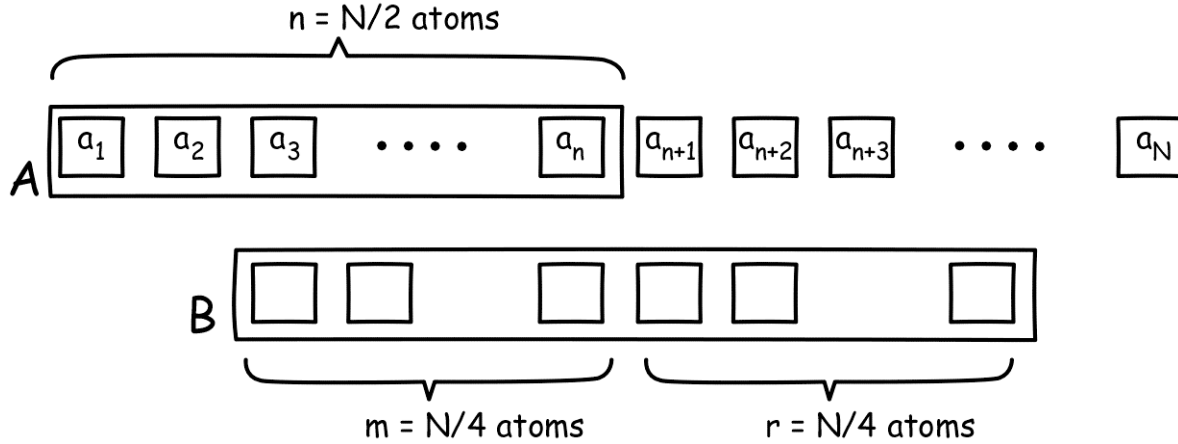


Figure 9. Most frequent distribution of atoms in propositions A and B

The same sort of combinatorial argument leads to this result. If there are n atoms in A and B such that it comprises m atoms that are also in A and r atoms that are not, then the number of ways of choosing pairs of propositions A and B are

$$\frac{N!}{n!(N-n)!} \cdot \frac{n!}{m!(n-m)!} \cdot \frac{(N-n)!}{r!(N-n-r)!}$$

For large N , this quantity has a very sharply peaked maximum at the most frequent case

$$n = N/2, m = r = N/4. \quad (F)$$

To convert these results into a conclusion concerning inductive independence, we need to define inductive independence. The natural definition is:¹³

Inductive independence. Proposition A is inductively independent of proposition B just if

$$[A|B] = [A|\sim B] \quad (I)$$

for then A acquires just the same inductive support from the truth of B as from its falsity.

Our concern is very large Boolean algebras. We have assumed that, through disjunctive refinement, we can enlarge the algebra arbitrarily. We have also assumed that there is no upper limit to the size of the inductively adapted partitions. That is, for any N , we can always find an inductively adapted partition with N or more atoms. Therefore we can choose an inductively adapted partition of sufficiently large size to drive most values of n , m and r as close as we like to the most frequent values above, $n = N/2$, $m = r = N/4$.

In an inductively adapted partition, (S) holds. It follows that, for the most frequently occurring propositions A and B, we have

$$[A|B] = f_N(\#A\&B, \#A\&\sim B, \#\sim A\&B) = f_N(m, n-m, r) = f_N(N/4, N/4, N/4)$$

$$[A|\sim B] = f_N(\#A\&\sim B, \#A\&B, \#\sim A\&\sim B) = f_N(n-m, m, N-n-r) = f_N(N/4, N/4, N/4)$$

Therefore, for these most frequent propositions, $[A|B] = [A|\sim B]$ so that A is inductively independent of B. For large N, the distribution of atoms in most pairs of propositions will be close to (F), so that most pairs of propositions will be close to satisfying the condition of inductive independence (I)—and arbitrarily close if we allow N to grow sufficiently large.

While the above arguments display the essential basis of the generic independence of propositions in a large Boolean algebra, they are not precise in so far as they depend upon the assumption that most cases are close to the most frequent cases and that the sense of closeness is sufficient to bring independence. Making these inferences more precise requires a more careful analysis of the distribution of propositions, as given in *Appendix 3. Inductive Independence in Large Boolean Algebras*. That analysis also requires a notion of continuity of the assignment of inductive strengths, so that if the distribution of atoms in propositions A and B is close to (F), we can say that the inductive strengths are close to the values that satisfy the condition of inductive independence (I). The requisite notion of continuity is developed in *Appendix 4. A Notion of Continuity for Strengths of Inductive Support*. In sum, the result established is:

Inductive independence is generic.

In inductive logics that are deductively definable in adapted partitions, continuous and asymptotically stable, we can always find a partition of sufficiently large size, such that there are subsets of pairs of proposition comprising a fraction as close to one as desired of all pairs, in which any pair of propositions A, B satisfies the condition of independence $[A|B] = [A|\sim B]$ to within any nominated the degree of closeness of strengths.

9. A Limit Theorem for Narrow Inductive Logics

9.1 The Result in the Probability Calculus

There is a striking limit theorem discussed in the Bayesian literature. (Earman,1992, §4.2, 4.3) Let H be any hypothesis that deductively entails a countable infinity of consequences E_1, E_2, \dots , where we assume that H has non-zero prior probability, $P(H|\Omega) > 0$. We have that

$$\begin{aligned}
P(H \& E_1 \& E_2) &= P(H | E_1 \& E_2) \cdot P(E_1 \& E_2) = P(H | E_1 \& E_2) \cdot P(E_2 \& E_1 | E_1) \cdot P(E_1) \\
&= P(H | E_1 \& E_2) \cdot P(E_2 | E_1) \cdot P(E_1) = P(H)
\end{aligned}$$

since $H \& E_1 \& E_2$ is just H . Repeating this calculation for the first n consequences, we recover

$$P(H) = P(H | E_1 \& \dots \& E_n) \cdot P(E_n | E_1 \& \dots \& E_{n-1}) \cdot P(E_{n-1} | E_1 \& \dots \& E_{n-2}) \cdot \dots \cdot P(E_2 | E_1) \cdot P(E_1)$$

Since $P(H) > 0$, it follows that

$$\lim_{n \rightarrow \infty} P(E_n | E_1 \& \dots \& E_{n-1}) = 1$$

That is, we can become arbitrarily confident of the truth of the next as yet untested consequence of hypothesis H , once we have seen that sufficiently, but finitely, many of its deductive consequences prove true. I have remarked elsewhere (manuscript, §5.7) on the extraordinary immodesty of this result. Jeffrey (1983, p. 194) thought the result so close to “jumping to conclusions” that it enticed him to renounce non-zero prior probabilities on universal hypotheses.

9.2 Its Analog Here

This limit theorem depends more upon the deductive structure of the Boolean algebras than it does on the probability calculus. The particular feature of the deductive structure of the Boolean algebras is shown in the Venn diagrams of Figure 10.

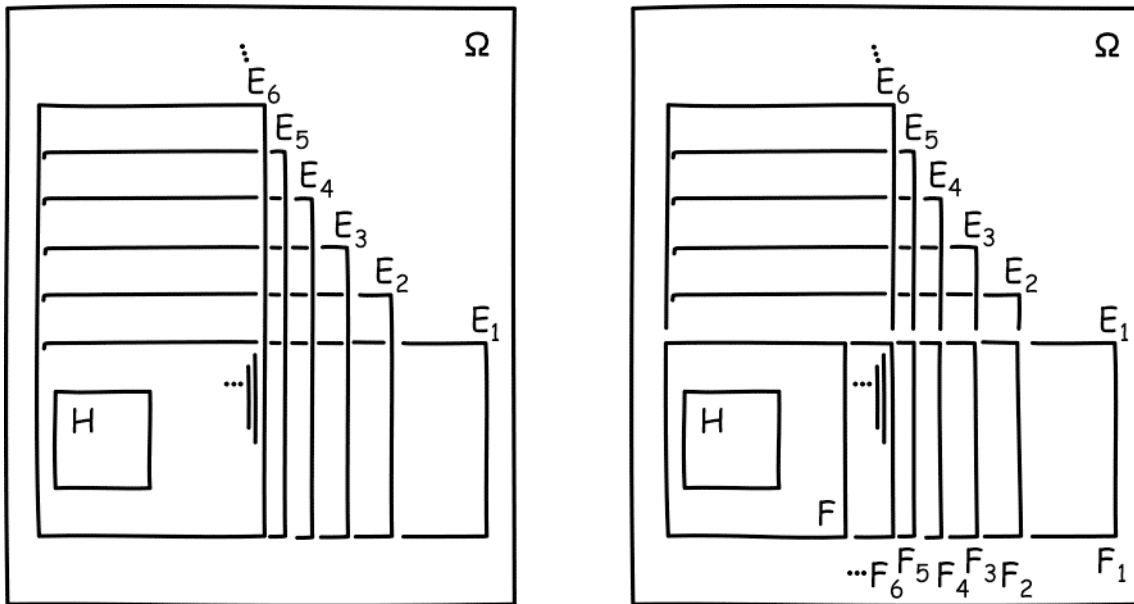


Figure 10. The Sequence of Propositions of the Limit Theorem

It is that the sequence of propositions entailed by $H, F_1 = E_1, F_2 = E_1 \& E_2, F_3 = E_1 \& E_2 \& E_3, \dots$ approaches a limit proposition F that is not \emptyset . The limit theorem arises by reflecting this approach to a limit in the probability measure. It is expressed as the existence of a limiting value of one for $P(F|E_1 \& \dots \& E_{n-1})$ as n grows large. That in turn ensures that the quantity $P(E_1 \& \dots \& E_n | E_1 \& \dots \& E_{n-1})$ also approaches one in the limit. All that is then needed to arrive at the final limit theorem is the fact that the conditional probability $P(A|B)$ ignores $A \& \sim B$, so that $P(A|B)$ is unaffected by the value of $P(A \& \sim B)$. Elsewhere (2007, §5.1), I have called this property “Narrowness”: $P(A|B) = P(A \& B|B)$. Applied here, it asserts that

$$P(E_n | E_1 \& \dots \& E_{n-1}) = P(E_1 \& \dots \& E_n | E_1 \& \dots \& E_{n-1}).$$

We arrive at the more general version of the limit theorem by defining the extension of this notion to more general logics:

Narrowness. $[A|B] = [A \& B|B]$, for any A and B , where the strengths mentioned are defined.

If a more general logic conforms to this general notion of Narrowness, it can also capture this deductive feature of the Boolean algebras in analogous ways and arrive at an analogous limit theorem.¹⁴ Indeed it turns out that, excepting a technicality concerning continuity, this characteristic is all that is needed to arrive at the limit theorem.

Limit theorem for narrow inductive logics.

If (a) an inductive logic that is asymptotically stable and deductively definable in preferred partitions is also narrow;

(b) it admits an unbounded sequence of propositions E_1, E_2, E_3, \dots , in inductively adapted partitions, where the finite partitions are expanded as needed by disjunctive refinement to admit later members of the series, so that, for any particular value of n , there will always be an inductively adapted, finite partition in which the quantities E_1, E_2, \dots, E_n are defined;

(c) E_1, E_2, E_3, \dots are all deductive consequences of H ;

(d) the limiting value of $[H|\Omega]$ under asymptotic stability is not equal to the minimal strength $[\emptyset|\Omega]$;

(e) and the logic is continuous in the sense of Appendix 4;

then $\text{Lim}_{n \rightarrow \infty} [E_n | E_1 \& \dots \& E_{n-1}] = [\Omega|\Omega] = \text{Certainty}$.

The proof is supplied in Appendix 4.

10. Scale-Free Inductive Logics

There are many interesting examples of structures that are scale free. The best known are fractal curves. They “look the same” when viewed normally or when we zoom in and magnify just one portion. Closer to graphs, this idea of self-similarity appears in the literature on random graphs through the notion of “scale free networks” (Newman, 2003, III.1.) The defining characteristic of these networks is that the frequency of nodes of order k is proportional to $k^{-\alpha}$, where α is a constant characteristic of the network.¹⁵

The notion of scale freedom has a natural and interesting application in inductive logic in science. We may take a large scale view and investigate the inductive relations holding between groups of propositions. Our record of astronomical observations provides good evidential support for Newtonian celestial physics. Does it also thereby provide good support for Newtonian terrestrial physics? Or we may change scales and investigate relations between very specific propositions within astronomy. Are this and that observation of the position of Mars on such and such a date good evidential support for this particular orbit of Mars? If the same inductive logic prevails at both scales (and all others) then we have a scale-free inductive logic.

To implement the notion of a scale-free logic the present framework, we seek a logic that uses the same rule while we rescale. We need to find precise expressions of the notions of “rule” and “rescaling.” The notion of rule used so far is the defining of the strengths $[A|B]$ in inductively adapted partitions of Section 7 by the function f in (S) of the symmetry theorem of Section 4. The notion of rescaling employed here is that of disjunctive refinement. An algebra with few atoms is the large scale view described above. A disjunctive refinement of the algebra with many atoms is the small scale view. The inductively adapted disjunctive refinements of Section 7 take us to new partitions in which the rule for computing strengths by means of the function f in (S) still obtains. So scale freedom can be expressed by requiring that the *same* function f is used to compute the strengths $[A|B]$ before and after the refinement. We only slightly restrict the definition, but greatly simplify the analysis, if we presume that the inductively adapted disjunctive refinements are limited to α -fold uniform refinements, as defined in Section 2, for α any natural number.

These considerations yield the following definition:

In a *scale-free inductive logic*,

- (a) the logic admits inductively adapted partitions, such that, for any such partition and any α , there exists¹⁶ α -fold inductively adapted uniform refinements;
- (b) the strengths $[A|B]$ are computed by the same function f of the Symmetry Theorem (S) in every inductively adapted partition;
- (c) the strengths $[A|B]$ computed via (b) are the same in every partition, inductively adapted or not, in which A and B are present; and
- (d) the notion of scale-freedom applies only to strengths computed according to (b) and (c).

Scale-free inductive logics will automatically satisfy the requirement of asymptotic stability under disjunctive refinement. Indeed these logics satisfy the requirement in the strongest sense that they have already settled on the limiting, stable strengths.

These scale-free logics have a special importance: their assignments of inductive strengths are the ones approached asymptotically by other inductive logics. It is natural when choosing an inductive logic to give them precedence over these other inductive logics that are not scale free. For these other logics change their assignments of inductive strength when the algebra of propositions is enlarged. That indicates that, according to their own standards, these logics are still learning from the expansion in the process of settling in on the scale-free values.

In a scale-free inductive logic, the function f of (S) reduces to a simpler form. The reduction is derived by recalling that the strengths $[A|B]$ remain unchanged under an α -fold uniform refinement. That amounts to asserting that, for each pair of propositions A and B present in both partitions,

$$[A|B]_N = [A|B]_{\alpha N}$$

where $[.]_N$ indicates strengths in the original partition with N atoms and $[.]_{\alpha N}$ indicates strengths in the new partition with αN atoms. If the two partitions are inductively adapted, this last equality becomes

$$f_N(x, y, z) = f_{\alpha N}(\alpha x, \alpha y, \alpha z)$$

where the counts of atoms in the original partition have been abbreviated as $x = \#A\&B$, $y = \#A\&\sim B$ and $z = \#\sim A\&B$. For arbitrary N , x , y , z and α , this functional equation has a unique solution:

$$[A|B] = g(x/N, y/N, z/N) = g(\#A\&B/N, \#A\&\sim B/N, \#\sim A\&B/N) \quad (\text{SF})$$

for some function g from triples of rational numbers to inductive strengths, where that function g is characteristic of the particular inductive logic. (For the proof, See *Appendix 6. Proof of the Scale-Free Condition.*)

11 Sample Scale-Free Logics

11.1 Probabilistic Logic

There are very many scale-free logics. Each selection of a function g in (SF) amounts to the selection of a new scale-free logic. The best-known example arises when inductive strengths are probabilities, as in (P) above. Then we have

$$g(\#A\&B/N, \#A\&\sim B/N, \#\sim A\&B/N) = \frac{\#A\&B/N}{\#A\&B/N + \#\sim A\&B/N} = \frac{\#A\&B}{\#B} \quad (\text{P}')$$

In a more familiar guise, this relation is just $P(A|B) = P(A\&B)/P(B)$. What is evident when this familiar result is expressed more fully in (P') is that the conditional probability $P(A|B)$ ignores $A\&\sim B$, so it conforms to the property of "Narrowness" defined in Section 9.2.

This property merits scrutiny. A proposition A can be divided into two disjuncts, $A = (A\&B) \vee (A\&\sim B)$, so that $\#A = \#A\&B + \#A\&\sim B$. Narrowness asserts that the first disjunct only affects the value of $[A|B]$. The second, $\#A\&\sim B$, may be varied at will without any change in $[A|B]$ ensuing. See Figure 11 for a pictorial view of these disjunctive parts and their relation to B .

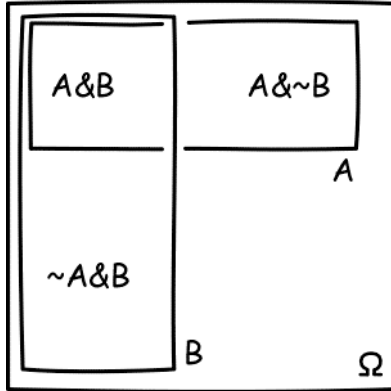


Figure 11 Proposition A in relation to the total evidence $B = B \& \Omega$

Narrowness can produce some curious effects. Imagine that we seek to identify some unknown living being. On the evidence that the being is a bird, we find from Narrowness for conditional probabilities that

$$P(\text{canary v whale} \mid \text{bird}) = P(\text{canary} \mid \text{bird})$$

In one sense, this is a comfortable result. The frequency with which we find “canary v whale” verified in the bird population exactly coincides with the frequency with which we find “canary” verified. However there does seem to be more to be said in discerning the import of the evidence at hand. Indeed the evidence points more specifically to “canary” than “canary v whale”, yet the conditional probability measure is indifferent to the presence of “whale” since that disjunct corresponds the “A & ~B” disjunct mentioned.

11.2 Specific Conditioning

This raises the possibility that we formulate an alternative logic that penalizes a proposition A in so far as it extends beyond the totality of evidence $B = B \& \Omega$. We shall say such a logic implements “specific conditioning” since it assigns a greater inductive strength $[A|B]$ to those A that come closer to the total evidence $B = B \& \Omega$ specifically. Deviations in any direction are penalized. An example arises if we choose a function g in (SF)

$$[A|B] = g(\#A \& B/N, \#A \& \sim B/N, \#\sim A \& B/N) = \frac{(\#A \& B/N)^2}{(\#A/N)(\#B/N)}$$

which reduces to the simpler expression

$$[A|B] = \frac{\#A \& B}{\#B} \cdot \frac{\#A \& B}{\#A} \tag{SC}$$

The two terms of (SC) have simple interpretations:

- The first factor $\#A\&B/\#B$ penalizes A in so far as it fails to exhaust the total evidence $B = B\&\Omega$. If used alone, it would yield the familiar probabilistic rule (P’).
- The second factor $\#A\&B/\#A = \#A\&B/(\#A\&B + \#A\&\sim B)$ penalizes A in so far as A extends beyond the total evidence $B = B\&\Omega$; that is, in so far as $\#A\&\sim B$ is not zero.

The second factor introduces a dependency on $\#A\&\sim B$ in (SC) that violates Narrowness. If proposition A does not extend beyond the total evidence $B = B\&\Omega$, then the inductive strengths $[A|B]$ of (SC) will coincide with the conditional probabilities of (P’).

In this specific conditioning logic, $[A|B] = 1$ just in case $A = B$, whereas that is not the case with the probabilistic conditioning of (P’). To continue the “bird” example, for (P’) we have

$$P(\text{canary } \vee \text{ whale} | \text{canary}) = p(\text{canary} | \text{canary}) = 1$$

However, for the logic (SC) we have

$$[\text{canary } \vee \text{ whale} | \text{canary}] < [\text{canary} | \text{canary}]$$

This reflects (SC)’s assessment that the evidence “canary” more specifically supports just “canary” than it does “canary \vee whale”.

A characteristic property of the logic (SC) is a striking symmetry in the inductive strengths:

$$[A|B] = [B|A]$$

This symmetry is not fortuitous. As shown in *Appendix 7. Sample Scale-Free Logics: Further Properties*, this symmetry is a consequence of a symmetry in the structure of the logic. A symmetric logic such as (SC) penalizes a proposition in equal measures for its failure to exhaust the total evidence $B = B\&\Omega$ and for its extending beyond the total evidence $B = B\&\Omega$. (SC) is not the only logic that exacts these penalties symmetrically; however it is one of the simplest.

This logic does *not* manifest a different symmetry. If A is independent of B so that $[A|B] = [A|\sim B]$, it does *not* follow that B is independent of A; that is, the equality $[B|A] = [B|\sim A]$ may fail.¹⁷

11.3 Comparison to the Probabilistic Account of Incremental Confirmation

Bayesian theory includes incremental measures of inductive support in which evidence is seen specifically to support a proposition. Of the measures most discussed, the ratio measure of

support $r(A,B) = P(A|B)/P(A) = P(A\&B)/P(A).P(B)$ is closest to (SC) and distinctive in that it also obeys the symmetry condition above in that $r(A,B) = r(B,A)$ (Eells and Fitelson, 2002). According to it, proposition B lends less support to a proposition A the more A extends beyond B. Unlike (SC), this ratio measure does not coincide with $P(A|B)$ when A entails B, except trivially when $A = B = \Omega$.

The conception of these incremental measures is very different, however, from the scale-free logics developed here. The probability $P(A|B)$ is conceived as expressing the total support that A receives from all of the evidence; the ratio measure $r(A,B)$ is designed to delineate that part of this total support that comes from B specifically. The logic (SC) asserts that the total support from all the evidence is given by the formula (SC).

Nonetheless, if we define $P(A|B) = \#A\&B/\#B$, then we can rewrite (SC) as

$$[A|B] = P(A|B).P(B|A)$$

Some Bayesians will be unable to resist reading this as the “real meaning” of the logic (SC), which, they may think, has now had its true foundation revealed as essentially probabilistic. I do not see it that way at all. (SC) is an independent logic and an alternative to a probabilistic logic. One might as well note that the conditional probability of (P') can be re-expressed as

$$P(A|B)^2 = [A|B].[A|\Omega]/[B|\Omega]$$

and thereby infer that all probabilistic inductive logic is really at heart an (SC) logic. That inference would be just as mistaken.

11.4 Other Scale-Free Logics

Selection of other functions g in (SF) yield other scale-free logics. Another example is the partial neutrality inductive logic described in *Appendix 7. Sample Scale-Free Logics: Further Properties*. In it, as conditionalization proceeds, there is a gradual migration of strengths of support from superadditive measures, whose lesser values represent a neutral bearing of evidence, to additive measures, whose lesser values represent disconfirmation.

12. Conclusion

The goal of this paper has been to develop precise results concerning induction without committing to the notion that there is a single, universal logic of induction. This goal has been met by exploring common features in a large class of logics of induction. These inductive logics

have been delineated as those that are deductively definable in preferred partitions, asymptotically stable and, sometimes, continuous. These logics include the probability calculus and they admit a notion of a scale-free logic. Two theorems have been proved for them. Both theorems depend essentially on properties of the deductive structure of a Boolean algebra of propositions. The theorems are the reflections of these properties in inductive logics that are appropriately adapted to the deductive structure.

The first theorem asserts that inductive independence is generic. It reflects the fact of deductive structure that in a very large Boolean algebra with N atoms, virtually all propositions are disjunctions of roughly $N/2$ atoms. The second is a limit theorem that pertains to an unbounded sequence of consequences of some hypothesis. The relevant fact of the deductive structure is that conjunctions of these consequences will always remain a deductive consequence of the hypothesis and thus have no fewer atoms, no matter what the partition. That fact, re-expressed in a suitably adapted inductive logic is the limit theorem.

The results developed here do not exhaust those possible in this approach. Another that has not been developed here is the representation of complete neutrality of evidence or, in subjective terms, complete ignorance. I have urged elsewhere (2007a, 2008) that an epistemic state of complete ignorance must be invariant under a transformation in the deductive structure that replaces propositions by their negations. That condition and a presumption of monotonicity of the inductive logic suffice to fix a unique representation.

Appendix 1. Characterizing Finite Boolean Algebras

A Boolean algebra of propositions is a set of propositions closed under the operations of \sim (negation), $\&$ (conjunction) and \vee (disjunction). Implication is stronger than material implication; $A \Rightarrow B$ means that propositions A and B are so related that $\sim A \vee B$ is always true. The universally true proposition is Ω and the always false contradiction is \emptyset .

Our concern will be the deductive relations obtaining between the propositions of the algebras. Therefore we shall discard syntactic duplications of propositions generated by logical equivalence, for such syntactic duplicates enter into the same set of deductive entailment relations. (For example, each of the infinitely many $A \vee A$, $A \vee A \vee A$, ..., $A \& A$, and $A \& A \& A$... etc. are dropped since they are each logically equivalent to A , which is included in the algebra.)

The resulting reduced algebras will be assumed to have finitely many propositions. (Henceforth finiteness will always be understood to pertain to the number of propositions in the algebra after elimination of logically equivalent duplications.) We arrive at an inventory of all such finite algebras by identifying their atoms. An atom a is defined as a deductively strongest, non-contradictory proposition in the algebra; that is, if $A \Rightarrow a$ for any proposition A in the algebra, then $A = \emptyset$ or $A=a$.¹⁸ Each finite algebra must have finitely many atoms. Any two atoms must be pairwise mutually exclusive. That is, if an algebra has N distinct atoms a_1, a_2, \dots, a_N , if we choose any two, a_i and a_k we have that $a_i \& a_k = \emptyset$, for i and k unequal.¹⁹

We can now catalog the finite algebras merely by noting how many atoms there are in each algebra has. An algebra with N atoms, a_1, a_2, \dots, a_N , has a universal proposition

$$\Omega_N = a_1 \vee a_2 \vee \dots \vee a_N$$

This same symbol Ω_N will be used to designate the algebra generated by closing the set of atoms a_1, a_2, \dots, a_N under the Boolean operations \sim, \vee and $\&$, with syntactic duplications discarded. The algebra Ω_1 is the trivial algebra with just two propositions, \emptyset and Ω . The first non-trivial algebra is Ω_2 with atoms a_1 and a_2 .

Appendix 2. Proof of “No-go” Result of Section 6

Start with two mutually exclusive, contingent propositions A and B , where “contingent” means that they are not equivalent to Ω or the contradiction \emptyset . Without loss of generality, we can assume that A and B each have the same number of atoms, so that, from (S), $[A|\Omega] = [B|\Omega]$ and, in this algebra, a symmetry of the deductive structure switches A and B .²⁰ Under continued disjunctive refinement, this equality of $[A|\Omega]$ and $[B|\Omega]$ may fail. Indeed we ought generally to expect that to be the case. Let us assume that we eventually stabilize at limiting values $[A|\Omega] \geq [B|\Omega]$. We have not specified just what limiting notions are needed to arrive at judgments of stability. That will not matter for the argument that follows. Whatever these notions may be, they must admit a sequence of inductive refinements to the Boolean algebras of increasing sizes, N, N', \dots , such that the inductive strengths $[A|\Omega]_N, [B|\Omega]_N, [A|\Omega]_{N'}, [B|\Omega]_{N'}, \dots$ in the algebras indicated approach arbitrarily closely to limiting values for which $[A|\Omega] \geq [B|\Omega]$. However, by an extension of the construction just used, for the above sequence of refinements, there will be a

“clone” sequence of refinements to algebras of size N, N', \dots that lead to the reverse judgment, $[B|\Omega] \geq [A|\Omega]$ in the limiting strengths. The clone sequence is arrived at by first applying the symmetry that switches A and B and then proceeding with the original sequence of disjunctive refinements. It follows that the only relation between $[A|\Omega]$ and $[B|\Omega]$ that is stable under all disjunctive refinement is $[A|\Omega] = [B|\Omega]$.

Propositions A and B are mutually exclusive. It is easy to extend the equality to pairs of propositions that are not mutually exclusive. From the above analysis, we have for contingent A that $[\sim A|\Omega] = [A|\Omega]$ in the limit of arbitrarily many refinements and also $[\sim A|\Omega] = [C|\Omega]$, where C is a proper disjunctive part of A.²¹ Therefore we have $[A|\Omega] = [C|\Omega]$ in the same limit, so that each contingent proposition A has the same limiting inductive support on Ω as any of its disjunctive parts. Finally, let A and B be any two contingent propositions such that $A \& B$ is not \emptyset . We have $[A|\Omega] = [A \& B|\Omega]$ in the same limit since $A \& B$ is either just A or one of its proper disjunctive parts. Similarly, we have $[B|\Omega] = [A \& B|\Omega]$ in the limit. Combining, for any two contingent propositions A and B, we have $[A|\Omega] = [B|\Omega]$ in the limit.

That is, the inductive logic assigns the same support to all contingent propositions in the limit of arbitrarily many disjunctive refinements. I have identified this distribution elsewhere as corresponding to complete ignorance. (Norton, 2008)

The analysis so far has treated the special case of contingent A and B conditioned on the background Ω . The same analysis could be mounted for more complicated cases with similar results. For example, consider mutually exclusive, contingent hypotheses H_1 and H_2 that both entail contingent evidence E (where E does not entail either of H_1 or H_2). The only inductive strengths that can be stable under arbitrary disjunctive refinement are $[H_1|E] = [H_2|E]$. The argument proceeds as before. If there is any sequence of disjunctive refinements that leads us to affirm that $[H_1|E] \geq [H_2|E]$ in the limit, then there will be a “clone” sequence of disjunctive refinements that leads us to affirm $[H_2|E] \geq [H_1|E]$ in the limit. The result can quickly be extended from mutually exclusive H_1 and H_2 to any contingent H_1 and H_2 that entail E by reasoning analogous to that used in the previous case. That is, in the limit of arbitrarily many disjunctive refinements, evidence E confirms equally all contingent hypotheses that deductively entail it.

Collecting these results yields the no-go result.

Appendix 3. Inductive Independence in Large Boolean Algebras

Consider a Boolean algebra Ω_N , with N atoms a_1, \dots, a_N , where N is very large. Propositions are formed by taking the disjunction of n atoms. Since each atom may or may not be selected for inclusion, each proposition is generated by N binary choices, so that there are, overall, 2^N propositions, including Ω and \emptyset . The number of proposition with n atoms is given by the combinatorial factor, $\frac{N!}{n!(N-n)!}$. Therefore the proportion of propositions formed as the disjunction of n atoms is

$$c(n) = \frac{N!}{n!(N-n)!} \cdot \frac{1}{2^N}$$

This corresponds to a binomial distribution with $p=1/2$. By the central limit theory we know that, for large N , this distribution is approximated closely by a normal distribution with mean $N/2$ and variance $\sqrt{N}/2$. We infer immediately that 95.4% of propositions will lie in the interval $n = N/2 \pm \sqrt{N}$, that is, within the two standard deviation interval. This interval grows in proportion to \sqrt{N} as the size N of the algebra increases. Therefore the interval does not grow as fast as N . The interval shrinks if measured as a fraction of the full range of values of n from 0 to N . As N grows, most propositions become more concentrated around the most common value of $n/N = 1/2$. That is, 95.4% of propositions will lie in the interval $n/N = 1/2 \pm 1/\sqrt{N}$. This interval can be made arbitrarily small by selecting a sufficiently large N . This concentration of propositions around $n/N=1/2$ is illustrated in Figure 12, where $c(n)$ is now re-expressed as a function of n/N :

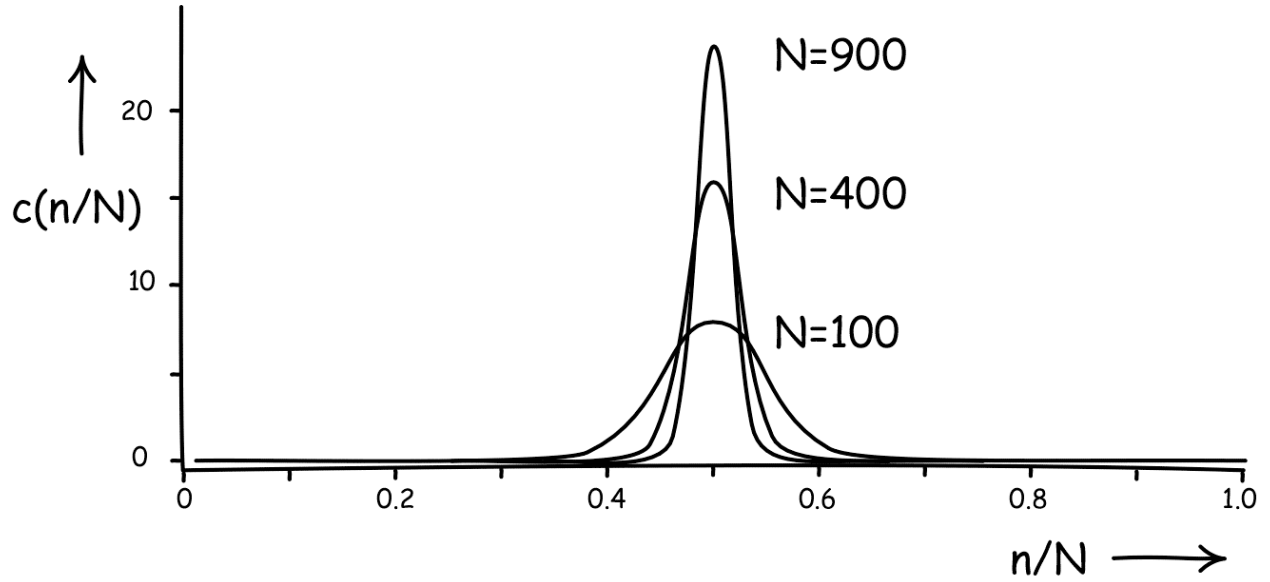


Figure 12. Concentration of Propositions around $n/N = 1/2$

Now consider the distributions of atoms in two propositions A and B. As before, the fraction of propositions A with just n atoms among the 2^N possible propositions is

$$c_A(n) = \frac{N!}{n!(N-n)!} \cdot \frac{1}{2^N}$$

For each proposition A, proposition B will be comprised of a disjunction of m atoms in

proposition A and r atoms not in A. There are $\frac{n!}{m!(n-m)!}$ ways of choosing these m atoms from

the n of proposition A; and there are $\frac{(N-n)!}{r!(N-n-r)!}$ way of choosing r atoms from the $N-n$ atoms

not in A. Therefore the fraction of the 2^N possible propositions for B that have specific values of m and r is given by

$$c_B(m,r) = \frac{n!}{m!(n-m)!} \frac{1}{2^n} \frac{(N-n)!}{r!(N-n-r)!} \cdot \frac{1}{2^{N-n}}$$

As before, $c_A(n)$ is a binomial distribution with $p=1/2$ and N trials. $c_B(m,r)$ is a product of two independent binomial distributions, each with $p=1/2$ and with trials n and $N-n$. Suppressing m and r dependence successively, they can be written as

$$c_B(m) = c_B(m,\cdot) = \frac{n!}{m!(n-m)!} \frac{1}{2^n} \quad c_B(r) = c_B(\cdot,r) = \frac{(N-n)!}{r!(N-n-r)!} \cdot \frac{1}{2^{N-n}}$$

For large N , n and $N-n$, each of these three distributions can be closely approximated by normal distributions, expressed as follows, using the notation “Normal(mean, variance²)”:

$$c_A(n) \sim \text{Normal}(N/2, N/4) \quad c_B(m) \sim \text{Normal}(n/2, n/4) \quad c_B(r) \sim \text{Normal}((N-n)/2, (N-n)/4)$$

Hence, employing three standard deviation confidence intervals, we know that

$$99.7\% \text{ of } n \text{ lie in } N/2 \pm (3/2)\sqrt{N}$$

and, for each value of n ,

$$99.7\% \text{ of } m \text{ lie in } n/2 \pm (3/2)\sqrt{n}$$

$$99.7\% \text{ of } r \text{ lie in } (N-n)/2 \pm (3/2)\sqrt{N-n}$$

Combining by substituting for n and neglecting constant and other terms that become arbitrarily small in relation to \sqrt{N} , we have that

$$99.5\% (=99.7\% \times 99.7\%) \text{ of } m \text{ lie in } N/4 \pm 3(1+\sqrt{2})/4 \cdot \sqrt{N}$$

$$99.5\% (=99.7\% \times 99.7\%) \text{ of } r \text{ lie in } N/4 \pm 3(1+\sqrt{2})/4 \cdot \sqrt{N}$$

That is, most values—here meaning 99.5%-- of n , m and r lie close to the values of $N/2$, $N/4$ and $N/4$ in an interval of size (constant $\cdot \sqrt{N}$). Repeating the analysis for confidence intervals of four, five, etc. standard deviations gives different senses of what portion of values lie in “most values” and these senses come arbitrarily close to 100%, while only adding multiples of \sqrt{N} to the intervals.

We have from Section 8 above that these most frequent values are the ones for which A and B are inductively independent.

$$[A|B] = f_N(\#A\&B, \#A\&\sim B, \#\sim A\&B) = f_N(m, n-m, r) = f_N(N/4, N/4, N/4)$$

$$[A|\sim B] = f_N(\#A\&\sim B, \#A\&B, \#\sim A\&\sim B) = f_N(n-m, m, N-n-r) = f_N(N/4, N/4, N/4)$$

That is, we have inductive independence when the reduced variables take values

$$n/N = 1/2 \quad m/N = 1/4 \quad r/N = 1/4$$

We have from the above that n/N , m/N and r/N for most pairs of propositions differ from these values by a factor in $1/\sqrt{N}$. Therefore, by selecting N large enough, we can bring n/N , m/N and r/N for most pairs of propositions arbitrarily close to these independence values.

Assuming that the assignment of inductive strengths is continuous in the sense of the Appendix that ensues, it follows, that, for most pairs of propositions A and B , $[A|B]$ can be brought arbitrarily close to $[A|\sim B]$ by selecting a large enough N . That is, for any standard of

nearness of strengths we may set, we can select an algebra large enough for most pairs of propositions to be independent to within that standard.

Appendix 4. A Notion of Continuity for Strengths of Inductive Support

The analysis of generic inductive independence among propositions of the previous appendix and the limit result of Section 9 depend essentially on the continuity of the applicable inductive logic.

The essential notion of continuity sought is familiar from the theory of real functions. If the real-valued function $f(x)$ is continuous in its real argument x at $x=X$, we mean this. If x_1, x_2, x_3, \dots is a sequence of values of x that approaches $x=X$ arbitrarily closely, then we can bring the value of $f(x_i)$ arbitrarily close to that of $f(X)$ by proceeding far enough along the sequence x_1, x_2, x_3, \dots . More precisely, the sequence x_i approaches X in the sense that for any $\delta > 0$, there exists a natural number k such that for all $i > k$, $|x_i - X| < \delta$. Then, for any $\varepsilon > 0$, there is a natural number k , such that for all $i > k$, $|f(x_i) - f(X)| < \varepsilon$.

The corresponding notion of continuity in the inductive logic in the simplest case will mimic this definition. In place of the sequence x_1, x_2, x_3, \dots , it has a sequences of pairs of propositions, $(A_1, B_1), (A_2, B_2), (A_3, B_3), \dots$ that approach the limiting propositions (A, B) arbitrarily closely in some measure. While it does this, the corresponding sequence of inductive strengths $[A_1|B_1], [A_2|B_2], [A_3|B_3], \dots$ approaches $[A|B]$ arbitrarily closely.

Since the propositions must all belong to finite algebras, there is no assurance that every otherwise well-behaved sequence of propositions $[A_i|B_i]$ will approach a single limit pair $[A|B]$ in some algebra.²² The more general and useful notion concerns the inductive strengths of two sequences of pairs : $[A_1|B_1], [A_2|B_2], [A_3|B_3], \dots$ and $[C_1|D_1], [C_2|D_2], [C_3|D_3], \dots$ which we will expect to approach one another arbitrarily closely when continuity obtains.

All the Boolean algebras considered here are finite algebras, whereas there must be no upper limit on the numbers of members of these sequences of pairs of propositions. This difficulty can be resolved by recalling that we can form new algebras from old by disjunctive

refinement. If we have some sequence of propositions A_1, A_2, A_3, \dots and B_1, B_2, B_3, \dots that cannot be extended as desired in the present algebra, we can form a disjunctive refinement of that algebra to a larger algebra in which the same sequences appear but can now be extended.

The definition of continuity requires two measures of closeness: of propositions and of strengths. The first is easily identifiable. The sequence of pairs A_i, B_i approaches C_i, D_i in the sense that the atoms comprising A_i and B_i come to agree with those of C_i and D_i arbitrarily closely in the ratio of their atoms to the size of the algebra, N_i , at each stage of the series. That is, as i grows indefinitely large²³

$$\begin{aligned} \#(A_i \& C_i)/N_i &\rightarrow \#C_i/N_i & \#(A_i \& \sim C_i)/N_i &\rightarrow 0 \\ \#(B_i \& D_i)/N_i &\rightarrow \#D_i/N_i & \#(B_i \& \sim D_i)/N_i &\rightarrow 0 \end{aligned}$$

Note that each atom count $\#A$ is tacitly a function of the number of atoms in the algebra and, for one proposition, this count can change as the algebra is disjunctively refined. Thus the sequence consisting of the same proposition A, A, A, \dots may correspond to a changing sequence of atoms $\#_M A, \#_N A, \#_R A, \dots$ as the size of the algebra increases, $\Omega_M, \Omega_N, \Omega_R, \dots$

The notion of closeness of inductive strengths is harder to specify since the value set for the strengths has, in general, been left unspecified here. In the special case in which the inductive strengths are real valued, such as when they are probabilities, then the measure of closeness is just the norm $|[A_i | B_i] - [A | B]|$. If the inductive strengths are, say, real vector valued, we might choose the corresponding vector norm. In general, the notion of continuity requires some appropriate measure of closeness on strengths to be defined and we shall assume that when continuity is asserted some appropriate measure is available.

Combining, we have the following definition

Continuity of inductive strengths of an inductive logic.

Let A_i, B_i and C_i, D_i be any two sequences of pairs of propositions that approach one another in the above defined sense. Then, for any nominated degree of closeness of strengths, there exists a k such for all $i > k$, $[A_i | B_i]$ is within that nominated degree of closeness to $[C_i, D_i]$.

A reasonable notion of continuity depends on the existence of a reasonable measure of closeness.

Analogous notions enable a more precise statement of the property of asymptotic stability of Section 5.2.

Asymptotic stability under disjunctive refinement.

For any pair of propositions A,B, the sequence of strengths $[A|B]_M, [A|B]_N, [A|B]_R, \dots$ will eventually converge to a unique value $[A|B]$ as we proceed through any sequence of unlimited disjunctive refinements $\Omega_M, \Omega_N, \Omega_R, \dots$ in which $[A|B]$ are defined. That is, for any nominated degree of closeness of inductive strengths, there exists an N such that for all $N > M$, $[A|B]_N$ is within that degree of closeness to $[A|B]$.

For the logics introduced in Section 7 that are deductively definable in preferred partitions only, the above definition is restricted to the case of sequences of disjunctive refinements in inductively adapted partitions.

Appendix 5. Proof of the Limit Theorem of Section 9.

Define, as before, $F_1 = E_1, F_2 = E_1 \& E_2, F_3 = E_1 \& E_2 \& E_3, \dots$, for which we have

$$\frac{\#H}{N} = \frac{\#H}{\#F_n} \cdot \frac{\#F_n}{\#F_{n-1}} \cdot \frac{\#F_{n-1}}{\#F_{n-2}} \dots \frac{\#F_2}{\#F_1} \cdot \frac{\#F_1}{N} \quad (R)$$

In inductively adapted partitions, we have from (S) that

$$[H|\Omega] = f_N(\#H\&\Omega, \#H\&-\Omega, \#\sim H\&\Omega) = f_N(\#H, 0, \#\sim H)$$

$$[\emptyset|\Omega] = f_N(\#\emptyset\&\Omega, \#\emptyset\&\sim\Omega, \#\sim\emptyset\&\Omega) = f_N(0, 0, N)$$

Since the limiting value for large N of $[H|\Omega]$ does not equal the minimal $[\emptyset|\Omega]$, it follows that $\#H/N$ approaches a value larger than zero for the limit of large N.²⁴

For all n, $\#F_n \leq \#F_{n-1}, \#F_{n-1} \leq \#F_{n-2}, \dots, \#F_2 \leq \#F_1$. Therefore, as we proceed arbitrarily far along any sequence of the propositions $F_1, F_2, \dots, F_n, \dots$, the ratio $\#F_n/\#F_{n-1}$ must be either less than or equal to one. It may stabilize to the value of unity for all $n > n'$ for some n' . If not, then there are always larger values of n at which the ratio is strictly less than one. Then the size of the algebra hosting the propositions must increase without limit. In that case we have from asymptotic stability that $\#H/N$ will converge to some value greater than zero. Since $\#H \leq \#F_n$, $\#H/N$ can only attain this limit if, recalling (R), we have

$$\text{Lim}_{n \rightarrow \infty} (\#F_n/\#F_{n-1}) = 1$$

Now consider the two sequences of pairs of propositions F_n, F_{n-1} and F_{n-1}, F_{n-1} as n grows without limit. It follows from this limit result that the two approach one another in the ratio of

their atoms as defined in Appendix 4. Therefore, from continuity as defined in Appendix 4, the corresponding strengths converge. That is

$$\text{Lim}_{n \rightarrow \infty} [F_n | F_{n-1}] = \text{Lim}_{n \rightarrow \infty} [F_{n-1} | F_{n-1}]$$

Now, for all values of $n > 1$, $[F_{n-1} | F_{n-1}] = [\Omega | \Omega]$, which is Certainty. By Narrowness we have

$$[F_n | F_{n-1}] = [E_1 \& \dots \& E_n | E_1 \& \dots \& E_{n-1}] = [E_n | E_1 \& \dots \& E_{n-1}]$$

Combining, we infer that

$$\text{Lim}_{n \rightarrow \infty} [E_n | E_1 \& \dots \& E_{n-1}] = \text{Certainty}$$

which is the limit theorem.

What is striking is that the derivation of this result does *not* require Bayes' theorem or Bayes' theorem-like manipulation of conditional dependencies (such as the relation $P(A \& B) = P(A|B).P(B)$). The function that these dependencies served in the probabilistic derivation is here served by the relation (R), which only expresses a fact about the deductive structure of Boolean algebras.

Appendix 6. Proof of the Scale-Free Condition

The proof of (SF) proceeds as follows. Let $h(x, y, z, N)$ be a function of rationals that satisfies

$$f_N(x, y, z) = h(x, y, z, N)$$

for whole number values of x, y, z and N and also satisfies the scaling equation

$$h(\alpha x, \alpha y, \alpha z, \alpha N) = h(x, y, z, N).$$

Select any set of whole number values for x, y, z, N . Since we can choose any whole number value for α , also set $\alpha = N$. In $h(x, y, z, N)$ substitute α for N and multiply the arguments x, z and z by $1 = \alpha/N$. That is, we have

$$h(x, y, z, N) = h(\alpha x/N, \alpha y/N, \alpha z/N, \alpha).$$

By the scaling equation, the right hand side of this equality is

$$h(\alpha x/N, \alpha y/N, \alpha z/N, \alpha) = h(x/N, y/N, z/N, 1)$$

Combining, we have

$$f_N(x, y, z) = h(x/N, y/N, z/N, 1) = g(x/n, y/N, z/N)$$

where the second equality defines the function g introduced in (SF).

Appendix 7. Sample Scale-Free Logics: Further Properties

Scale free logics satisfy the condition (SF) and we implement specific scale-free logics by choosing particular functions g in (SF). Here are further properties of some scale-free logics discussed in the main text.

A7.1. A Specific Conditioning Inductive Logic

The specific conditioning logic (SC) was defined in Section 11.2 above as

$$[A|B] = \frac{\#A \& B}{\#B} \cdot \frac{\#A \& B}{\#A} \quad (\text{SC})$$

Its inductive strengths conditioned on the background are

$$[A|\Omega] = \frac{(\#A \& \Omega)^2}{\#A \cdot N} = \#A/N$$

so that $[A|B]$ can be expressed without mention of atom counts from any specific partition

$$[A|B] = \frac{[A \& B|\Omega]^2}{[A|\Omega] \cdot [B|\Omega]}$$

There is an analog of additivity, which reads, for mutually exclusive A and B

$$[A \vee B|C] \frac{[A \vee B|\Omega]}{[(A \vee B) \& C|\Omega]} = [A|C] \frac{[A|\Omega]}{[A \& C|\Omega]} + [B|C] \frac{[B|\Omega]}{[B \& C|\Omega]}$$

If $C = \Omega$, it reduces to the familiar rule of additivity, $[A \vee B|\Omega] = [A|\Omega] + [B|\Omega]$. Thus the logic (SC) supplies an additive measure $[\cdot|\Omega]$ on the background Ω and a non-additive measure $[\cdot|C]$ under conditionalization on proposition C. Hence the logic is characterized as a “specific conditioning” logic, since it differs from the probabilistic logic of (P’) only in its rule of conditionalization.

A7.2 The Symmetry Condition

In forming the inductive strength $[A|B]$, the logic (SC) penalizes A both for failing to exhaust the total evidence $B = B \& \Omega$ and for extending beyond it. If a logic such as (SC) applies these penalties in equal measures, then a symmetry condition

$$[A|B] = [B|A]$$

is satisfied. To arrive at this condition, note that $\#\sim A \& B$ measures how much A fails to exhaust B; and $\#A \& \sim B$ measures how much A extends beyond B. (See Figure 11 for a pictorial

representation). A precise statement that these two failings are punished (or rewarded) equally is that the function g in (SF) has the same dependence on both $\# \sim A \& B$ and $\# A \& \sim B$. That is, the value of g is unchanged if we switch these two arguments in g :

$$[A|B] = g(\#A\&B/N, \#A\&\sim B/N, \#\sim A\&B/N) = g(\#A\&B/N, \#\sim A\&B/N, \#A\&\sim B/N)$$

It now follows that $[A|B]$ is a symmetric function in A and B . For, replacing B by A and A by B in the last expression, we recover

$$[B|A] = g(\#B\&A/N, \#\sim B\&A/N, \#B\&\sim A/N)$$

Since the conjunction operator “&” is symmetric, we have

$$[B|A] = g(\#A\&B/N, \#A\&\sim B/N, \#\sim A\&B/N) = [A|B]$$

These inferences may be reversed so that the symmetry $[A|B] = [B|A]$ entails that g has the same dependence on both $\# \sim A \& B$ and $\# A \& \sim B$. Therefore, a logic conforms to $[A|B] = [B|A]$ just if it punishes and rewards deviations of A from the total evidence $B = B \& \Omega$ symmetrically.

The logic (SC) is one of many logics satisfying this symmetry condition. It is distinctive in that $[A|B]$ satisfies the additional condition of reducing to the probabilistic expression (P') when A does not extend beyond B , that is, when $\#A\&\sim B = 0$. However it is not the only logic to satisfy both conditions. Both are also satisfied by all logics

$$[A|B] = \frac{\#A\&B}{\#B} \cdot \frac{\#A\&B}{\#A} \cdot h(\#A\&B, \#\sim A\&B, \#A\&\sim B)$$

where $h(x, y, z)$ is symmetric in y and z and satisfies $h(x, 0, 0) = 1$ for all x .

A7.3 A Partial Neutrality Inductive Logic

If larger values of probability represent higher degrees of confirmatory support, then the additivity of probability measures entails that lower probabilities represent disconfirmation. For, a low probability $P(A|B)$, forces a high value of $P(\sim A|B) = 1 - P(A|B)$, meaning that we have high support for $\sim A$. This coupling of $P(A|B)$ and $P(\sim A|B)$ needs to be broken if we are to represent neutral evidence as opposed to disconfirmation. One popular approach (Shafer, 1976) employs superadditive measures in which each of the degrees of belief, $Bel(A)$ and $Bel(\sim A)$, assigned to A and $\sim A$ can be small, so that $Bel(A) + Bel(\sim A) \ll 1$, even though $Bel(A \vee \sim A) = 1$. These lower values are interpreted in the Shafer system as representing ignorance rather than disbelief. They are the subjective analogs of the objective notions of neutrality and disconfirmation employed here.

The following logic is highly superadditive when propositions are conditioned on the background Ω and it transforms towards one with additive degrees as we conditionalize on more specific propositions. Thus the process of conditionalization is accompanied by a further process of replacement of neutrality by disconfirmation. The choice of function g in (SF) is

$$[A|B] = \left(\frac{\#A \& B}{\#B} \right)^{m(B)}$$

where the exponent is defined as

$$m(B) = \frac{1 + \varepsilon}{1 + \varepsilon - \#B/N}$$

The positive parameter $0 < \varepsilon \ll 1$ is presumed to be very close to zero. It follows that

$[B|\Omega] = (\#B/N)^{(1+\varepsilon)/\varepsilon}$, where the exponent $(1+\varepsilon)/\varepsilon$ is very large.

We can render the definition of $[A|B]$ free from mention of any specific partition if we invert this last relation to recover $(\#B/N) = [B|\Omega]^{\varepsilon/(1+\varepsilon)} = [B|\Omega]^{1/m(\Omega)}$. We can now write

$$[A|B] = \left(\frac{[A \& B|\Omega]^{1/m(\Omega)}}{[B|\Omega]^{1/m(\Omega)}} \right)^{m(B)}$$

and also²⁵

$$m(B) = \frac{1 + \varepsilon}{1 + \varepsilon - [B|\Omega]^{\varepsilon/(1+\varepsilon)}}$$

We can quickly see that superadditivity obtains when we condition on Ω :

$$[A|\Omega] = (\#A/N)^{(1+\varepsilon)/\varepsilon} \ll (\#A/N) \quad [\sim A|\Omega] = (\#\sim A/N)^{(1+\varepsilon)/\varepsilon} \ll \#\sim A/N$$

so that

$$[A|\Omega] + [\sim A|\Omega] \ll 1 = [A \vee \sim A|\Omega]$$

since we still have that

$$[A \vee \sim A|\Omega] = (N/N)^{(1+\varepsilon)/\varepsilon} = 1^{(1+\varepsilon)/\varepsilon} = 1$$

If, however, we conditionalize on a proposition B for which $\#B/N \ll 1$, we find that $m(B)$ can be brought arbitrarily close to $(1+\varepsilon)/(1+\varepsilon) = 1$, so that the strengths can be brought arbitrarily close to satisfying the additivity condition

$$[A|B] + [\sim A|B] = 1 = [A \vee \sim A|B].$$

There is a theorem closely analogous to Bayes' theorem. First rewrite the definition of $[A|B]$, substituting H for A and E for B , and then E for A and H for B :

$$[H \& E | \Omega]^{1/m(\Omega)} = [H | E]^{1/m(E)}. [E | \Omega]^{1/m(\Omega)} = [E | H]^{1/m(H)}. [H | \Omega]^{1/m(\Omega)}$$

The analog of Bayes' theorem is then

$$[H | E]^{1/m(E)} = \frac{[E | H]^{1/m(H)}}{[E | \Omega]^{1/m(\Omega)}} \cdot [H | \Omega]^{1/m(\Omega)}$$

Acknowledgements

For helpful discussion, I thank the Visiting Fellows in the Center for Philosophy of Science in April 2008 (Michael Baumgartner, Delphine Chapuis-Schmitz, Richard Dawid, Isabelle Drouet, Mehmet Elgin, Ed Slowik and Derek Turner); the participants in FEW 2009, especially my commentator, David Etlin; Erik Curiel, Balazs Gyenis and James Woodward; and an anonymous referee for thoughtful and constructive recommendations.

Notes

¹ Narrowness is the property that $[H \& E | E] = [H | E]$.

² One can conceive the transformation in two ways. In one, we conceive the atoms as left fixed and the labels permuted over them; in the other, we conceive the labels as fixed and the atoms permuted over them. The difference between the two conceptions depends on the notion of what is left "fixed" with the other is rearranged. As far as the deductive structure is concerned, a decision as to which of label or atom is "fixed" in these transformations is purely conventional.

³ The alternative approach, as explored in Norton (2007), is to add cautiously further properties to the value set until the set has the full structure of the real numbers, but being prepared to halt at intermediate stages as a way of defining weaker inductive logics.

⁴ Proof: Lay out the atoms of $A \& B$, $A \& \sim B$, $\sim A \& B$, $\sim A \& \sim B$ sequentially and label them a_1, \dots, a_N . Correspondingly, lay out the atoms of $A' \& B'$, $A' \& \sim B'$, $\sim A' \& B'$, $\sim A' \& \sim B'$ sequentially and label them a'_1, \dots, a'_N . The permutation sought simply switches a_i and a'_i for $i = 1, \dots, N$.

⁵ One might expect $\# \sim A \& \sim B$ also to appear in the arguments of f . It is superfluous, however, since it is a function of the other arguments: $\# \sim A \& \sim B = N - \# A \& B - \# A \& \sim B - \# \sim A \& B$.

⁶ That is, the rule assigning the degree of support $[A|B]$ for propositions A and B implements a function that depends only on the deductive relations among A, B and the other propositions of the algebra.

⁷ The fullest account of how the notion of instance can be formulated in predicate logic was given by Hempel (1943). The definition is recursive and quite complicated.

⁸ This incompleteness will not affect the demonstration of the no-go result of the next section. That demonstration merely requires that some sense of converge be applicable and is insensitive to how it may be implemented.

⁹ How does the no go result here relate to my material theory of induction (2003), which asserts that inductive inferences are licensed by material postulates that may themselves be part of the algebra of propositions? It shows that if the licensing is to authorize non-trivial inductive relations, it must in addition to the material postulate call upon inductive notions from outside the deductive structure of the algebra. This means that no algebra of propositions can itself authorize the inductive inferences that proceed within it solely in terms of the deductive relations within the algebra.

¹⁰ This requirement of arbitrarily large size does not apply directly to the simple cases of a coin toss or a die throw, since each has a small number of outcomes. It applies to them indirectly, in so far as they arise in larger algebras, such as those able to treat the cases of arbitrarily many coin tosses and die throws.

¹¹ The probability measures defined by (P) are a subset of the probability measures ordinarily employed, for they require equal probability for each atom. If a probability distribution comprises only the assigning rational probabilities to finitely many outcomes, then a refinement will be possible to a larger algebra that gives us this distribution. Otherwise more oblique strategies are needed that mimic the transition from the historical introduction of probabilities as ratios of counts of equally likely cases to the more general additive measures. For example the atom count operator $\#(.)$ in (P) can be replaced by an additive measure $\mu(.)$. A similar replacement of $\#(.)$ by such a measure $\mu(.)$ generalizes the “specific conditioning logic”(SC) and the “partial ignorance logic” defined below.

¹² Beware of the obvious trap. While that one proposition does not enable inductive discrimination among most propositions, it may still allow inductive discrimination among most

interesting propositions! Of course we need further structures to determine just which few of the many propositions are favored as “interesting.”

¹³ The definition of independence is not symmetric in A and B. The equality $[A|B] = [A|\sim B]$ does not in general entail the equality $[B|A] = [B|\sim A]$. We could employ a logically stronger, symmetric definition of independence that requires both equalities. However the weaker definition lets us see that the symmetry of independence is not assured and should be derived in particular logics. It is deducible in the probability calculus and also for the most frequent propositions of large Boolean algebras in inductively adapted partitions. The symmetry fails for the specific conditioning logic of Section 11.2.

¹⁴ A technical point: The inferences must proceed a little more cautiously, however. In the usual development of the probabilistic result, we can be assured that the limit proposition F is in the algebra by locating the derivation in a single, infinite σ -algebra of propositions. In the more general inductive logics, we are assuming only finite algebras of propositions and we cannot be assured that any contain the proposition F. The fact of deductive structure corresponding to the existence of the limit proposition F is this. No matter the partition and how far we have proceed along the sequence, F_1, F_2, F_3, \dots , the atom count $\#F_n$ will always be greater than or equal to $\#H$.

¹⁵ The sense of scale freedom is forced. If we rescale our view of the node order by rescaling the way we count node order—replace k by $k' = c.k$, say—then the frequency distribution remains a power law with the same exponent. There is some size dependence, however, since the constant in the power law will change through $(k')^{-\alpha} = c^{-\alpha} k^{-\alpha}$.

¹⁶ The definition asserts the existence of α -fold inductively adapted uniform refinements, but *not* that all α -fold uniform refinements are inductively adapted. Indeed it is easy to find counter-examples in which not all α -fold uniform refinements are inductively adapted. Consider Ω_2 with atoms a_1 and a_2 ; and the four-fold inductive refinement to Ω_8 where $a_1 = c_1 \vee c_2 \vee c_3 \vee c_4$ and $a_2 = c_5 \vee c_6 \vee c_7 \vee c_8$. Assume both are inductively adapted and admit probabilistic degrees by means of (P), which can be used in a scale-free logic, so that c_1, c_2, c_3 and c_4 are equiprobable. Now consider the two-fold uniform refinement of Ω_2 to Ω_4 in which $a_1 = b_1 \vee b_2$ and $a_2 = b_3 \vee b_4$, where

$b_1 = c_1$ and $b_2 = c_2 \vee c_3 \vee c_4$ and $b_3 = c_5$ and $b_2 = c_6 \vee c_7 \vee c_8$. This two-fold uniform refinement cannot be inductively adapted because b_1 and b_2 cannot be equiprobable.

¹⁷ The failure is illustrated in an example: $\#A \& B = \#\sim A \& B = 4$; $\#A \& \sim B = 6$; $\#\sim A \& \sim B = 12$. Then $[A|B] = [A|\sim B] = [B|A] = 1/5$, but $[B|\sim A] = 1/8$.

¹⁸ There must be at least one such atom. Otherwise, for each proposition A , there exists a logically stronger B such that $B \Rightarrow A$ but not $A \Rightarrow B$. Repeating, we trigger an infinite regress of successively stronger propositions logically and that contradicts the assumed finiteness of the algebra.

¹⁹ This must follow, otherwise we would have a proposition $A = a_i \& a_k$ that entails a_i but is neither a_i nor \emptyset , contradicting the assumption that a_i is an atom. (Of course, if $a_i \& a_k = a_i$, then a_k is not an atom since then $a_i \Rightarrow a_k$.)

²⁰ If we have $\#A > \#B$ in the algebra, we form a new algebra by disjunctively refining the atoms of B until $\#A = \#B$. The case of $\#B > \#A$ is treated analogously.

²¹ That is, $A = C \vee D$, where A is equivalent to neither C nor D individually.

²² To see this, imagine that the ratio $\#A_i/\#B_i$ approaches some irrational number. No pair of proposition can instantiate this ratio if they reside in a finite algebra.

²³ One may be tempted by the alternative $\#(A_i \& C_i) \rightarrow \#C_i$ and $\#(A_i \& \sim C_i) \rightarrow 0$. This alternative fails since the values of $\#(A_i \& C_i)$, etc. are integers. So, if $\#(A_1 \& C_1) = 1$, $\#C_1 = 2$; $\#(A_2 \& C_2) = 9$, $\#C_2 = 10$; $\#(A_3 \& C_3) = 99$, $\#C_3 = 100$; ..., we would not judge that $\#(A_i \& C_i)$ is coming any closer to $\#C_i$, for throughout $\#(A_i \& C_i)$ remains one atom away from $\#C_i$.

²⁴ Otherwise, if $\#H/N$ were to approach zero and, therefore, $\#\sim H/N$ were to approach one in this limit, continuity as defined in Appendix 4 would force the sequences of inductive strengths $[H|\Omega]_M, [H|\Omega]_N, [H|\Omega]_R, \dots$ and $[\emptyset|\Omega]_M, [\emptyset|\Omega]_N, [\emptyset|\Omega]_R, \dots$ to converge as we pass through the infinite sequence of disjunctive refinements $\Omega_M, \Omega_N, \Omega_R, \dots$. That would contradict the requirement that the asymptotically stable value of $[H|\Omega]$ is not equal to the minimal strength $[\emptyset|\Omega]$.

²⁵ The parameter ϵ exercises a controlling influence here directly and indirectly through $[B|\Omega]$ If, for example, B is one of the members of an m-fold partition of Ω in into equally well supported parts, then $[B|\Omega] = (1/m)^{(1+\epsilon)/\epsilon}$.

References

- Earman, J. (1992). *Bayes or Bust*. (Cambridge, MA: Bradford-MIT.)
- Eells, E. & Fitelson, B. *Symmetries and Asymmetries in Evidential Support* (2002).
Philosophical Studies, 107, 129-142.
- Hempel, C. G. (1943). A Purely Syntactic Definition of Confirmation. *Journal of Symbolic Logic*, 8, 122-43.
- Hempel, C. G. (1945). *Studies in the Logic of Confirmation*. *Mind*, 54, 1-26, 97-121 (revised as Ch.1 in *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press, 1965).
- Jeffrey, R. (1983). *The Logic of Decision*. (2nd ed. Chicago: University of Chicago Press.)
- Kelly, K. (1996). *The Logic of Reliable Inquiry*. (Oxford: Oxford University Press.)
- Newman, M. E. J. (2003) The structure and function of complex networks. *SIAM Review*, 45(2), 167- 256.
- Norton, J. D. (2003). A Material Theory of Induction. *Philosophy of Science*, 70, 647-70.
- Norton, J. D. (2005). A Little Survey of Induction. (In P. Achinstein (Ed) *Scientific Evidence: Philosophical Theories and Applications*. (pp. 9-34) Johns Hopkins University Press.)
- Norton, J. D. (2007). Probability Disassembled. *British Journal for the Philosophy of Science*, 58, 141-171.
- Norton, J. D. (2007a). Disbelief as the Dual of Belief. *International Studies in the Philosophy of Science*, 21, 231-252.
- Norton, J. D. (2008). Ignorance and Indifference. *Philosophy of Science*, 75, 45–68.
- Norton, J. D. (manuscript). “Challenges to Bayesian Confirmation Theory,” Prepared for P. S. Bandyopadhyay and M. Forster (Eds.), *Philosophy of Statistics: Vol. 7 Handbook of the Philosophy of Science*. (Elsevier.)
- Shafer, G. (1976). *A Mathematical Theory of Evidence*. (Princeton: Princeton Univ. Press.)