Bruno de Finetti

The “logic of plausible inference” according to Polya (La “logica del plausibile” secondo la concezione di Polya), presented in 1949 and published in 1951, in Atti della XLII Riunione, Società Italiana per il Progresso delle Scienze, pages 227-236.

English translation by Paul Snow, who gratefully acknowledges, without implicating, generous discussions with Marco Ramoni about some points. This translation copyright 2012, all rights reserved by the translator. The intellectual property status of the underlying work is unknown.

Page breaks in this translation follow closely the pagination of the original work. The proceedings pages were “double numbered,” both 227-236 (at the bottom) and 2-10 (at the top, beginning with the second page). This sometimes causes confusion in citing the work. What follows is numbered at the top, 2-10, beginning after the initial page below. Figures are images of the originals. De Finetti’s use of “>” for “greater than or equal to” is preserved throughout. However, his use of “+” for logical disjunction has been translated to “or.”
The “Logic of Plausible Inference” According to Polya

Dr. Prof. Bruno de Finetti

Abstract: Polya has recently discussed the need for and characteristics of a proposed “logic of plausible inference.” The current author relates this problem to his own axiomatic foundations for the probability calculus based on qualitative comparisons, showing that (except for some still open difficulties) it is enough to omit one of these axioms to obtain a formulation responsive to Polya’s desiderata and constitute an intermediate step between the logic of implication and the (quantitative) theory of probability.

1. Polya’s Concept

In a recent issue of *Dialectica* devoted to the foundations of probability theory [Vol. 3, n. 9-10, 1949], an article by G. Polya, “Preliminary Remarks on a Logic of Plausible Inference,” contained interesting examples and discussion to enliven the arguments between the two fundamentally opposite concepts of probability theory which contest the territory: the one which is limited to a “theory of random mass phenomena” and the other which includes a “logic of plausible inference.”

Not just taking part in the dispute, Polya busies himself, in the cited article [also in two others cited there, but which I was unable to consult: Heuristic reasoning and the theory of probability, *Amer. Math. Monthly* 48, 1941; On patterns of plausible inference, *Studies and Essays Presented to R. Courant*, 1948] to illustrate with apt examples the need for consideration related to the “logic of plausible inference” in many varied fields, and delineates certain attributes which such a theory ought to have, whether or not identified with probability theory.

It would have: (1) to be broad enough to include the applications presented in the examples, particularly those related to the plausibility of mathematical conclusions, (2) to include the “heuristic syllogism,” and (3) to be limited to qualitative appraisal, not translatable to the attribution of a numerical value to the “plausibility.”
2. On the “Qualitative” Formulation

A qualitative formulation should be based, rather than upon a numerical function $P(A)$, the probability of $A$, defined on a field of propositions (or statements or events in the more usual terminology of probability), and subject to certain restrictions (“axioms” of probability), instead upon the consideration of simple inequalities of plausibility, of the type $A > B$, which says that the statement $A$ is “more plausible” than $B$, subject to some restrictions imposed by the “axioms” of the relation “$>$”.

A formulation of this sort is found in my paper *Sul significato soggettivo della probabilità* in *Fundamenta Mathematicae* [vol. XVII, Warsaw, 1931], but which proceeds from a qualitative foundation to achieve the ordinary quantitative formulation. I propose to analyze in the current communication whether that really, or only apparently, differs from Polya’s formulation. The importance of the question lies in the alternative it poses: the proposed “logic of plausible inference” would in the first case conflict with probability theory and not be reducible to it; in the second case, it would be consistent with it, but would be a “weaker” version. In axiomatic terms, it would demand either different axioms or else weaker axioms, and so would stand to probability theory as Euclidean geometry stands to, respectively, non-Euclidean geometry or else projective geometry.

The axioms of the aforementioned formulation (glossing over the details) were the following: (1) a certainly true (respectively, impossible) statement is more (less) probable than any other, (2) the transitivity property, (3) an “additive” property in the sense that if $A$ and $B$ each comprise two incompatible events, $A = A'$ or $A''$ and $B = B'$ or $B''$, then if $A' > B'$ and $A'' > B''$, then $A > B$, (4) any two events are always comparable in probability (and therefore always $A > B$ or $B > A$; to include in a fastidious sense the “equal probability” case, it is convenient to interpret $A > B$ as $A$ is *not less* probable than $B$, so $A > A$, and if both $A > B$ and $B > A$, then $A$ and $B$ are of equal probability.)

The first two properties are trivial and the third expresses, I think, a peculiar requirement which a logic of probable inference or of plausible inference needs to have; it is difficult to be able to have interest in a prospective theory which lacks it. The last may very well come to be omitted, not in a sense which asserts the existence of pairs of events $A$ and $B$ which are incomparable in principle, but rather in the sense that we are prevented from drawing plausible conclusions by reasoning in a situation where the sense of the inequality is “known” or “has been established” for only some pairs and not for all (and
we specify “known” or “has been established” in order to include without distinction every objective or subjective meaning of such a “comparison of plausibility”).

3. On the Passage to a “Quantitative” Formulation

In the cited work, the passage to a quantitative formulation relied upon the four axioms already discussed and two additional ones admitted implicitly: (a) the applicability of these axioms subtends not only a specific predetermined field \( C \) of events, but also any larger field \( C' \), and (b) \( C' \) is imagined always to be able to include a set of “incompatible equally probable cases” of number \( n \) however large (for example, arranged in conditions of equal probability between the numbers 1, 2, 3, ..., \( n - 1, n \)). In such circumstances, it turns out to be possible to introduce the quantity \( \frac{m}{n} \) as a measure of probability in cases corresponding to the classical definition, and therefore to determine by comparison the numerical value of any probability (necessarily satisfying the theorem of total probability).

If, holding on to the first three axioms, we can see from the fourth and the two additional admissions just discussed (heeding the counsel of caution no more than on the previous occasion), that knowing the relation “>” will not be sufficient to determine a value for the probability. But the earlier doubt remains, and comes not so much in a vague form (depending on the choice of axioms), but like an underdetermined mathematical problem. Precisely: given a system of events \( A_1, A_2, ..., A_n \), we know that however we assign values \( p_1, p_2, ..., p_n \) which are admissible as probabilities, and taking \( A_h > A_k \) whenever \( p_h > p_k \), we thus obtain a system of inequalities satisfying those axioms. The question is the converse: if every system of inequalities which satisfies such axioms can be obtained in this way (with appropriate choice of the numbers \( p_h \)).

I limit myself to thus setting up the problem, but observe that if the answer were negative, there would still remain the choice between two attitudes: to accept the solutions compatible with the axioms even if they are inconsistent with any possible valuation, or alternatively to add new axioms in order to make up for the effects of the intentional light specification in which the formulation comes conceived. For my part, I am inclined toward the second alternative (wishing to eliminate from the solutions the presence of a latent contradiction with the axioms that is revealed when the field \( C \) is enlarged).

4. Plausibility Structures

We see to understand concretely, by taking particular care with the simpler cases, with what meaning and with what degree of arbitrariness
it is possible to establish in a given field of events \( C \) a plausibility structure \( S \), based on the relation “\( > \)” which satisfies postulates (1), (2), and (3). If there are \( n \) events, then there are \( n(n-1)/2 \) distinct pairs of events, and for each pair we suppose either to choose the sense of the inequality or not fix it at all. We thus have a great number of structures, of which only those satisfying the axioms are of interest, and these will be total or partial structures, depending upon whether the sense of the inequality is fixed for all pairs or only for some, and complete or incomplete (in the field \( C \) being considered) depending on whether all the inequalities are deducible (by means of the axioms) from those data, or not.

Two structures (for simplicity we refer here to a complete structure) will be called compatible if their sum (the structure formed from all of the inequalities of the one and of the other) is also a plausibility structure. We say that a structure is a substructure of another, or that it is weaker than another, if it consists of a part of the inequalities of the other.

It is instructive to observe that in every case, by axioms (2) and (3), \( A > B \) when \( A \) is a consequence of \( B \). Therefore, an implication structure is always a substructure of another structure. In other words: any plausibility structure appears as an elaboration of a purely logical structure of implications; only when \( A \) implies \( B \) can one say logically that that \( B \) is “no less plausible” than \( A \) - in every other case, one must introduce some additional basis other than the purely logical for the valuation.

If, on the other hand, the doubt of part 3 resolves affirmatively (as it seems it should), a plausibility structure will always be coincident with (or be contained in) some correspondence to a possible valuation of a numerical probability with a degree of arbitrariness which will be made clear in section 5). It would seem, and that was what we wanted to call attention to earlier, that a plausibility structure can be interpreted as an intermediate step between the implication structure in which comparisons are limited to the case of two events, one of which implies the other, and that of a quantitative theory in which, thanks to the numerical appraisal, the comparisons are totally specified.

5. Probability Structures

A plausibility structure will be said to be a probability structure when it corresponds to a possible valuation of numerical probability (and the supposition is that any plausibility structure is a probability structure). I exclude always the limiting case of equality, and deal
with delimiting the region of probability appraisals which satisfy the desired inequalities which constitute a structure.

We consider the field of events \( C \) formed from \( s \) “possible cases” or “constituents” (that is, events of which one and only one is true). There are thus \( 2^s \) events (disjunctions of the constituents; \( 2^s - 2 \) if one wishes to exclude the certainly true and the impossible events, which are the disjunctions, respectively, of all or none of the constituents). A probability distribution on \( C \) is defined by assigning the \( s \) probabilities \( x_1, x_2, \ldots, x_s \) to the constituents (each number non-negative and all summing to unity) and corresponds geometrically to a point in a simplex of \( s - 1 \) dimensions with vertices \( P_1, P_2, \ldots, P_s \) (precisely, the point \( P = \sum x_h P_h \) in barycentric coordinates \( x_h \)).

Every inequality between the probabilities of two events in \( C \) translates into an inequality between two sums of the elementary probabilities \( x_h \), leading to an inequality of the form \( a_1 x_1 + a_2 x_2 + \ldots + a_s x_s > 0 \) with coefficients \( a_h \) of 1, -1, or 0 (among which there is at least one 1 and at least one -1). Of such inequalities there are \( \frac{1}{2} (3^s - 1) - 2^s \); thus defining the hyperplane which divides the points representative of satisfying and non-satisfying distributions for the inequalities of the type considered.

A probability structure corresponds to the region belonging to some number of half-spaces thus determined (completely by the total structures, incompletely by the partial structures).

6. Examples: the Cases \( s = 2, 3, 4 \)

The case of \( s = 2 \) is altogether trivial: the two possible cases are \( A \) and \( B = \) not \( A \), and only two structures are possible (\( A \) more or less probable than \( B \)), corresponding to the division of a line into two segments of length \( x \) and \( 1-x \).

The case of \( s = 3 \) is also simple, but somewhat instructive. In the triangle \( ABC \) any point (with barycentric coordinates \( x_1, x_2, x_3 \)) represents a probability distribution among the three possible events (see figure 1), the network of lines which join the centers of the sides or the centers with the
vertices separates the regions corresponding to the inequalities (there are six such lines, \(6 = 1/2 \left( 3^2 + 1 \right) - 2^3\)), the 12 triangular regions so separated correspond to the possible probability structures (total structures: the partial ones can be depicted as various groupings of the regions just mentioned). Note that the three lines through the vertices correspond to inequalities between individual constituents, and these separate out \(6 = 3!\) regions corresponding to the permutations of \(A, B, C\) according to increasing probability; there are then two further cases for each choice of which the constituent is the most probable among the three events and more or less probable than the sum of the other two (in other words, has probability more or less than \(1/2\)), and so altogether there are \(2 \cdot 3! = 12\) cases.

For \(s = 3\) (as is obvious for \(s = 2\)), one can check that these are the only possible plausibility structures, and the hypothesis turns out to be verified. In fact, suppose that \(A-B-C\) are in increasing order of plausibility, then from the axioms so are \(AB-AC-BC\) (for ease of writing, we write \(AB\) for \(A\) or \(B\), and the hyphen indicates increasing order) and \(A-B-AC-BC\). The only thing in doubt is \(C\) and \(AB\), which come between \(B\) and \(AC\), should be arranged in the order \(C-AB\) or \(AB-C\), and so there are two possible orderings: (1) \(A-B-C-AB-AC-BC\), or (2) \(A-B-AB-C-AC-BC\) (and therefore there are \(2 \cdot 3! = 12\) structures found, just as with the probability structures).

Passing to \(s = 4\), the treatment already becomes somewhat tedious, but it is not difficult to exhaustively enumerate and to verify that the same conclusion holds. Geometrically, we consider a tetrahedron (see figure 2) \(ABCD\); each interior point represents by means of its barycentric coordinates \(x_1, \ldots, x_4\) a probability distribution on the four possible cases \(A, B, C,\) and \(D\). There are \(25 = 1/2 \left( 3^4 + 1 \right) - 2^4\) inequalities which correspond to planes passing (a) through the center of one edge and the edge
opposite (there are 6 edges), (b) through the centers of two intersecting edges and
the remaining vertex (there are 12 pairs of intersecting edges), (c) through the
centers of three edges which share one vertex (there are 4 vertices), and (d) through
the centers of four edges which are pair-wise intersecting (there are 3 separate pairs
of edges). One can say generally “planes among three edge-centers and vertices,”
but the division into cases can also be described and seen to come to 25 (6 + 12 + 4
+ 3), to aid intuition and visualization, as corresponding to the different kinds of
inequalities: (a) between single events, like A and B, (b) and (c) between single
events and respectively, the disjunction of two or three other cases, like A and BC
or BCD, (d) between disjunctions of two events, like AB and CD (as before, we
write the disjunctions tersely, omitting the operator +: AB for A or B, etc.).

The total number of structures, which I compute as shall be explained in a
moment, turns out to be 14 ⋅ 4! = 336, because 4! are the permutations, and we
suppose that the succession AB-BC-CD in increasing order of plausibility, and 14 is
the number of structures compatible with each permutation (and essentially
realized like the probability structures, and similarly checked with numerical
examples).

Considering the disjunctions of the events two by two, and looking at how
they will follow necessarily in a given ordering, except for the possibility of
reversal in the order of AD and CD: after all, it is either (1) AB-AC-AD-BC-BD-CD
or else (2) AB-AC-BC-AD-BD-CD. Then there is the matter of how to insert A-B-
C-D individually and “1/2” to obtain the whole structure. Let us observe first of all
that “probability 1/2” is not being introduced as an illicit quantitative notion: rather
it reflects the qualitative decision of whether the event is more or less plausible
than its opposite (so we might say that “1/2” occurs between AD and BC, which
are complementary events). After that, all that remains is to observe that the only
kind of event which has not explicitly been placed into increasing order of A, B, ...
and AB, AC, ... is that involving three events, for example ABC, but the only
relevant comparison with ABC is the one with D, the complement of ABC,
therefore the significance is none other than the comparison of D with “1/2”.

There turn out to be 5 ways to insert A-B-C-D and “1/2” into (1) and 9 ways
for (2): we show them on the next page, indicating off to the side the probabilities
(in %) of A, B, C, D that realize each such ordering, thus assuring that all of the
structures which are found are probability structures.
1) a. $A-B-C-D-AB-AC-AD-1/2-BC-BD-CD$  
   (19, 26, 27, 28)

b. $A-B-C-AB-D-AC-AD-1/2-BC-BD-CD$  
   (10, 21, 30, 39)

c. $A-B-C-AB-AC-D-AD-1/2-BC-BD-CD$  
   (10, 25, 26, 39)

d. $A-AB-C-D-AC-AD-1/2-BC-BD-CD$  
   (10, 20, 31, 39)

e. $A-AB-C-AC-D-AD-1/2-BC-BD-CD$  
   (8, 20, 31, 41)

2) a. $A-B-C-D-AB-AC-BC-1/2-AD-BD-CD$  
   (20, 24, 25, 31)

b. $A-B-C-AB-D-AC-BC-1/2-AD-BD-CD$  
   (16, 17, 32, 35)

c. $A-B-C-AB-AC-BC-1/2-AD-BD-CD$  
   (10, 24, 25, 41)

d. $A-B-AB-C-BC-D-1/2-AD-BD-CD$  
   (10, 20, 21, 49)

e. $A-B-AB-C-AC-BC-1/2-D-AD-BD-CD$  
   (8, 20, 21, 51)

f. $A-AB-C-D-AC-BC-1/2-AD-BD-CD$  
   (10, 13, 36, 41)

g. $A-AB-C-AC-D-BC-1/2-AD-BD-CD$  
   (10, 19, 30, 41)

h. $A-AB-C-AC-BC-D-1/2-AD-BD-CD$  
   (10, 11, 30, 49)

i. $A-AB-C-AC-BC-1/2-D-AD-BD-CD$  
   (8, 12, 29, 51)

7. The General Case

As $s$ grows, not only does the problem come to lack any immediate representation in three-dimensional space, but the degree of complication increases breathtakingly. For $s = 5$ there are 90 inequalities and surely for each permutation there are more than a thousand and perhaps several thousand (I did not finish a computation); in all, therefore hundreds of thousands or millions. For $s = 6, 7, 8, 9, 10, ...$ already the inequalities increase to 301, 966, 3025, 9332, 28501, ..., the determination of the rest of these numbers doesn’t
hold interest beyond arithmetical curiosity: clearly we note that the direct verification surely cannot be carried out except in the first trivial cases, and what is needed is some general demonstration (by recursion? something recoverable from an analysis of the axioms?) which is not necessarily difficult but which I have not yet found.

The same kinds of study, what is the qualitative interpretation of the quantitative one, can be interpreted and be interesting from various points of view (and the issue of the reconcilability of one point of view to the other, which is the fundamental problem here, has interest for all these cases).

From the qualitative point of view, beyond that of the theory of plausible inference, the problem may be interesting for the general theory of ordered structures and for a qualitative “measure” theory. In the first case, one might study how the introduction of a restriction like “additivity” might influence the issues of ordering. In the second, one might try to apply a condition of ordering between some sets based on a comparison (“greater” “smaller”) leading to a “measurability” in a qualitative sense (and so, eventually, a quantitative one).

From the quantitative point of view, beyond that for probability theory, the problem can hold the same interest either for the geometric interpretation, or for the corresponding arithmetic interpretation, and that would consist in the examination of the various ways in which the $2^s-2$ numbers formed from $s$ given (positive) numbers can be summed and arranged in order of magnitude, and also their sums two to two, three to three, ..., $s -1$ to $s -1$.

8. Plausibility and Probability

To summarize and conclude: there is a technical problem (the sufficiency of certain axioms) that is stated which arises in relation to the “logic of plausible inference” of Polya. Whatever the answer is, I think that (apart from the completeness of the axioms) the logic of plausible inference should coincide with the logic of probable inference (the calculus of probability), subject to the limitation of considering some qualitative conclusions which are obtainable from the exact numerical determination of the probabilities based on noting the inequalities between their values.

Except for such specification, which can doubtless succeed to be explained explicitly, the theory of plausibility coincides perfectly with probability theory conceived from the subjective point of view, and satisfies the requirements set by Polya (see citation of the *Dialectica* article) because:

1) it is applicable to whatever examples;
2) it includes the “heuristic syllogism,” which reduces to “Bayes’ theorem” (or eventually, in a weaker form, to some translation of it within the qualitative formulation);

3) it does not require (when restricted to postulates 1-3) a quantitative evaluation.

As for the other conception of the probability calculus as a “theory of random mass phenomena,” all the arguments which I have carried out tend to deny the legitimacy of every supposed autonomous construction with such understandings and to prove the perfect futility of any effort in such sense, even insofar as statements concerning frequencies and similar statistical applications enter in the general case statements of whatever nature is applicable the “logic of plausible inference” or “of probable inference” (I take these as synonymous). And such application of the general theory to the specific case not only responds to the requirements in full, but it is essential therefore that just the attempts to eliminate it in order to construct supposedly autonomous theories are responsible for the flourishing of those juniper bushes of pseudoproblems for which the idea of probability has drawn the undesirable and undeserved reputation for being dark and uncertain.