Epistemology Probabilized

Richard Jeffrey
Department of Philosophy
Princeton University
Princeton, NJ 08544
USA
dickjeff@princeton.edu

Abstract

Here is a framework for judgment in terms of a continuum of "subjective" probabilities, a framework in which probabilistic judgments need not stand on a foundation of certainties. In place of propositional data bases, this *radical* probabilism ("probabilities all the way down to the roots") envisages full or partial probability assignments to probability spaces, together with protocols for revising those assignments and their interconnections in the light of fresh empirical or logico-mathematical input. This input need not be of the limiting 0-or-1 sort. Updating by ordinary conditioning is generalized (sec. **2.2**) to *probability kinematics, where an observation on a random variable X need not single out one value, but may prompt a new probability distribution Q over all values of X.*

The effect of an observation itself, apart from the influence of prior probabilities (sec. **3**), is given by the ("Bayes") factors $\frac{\textit{new odds}}{\textit{old odds}}$ by which the observer's odds between hypotheses are updated. We are not generally interested in adopting an observer's new odds as our own, for those are influenced by the observer's old odds, not ours. It is rather the observer's Bayes's factors that we need in order to use that observation in our own judgments. An account of collaborative updating is presented in these terms.

Jon Dorling's bayesian solution of the Duhem-Quine "holism" problem is sketched in sec. **4**.

We finish with a brief look at the historical setting of radical probabilism (sec. **5**), and an indication of how "real" probabilities can be accomodated in subjectivistic terms (sec. **6**).

(Apendix.) "Interactive epistemology" is generally predicated on updating by conditioning. In sec. **7**, Aumann's Agreement Theorem is extended to cases where updating is by generalized conditioning, and beyond—the only restriction being Michael Goldstein's condition of temporal coherence.

1 Judgmental ("Subjective") Probability

Your "subjective" probability is not something fetched out of the sky on a whim; it is your actual judgment, normally representing what you think your judgment *should* be, even if you do not regard it as a judgment that everyone must share on pain of being wrong in one sense or another.

1.1 Probabilities from statistics: Minimalism

Where do probabilistic judgments come from? Statistical data are a prime source; that is the truth in frequentism. But that truth must be understood in the light of certain features of judgmental probabilizing, e.g., persistence, as you learn the relative frequency of truths in a sequence of propositions, of your judgment that they all have the same probability. That is an application of the following theorem of the probability calculus.[1]

> **Law of Little Numbers.** In a finite sequence of propositions that you view as equiprobable, if you are sure that the relative frequency of truths is p, then your probability for each is p.

Then if, judging a sequence of propositions to be equiprobable, you learn the relative frequency of truths *in a way that does not change your judgment of equiprobability,* your probability for each proposition will agree with the relative frequency.[2]

The law of little numbers can be generalized to random variables:

> **Law of Short Run Averages.** In a finite sequence of magnitudes for which your expections are equal, if you know only their arithmetical mean, then that is your expectation of each.

Then if, while requiring your final expectations for a sequence of magnitudes to be equal, you learn their mean value in a way that does not lead you to change that requirement, your expectation of each will agree with that mean.[3]

[1] See Jeffrey (1992) pp. 59-64. The name "Law of Little Numbers" is a joke, but I know of no generally understood name for the theorem. That theorem, like the next (the "Law of Short Run", another joke) is quite trivial; both are immediate consequences of the linearity of the expectation operator. Chapter 2 of de Fineti (1937) is devoted to them. In chapter 3 he goes on to a mathematically deeper way of understanding the truth in frequentism, in terms of "exchangeability" of random variables (sec. **1.3**, below).

[2] To appreciate the importance of the italicized caveat, note that if you learn the relative frequency of truths by learning which propositions in the sequence are true, and which false, and as you form your probabilities for those propositions you remember what you have learned, then those probabilities will be zeros and ones instead of average of those zeros and ones.

[3] If you learn the individual values and calculate the mean as their average without forgetting the various values, you have violated the caveat (unless it happens that all the values were the same), for what you learned will have shown you that they are not equal.

Example: Guessing Weight. Needing to estimate the weight of someone on the other side of a chain link fence, you select ten people on your side whom you estimate to have the same weight as that eleventh, persuade them to congregate on a platform scale, and read their total weight. If the scale reads 1080 lb., your estimate of the eleventh person's weight will be 108 lb.—if nothing in that process has made you revise your judgment that the eleven weights are equal.[4]

This is a frequentism in which judgmental probabilities are seen as judgmental *expectations* of frequencies, and in which the Law of Little Numbers guides the recycling of observed frequencies as probabilities of unobserved instances. It is to be distinguished both from the intelligible but untenable finite frequentism that simply identifies probabilities with actual frequencies (generally, unknown) when there are only finitely many instances overall, and from the unintellible long–run frequentism that would see the observed instances as a finite fragment of an infinite sequence in which the infinitely long run inflates expectations into certainties that sweep judgmental probabilities under the endless carpet.[5]

1.2 Probabilities from statistics: Exchangeability[6]

On the hypotheses of (a) equiprobability and (b) certainty that the relative frequency of truths is r, the the Law of Little Numbers identified the probability as $=r$. Stronger conclusions follow from the stronger hypothesis of EXCHANGEABILITY:

> You regard a set of propositions as exchangeable when, for any two disjoint subsets, your probability that those in the first are all true and those in the second are all false depends only on the sizes, t, f, of the two sets.[7]

Here, again, as in sec. **1.1**, probabilities will be seen to come from statistics—but, again, only under probabilistic hypotheses.

In the presence of exchangeability of a set $\{H_1, \ldots H_n\}$, your probabilities for all $2^{(2^n)}$ of their Boolean compounds are determined by your probabilities for the following $n+1$ of those compounds:

$$H^n_t =_{df} \text{the number of truths among } H_1, \ldots, H_n \text{ is } t \ (= 0, \ldots, n).$$

[4]Note that turning statistics into probabilities or expectations in this way requires neither conditioning nor Bayes's theorem, nor does it require you to have formed particular judgmental probabilities for the propositions or particular estimates for the random variables prior to learning the relative frequency or mean.

[5]See Jeffrey (1992) chapter 11.

[6]See chapters 3 and 5 of Finetti (1937), (1980).

[7]This comes to the same thing as invariance of your probabilities for Boolean compounds of finite numbers of the H_i under all finite permutations of the positive integers, e.g., $P(H_1 \wedge (H_2 \vee \neg H_3)) = P(H_{100} \wedge (H_2 \vee \neg H_7))$.

DE FINETTI'S RULE OF SUCCESSION.[8]

> If you regard $\{H_1, \ldots, H_{n+1}\}$ as exchangeable, your probability for truth of H_{n+1}, given a conjunction ("H") of a particular t of them with the denials of the $f = n - t$ others will be
> $$P(H_{n+1}|H) = \frac{(t+1)\,P(H^{n+1}_{t+1})}{(n+1)\,P(H^n_t)} = \frac{t+1}{n+2+(f+1)(\frac{P(H^{n+1}_t)}{P(H^{n+1}_{t+1})}-1)}$$

EXAMPLE: THE UNIFORM DISTRIBUTION, $P(H^{n+1}_0) = \ldots = P(H^{n+1}_{n+1})$. In this ("Bayes-Laplace-Johnson-Carnap") case, the denominator at the end of de Finetti's rule reduces to $n+2$, and we have $P(H_{n+1}|H) = \frac{t+1}{n+2}$.

2 Updating on your own Observations

In this section we determine the conditions of applicability of certain maps $P \mapsto Q$ that update your prior probability function P to a posterior probability function Q. The maps are:

$\overset{A,1}{\longmapsto}$, Conditioning on a data proposition A for which $Q(A) = 1$, and

$\overset{\vec{A},\vec{q}}{\longmapsto}$, Generalized Conditioning (or "Probability Kinematics") on a partition $A_1, \ldots, A_n$ on which you have new probabilities $Q(A_i) = q_i$.

The discussion addresses "you," a medical oncologist who is also a histopathologist. The A_i are diagnoses, whose probabilities have been driven to new values q_i by your observations. The B_s are prognoses, "s-year survival". Your problem is to determine your new probabilities for them.

On the basis of a microscopic examination of cells wiped from your patient's bronchial tumor you have updated your probabilities on a set of mutually exclusive, collectively exhaustive diagnoses A_i:

A_1, Benign; A_2, Adeno ca.; A_3, Small cell ca.; A_4, none of the foregoing.

How are you to extend the map $P(A_i) \mapsto Q(A_i)$ on the diagnoses to a map $P(B_s) \mapsto Q(B_s)$ on the prognoses? There is no general answer, but there are answers in the two special cases defined above: $\overset{A,1}{\longmapsto}$ and $\overset{\vec{A},\vec{q}}{\longmapsto}$.

2.1 A special case: Conditioning

Suppose an observation drives your probability for some diagnosis (say, A_2) all the way to 1:

$Q(A_2) = 1$, CERTAINTY.

Question. When is it appropriate to update by conditioning on A_2?

$P(B_s) \mapsto Q(B_s) = P(B_s|A_2)$, CONDITIONING.

[8] de Finetti, "Foresight," pp. 104-5. The best thing to read on this is Zabell (1989).

Answer. When the change $P(A_2) \mapsto Q(A_2) = 1$ leaves conditional probability given truth of A_2 invariant:[9]

$$Q(B_s|A_2) = P(B_s|A_2), \text{ INVARIANCE.}$$

The invariance condition is equivalent to uniformity of expansion within A_2, i.e., constancy of odds between propositions $C, D \subseteq A_2$: $\frac{P(C)}{P(D)} = \frac{Q(C)}{Q(D)}$, or

$$P(C)Q(D) = P(D)Q(C) \text{ for all } C, D \subseteq A_2, \text{ UNIFORMITY.}$$

Certainty alone is not enough to license conditioning, for one observation will generally yield many new certainties, on which conditioning would lead to different updated probability functions.

EXAMPLE. Drawing a card from a well-shuffled deck, you see that it is a heart. Conditioning on that certainty yields $P(\text{Queen of hearts}|\text{heart}) = \frac{1}{13}$. But in seeing that the card is a heart, you have also seen that it is red, and conditioning on *that* certainty yields $P(\text{Queen of hearts}|\text{red}) = \frac{1}{26}$.

2.2 A less special case: Probability Kinematics

Suppose an observation changes your probability distribution over a partition of diagnoses $A_1, \ldots, A_n$, without necessarily changing any of the $P(A_i)$ to $Q(A_i) = 1$.

Question. When is it appropriate to update by "probability kinematics"?

$$Q(B) = \sum_{i=1}^{n} Q(A_i)P(B|A_i), \text{ PROBABILITY KINEMATICS}$$

Answer. When the invariance condition holds for each of the A_i's:

$$Q(B|A_i) = P(B|A_i) \text{ for } i = 1, \ldots, n, \text{ INVARIANCE}$$

Notes:

• By the law of total probability in the form $Q(B) = \sum_i Q(A_i)Q(B|A_i)$, invariance relative to all the A_i is *equivalent* to probability kinematics.

• Conditioning is the special case in which some $Q(A_i) = 1$.

• On the native ground of probability kinematics, you are your own probability meter. In the context of *your* prior judgments, *your* new observation urges new probabilities $Q(A_i) = q_i$ upon *you*. In the most highly prized cases, you are able to explain these urges in terms of considerations which would weigh with other experts as well. But the urge is there, nudging the needle of your inner probability meter, even in the absence of such an explanation.

[9] It is easy to verify that conditioning is equivalent to invariance together with certainty.

3 Collaborative Updating

We now move outside the native ground of probability kinematics into a region where your new $Q(A_i)$'s are based on other people's observations. You are unlikely to simply adopt such alien probabilities as your own, for they are a confusion of the bare alien observation, which you would like to use, with the alien prior judgmental state, for which you may prefer to substitute your own.

We continue in the medical setting. You are a clinical oncologist, but no longer a histopathologist. You want to make the best use you can of the observations of a histopathologist whom you have consulted.

3.1 Adopt the Expert's New Probabilities?

P and Q : Your probabilities *before* and *after* the histopathologist's observation has replaced her prior probabilities $P'(A_i)$ for the diagnoses by her posterior values $Q'(A_i)$.

Will you simply adopt her new probabilities for the diagnoses, setting your $Q(A_i)$'s = her $Q'(A_i)$'s? If so, you can update by probability kinematics even if you had no prior diagnostic opinions $P(A_i)$ of your own; all you need are *her* new $Q'(A_i)$'s and *your* invariant conditional prognoses $P(B|A_i)$.

Note that she may have conditional prognoses $P'(B|A_i)$ different from yours and invariant as yours. No matter. What concern you are her diagnoses, not her prognoses.

3.2 Dissecting out the Purely Observational Part

But suppose you have priors $P(A_i)$ which you take to be well-founded, and although you have high regard for the histopathologist's ability to interpret histographic slides, you view her prior probabilities for the various diagnoses as arbitrary and uninformed. (Perhaps she has told you that she had no prior judgment in the matter, but for the purpose of formulating her report adopted convenient flat priors, $P'(A_i) = \frac{1}{n}$ for all i.)

Here you would like to dissect out of the histopathologist's report the components that represent what she has actually seen, and combine them with your own priors. These components will be her BAYES FACTORS,[10]

$$\beta'_{i1} = \frac{\text{her old odds on } A_i \text{ against } A_1}{\text{her new odds on } A_i \text{ against } A_1}.$$

In general, the factor $\beta(A:B)$ updates your odds on hypothesis A against hypothesis B:

$$\beta(A:B) =_{df} \frac{Q(A)}{Q(B)} / \frac{P(A)}{P(B)}.$$

[10]The choice of A_1 as a reference point is arbitrary, since the ratios $\beta'_{i,k} : \beta'_{j,k}$ are the same for all choices of k from 1 to n. See Schwartz et al. (1981).

Where the hypotheses are diagnoses A_i, A_j, we abbreviate

$$\beta_{i,1} =_{df} \beta(A_i : A_1) = \frac{Q(A_i)}{Q(A_1)} / \frac{P(A_i)}{P(A_1)}.$$

The histopathologist's probabilities and Bayes factors will be written P', Q', β'

According to Bayesian lore it is the histopathologist's β''s that tell you what she has learned from the observation itself, with her prior probabilities factored out.

Now you can simply multiply your prior odds on A_i against A_1 by her Bayes factors $\beta'_{i,1}$ to get your posterior odds in the light of her observations. It is straightforward to verify that you can then update your probability for a prognosis B by using the formula for probability kinematics in sec. **2.2**, but with your $Q(A_i)$'s computed as follows from her β''s and your $P(A_i)$'s:

$$Q(A_i) = \frac{P(A_i)\beta'_{i1}}{\sum_i P(A_i)\beta'_{i1}}, \text{ so that } Q(B) = \frac{\sum_i P(A_i)\beta'_{i1} P(B|A_i)}{\sum_i P(A_i)\beta'_{i1}}.$$

3.3 Updating Twice: Commutativity

Here we consider the outcome of successive updating on the reports of two different experts—say, a histopathologist and a radiologist.

If you update twice, should order be irrelevant?

Should $\overset{1}{\mapsto}\overset{2}{\mapsto} = \overset{2}{\mapsto}\overset{1}{\mapsto}$?

The answer depends on particulars of

(1) the partitions on which $\overset{1}{\mapsto}$ and $\overset{2}{\mapsto}$ are defined;
(2) the mode of updating (by probabilities? Bayes factors?); and
(3) your starting point, P.

3.3.1 Updating on new probabilities for diagnoses

A propos of (2), suppose you accept two new *probability* assignments to the same partition—first one, then another.

- Can order matter?

Certainly. Since the second assignment simply replaces the first, the result of accepting first one and then the other is the same as the result that just accepting the second would have had, by itself.

- When is order immaterial?

When there are two partitions, and updating on the second leaves probabilities of all elements of the first unchanged. This happens when the two partitions are independent relative to P.[11]

[11] For more about this, see Diaconis and Zabell (1982), esp. 825-6.

3.3.2 Updating on Bayes Factors for diagnoses

In updating by Bayes factors, order never matters.[12]

Example 1: One partition. Adopting both a pathologist's Bayes factors $\vec{\beta}^1$ and a radiologist's Bayes factors $\vec{\beta}^2$ on the same partition as your own—in either order—you come to the same result: your overall Bayes factors will be products $\beta^1_{i1}\beta^2_{i1}$. Putting it differently, your final probabilities for the diagnoses will be

$$Q(A_i) = \frac{P(A_i)\beta^1_{i1}\beta^2_{i1}}{\sum_i P(A_i)\beta^1_{i1}\beta^2_{i1}}.$$

Example 2: Two partitions.[13] the mappings $\stackrel{1}{\longmapsto}$ and $\stackrel{2}{\longmapsto}$, with 1 as $\vec{A}^1, \vec{\beta}^1$ and 2 as $\vec{A}^2, \vec{\beta}^2$, always commute, and are equivalent to a single mapping: $\stackrel{3}{\longmapsto}$, with 3 as $\vec{A}^1 \times \vec{A}^2$, $\vec{\beta}^1\vec{\beta}^2$.

Notation: $\vec{A}^1 \times \vec{A}^2$ is the partition $\{A_{ij}\}$ whose $\leq nm$ elements $A^1_i \wedge A^2_j$ $(i = 1, \ldots, n;\ j = 1, \ldots, m)$ are the conjunctions of elements of the two partitions $\vec{A}^1$, $\vec{A}^2$ for which $P(A^1_i \wedge A^2_j) > 0$. Your $=_{df}$ $P(A^1_i)$ and $P(A^2_j)$ are all to be positive. Now we have

$$Q(A_i \wedge A_j) = \frac{P(A_i \wedge A_j)\beta^1_{i1}\beta^2_{j1}}{\sum_{ij} \beta^1_{i1}\beta^2_{j1}}.$$

4 Dorling on the Duhem-Quine Problem

Skeptical conclusions about the possibility of scientific hypothesis-testing have been drawn from the presumed arbitrariness of answers to the question of which to give up—a theory (e.g., in **4.2** below, general relativity), or an auxiliary hypothesis ('The equipment was in working order')—when they jointly contradict empirical data. The problem, posed by Pierre Duhem in the first years of the 20th century, was reanimated by W.V. Quine in mid-century.[14] But the holistic conclusion depends on the assumption that deductive logic is our *only* tool for confronting theories with empirical data. That would leave things pretty much as Descartes saw them, just before the mid-17th century emergence of the probabilistic ("Bayesian") methodology that Jon Dorling has brought to bear on various episodes in the history of science. Here is an introduction to Dorling's work, using extracts from his important but still unpublished 1982 paper[15]

[12] Proofs are straightforward. See my Petrus Hispanus Lectures, II: 'Radical Probabilism', *Actas da Sociedade Portuguesa da Filosofia* (forthcoming; and currently available in http://www.princeton.edu/ bayesway/pu/Lisbon.pdf).

[13] Adapted from Field (1978) 361-7.

[14] Quine (1953), p. 41: 'our statements about the external world face the tribunal of sense experience not individually but as a corporate body.'

[15] This section is based on Dorling (1982). His work is also discussed in Howson and Urbach (1993). See also Dorling (1979) and Redhead (1980).

4.1 Setting the Stage

Were, as in Dorling's analysis, updating is by conditioning on a data statement D, the Bayes factor for a theory T against an alternative theory S equals the likelihood ratio,

$$\beta(T:S) = \frac{P(D|T)}{P(D|S)}.$$

The empirical result D is not generally deducible or refutable by T alone, or by S alone, but in interesting cases of scientific hypothesis testing D is deducible or refutable on the basis of the theory and an auxiliary hypothesis A (e.g., the hypothesis that the equipment is in good working order). To simplify the analysis, Dorling makes an assumption, prior independence, that can generally be justified by appropriate formulation of A:

$$P(A \wedge T) = P(A)P(T), \;\; P(A \wedge S) = P(A)P(S),$$

Generally speaking, S is not simply the denial of T, but a definite scientific theory in its own right, or a disjunction of such theories, all of which agree on the phenomenon of interest, so that, as an explanation of that phenomenon, S is a rival to T. In any case Dorling uses the independence assumption to expand the right-hand side of the Bayes Factor = Likelihood Ratio equation:

$$\beta(T:S) = \frac{P(D|T \wedge A)P(A) + P(D|T \wedge \neg A)P(\neg A)}{P(D|S \wedge A)P(A) + P(D|S \wedge \neg A)P(\neg A)}$$

To study the effect of D on A, he also expands $\beta(A : \neg A)$ with respect to T (and similarly with respect to S, although we do not show that here):

$$\beta(A:\neg A) = \frac{P(D|A \wedge T)P(T) + P(D|A \wedge \neg T)P(\neg T)}{P(D|\neg A \wedge T)P(T) + P(D|\neg A \wedge \neg T)P(\neg T)}$$

4.2 Einstein/Newton, 1919

In these terms Dorling analyzes two famous tests that were duplicated, with apparatus differing in seemingly unimportant ways, with conflicting results: one of the duplicates confirmed T against S, the other confirmed S against T. Nevertheless, the scientific experts took the experiments to clearly confirm one of the rivals against the other. Dorling explains why the experts were right:

"In the solar eclipse experiments of 1919, the telescopic observations were made in two locations, but only in one location was the weather good enough to obtain easily interpretable results. Here, at Sobral, there were two telescopes: one, the one we hear about, confirmed Einstein; the other, in fact the slightly larger one, confirmed Newton. Conclusion: Einstein was vindicated, and the results with the larger telescope were rejected." (1982, sec. 4)

NOTATION

T: Einstein: light-bending effect of the sun

S: Newton: no light-bending effect of the sun

A: Both telescopes are working correctly

D: The conflicting data from both telescopes

In the Bayes factor $\beta(T:S)$ above, $P(D|T \wedge A) = P(D|S \wedge A) = 0$ since if both telescopes were working correctly they would not have given contradictory results. Then the first terms of the sums in numerator and denominator vanish, so that the factors $P(\neg A)$ cancel and we have

$$\beta(T, S) = \frac{P(D|T \wedge \neg A)}{P(D|S \wedge \neg A)}$$

Dorling continues: "Now the experimenters argued that one way in which A might easily be false was if the mirror of one or the other of the telescopes had distorted in the heat, and this was much more likely to have happened with the larger mirror belonging to the telescope which confirmed S than with the smaller mirror belonging to the telescope which confirmed T. Now the effect of mirror distortion of the kind envisaged would be to shift the recorded images of the stars from the positions predicted by T to or beyond those predicted by S. Hence $P(D|T \wedge \neg A)$ was regarded as having an appreciable value, while, since it was very hard to think of any similar effect which could have shifted the positions of the stars in the other telescope from those predicted by S to those predicted by T, $P(D|S \wedge \neg A)$ was regarded as negligibly small, hence the result as overall a decisive confirmation of T and refutation of S." Thus the Bayes factor $\beta(T, S)$ is very much greater than 1.

5 Radical Probabilism

Descartes sought to refute skepticism about experience by proving the existence of a God ($=_{df}$ the perfect being) who surely does not deceive us. On this "dogmatic" foundation of certainty he would build empirical science.

Shortly after Descartes's death, probabilistic thinking—in much the same form in which we know it today—emerged from a famous correspondence between Fermat and Pascal. It was given the place of honor at the end of the best-selling "How to Think" book known as "The Port-Royal Logic" (Arnauld, 1662):

> "To judge what one must do to obtain a good or avoid an evil, it is necessary to consider not only the good and the evil in itself, but also the probability that it happens or does not happen; and to view geometrically the proportion that all these things have together."

But a quasi-Cartesian dogmatism figures prominently in 20th century thought about the foundations of probabilistic thinking. Here is an example from the pragmatist–empiricist philosopher C. I. Lewis (1946, p. 186):

> "If anything is to be probable, then something must be certain. The data which themselves support a genuine probability, must themselves be certainties. We do have such absolute certainties, in the sense data initiating belief and in those passages of experience which later may confirm it."

What I call "radical probabilism" denies this claim. The claim itself seems to be based on the thought that conditioning on certainties is the only way to update probabilities. That basis for dogmatic probabilism is undermined by the existence of a generalized conditioning—probability kinematics—as a way of updating on mere probabilities.

Radical probabilism offers Bayes factors as a surrogate for Lewis's absolute certainties. Lewis continued: these certainties cannot

> "be phrased in the language of objective statement — because what can be so phrased can never be more than probable. Our sense certainties can only be formulated by the expressive use of language, in which what is signified is a content of experience and what is asserted is the givenness of this content."

In radical probabilism the Bayes factors β_{i1} do the job of your ineffable sense certainties. Like those ineffables, your Bayes factors lie outside the Boolean algebra of objective statements on which your P and Q are defined. But where Lewis can give no intelligible account of your sense certainities, radical probabilism can identify the β's as ratios of your new to old odds on items that *are* expresible in the language of objective statement.

6 Mad-Dog Subjectivism

Radical probabilism gets along without objective probabilities, real chances (R). The thought is that these are nothing but projections of judgmental probabilities P out into the world, whence we hear them clamoring to be let back in. The following equation ("Miller's Principle") could be their return ticket:[16]

$$P[H|R(H) = r] = r, \text{ where 'r' is a purely mathematical designator.}$$

(If you know the real chance, that will be your judgmental probability.) But they don't need a return ticket; they never really left.

[16]Here, 'r' might be '.7' or '$1/\pi$', but not '1/(my mass in Kg)', and not '$R(H)$'. The thought is that you must be able to tell what number 'r' denotes without recourse to any empirical facts. The need for this restriction is seen when we put '$R(H)$' for 'r'. We then have $P[H|R(H) = R(H)] = R(H)$, which is equivalent to $P(H) = R(H)$. Unrestricted, Miller's Principle implies that your subjective probabilities always agree with the objective probabilities. In effect, Miller (1966) rejected the restriction and welcomed the result as a reductio ad absurdum of the concept of judgmental probability. (Note that since 'R' means real chance, not future probability, van Fraassen's reflection principle is only formally identical with Miller's principle.)

There certainly *are* numbers "out there"—numbers like the fraction of 70-year-olds who live to be 80—and it may well be that if you knew that number you would adopt it as your judgmental probability that your 70-year-old uncle, Bob, will live to be 80. But maybe not. Maybe you know that Bob comes from remarkably long-lived stock, in which case your probability for his reaching 80 might be higher than the statistics on 70-year-olds would suggest. (But maybe you also know that Bob has pancreatic cancer, so that your probability is below the statistical average for 70-year-olds.) This is the problem of the dreaded reference-class. According to radical probabilism, this problem can be solved by putting the horse before the cart, using your probabilistic judgments to choose among the various numbers out there. But the point is that there are enough familiar numbers out there—statistics, fractions of green balls in urns, etc.—to do the jobs that objectivists send "real probabilities" out there to do.

Once such numbers have been chosen, formulas that look rather like the return ticket may come into play:

$$P(H|X = x) = x, \text{ where } X \text{ is an ordinary random variable}$$

EXAMPLE. P(A green ball will be drawn|70% are green)=.7 if you think the balls are well mixed, etc.

To hypostasize $R(H)$ as a physical magnitude is to sweep the subjective element in "objective" probability under the carpet. We do better to identify the parameter X case by case, as the fraction of green balls in the urn, or of septuagenerian men who live to be 80, or whatever. In all of these quests for a suitable X the great clue, the sticky subjective core, is that, whatever X turns out to be, it must satisfy the equation $P(H|X = x) = x$.

7 Appendix: Aumann's Theorem generalized[17]

In "Agreeing to Disagree" (1976) Robert Aumann proves that a group of agents who once agreed about the probability of some proposition for which their current probabilities are common knowledge must still agree, even if those probabilities reflect disparate observations. Perhaps one saw that a card was red and another saw that it was a heart, so that as far as that goes, their common prior probability of 1/52 for its being the Queen of hearts would change in the one case to 1/26, and in the other to 1/13. But if those are indeed their current probabilities, it cannot be the case that both know it, and both know that both know it, etc., etc.

In Aumann's framework new probabilistic states of mind can only arise by conditioning old ones on new knowledge. In such a framework, current probabilities must derive from what is in effect knowledge, i.e., true full belief.

[17]This is joint work with Matthias Hild (CalTech) and Mathias Risse (Yale): see Hild, Jeffrey and Risse (1998), (1999). I am unsure of its relationship to Dov Samet's work: see http://ideas.uqam.ca/ideas/data/Papers/wpawuwpga9902004.html

But here we derive Aumann's result from common knowledge of the shared value of a probability, however arrived at. We work with possible worlds in which the agents' probabilities and their evolution are matters of fact, represented within the model.

Independence of particular update rules is a central feature of the new framework. But of course we need some constraint on how agents update their probabilities. For this we use Michael Goldstein's (1983) requirement that current expectations of future expectations equal current expectations.[18] This is the workhorse for our proof of the Generalized 'No Agreement' Theorem.

Notation. $1, \ldots, N$ are the agents, Ω is a non-empty set of "possible worlds"—each element of which specifies a complete history, past, present, and future—and $\mathcal{A}$ is a σ–field over Ω. So far, our framework is like Aumann's; but now the change comes. In our framework agents belong to worlds, and as time goes by, their probabilities concerning their own and other agents' probabilities evolve along with their probabilities concerning other matters—e.g., in the year 2001, my probability for your probability for my probability for a Republican President in 2004.

We model this via probability measures $P_{i\omega}$ and $Q_{i\omega}$ representing ideally precise probabilistic states of mind of agents i in worlds ω before (P) and after (Q) updating.[19]

After defining posterior "belief" as 100% Q-probability and "knowledge" as true belief, we define posterior mutual and common knowledge as usual.

BELIEF (new, that A by i in ω): $B_iA = \{\omega : Q_{i\omega}(A) = 1\}$

KNOWLEDGE (new, that A by i in ω): $K_iA = A \cap B_iA$

MUTUAL KNOWLEDGE of degree 0: $M_0A = A$

MUTUAL KNOWLEDGE OF DEGREE $n+1$: $M_{n+1}A = \bigcap_{i=1}^{N} K_iM_nA$

COMMON KNOWLEDGE: $\kappa A = \bigcap_{n=0}^{\infty} M_nA$

In the generalized theorem $P_{i\omega}(H)$ and $Q_{i\omega}(H)$ are agent i's old and new probabilities for some hypothesis $H \in \mathcal{A}$ as they would be in world ω. We now write these simply as $P_\omega(H)$ and $Q_\omega(H)$, respectively, with the subscript i understood:

SHORTHAND: P_ω for $P_{i\omega}$, Q_ω for $Q_{i\omega}$

[18] This was independently argued for by Bas van Fraassen (1984) in the form: Current expectations of future probabilities equal current probabilities ("Reflection Principle").

[19] In the present framework there might be different common priors in different worlds ω, ω': perhaps the measures $P_{i\omega}$ are the same for all i, and so are the measures $P_{i\omega'}$, but $P_{i\omega}(A) \neq P_{i\omega'}(A)$ for some A.

In the generalized theorem the proposition C specifies $q_1, \ldots, q_N$ as agent i's new probabilities for H:

$$C =_{df} \bigcap_{i=1}^{N} \{\omega : Q_\omega(H) = q_i\}$$

The crucial hypothesis of the generalized theorem is Goldstein's (1983) principle **G** that *old probabilities = old expectations of new probabilities.* Here $Q(A)$ is a random variable, a P_ω-measurable function with worlds $\omega' \in \Omega$ as arguments (subscripts), to which it assigns real values $Q_{\omega'}(A)$:

$$\mathbf{G}: \quad P_\omega(A) = E(Q(A)) = \int_\Omega Q(A)\, dP_\omega$$

The second hypothesis says that whenever A is in $\mathcal{A}$, so are the N propositions saying ('BA') that the several agents Q-believe that A is true. This guarantees that $\mathcal{A}$ is closed under all the operations K_i, M_n, and κ.

In proving the theorem we use two lemmas.

LEMMA 1

While something is common knowledge, everyone is sure it is: $\kappa A \rightarrow B\kappa A$.

Proof. By the definitions of CK^2, MK^2, and K^2.

LEMMA 2

If **G** holds, then $\int_{\Omega-\kappa C} Q(\kappa C)\, dP_\omega = 0$

Proof. $P_\omega(\kappa C) = \int_\Omega Q(\kappa C)\, dP_\omega = \int_{\kappa C} Q(\kappa C)\, dP_\omega + \int_{\Omega-\kappa C} Q(\kappa C)\, dP_\omega$ by **G**. By lemma 1 the first term of this sum $= \int_{\kappa C} 1\, dP_\omega = P_\omega(\kappa C)$, so the second term $= 0$.

AUMANN'S AGREEMENT THEOREM GENERALIZED

FOUR HYPOTHESES:

G holds, $\quad B : \mathcal{A} \mapsto \mathcal{A}, \quad P_\omega$ is the same for all $i, \quad P_\omega(\kappa C) > 0.$

CONCLUSION:

$P_\omega(H|\kappa C) = q_1 = \ldots = q_N.$

Proof. $P_\omega(H|\kappa C) = \dfrac{\int_\Omega Q(H \cap \kappa C)\, dP_\omega}{\int_\Omega Q(\kappa C)\, dP_\omega}$ by **G**,

$$= \frac{\int_{\kappa C} Q(H \cap \kappa C)\, dP_\omega + \int_{\Omega-\kappa C} Q(H \cap \kappa C)\, dP_\omega}{\int_{\kappa C} Q(\kappa C)\, dP_\omega + \int_{\Omega-\kappa C} Q(\kappa C)\, dP_\omega},$$

$= \dfrac{\int_{\kappa C} Q(H \cap \kappa C)\, dP_\omega}{\int_{\kappa C} Q(\kappa C)\, dP_\omega}$ by lemma 2 since $Q(A \cap \kappa C) \leq Q(\kappa C)$,

$= \dfrac{\int_{\kappa C} Q(H)\, dP_\omega}{\int_{\kappa C} 1\, dP_\omega}$ by lemma 1,

$$= \frac{\int_{\kappa C} q_i \, dP_\omega}{\int_{\kappa C} dP_\omega} \text{ by definition of } C,$$
$= q_i$, i.e. the same for $i = 1, \ldots, N$.

References

Arnauld, Antoine (1662, 1964) *Logic, or, the Art of Thinking* ("The Port-Royal Logic") tr. J. Dickoff and P. James, Bobbs-Merril, Indianapolis.

Aumann, R.J. (1976), Agreeing to Disagree, *Annals of Statistics* **4**, 1236–1239.

Bayes, Thomas (1763) An essay towards solving a problem in the doctrine of chances, *Philosophical Transactions of the Royal Society* **53**, 370-418.

Carnap, R. (1950), *Logical Foundations of Probability*, Chicago: University of Chicago Press. (2nd ed., 1962)

de Finetti, Bruno (1937, 1980) La prévision, ses lois logiques, ses sources subjectives, *Annales de l'institut Henri Poincaré* **7**, 1-68. Translation ("Foresight ...") in *Studies in Subjective Probability,"* Henry Kyburg, Jr. and Howard Smokler (eds.), Krieger, New York, 53-118.

Diaconis, Persi and Sandy Zabell (1989) Updating Subjective Probability, *Journal of the American Statistical Society* **77**, 822-30.

Dorling, Jon (1979) Bayesian personalism, the methodology of research programmes, and Duhem's problem, *Studies in History and Philosophy of Science* **10**, 177-87.

Dorling, Jon (1982) Further illustrations of the Bayesian solution of Duhem's problem (29 pp., photocopied); accessible on *http* : *www.princeton.edu*\~*bayesway*

Duhem, Pierre (1906, 1914. 1954) *La Théorie Physique: Son Object, Sa Structure*, Paris. Translation, *The Aim and Structure of Physical Theory*. Princeton University Press.

Field, Hartry (1978) A Note on Jeffrey Conditionalization, *Philosophy of Science* **45**, 361-7.

Geanakoplos, J.D., Polemarchakis, H.M. (1982), "We Can't Disagree Forever", *Journal of Economic Theory* **28**, 192–200.

Goldstein, Michael (1983), The Prevision of Prevision, *Journal of the Amer-*

ican Statistical Association **78**, 817-819.

Howson, Colin and Peter Urbach (1993) *Scientific Reasoning: the Bayesian approach,* Open Court, La Salle, Illinois, 2nd ed.

Hild, Matthias, Jeffrey, R., and Risse, Mathias (1998), Agreeing to disagree: Harsanyi and Aumann, in *Game Theory, Experience, Rationality,* Werner Leinfellner (ed.), Kluwer, 1998.

Hild, Matthias, Jeffrey, R., and Risse, Mathias (1999), Aumann's 'No Agreement' Theorem Generalized, *The Logic of Strategy,* C. Bicchieri, R. Jeffrey, and B. Skyrms, eds.: Oxford University Press.

Jeffrey, Richard (1970) [untitled review of Miller (1966) and other papers] *Journal of Symbolic Logic* **35**, 124-127.

Jeffrey, Richard (1992) *Probability and the Art of Judgment,* Cambridge U. P.

Lewis, C. I. (1946) *An Analysis of Knowledge and Valuation,* Open Court.

Miller, David (1966) A Paradox of Information, *British Journal for the Philosophy of Science* **17**, 59-61.

Neurath, Otto (1932), Protokollsätze, *Erkenntnis* **3**, 204-214. Translated in his *Philosophical Papers 1913-1946*, Dordrecht: Reidel.

Quine, Willard van Orman (1951. 1953) Two Dogmas of Empiricism, *Philosophical Review* **60**, 20-43. Reprinted in *From a Logical Point of View,* Harvard University Press.

Redhead, Michael (1980) A Bayesian reconstruction of the methodology of scientific research programmes. *Studies in History and Philosophy of Science* **11**, 341-7.

Schwartz, W.B., Wolfe, H.J., and Pauker, S.G. (1981) Pathology and Probabilies: A New Approach to Interpreting and Reporting Biopsies, *The New England Journal of Medicine* **305**917-23.

van Fraassen, Bas (1984) Belief and the will, *J. Philosophy* **65** 243-60.

Zabell, Sandy (1989) The Rule of Succession. *Erkenntnis* **31**, 283-321.