A Review of the Lottery Paradox

Gregory Wheeler

Artificial Intelligence Center - CENTRIA
Department of Computer Science, Universidade Nova de Lisboa
2829-516 Caparica, Portugal
`greg@di.fct.unl.pt`

Penultimate draft. Final version to appear in *Probability and Inference: Essays in honor of Henry E. Kyburg, Jr.*, edited by W. Harper and G. Wheeler.

Henry Kyburg's *lottery paradox* (1961, p. 197)[1] arises from considering a fair 1000 ticket lottery that has exactly one winning ticket. If this much is known about the execution of the lottery it is therefore rational to accept that one ticket will win. Suppose that an event is very likely if the probability of its occurring is greater than 0.99. On these grounds it is presumed rational to accept the proposition that ticket 1 of the lottery will not win. Since the lottery is fair, it is rational to accept that ticket 2 won't win either—indeed, it is rational to accept for any individual ticket i of the lottery that ticket i will not win. However, accepting that ticket 1 won't win, accepting that ticket 2 won't win, ..., and accepting that ticket 1000 won't win entails that it is rational to accept that no ticket will win, which entails that it is rational to accept the contradictory proposition that one ticket wins and no ticket wins.

The lottery paradox was designed to demonstrate that three attractive principles governing *rational acceptance* lead to contradiction, namely that

1. It is rational to accept a proposition that is very likely true,
2. It is not rational to accept a proposition that you are aware is inconsistent, and
3. If it is rational to accept a proposition A and it is rational to accept another proposition A', then it is rational to accept $A \wedge A'$,

are jointly inconsistent.

The paradox remains of continuing interest because it raises several issues at the foundations of knowledge representation and uncertain reasoning: the relationships between fallibility, corrigible belief and logical consequence; the

[1] Although the first published statement of the lottery paradox appears in Kyburg's 1961 *Probability and the Logic of Rational Belief*, the first formulation of the paradox appears in "Probability and Randomness," a paper delivered at the 1959 meeting of the Association for Symbolic Logic and the 1960 International Congress for the History and Philosophy of Science, but published in the journal *Theoria* in 1963. It is reprinted in Kyburg (1983).

roles that consistency, statistical evidence and probability play in belief fixation; the precise normative force that logical and probabilistic consistency have on rational belief.

1 Rational Acceptance

Nevertheless many commentators think the lottery paradox is a blunder in reasoning rather than a genuine puzzle about the structure of collections of rationally accepted statements. Some suggest that the challenge is to explain away the initial appeal of rationally accepting a statement—the assumption being that rational acceptance is either untenable or incompatible with other entrenched epistemic principles. Others think that the notion of rational acceptance is underspecified in the setup for the lottery paradox. This view has motivated some to argue that there are substantive differences between the lottery paradox and other epistemic paradoxes, notably David Makinson's *paradox of the preface* (1965).[2]

Although there is a consensus among this group that one should deny that it is rational to accept that a ticket of the lottery loses, there is less agreement over why this should be so. Some think that there are features of "lottery propositions"[3] that clash with norms for accepting, believing or claiming to know a proposition.[4] Another view is that "lottery contexts"[5] are self-undermining,[6] or require an agent to reject all alternatives before accepting that a particular ticket will lose.[7] Still another theme is that statistical evidence is an insufficient basis for rational belief,[8] either because of missing explanatory information,[9] missing causal information,[10] or because logical consistency is thought fundamental to the notion of rationality but cannot be guaranteed on the basis of statistical evidence alone.[11] Finally, the notion of rational acceptance itself, when it is conceived in terms of high, finite, *subjective* probability is often shown to

[2] For instance, John Pollock (1986, 1993) has advanced the claim that the lottery paradox and the paradox of the preface are fundamentally different paradoxes of rational acceptance. See Neufeld and Goodman (1998) for a critical discussion of Pollock's approach, and Kvanvig (1998) for an overview of the epistemic paradoxes. Finally, an historical note: Makinson formulated the paradox of his preface without any prior knowledge of Kyburg's lottery paradox (Makinson, July 2005 conversation).

[3] Harman (1973), Cohen (1988), Vogel (1990), Baumann (2004), Hawthorne (2004)

[4] Harman (1967), Stalnaker (1987), DeRose (1996), Williamson (1996, 2000), Christensen (2004).

[5] Levi (1967), Cohen 1988, DeRose (1996), Lewis (1996).

[6] Bonjour (1985), Pollock (1986), Ryan (1996), Evnine (1999), Douven (2002).

[7] Goldman (1986).

[8] Harman (1967), Cohen (1988), Nelkin (2000).

[9] Harman (1968)

[10] See Shoeman's (1987) lucid discussion of a version of this proposal in Thompson (1984).

[11] Levi (1967), Lehrer (1980), Stalnaker (1987).

be problematic for accepted sentences whose probability is any value less than unity.[12]

So although there is a common assumption that the lottery paradox is the result of a confusion over rational acceptance, there is little agreement over precisely what that confusion is. Part of the reason for this impasse is that most of the contemporary literature has developed without engaging Kyburg's original motivations for a theory of rational acceptance. This isn't necessarily a bad development, as we shall soon see. But one consequence is a consensus in the literature that simultaneously under- and over-constrains the lottery paradox. The paradox has been under-constrained by its popularity in traditional epistemology, where intuitions about knowledge ascriptions and ordinary language arguments hold sway, giving theorists a wide berth. The paradox has been over-constrained by well-meaning, but misguided attempts to impose formal constraints on the paradox to keep in check the excesses of common language epistemology. The problem is that most of the time the formal constraints that are proposed are explicitly rejected by Kyburg's view of rational acceptance.

The lottery paradox is neither motivated by ordinary language epistemology, nor is it the result of a misunderstanding over the structural properties of probability. The theory of rational acceptance was devised as an integral component for a theory of uncertain inference. Kyburg views his lottery ticket thought experiment not as a paradox about rational acceptance, but rather as an argument highlighting the trouble with unrestricted aggregation of well-supported but less than certain statements.

The problem that aggregation poses to an account of uncertain inference is a surprisingly deep one—or so I hope to demonstrate with this review. In this section we revisit the notion of rational acceptance to highlight what the theory is conceived to do. In the next section we consider how to approach the paradox if one doesn't wish to adopt Kyburg's theory of rational acceptance. One conclusion of this section is that even if you reject Kyburg's account of rational acceptance, you are still left with the basic set of problems that the theory is designed to solve. The third section demonstrates the force of this point by considering a recent argument against treating statistical evidence as a basis for reasonable belief.

1.1 Practical Certainty and Full Belief

For Kyburg uncertainty is a fundamental feature of statements with empirical content. The aim of his program is not to provide a "logic of cognition" that specifies complete and precise rules for belief change, but instead to formulate the logical structure of the set of conclusions that may be obtained by non-demonstrative inference from a body of evidence. There are several important points that distinguish Kyburg's program from other theories of epistemic probability.

[12] Hintikka and Hilpinen (1966), Kaplan (1981), Halpern (2003), Douven and Williamson (2006).

First, and most generally, *full belief* on Kyburg's view is a disposition to "act as if" an accepted statement p were true. It is not a disposition to make (or fail to make) a collection of bets on the truth of p. If an agent bets at even money on a coin toss landing heads, the agent is not acting as if the statement *the coin will lands heads* is true. The agent's betting on this outcome of the toss is an action, which is entirely different than acting as if the outcome of the toss is in his favor. The agent is performing an action (the placing of a bet) based upon his beliefs about the uncertainty mechanism generating the coin toss event (e.g., that the outcome of the toss is independent of his betting on the outcome, that the coin is fair, that outcomes are determinate).

However, the disposition to act on a set of beliefs is contingent upon what is at stake for the agent. In modern parlance, full belief is contextually determined: an agent's disposition to fully belief p may vary across cases in which the epistemic status of that belief is constant.

But the disposition to fully believe p is not determined by the total magnitude of the stake, but instead is determined by the ratio of the amount risked to the amount gained. Under ordinary circumstances, an agent will fully believe a factual statement p, such as *Lisbon is the capital of Portugal*. However, there are circumstances under which that same agent will neither believe p nor believe the negation of p—such as when an action based upon believing p presents a penalty for being wrong that is judged to be too high.

The idea is that an agent @ believes p when the stakes for any relevant action based upon @ are below an acceptable level of economic risk. In other words, we say that @ *[r/s]-believes* p iff for any action A,

(i.) @ evaluates A to cost r' if p is false,
(ii.) @ evaluates A to return s' if p is true,
(iii.) $r'/s' < r/s$, and
(iv.) @ acts as if p were true.

The candidates for r/s belief are rationally accepted statements. *Rational acceptance* is determined by the epistemic risk of accepting p, which is understood as the chance of accepting p when p is false. Probability is assigned to a statement p based entirely upon what is known about p, including the limit case when *nothing* is known, in which case $[0, 1]$ is assigned to represent our total ignorance about p.

So, the set of r/s beliefs are those rationally accepted statements that an agent is willing to let ground some set of actions. Based upon the evidence in your corpus, you may rationally accept that Lisbon is the capital of Portugal and be rewarded modestly for some set of actions based upon this belief. But if the penalty for being mistaken about the name of Portugal's capital is high enough, such as the sacrifice of your life, that risk-to-reward ratio may exceed your willingness to fully believe that Lisbon is the capital of Portugal.[13] You may have sufficiently good reasons for believing that p is true, and sufficiently good

[13] See Bill Harper's contribution for a discussion of Kyburg's theory of full belief. See also Kyburg (1988; 1990, p. 245).

reasons for refusing to act as if you had sufficiently good reasons for believing that p is true.

The lottery paradox concerns rational acceptance, not full [r/s]-belief.

There are important differences between rationally accepting p because the evidence makes it very probable that p, and rationally accepting p because an agent assigns a high probability to the event (a degree of belief that) p. On the first reading, a statement is assigned a (possibly) imprecise probability based upon all evidence available about p. All parameters for making this inference are presumed to be objective; the result is a unique, if imprecise, probability assignment.[14] This outline conforms to Kyburg's conception of rational acceptance. On the second view each agent represents the statements he believes about a problem domain as elements of a σ-algebra over which a probability distribution is defined. Changes to probability assignments are effected by changing this distribution, subject to the mathematical constraints of the probability calculus. Changes are then propagated through the entire algebra. These probability assignments are traditionally construed as precise point values; the assignment is determined entirely by an individual agent, so probability assignments for the same problem domain may vary among agents. We will return to the differences between these two conceptions of probability in the next subsection.

The basic idea behind Kyburgian rational acceptance is to represent the maximum chance of error that is considered tolerable by a real value ϵ close to zero,[15] and to judge rationally acceptable those statements whose chance of error is less than ϵ. To disambiguate between these two readings of high probability, a statement whose chance of error is less than ϵ is considered to be a statement of *practical certainty*. A threshold for acceptance may then be defined in terms of ϵ, namely $1 - \epsilon$.

Kyburg's strategy is to view non-demonstrative inference as a relation from a set of accepted evidence statements, called *evidential certainties*, and a practical certainty. The set of practical certainties is the *knowledge* that is based upon that evidence. The evidence itself may be uncertain, but it is supposed that the risk of error is less than the knowledge derived from it. Membership to the evidence set is determined by a threshold point δ for acceptance that is less than or equal to the threshold point ϵ for accepting practical certainties. Thus, the notion of rational acceptance is fundamental to constructing the distinction between evidence and knowledge on this account.

[14] See Choh Man Teng's contribution for a discussion of the conditions employed for maintaining consistency within *evidential probability* while making a unique probability assignment.

[15] Note that "ϵ" is to be construed as a fixed number, rather than a variable that approaches a limit. This is in sharp contrast to an approach advanced by Ernest Adams (Adams 1966, 1975) and Judea Pearl (1988, 1990) designed to draw a connection between high probability and logic by taking probabilities *arbitrarily* close to one as corresponding to knowledge. For a discussion of the comparison between these two approaches with respect to preferential semantics (System P), see (Kyburg, Teng and Wheeler, 2006).

One consequence of this stratified view of evidence and knowledge is that the rule of adjunction is unsuitable for aggregating practical certainties.

> Suppose we consider that there is a finite basis for our reasonable beliefs [practical certainties]....If we accept these statements individually as reasonable beliefs, we are led by adjunction to accept their full conjunction as a reasonable belief. But if those statements have empirical content, they must be (perhaps only slightly) uncertain. Their conjunction, then, must be even more uncertain (Kyburg 1997, p. 117).

Another consequence is that there is a sharp distinction drawn between evidential relations and classical closure operations.

> It is not ... that the premises of the lottery paradox (ticket i will not win; the lottery is fair) do not *entail* its contradictory conclusion, but that the grounds we have for accepting each premise do not yield grounds for accepting the conjunction of the premises, and thus do not force us to accept the proposition that no ticket will win. (Kyburg 1997, p. 124)

These two points—the denial that the rule of adjunction is suitable for aggregating rationally accepted statements, and that evidential relations are not isomorphic to logical closure operations—are important. Although there are commentators who likewise reject one or both of these conditions,[16] Kyburg's position is rooted in his theory of uncertain inference.

The lottery paradox is Kyburg's shorthand for the perils that aggregation properties of boolean algebras present to formal models of evidential relations: for him it is a bumper sticker, not a puzzle. Nevertheless, some have viewed the clash between these logical properties and the minimal conditions for rational acceptance encoded in the paradox as a *reductio* against the very concept of rational acceptance. When evaluating an argument of this type, however, it is important not to give short shrift to the theory of rational acceptance by simply labeling it "non-bayesian". For doing so runs the risk of over-constraining the paradox by saddling the theory of rational acceptance with properties that Kyburg explicitly rejects. Let's now review the motivations for the theory of rational acceptance.

1.2 Uncertain Inference

The first exchanges between Kyburg and his critics over the lottery paradox pitted the notion of rational acceptance against the constraints imposed by probabilism.[17] Rudolf Carnap (1968) claimed that acceptance of a statement

[16] For instance, (Dretzky 1971), (Nozick 1981), (McGinn 1984), (Sorensen 1988), (Kitcher 1992), (Korb 1992), (Foley 1993), (Hunter 1996), (Brown 1999), (Hawthorne and Bovens 1999), (Wheeler 2005, 2006).

[17] In particular, see Richard Jeffrey's (1956; 1965; 1970); Rudolf Carnap's (1962, 1968); Isaac Levi's (1967); Henry Kyburg's (1961; 1970a; 1970b).

should be viewed as a rough approximation of assigning a high probability to it, and Richard Jeffrey (1956, 1970) pressed this point by charging that developing a theory to accommodate rational acceptance amounts to building a rigorous theory around “loose talk”.

The Carnap-Jeffrey view of rational acceptance as an approximation of a particular agent’s assignment of high, precise probability is tacitly accepted by many contemporary commentators on the lottery paradox. Indeed, one frequently finds the claim that the lottery paradox is a mistake that results from a failure to grasp the basic constraints of probabilism. The aim of this section is to highlight some features of the Carnap-Jeffrey conception of rational acceptance, and to highlight why Kyburg’s rejection of this view is based upon a deep rejection of probabilism.

Probabilism maintains that the mathematical structure of the probability calculus reveals the structure of the problem one intends to model with probabilities—or, more specifically, it maintains that the probability calculus reveals the proper structural constraints on our uncertain beliefs about the problem we are considering.

The appeal of Bayesianism is the apparent conceptual simplicity of putting probabilism into practice. This assumption is questionable when applied to models of rational belief,[18] but set aside issues of computational and conceptual complexity. Probabilism faces another problem, for the picture it provides of statistical inference is fundamentally at odds with what is conceived and performed by most statisticians.

Statisticians occupy themselves with developing and applying techniques that facilitate the reduction of data to a minimal set of features that captures the information of interest in that data. There are two basic forms of statistical inference that may be drawn. The first, which we call *statistical estimation*, is the drawing of an inference from a sample to a population. An argument drawing the conclusion that “The proportion of Fs that are Gs is r” from the premise that “The proportion of sampled Fs that are Gs is r” is an example of a statistical estimation. The second form, which we call *direct inference*, draws a conclusion about a particular instance from a statistical generalization. Drawing the conclusion that “This F is a G” from the premise “Nearly all F are G” is an example of direct inference.[19]

One reason that both subjective and objective versions of Bayesianism have difficulty here is that they do not provide an account of the relationship between probability and evidence; rather, Bayesian techniques are designed to get around accounting for the relationship between evidence and probability assignment. Bayesianism requires an assignment of precise probability to a statement even when there is no evidence about that statement whatsoever. This assump-

[18] See Clark Glymour’s contribution in this volume for a critical discussion of this assumption within Bayesian models of causal reasoning.

[19] See John Pollock’s contribution for an alternative account of direct inference to Kyburg’s. See also the contributions by Isaac Levi, Teddy Seidenfeld and Henry Kyburg.

tion is necessary because the account must start with a distribution of some kind; the method is one for moving a probability mass around an algebra, not one that guides in the assignment of a distribution. Strictly speaking, any probability distribution is as good as the next; it is the extra-logical requirements of fixing an initial distribution according to betting rates or maximization of entropy that are designed to select an appropriate distribution. Note however that betting quotients are constraints on decision, not epistemic constraints on belief, whereas *the principle of maximum entropy* is a method for generating testable hypotheses about physical probabilities rather than a principle for generating epistemic probabilities. It is one thing to argue that one does not *need* a theory of probability assignment based upon evidential constraints; it is quite another to confuse a decision theoretic account for a theory of evidential probability assignment.

Another problem is that the underlying structure of these two forms of statistical inference are not directly reflected in the structure of the probability calculus. The measure theoretic properties of the calculus impose assumptions on the problem domain that are not always warranted. For example, in cases where there is a basis for assigning probability to a statement (event) p, and a basis for assigning probability to another statement (event) p', Bayesianism assumes that there is always a joint distribution to provide a basis for assigning probability to the joint occurrence of these two statements (events), regardless of whether in fact one exists. Note that this assumption amounts to projecting an aggregation property onto the problem domain, which here is effected by a projection onto the structure of the agent's beliefs about the problem domain. But this aggregation property is neither a necessary feature of reasonably held beliefs, a necessary feature of data, nor a necessary feature of the world from which this data is drawn; rather, it is a necessary feature of any extended, real-valued, non-negative, countably additive set function defined on a ring.

The point to highlight here is that standard incoherence arguments run against the evidential notion of rational acceptance presuppose that the mathematical properties of measure functions are necessary features of rational belief. But the lottery paradox was conceived to spotlight the problems inherited by construing evidential probability in measure theoretic terms.[20] The failure of the evidential notion of rational acceptance to agree with the tenants of probabilism is a consequence of rejecting probabilism, not a failure to see the consequences of probabilism.

There have been attempts to relax Bayesian constraints to accommodate a more realistic model of statistical reasoning in general, and a more realistic model of the relationship between evidence and probability in particular. But moves in this direction quickly take the theory beyond the philosophical grounds

[20] See Neufeld and Goodman's (1998) and Hunter (1996) for insightful discussions of this point.

that support orthodox Bayesianism, which is to say that one quickly runs afoul of the basic tenants of probabilism.[21]

This summary highlights two particular weaknesses of Bayesianism. But the reader may wonder whether there are positive grounds for why a logic for evidential probability should avoid probabilistic coherence constraints. The short answer is that by doing so one can get a unified theory of statistical inference that includes both direct inference and statistical estimation as special instances (Kyburg 1961, 1974; Kyburg and Teng 2001). Kyburg's theory yields a genuine inductive logic, one that is built atop a non-monotonic consequence relation, which allows for a theory of direct inference whereby new evidence can increase the precision in upper and lower probability assignment of a statement, *even in cases where there is no prior empirical evidence whatsoever* about that statement, that is even in the case where the prior assignment to the statement is the unit interval [0,1]. There are several novel features of Kyburg's view that are a consequence of these features, including a theory of measurement that, among other things, provides an account of theory selection based upon epistemic criteria in addition to practical considerations, such as simplicity and elegance. Kyburg also articulates the rudiments of an imprecise decision theory.[22]

The irony of Jeffrey's "loose talk" remark is that he's right: Kyburg's program is specifically designed to remove the illusion of rigor created by precise but arbitrarily assigned prior values, and to provide a theory of uncertain inference, i.e., a theory for the assignment of probability based upon evidence. Moreover, this entire program is constructed atop the evidential notion of rational acceptance.

Kyburg may be wrong about all of this, but the reason is not because he has failed to recognize the coherence constraints imposed by the Carnap-Jeffrey conception of rational acceptance: Kyburg has persistently maintained that probabilistic coherence plays no central role in assigning probability based upon evidence. The issue is not that evidential probability fails to abide by probabilism; the issue is why evidential probability should be so constrained.

2 Formal Solutions

Suppose that we divorce the lottery paradox from its original motivations—as many commentators are inclined to do. One question we might ask is whether there are formal constraints on the candidate solutions described informally in the literature. If it is impossible to formally represent a particular strategy, then this impossibility result would seem to discredit that strategy.

However, addressing this question is subtle, for it raises both methodological and substantive issues about formal knowledge representation. The reason why

[21] See Peter Walley's (1991) for critical discussion. See Jon Williamson's contribution for a discussion of these issues from an objective Bayesian point of view, and Jan-Willem Romeyn (2005) for a subjective Bayesian view of statistical reasoning.

[22] See Kyburg's (1990) for a programmatic overview of his philosophy of science built atop this approach to uncertain inference, and Kyburg and Teng (2002) for recent formal epistemology.

this is delicate is that the lottery paradox is an applied logic problem rather than a pure logic problem. So solutions will be judged by the fit between the problem domain and the formal framework, which will be guided in part by non-logical properties of rational acceptance. Since applied logic is a branch of applied mathematics, three factors influence this evaluation. The first two are the properties of the formal framework and the properties of the problem domain. The third factor is the intended purpose of the model. What we intend to *do* with the solution must be specified before we can evaluate it. For example, a framework that is illuminating from a semantic point of view may be useless for computational modeling.

In sum, a complete standard must identify the appropriate features of rational acceptance to be captured by the syntax and semantics of the representation language and we must specify what we would like the framework to do.

In this section we will see how these three factors impact one another by discussing the relationship between a pair of properties that are often mentioned in informal treatments of the lottery paradox, *defeasible inference* and *conjunctive closure*, and their formal counterparts, namely *nonmonotonic consequence relations* and *aggregation properties* of a particular class of logics that we will specify later. Our treatment will be far from exhaustive. Rather, we will look at two recent impossibility results, one that targets nonmonotonic logics, the other which (in effect) targets logics with weakened aggregation properties. The goal here is to pull together some formal results that are relevant to the lottery paradox, and to demonstrate the subtle interplay between candidate logics, the features we think are constitutive of the concept of rational acceptance, and how our intended purpose for the framework impacts our preference ordering of the constraints that should be satisfied.

2.1 Logical Structure and Nonmonotonicity

The view that there are nonmonotonic argument structures goes at least as far back as R. A. Fisher's (1922; 1936) observation that statistical reduction should be viewed as a type of *logical*, non-demonstrative inference. Kyburg, following Fisher, has long stressed the nonmonotonic character of uncertain inference in his assaults on Bayesianism, along with stressing the importance of imprecise and non-comparable probabilities, and the weak force that probabilistic coherence constraints exhibit within models of evidential relations.

Unlike demonstrative inference from true premises, the validity of a non-demonstrative, uncertain inference can be undermined by additional premises: a conclusion may be drawn from premises supported by the total evidence available now, but new premises may be added that remove any and all support for that conclusion.

This behavior of non-demonstrative inference does not appear in mathematical arguments, nor should it. If a statement is a logical consequence of a set of premises, that statement remains a consequence however we might choose to augment the premises. Once a theorem, always a theorem, which is why a theorem may be used as a lemma: even if the result of a lemma is misused in an incorrect proof, the result of the lemma remains. We do not have to start the argument again from scratch.

Formal approaches to the lottery paradox often mention the *structure* of rational acceptance, the *logic* of rational acceptance, or the *formal properties* of rational acceptance. But observe that logical structure can mean many things. Logical structure may mean what mathematical logicians mean by classical consequence preserving uniform substitution of arbitrary formulas for elementary formulas within any formula in the language. In other words, logical structure may refer to any substitution function with domain and range on the free algebra of boolean formulas that is an endomorphism.

This sense of logical structure marks an important limit point. Let $\mathcal{L}$ be the set of Boolean formulas, $\Gamma \subseteq \mathcal{L}$, $\phi \in \mathcal{L}$ and '$\vdash$' denote the relation of logical consequence. Then define the operation of logical consequence Cn as $Cn\{\Gamma\} = \{\phi : \Gamma \vdash \phi\}$. A consequence relation $\mid\!\sim$ is *supraclassical* if and only if $\vdash \subseteq \mid\!\sim \subseteq 2^{\mathcal{L}} \times \mathcal{L}$, and a consequence operation C is supraclassical if and only if $Cn \leq C$.

The reason that the mathematical logician's notion of logical structure is important is that logical consequence is maximal with respect to uniform substitution: there is no nontrivial supraclassical closure relation on a language $\mathcal{L}$ that expresses logical consequence that is closed under uniform substitution except for logical consequence.[23]

We will consider other senses of logical structure shortly, but first let's consider whether it makes sense to classify non-monotonic logics as systems of logic. After all, how can a logic fail to be monotone?

Some authors who address this question appear to have this strong notion of logical structure in mind when arguing that nonmonotonic logics are not properly classified as logics. For example, Charles Morgan (1998, 2000) has argued that any consequence relation must be weakly, positively monotone.[24]

Let's study Morgan's argument in some detail, for doing so will reveal a point about how nonlogical properties of epistemic notions—in this case, the concept of joint acceptability—impacts our evaluation of formal results.

Rather than build his argument directly over Boolean languages, Morgan aims for a more general result independent of the structure of the logical connectives of a particular language by building his main argument in terms of *belief structures*. A belief structure is a semi-ordered set whose elements are sets of sentences of a language $\mathcal{L}$. The focus of interest is the ordering relation LE defined over arbitrary sets, Γ and Δ. The idea is that a set of sentences is intended to represent "an instantaneous snapshot of the belief system of a rational agent" while $\Gamma\ LE\ \Delta$ expresses that "the degree of joint acceptability of the members [sentences] of Γ is less than or equal to the degree of joint acceptability of the

[23] See Makinson (2005, p. 15) for discussion of this theorem.

[24] In Morgan (2000) there are actually 4 theorems given which aim to establish this impossibility result, viewed as an argument by four cases. The result we discuss here is the first of these theorems, and is first offered in (1998). It is the most general of his arguments.

members [sentences] of Δ" (2000, p. 328). On Morgan's view, a logic **L** is a set of arbitrary *rational* belief structures $\{LE_1, LE_2, ...\}$, where a rational belief structure is a subset of $\mathcal{P}(\mathcal{L}) \times \mathcal{P}(\mathcal{L})$ satisfying the structural properties of

> **Reflexivity**: $\Gamma\ LE\ \Gamma$,
> **Transitivity**: If $\Gamma\ LE\ \Gamma'$ and $\Gamma'\ LE\ \Gamma''$, then $\Gamma\ LE\ \Gamma''$, and
> **The Subset Principle**: If $\Gamma \subseteq \Delta$, then $\Delta\ LE\ \Gamma$.

Soundness and completeness properties for **L** with respect to a provability relation $\vdash_b$ are defined as follows:

> **Soundness**: If $\Gamma \vdash_b A$, then $\Gamma\ LE\ \{A\}$ for all (most) rational belief structures $LE \in \mathbf{L}$.
> **Completeness**: If $\Gamma\ LE\ \{A\}$, then $\Gamma \vdash_b A$ for all (most) rational belief structures $LE \in \mathbf{L}$.[25]

Morgan's proposal, then, is for **L** to impose minimal prescriptive restrictions on of belief structures, selecting the most general class that are "rational":

> Each distinct logic will pick out a different set of belief structures; those for classical logic will be different from those for intuitionism, and both will be different than those for Post logic. But the important point is that from the standpoint of Logic **L**, all and only the belief structures in **L** are rational (Morgan 2000: 329).

Hence, if every arbitrary set of rational belief structures is monotonic, then every logic must be monotonic as well—which is precisely the result of Morgan's Theorem 1.

Theorem 1. *Let **L** be an arbitrary set of rational belief structures which are reflexive, transitive, and satisfy the subset principle. Further suppose that logical entailment $\vdash_b$ is sound and complete with respect to the set **L**. Then logical entailment is monotonic; that is, if $\Gamma \vdash_b A$, then $\Gamma \cup \Delta \vdash_b A$.*

Proof. (Morgan 2000):

1. $\Gamma \vdash_b A$	given.
2. $\Gamma\ LE\ \{A\}$ for all $LE \in L$.	1, soundness.
3. $\Gamma \cup \Delta\ LE\ \Gamma$ for all $LE \in L$.	subset principle
4. $\Gamma \cup \Delta\ LE\ \{A\}$ for all $LE \in L$.	2, 3 transitivity.
5. $\Gamma \cup \Delta \vdash_b A$	4, completeness.

The normative force of Theorem 1 rests upon the claim that *reflexivity, transitivity* and *the subset principle* are appropriate rationality constraints for belief structures.

Consider Morgan's case for the subset principle:

[25] The inclusion of 'most' in each construction appears only in (2000). Morgan states that a corollary to Theorem 1 may be established if 'most' means more than 50% (2000: 330).

> [The subset principle] is motivated by simple relative frequency considerations....[A] theory which claims both A *and* B will be more difficult to support than a theory which claims just A. Looking at it from another point of view, there will be fewer (or no more) universe designs compatible with both A and B than there are compatible with just A. In general, if Γ makes no more claims about the universe than Δ, then Γ is at least as likely as Δ (2000: 329).

Morgan's argument appears to be that a set of sentences Δ has fewer compatible universe designs than any of its proper subsets, so Δ is less likely to hold than any of its proper subsets. Hence, Δ is harder to support than any of its subsets. Therefore, the degree of support for a set Δ is negatively correlated with the number of possible universe designs compatible with Δ.

There are, however, two points counting against this argument. First, it is misleading to suggest that the subset principle is motivated by relative frequency considerations. Morgan is not referring to repetitive events (such as outcomes of a gaming device) but rather to the likelihood of universe designs, for which there are no relative frequencies for an epistemic agent to consider. Second, even though there are fewer universe designs that satisfy a given set of sentences than with any of its proper subsets, this semantic feature of models bears no relationship to the degree of joint acceptability that may hold among members of an arbitrary set of sentences: "Harder to satisfy" does not entail "harder to support". Relations such as *prediction* and *justification* are classic examples of epistemic relations between sentences that bear directly on the joint epistemic acceptability of a collection of sentences but do not satisfy the subset principle.[26]

To see this last point, recall the notion of joint acceptability that underpins the interpretation of LE. The subset principle is equivalent to the following proposition:

Proposition 1. *If a set of sentences Γ is a subset of the set of sentences Δ, then the degree of joint acceptability of the members of Δ is less than or equal to the degree of joint acceptability of Γ.*

Proposition 2, however, is false. Suppose that a hypothesis H predicts all and only observations $o_1, ..., o_n$ occur for some $n > 1$. Then H receives maximal evidential support just when o_1 occurs and ... and o_n occurs, which is represented by sentences $O_1, ..., O_n$, respectively. Hence, it is more rational to accept a set Δ consisting of $H, O_1, ..., O_n$ than Δ without some observation statement O_i, since the set $\{O_1, ..., O_n\}$ is better support for H than any of its proper subsets. Let $\Gamma = \{H, O_1, ..., O_{n-1}\}$. Then $\Gamma \subseteq \Delta$ but the joint degree of acceptability of Δ is not less than the joint degree of acceptability of Γ.

[26] Note also that these particular notions are not outside the scope of Morgan's program as he conceives it: "[O]ur rational thought processes involve modeling and predicting aspects of our environment under conditions of uncertainty. Here we will assume that our formal language is adequate for such modeling and for the expression of claims about our world that are important to us" (2000, p. 324).

The behavior of joint acceptability of sentences evoked by this example is common and reasonable. Ångström measured the wavelengths of four lines appearing in the emission spectrum of a hydrogen atom (410 nm, 434 nm, 486 nm, and 656 nm), from which J.J. Balmer noticed a regularity that fit the equation

$$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right), \text{when } n = 3, 4, 5, 6,$$

$R = 1.097 \times 10^7 m^{-1}$ is the *Rydberg* constant, and λ is the corresponding wavelength. From the observations that $\lambda = 410$ nm iff $n = 3$, $\lambda = 434$ nm iff $n = 4$, $\lambda = 434$ nm iff $n = 5$, and $\lambda = 656$ nm iff $n = 6$, Balmer's hypothesis—that this equation describes a series in emission spectra beyond these four visible values, that is for $n > 6$—is predicated on there being a series of measured wavelengths whose intervals are specified by $\frac{1}{\lambda}$. Later experimenters confirmed over time that the Balmer series holds for values $n > 6$ through the ultraviolet spectrum. (We now know that it does not hold for all of the non-visible spectrum, however.) The point behind this historical example is that the grounds for a Balmer series describing the emission spectrum of a hydrogen atom *increased* as the set of confirmed values beyond Ångström's initial four measurements increased.

Returning to Morgan, the normative force of his impossibility result rests on a particular view about the structure of joint acceptability. What we have demonstrated in the discussion is that it is unreasonable to interpret the formal constraints of a belief structure to be normative constraints on joint acceptability. In other words, belief structures are not a good formal model for joint acceptability. So, it is irrelevant that belief structures are necessarily monotone.

2.2 System P and Aggregation

Nevertheless we might wonder whether there is any logical structure to nonmonotonic logics. Even if Morgan's views on joint acceptability are mistaken, perhaps he's right about the broader point that nonmonotonic logics fail to have enough structure to be properly classified as logics. To attack this question, first observe three standard properties that classical consequence $\vdash$ enjoys

Reflexivity ($\alpha \vdash \alpha$),
Transitivity (If $\Gamma \vdash \delta$ for all $\delta \in \Delta$ and $\Delta \vdash \alpha$, then $\Gamma \vdash \alpha$), and
Monotonicity (If $\Gamma \vdash \alpha$ and $\Gamma \subseteq \Delta$, then $\Delta \vdash \alpha$),

which were thinly disguised in Morgan's constraints on belief structures. What we want to investigate is whether these three properties are necessary to generate a non-trivial consequence relation. Or, more specifically, we wish to investigate whether there are any non-trivial consequence relations that do not satisfy the monotonicity condition.

Dov Gabbay (1985) noticed that a restricted form of transitivity was helpful in isolating a class of nonmonotonic consequence relations which nevertheless enjoy many properties of classical consequence. The result is important because it reveals that, *pace* Morgan, monotonicity isn't a fundamental property of consequence relations.

Gabbay observed that the general transitivity property of $\vdash$ is entailed by reflexivity, monotonicity, and a restricted version of transitivity he called *cumulative transitivity*:

Reflexivity ($\alpha \vdash \alpha$),
Cumulative Transitivity (If $\Gamma \vdash \delta$ for all $\delta \in \Delta$ and $\Gamma \cup \Delta \vdash \alpha$, then $\Gamma \vdash \alpha$), and
Monotonicity (If $\Gamma \vdash \alpha$ and $\Gamma \subseteq \Delta$, then $\Delta \vdash \alpha$).

This insight opened a way to systematically weaken the monotonicity property by exploring relations constructed from reflexivity and cumulative transitivity, which yields the class of cumulative nonmonotonic logics.[27]

Since Gabbay's result, many semantics for nonmonotonic logics and conditional logics have been found to share a core set of properties identified by Kraus, Lehmann and Magidor (1990), which they named axiom System P. System P is defined here by six properties of the consequence relation $\mid\!\sim$:

Reflexivity	$\alpha \mid\!\sim \alpha$
Left Logical Equivalence	$\dfrac{\models \ulcorner \alpha \leftrightarrow \beta \urcorner; \ \ \alpha \mid\!\sim \gamma}{\beta \mid\!\sim \gamma}$
Right Weakening	$\dfrac{\models \ulcorner \alpha \rightarrow \beta \urcorner; \ \ \gamma \mid\!\sim \alpha}{\gamma \mid\!\sim \beta}$
And	$\dfrac{\alpha \mid\!\sim \beta; \ \ \alpha \mid\!\sim \gamma}{\alpha \mid\!\sim \ulcorner \beta \wedge \gamma \urcorner}$
Or	$\dfrac{\alpha \mid\!\sim \gamma; \ \ \beta \mid\!\sim \gamma}{\ulcorner \alpha \vee \beta \urcorner \mid\!\sim \gamma}$
Cautious Monotonicity	$\dfrac{\alpha \mid\!\sim \beta; \ \ \alpha \mid\!\sim \gamma}{\ulcorner \alpha \wedge \beta \urcorner \mid\!\sim \gamma}$

System P is commonly regarded as the core set of properties that every nonmonotonic consequence relation should satisfy.[28] This assessment is based primarily on the observation that a wide range of nonmonotonic logics have been found to satisfy this set of axioms,[29] including probabilistic semantics for conditional logics.[30]

[27] See Makinson (2005) for discussion.
[28] See for instance the review article by Makinson (1994) and textbook treatment in Halpern (2003).
[29] An important exception is Ray Reiter's default logic (1980). See Marek and Truszczynski (1991) and Makinson (2005) for textbook treatments of non-monotonic logic.
[30] See Judea Pearl (1988, 1990) which is developed around Adams' infinitesimal ϵ semantics, and Lehman and Magidor (1990) which is built around non-standard probability.

However, some authors have drawn a stronger conclusion, namely that System P marks minimal *normative* constraints on nonmonotonic inference.[31] For instance, an impossibility argument may be extracted from Douven and Williamson (2006) to the effect that no coherent probabilistic modeling of rational acceptance of a sentence p *can* be constructed on a logic satisfying axiom System P for values of $p < 1$, and that any "formal" solution to the lottery paradox must have at least this much structure.[32]

But it is one thing to assert that many non-monotonic logics are cumulative, or that it is formally desirable for non-monotonic logics to satisfy system P to preserve horn rules, say, or that there is no coherent and non-trivial probability logic satisfying System P, and it is quite another matter to say that non-monotonic argument forms *should* satisfy System P, or to say that a logical account of rational acceptance *must* minimally satisfy the [**And**] rule of System P.

There are good reasons to think that statistical inference forms do not satisfy the [**And**], [**Or**] and [**Cautious Monotonicity**] axioms of System P. See (Kyburg, Teng and Wheeler 2006) for the KTW axioms for evidential probability, which includes counterexamples to these three properties of System P and discussion. And there are logically interesting probabilistic logics that are weaker than P. Among the weakest systems is System Y (Wheeler 2006), which preserves *greatest lower bound* on arbitrary joins and meets of probabilistic events, each of whose marginal lower probability is known but where nothing is known about the probabilistic relationship between the collection of events. This logic preserves *glb* by inference rules, called *absorption*, for combining conjunctive and disjunctive events, and it preserves this bound purely on the structural features of the measure.

An interesting feature of System P is that it weakens the link between monotonicity and demonstrative inference: **Cautious Monotonicity** and the [**And**] rule together specify the restricted conditions under which nonmonotonically derived statements are aggregated.

Nevertheless, several authors reject monotone consequence operations *because* of the aggregation properties that are preserved under System P. This has led to several studies of non-aggregative logics that in effect reject or restrict the [**And**] rule.[33]

Rather than repeat the arguments for why a logic for rational acceptance *should* be weakly aggregative, let us instead highlight three issues that sub-P

[31] See Makinson (1994).

[32] Although Douven and Williamson do not mention System P nor Gabbay's result, the weakened form of transitivity they discuss in their footnote 2 is cumulative transitivity, and the generality they gesture toward here suggests that they are discussing the class of cumulative nonmonotonic logics.

[33] The history of non-aggregative logics precedes and has developed independently from the study of non-monotonic logics. See particularly (Jaskowski 1969), da Costa (1974), Schotch and Jennings (1989), Brown (1999) Arló-Costa (2002, 2005), Wheeler (2004; 2006), Kyburg, Teng and Wheeler (2006), Hawthorne (2006), Hawthorne and Makinson, forthcoming.

logics face. The first concerns the syntactic functions one would like the logic to enjoy. Minimally we would like facilities in the object language for manipulating sentences that are rationally accepted. On this view rational acceptance is treated as a semantic value (or operator) assigned (attached) to sentences of the language, and our interest is to formulate rules for manipulating sentences purely on the basis of logical combinations of statements having this semantic value or being under the scope of such an operator. This is to be contrasted to a model theoretic approach under which one constructs all possible combinations of sentences that are probabilistically sound. One of the general challenges to formulating an adequate probabilistic logic is that attempts to introduce genuine logical connectives into the object language often erode the precision of the logic to effect proofs of all sound combinations of formulas. The short of it is that an inductive logic enjoying minimally interesting modularity properties in the object language may not be complete.[34]

Another issue facing sub-P logics is to specify the relationship that weakly aggregative logics have to the notion of rational acceptance we wish to model. For instance, the KTW axioms (Kyburg, Teng and Wheeler 2006) articulate a weakly aggregative logic that comes closest to axiomatizing evidential probability, although disjunction in this system is weaker than what one would expect from a standard probabilistic logic, and also different from non-aggregative modal logics. There are other conceptions of rational acceptance that authors have attempted to model as well.[35]

Even if we fix upon a structure for rational acceptance, it should be noted that while a sub-P logic may offer the most accurate representation of that structure from a knowledge representation point of view, this logic will be too weak on its own to be of much inferentially use. This is not an argument for interpreting System P as normative constraints on nonmonotonic reasoning. But it is a recognition that the "given" structure of a problem domain, in this case the features of rational acceptance, may be considerably weaker than the minimal structure of an effective framework. Nevertheless, it is crucial to formally articulate what the given structure of a problem domain *is* in order to understand what structural properties must be added to, or assumed to hold within, a problem domain in order to generate an effective representation. Approaching

[34] An example of a purely model theoretic approach to the paradox is (Bovens and Hawthorne 1999), whereas (Brown 1999) and (Wheeler 2006) are examples of syntactic approaches to the paradox. See (Wheeler 2005) for a discussion of Bovens and Hawthorne and arguments in favor of syntactic approaches.

[35] Erik Olsson (2004) views Hawthorne and Bovens's logic of belief (1999) as giving us a picture of Kyburg's theory in subjectivist terms, but Olsson remarks that the difficulty facing the logic of belief "is to explain why in the first place the agent should take the trouble of separating from other propositions those propositions in whose veracity she has a relatively high degree of confidence. What is the point of this separation business?" In reply, we've seen the motivation for rational acceptance in the first section. So, the question then is whether the logic of belief (or any of its competitors) represent the Kyburgian notion of rational acceptance? The answer here is clearly, No.

the problem in this manner puts us in a position to articulate precisely what assumptions stand between an accurate formal representation of the problem and an effective model of that problem. And having the capacities to pinpoint these assumptions in this manner is necessary to evaluate when it is reasonable to make them.

The moral of this section is that formal solutions to applied logic problems are influenced both by the relationship between the formal language and the non-logical structure of the problem, and by the purpose or intended application of the formal model. Simply because a formal framework has a certain set of attractive mathematical or computational properties does not entail that the problem domain it is intended to model must have that structure. Assessing the basic structure of the problem domain is necessary before making a claim that a formal framework characterizes normative constraints for that problem domain. Furthermore, we may have good reason to select a less accurate formal representation of a problem domain, such as when that representation is enables computation and we have a measure or control of the error that the framework introduces.

These remarks might suggest to some readers that formal approaches are prone to failure, particularly for philosophically contentious notions like rational acceptance. Nevertheless, it is very important to construct formal solutions that may not directly match the concept or relation to be modeled. The reason is that often we need to understand the space of possible solutions offered by particular frameworks in order to understand the problem and how best to pick apart and recombine properties of various formal languages. The solution to this paradox, like many solutions to applied logic problems, will be in the form of an optimization problem. It will pick the salient properties of rational acceptance to model, and capture these features in a formal framework that balances precision against usefulness.

2.3 Applied Logic and Psychologism

We've seen that logical structure comes in degrees, and that it is often important for a framework to have the capabilities to represent very weak structures. A generalized view of logical structure that is more suitable for applied and philosophical logic problems is Arnold Koslow's (1992) notion of logical structure. Koslow considers logic to be concerned with the study of structures on arbitrary domains of elements that are generated by introducing a consequence relation of some kind or another over these elements. There is again a sense of closure under uniform substitution, but Koslow does not restrict the scope to only boolean languages, nor restrict the consequence relation to classical consequence.[36]

The trade-offs involved in settling on an appropriate applied logic impact what sense of logical structure one should demand. In addition to these considerations, there is the question of whether formal logic is appropriate for representing features of human "reasoning" or "inference". Gilbert Harman (1986,

[36] See Levi, this volume, and Wheeler (2005, 2006) for examples.

2002) has long argued that there is no relationship between logical consequence and the psychological process of drawing an inference, and has pointed out that serious confusion results from thinking of logic as a theory of reasoning. Harman is right about the confusion that results from viewing logic as a theory of reasoning, but he is wrong to assume that formal epistemology must adopt a psychologistic view of logic.

Logic is a branch of mathematics that, like many other branches of mathematics, lends itself to various applications. There is no conceptual confusion from applying logic to model features of a problem domain in general. Likewise, there is no conceptual confusion in using formal languages to model the structure of evidential relations that hold between evidence and probability assignment. The results from attempts to achieve an appropriate model may be disappointing; but we can readily make sense of the aim of the project and evaluate the outcome. The confusion Harman highlights is not a category mistake from failing to observe a distinction between *inference* and *implication*, but a methodological error in thinking that the properties of a logical system give us rules of thought for how to carry out the psychological process of drawing a reasonable inference. There is nothing about the application of logic to study epistemic relations that forces this methodological error upon us.

3 On Being Reasonable But Wrong

We started this essay listing a number of approaches to resolving the lottery paradox, but soon set them aside to review the original motivations for the paradox, what Kyburg's theory of rational acceptance achieves, some of what we've learned about non-classical logics that are relevant to formal solutions to the paradox, and *how* formal work on this paradox gives insight into the underlying structure of the problem. In light of these results, let us return to this literature and consider two issues. The first concerns an argument aimed to show that statistical evidence is an insufficient basis for rational belief, whereas the second highlights why the lottery paradox is about models of full but possibly false belief *simpliciter*, not about a lotteries per se.

3.1 Statistical evidence and rational belief

It is commonly held that statistical evidence is an insufficient basis for knowledge on grounds that the traditional analysis of knowledge is *factive*—known propositions are true propositions.[37] It turns out that Kyburg and Teng's conception of "risky knowledge" (2002) denies that knowledge is factive, but we may sidestep this point for the moment. For one might accept the standard analysis of knowledge and deny that statistical evidence is a sufficient basis for *knowledge*, but nevertheless maintain that statistical evidence can provide grounds to rationally believe a statement. So statistical evidence may be viewed as a sufficient basis for rational belief even if it is not a sufficient basis for knowledge.

[37] For instance, see Cohen (1988).

However, Dana Nelkin (Nelkin 2003) argues that statistical evidence is both an insufficient basis for knowledge and an insufficient basis for rational belief. Her thought is that an agent should grasp a causal or explanatory relationship between a target belief and the facts that make that belief true, which demands information that extends beyond what is provided by statistical evidence.

But Nelkin's position is based upon her commonsense analysis of why it is not rational to believe that a ticket of the lottery loses. And here we may harvest several of the points made above to evaluate her argument.

> If I believe that p (say, my ticket will lose) on the basis of a high statistical probability for p, and I find out that not-p (I won!), then there is nothing at all in my reasons to reject. I still believe the same odds were in effect, and I still believe that they made my losing extremely probable. I have no reason to think that my evidence failed to bear a connection to my conclusion that I previously thought it did. Learning that my belief is false puts no pressure on me to find some problem in my reasons. Thus, there is a way in which they are not "sensitive" to the truth, or at least to what I conceive of as the truth. It seems to me that, given the role of rationality as a guide toward the truth, this lack of sensitivity to the truth in the case of P-inferences [statistical inferences] might help explain why such inferences are not rational (Nelkin 2004, p. 401).

There are two points about this argument I'd like to address.

First, believing that evidence bears a connection to a conclusion is not to suppose that the evidence entails that conclusion. Sometimes perfectly good reasoning from strong evidence leads to a false conclusion, which is reflected in an *epistemic* distinction that is sometimes made between mistakes and errors. Statistical methods are designed to control error, but not to eliminate it. So when a false statement is accepted (or a true statement is rejected) on the basis of statistical evidence, it is perfectly sensible to ask whether a mistake has been made when an accepted statement is false or whether the accepted statement was reasonably drawn but resulted from bad luck. Often we can determine this by evaluating whether accepting the false statement occurred within the margin of error that was accepted from the start as an acceptable risk for performing this uncertain inference. We may ask whether the accepted but false statement was admitted due to a flaw in the experimental design or a mistake in the execution of the experiment—turning to techniques of data analysis to provide us an (uncertain!) answer.

Note that there are analogous methods used for correcting mistakes of common sense reasoning. After learning that a reasonably held belief is false, we often ask whether we thought of everything that could go wrong with the method or evidence base we used to draw this conclusion just as we investigate whether there was a flaw in experimental design. We ask whether we correctly assessed the risks of the evidence we accepted, just as we assess whether the statistical model correctly reflected the error probabilities. We ask whether we correctly thought through the consequences of those assumptions, just as we ask whether we correctly calculated the values within our experimental model. In all of these cases we are investigating whether a mistake was made when it is learned that an accepted conclusion from an uncertain inference is false.

One reason that error is controlled, rather than eliminated, stems from the fact that errors come in two forms: Errors of type (I), which result from accepting as true a falsity, and errors of type (II), which result from failing to accept a truth. It turns out that methods that reduce the frequency of one type of error typically typically increase the frequency of the other. So this failure to eliminate error or to avoid drawing false conclusions is in no way evidence that statistical evidence is "insensitive" to the truth. On the contrary, statistical methodology has developed from a keen sensitivity to how to mitigate the naturally occurring obstacles to the truth that noisy, real-world data and evidence present.

What is less clear in Nelkin's argument is what is meant by rationality guiding us toward the truth. Resolving the lottery paradox is to resolve how reasoning from evidence guides us to true probability assessments. The underlying problem with Nelkin's argument is that the idea of "rationality guiding us to the truth" that she thinks tells against statistical evidence being a cogent basis for acceptance presupposes a solution to the knot of problems that a theory of rational acceptance faces, which the lottery paradox highlights in outline. The difference between demonstrative and non-demonstrative reasoning is that there is no guarantee from a non-demonstrative inference that you will find the cause of a false belief to be a mistake in your reasoning. Nor should it be. If the proportion of F's that are G's is between 0.99 and 1, a is an F and there is no evidence in your corpus that says a isn't a G, it is practically certain that a is a G. Now a might in fact fail to be a G because of a mistaken classification of a among the F's or by over estimating the proportion of F's that are G's. But, a might in fact fail to be a G simply by bad luck.

3.2 Full but possibly false belief

The lottery paradox is a victim of its own success. Intended as a way to install in reader's minds that evidential relations are surprisingly tricky to model formally, it has spawned a literature and sub-culture of its own—much of it informal and focused on intuitions about buying lottery tickets in situations spelled out by the argument. But one problem with keying a solution of the lottery paradox to the particular details of the lottery ticket argument is that inessential features of the thought experiment can wind up as essential features of proposals for how rational acceptance ought to work. For instance, in describing the setup for the lottery paradox, some think that it doesn't make sense to ascribe knowledge or rational acceptability to the claim that ticket i won't win *before* the drawing because the agent knows in advance that one and only one ticket will win. The advice here is for us to recognize that we are in such a "lottery context" and to suspend judgment until the ticket is drawn.

But this would be bad advice to follow in general, since it isn't essential to rational acceptability that the agent know enough about the uncertainty mechanism to await a determinate outcome. When I book a flight from Philadelphia to Denver I judge it rational to accept that I will arrive without serious incident since I stand rough odds of $1 : 350,000$ of dying on a plane trip in the U.S. any given year[38] (and stand considerably better odds than this of arriving without incident if I book passage on a regularly scheduled U.S. commercial flight). I judge the uncertainty mechanism characterizing these odds to be applicable to my flight from Philadelphia to Denver. Nevertheless, I believe that there will be a fatality from an airline accident on a U.S. carrier in the coming year even though I don't believe that there must be at least one fatal accident each year. Indeed, 2002 marked such an exception.[39]

Notice that I am not speaking of my decision to board the plane, but my belief that I stand better than $1 : 350,000$ odds of getting off alive. My epistemological stance toward the belief that I land safely is distinct from what other actions I may be willing to take given this stance, such as booking a car and hotel. In short, I believe that I will survive this flight to Denver. My belief that I will survive the trip is part of my background knowledge I base my decisions about actions I will, or should, take while in Denver.

4 Conclusion

One issue that distinguishes Kyburg's view of uncertain inference from orthodox Bayesian approaches is his sharp distinction between thought and action, and his insistence that each be treated independently. We can find an analogue of this distinction within statistics between *informative inference* and *decision*. We close with a quotation from "Two World Views" that remarks on the importance of this distinction.[40]

> To have a coherent approach to decision and action in the world which is not self-defeating, is important, and the personalistic approach to statistics provides an enormous amount of clarification in this regard. But, we want to understand the world, as well as to act in it, and it is in connection with the mechanisms for rejecting and accepting hypotheses that other approaches to statistics are important. That we do not have to understand the world in order to act coherently in it is true but irrelevant. One might also claim that one need not act coherently in order to understand the world. In point of fact, we want both to act in the world and to understand it ... (Kyburg 1970b, p. 348).

[38] According to the *National Safety Council*

[39] According to the 2002 *National Transportation Safety Board* there were 34 accidents on U.S. commercial airlines during 2002 but zero fatalities, a first in twenty years.

[40] This essay is based upon work supported by FCT award SFRH/BPD-13699-2003 and the Leverhulme Trust. Thanks to Marco Castellani and Stephen Fogdall for their comments.