JAMES HAWTHORNE

THREE MODELS OF SEQUENTIAL BELIEF UPDATING ON UNCERTAIN EVIDENCE

Received in revised version on 16 July 2003

ABSTRACT. Jeffrey updating is a natural extension of Bayesian updating to cases where the evidence is uncertain. But, the resulting degrees of belief appear to be sensitive to the order in which the uncertain evidence is acquired, a rather un-Bayesian looking effect. This order dependence results from the way in which basic Jeffrey updating is usually extended to sequences of updates. The usual extension seems very natural, but there are other plausible ways to extend Bayesian updating that maintain order-independence. I will explore three models of sequential updating, the usual extension and two alternatives. I will show that the alternative updating schemes derive from extensions of the usual *rigidity* requirement, which is at the heart of Jeffrey updating. Finally, I will establish necessary and sufficient conditions for order-independent updating, and show that extended rigidity is closely related to these conditions.

KEY WORDS: Bayesian updating, Jeffrey conditionalization, probability kinematics, probabilistic logic, uncertain evidence

1. INTRODUCTION

Evidence claims are often uncertain. Perhaps they should always be regarded as such. For, sensory experiences must play a prominent role in providing whatever confidence we may legitimately have in the truth of contingent claims. But experiences are non-propositional states. And although some statements are tied to them more directly than others, we have no purely phenomenalistic *observation language* available to precisely express their content. So, there is always at least a bit of epistemic distance between statements that express evidence and the experiences or observations that ground our confidence in them. Evidence statements are fallible and an agent's doxastic state should reflect this by marking them as less than certain.

Although evidence statements may be uncertain, their credibility or empirical support comes more directly from observations and experiences than does the support of other contingent claims. Evidence statements function as conduits through which empirical support is passed to more remote claims. How this transmission may occur is what logical theories

Journal of Philosophical Logic **33:** 89–123, 2004.

of inductive inference are all about, and there is disagreement about the nature of such inferences. But, prima face, the most plausible way to develop a logical model of empirical support that explicitly represents the role of experience is to model it in terms of an evidential link of some sort from non-propositional (non-doxastic, experiential) states of the agent to defeasible belief strengths in some evidence statements that may in turn provide support for the rest of the agent's empirical corpus. In this paper I'll explicate and extend a Bayesian model of belief updating on uncertain evidence, first developed by Richard Jeffrey (1965), that fits this bill extremely well.

The Jeffrey model of empirical support also has a more practical side. Increasingly, the logic of Bayesian inference is utilized in automated expert systems in the form of Bayesian networks. The evidential updating of such systems with claims that are less than certain for one reason or another should prove to be quite useful. Consider, for example, how an automated inference aid for medical diagnosis might function. It has a hierarchy of hypotheses and sub-hypotheses about possible diseases and syndromes that branches down to an evidence level consisting of claims about observed symptoms and test results that may be fairly directly known. Updating the system amounts to telling it which of the possible evidence claims exhibited at its bottom level have been observed in the patient. The system then uses Bayesian inference to update the systems plausibility ratings or "belief strengths" for all hypotheses up the network hierarchy. In the design of such systems it will often be impractical to construct a hierarchy that runs all the way down to the kind of detailed, low level evidence claims that a physician may observe directly, and report with certainty – e.g., to such claims, 'a report from the lab asserts that in a specific sample of this patient's blood the white cell count was 2317 cells per microliter.' Rather, it may be more practical to have the system terminate at a higher level – e.g., 'the patient is very anemic' – call statements at this level "the evidence statements," and require the user to update the system by inputting her belief strengths for these higher level "evidence statements", based on whatever she observes more immediately (e.g., detailed but fallible lab reports). Also, some kinds of evidence may be explicitly reported to the physician with confidence levels attached – e.g., 'the blood test indicates, with degree of confidence .75, the presence of traces of a weak neuron-toxin.' Bayesian networks may accommodate such evidence by permitting the user to attach belief strengths or confidence levels to evidence statements.

Whether the probabilities acquired by uncertain evidence statements result from inexpressible experiences of an agent, or from propositionally expressible reports that fail to be expressible within a given automated system, Jeffrey's model is applicable. For convenience I will refer to inexpressible information that directly affects evidence statements, for an agent or for an expert system, simply as *experiences* or as *non-propositional states* of the agent. This terminology may seem strained when applied to expert systems. The inexpressible *states of the system* used to update its evidential probabilities usually do not reside in the system itself, but in the user or in experts the user employs. In a sense the *agent* in this case is the expert system together with its "sensory apparatus", which consists of the user and his experts.

One of the most controversial issues for the Jeffrey model is whether the *order* in which updating occurs should legitimately influence the values of the resulting belief-strengths of the agent. That is, when a sequence of states brings about a sequence of updates, should the order in which the states are acquired make a difference? Jeffrey realized from the start that the way in which he originally extended his approach to sequences of updates tends to produce a high degree of order-dependence (1965, pp. 161–162). Some investigators, convinced by Jeffrey's work, have concluded that update order should indeed matter – that states responsible for updating should fail to commute. However, such order dependence can lead to highly undesirable effects.

Imagine two Bayesian medical diagnosticians that initially share the same epistemic state. The results of various tests on the patient are delivered to both agents, but arrive in a different order. As a result they acquire quite different belief strengths regarding the various diseases that the patient may have; so they arrive at very different diagnoses. If this kind of problem is to be avoided, Jeffrey updating will need to be re-formulated in a way that produces order independent results.

The extent to which Jeffrey updating depends on order, and whether it should, remains highly controversial. Jeffrey's original scheme for performing sequential updates is clearly order dependent. So, if by 'Jeffrey Updating' one just means Jeffrey's original extension, then the issue is whether that original way is the right way for agents to update their belief functions. However, Jeffrey himself seems to have given up his original proposal for extending updating to sequences. His more recent work embraces an alternative updating scheme first proposed by Hartry Field (1978). So let us use the term 'Jeffrey Updating' broadly, to apply to Basic (non-sequential) Jeffrey Updating and to any of the various ways of extending it to sequences of updates. Let's call Jeffrey's original extension of Basic Updating to sequences 'Standard Sequential Updating'. And let's call Field's alternative extension of Basic Jeffrey Updating, 'Field Updating'. The issue, then, is this: what is the most plausible way to extend Basic Jeffrey Updating to sequences of states? Are there reasons to prefer Standard Sequential Updating, or Field Updating, or some other approach?

Although Jeffrey has revised his view,[1] other researchers still defend Standard Sequential Updating as the correct approach, or as superior to Field Updating.[2] The controversy largely turns on precisely what kind of *factor* at the evidence level should most directly represent the impact of non-propositional states? That is, when a sequence of propositionally inexpressible states brings about a sequence of uncertain updates, should these updates most immediately take the form of *direct changes in probability* for evidence statements, or should they take the form of *update factors* of some other kind that then bring about changes in the probabilities of evidence statements? This issue may sound arcane, but its resolution largely determines the extent to which update order may leave its mark on evidential probabilities, and on the probabilities of all higher level statements that the evidence may influence.

I will explicate the formalism of Jeffrey Updating in a way that shows precisely how Standard Sequential Updating compares with two alternative approaches, one of them a generalization of Field's approach. I will show how these two alternatives to Standard Sequential Updating derive from extensions of the usual *rigidity* requirement, which is the heart of Basic Jeffrey Updating. We will see that on the two alternative schemes updating is independent of order. I will also show how the main objection to Field Updating, due to Garber (1980), may be overcome. Finally, I will establish necessary and sufficient conditions for order-independent updating, and show how closely related these conditions are to the extensions of rigidity employed by the two order-independent update schemes.

2. Basic Jeffrey Updating

Jeffrey's scheme for belief updating is a natural extension of Bayesian updating to cases involving uncertain evidence. The idea is that although important aspects of an agent's experiences may fail to be expressible propositionally, the experiences may, nevertheless, bring about a direct change in her belief strengths in some propositions, which may in turn provide evidence for other propositions.

To express this idea formally, let Q be the agent's degree-of-belief function at a given time. We employ the usual Bayesian idealization. The agent is supposed to be an ideal model of rational belief in the sense that her degrees of belief may be represented by a function from sentences of her language to real numbers between 0 and 1 that satisfies the usual probability axioms. Let e be a propositionally inexpressible state for the agent. Suppose that, corresponding to e, there is a set of sentences, $\{E_i\}$ (associated with e by the agent) that are influenced more directly by e than are other sentences. The influence of state e on the agent's belief strengths for the associated evidences sentences $\{E_i\}$ is represented by a transformation of the agent's previous degree-of-belief function Q to a new belief function Q_e that assigns new belief strengths $Q_e[E_i]$ to sentences in $\{E_i\}$. I will call $\{E_i\}$ the *evidence basis directly affected by e*. An evidence basis is a partition – i.e. a mutually exclusive and exhaustive set of sentences.[3] The evidential import of an experience or observation e is supposed to be completely captured by the influence it has on the belief strengths of sentences of its evidence basis. The agent's belief strengths for other sentences may only be influenced by e through the mediation of e's evidence basis.

The evidence basis captures the evidential import of e through its ability to "screen off" the rest of the agent's beliefs from direct influence by e. That is, for any sentence S in the agent's language, each member of the basis $\{E_i\}$ affected by e satisfies the relationship $Q_e[S \mid E_i] = Q[S \mid E_i]$ (provided $Q_e[E_i] > 0$).[4] This means that each sentence S is independent of state e, given E_i. This screening off condition is sometimes called *rigidity*. It represents the idea that the sentences E_i in the evidence basis directly affected by e already carry all of the evidential import of e that is relevant to the agent's other sentences S. Given that any one of the basis sentence E_i holds, e adds nothing more to the agent's degree of confidence in S.

When rigidity is satisfied, the belief strength for a sentence S due to a state e is determined by the influence of e on sentences of its evidence basis together with the evidential support that each basis sentence supplies S prior to e's introduction: $Q_e[S] = \sum_i Q[S \mid E_i] \cdot Q_e[E_i] = \sum_{\{i:Q[E_i]>0\}} Q[S \cdot E_i] \cdot (Q_e[E_i]/Q[E_i])$. Given rigidity, this formula is just a theorem of probability theory.[5]

This formula implies that if state e makes some sentence E_j in its basis certain (i.e. if $Q_e[E_j] = 1$), then the new probability of a sentence S based on e is $Q_e[S] = Q[S \mid E_j]$, where Bayes' theorem governs $Q[S \mid E_j]$ in the usual way: $Q[S \mid E_j] = Q[E_j \mid S] \cdot Q[S]/Q[E_j]$. Thus, Jeffrey updating is a generalization of Bayesian updating to cases where the evidence remains uncertain.

3. Basic Sequential Jeffrey Updating

The agent may be subject to a sequence of experiences that influence her belief strengths, each through its own evidence basis. Suppose an agent whose present belief function is Q_e has an additional experience f that directly affects an evidence basis $\{F_j\}$, where rigidity holds relative to each F_j – i.e. $Q_{ef}[S \mid F_j] = Q_e[S \mid F_j]$. Then Jeffrey updating iterates to yield, for each sentence S, $Q_{ef}[S] = \sum_{\{j:Q_e[F_j]>0\}} \sum_{\{i:Q[E_i]>0\}} Q[S \cdot E_i \cdot F_j] \cdot (Q_e[E_i]/Q[E_i]) \cdot (Q_{ef}[F_j]/Q_e[F_j])$.[6]

The process of updating on additional uncertain evidence reiterates in the same way. That is, suppose that the agent's current belief function is a probability function Q. The agent's beliefs strengths may be (further) updated by a sequence of non-propositional states $a, b, \ldots, e, f, g$, affecting evidence bases $\{A_i\}, \ldots, \{E_i\}, \{F_i\}, \{G_i\}$, respectively. This produces a sequence of updated belief functions $Q_a, \ldots, Q_{a\ldots ef}, Q_{a\ldots efg}$. Then, supposing *rigidity* holds in each case relative to the respective bases (i.e., supposing for each $Q_{a\ldots b}$ and state c with basis $\{C_i\}$, $Q_{a\ldots bc}[S \mid C_i] = Q_{a\ldots b}[S \mid C_i]$) we have the following result:

BASIC SEQUENTIAL UPDATE FORMULA. For all S, $Q_{a\ldots efg}[S] = \sum_{\{i:Q_{a\ldots ef}[G_i]>0\}} \cdots \sum_{\{k:Q[A_k]>0\}} Q[S \cdot A_k \cdot \cdots \cdot F_j \cdot G_i] \cdot (Q_a[A_k]/Q[A_k]) \cdot \cdots \cdot (Q_{a\ldots efg}[G_i]/Q_{a\ldots ef}[G_i])$.

This formula holds even when some of the non-propositional states affect the same evidence basis. But when two states share a basis some of the terms in the formula reduce. For example, if g and b affect the same basis $\{G_i\}$ (so each B_k in the basis for b is one of the mem bers of $\{G_i\}$), the belief strengths $Q[S \cdot A_n \cdot B_k \cdot \cdots \cdot F_j \cdot G_i]$ must be 0 in all cases where B_k is not the same sentence as G_i. Thus, when states g and b share a basis the update formula becomes $Q_{a.b\ldots efg}[S] = \sum_{\{i:Q_{a.b\ldots ef}[G_i]>0\}} \sum_{\{j:Q_{a\ldots e}[F_j]>0\}} \cdots \sum_{\{n:Q[A_n]>0\}} Q[S \cdot A_n \cdot \cdots \cdot F_j \cdot G_i] \cdot (Q_a[A_n]/Q[A_n]) \cdot (Q_{ab}[G_i]/Q_a[G_i]) \cdot \cdots \cdot (Q_{ab\ldots ef}[F_j]/Q_{ab\ldots e}[F_j]) \cdot (Q_{ab\ldots efg}[G_i]/Q_{ab\ldots ef}[G_i])$.

Also notice that if a *final sequence* of states affect a common basis, they may be treated like a single state. That is, suppose ϵ is a sequence of states (e.g., ϵ is a sequence $e \ldots fgh$) that occurs just after a sequence $a, \ldots, c, d$. And suppose all states in ϵ affect the same basis $\{E_i\}$. Then the Basic Sequential Update Formula becomes $Q_{a\ldots cd\epsilon}[S] = \sum_{\{i:Q_{a\ldots cd}[E_i]>0\}} \sum_{\{j:Q_{a\ldots c}[D_j]>0\}} \cdots \sum_{\{k:Q[A_k]>0\}} Q[S \cdot A_k \cdot \cdots \cdot D_j \cdot E_i] \cdot (Q_a[A_k]/Q[A_k]) \cdot \cdots \cdot (Q_{a\ldots cd}[D_j]/Q_{a\ldots c}[D_j]) \cdot (Q_{a\ldots cd\epsilon}[E_i]/Q_{a\ldots cd}[E_i])$.[7]

The Basic Sequential Update Formula follows solely from the usually axioms for probabilities together with *rigidity* (applied to each basis). So it is just part of Basic Jeffrey Updating, not an *extension* of it. Standard Sequential Updating, which is the usual extension of Jeffrey updating to sequences, goes further. It requires an additional assumption about how probabilities of basis sentences are to be updated by non-propositional states. We are now prepared to explore it, and to investigate the two alternative extensions to Basic Jeffrey Updating.

4. Update Factors

An experience e updates a belief function Q_β, previously updated on a sequence of states β, by supplying new probabilities $Q_{\beta e}[E_j]$ for e's basis sentences $\{E_i\}$. Let us call these new probabilities *probabilistic update factors*. Does experience *directly* supply probabilistic update factors to the agent, or does it *more directly* supply some other kind of update factor, a factor which in turn produces the probabilistic updates of basis sentences. This issue has important implications for how sequential Jeffrey updating works.

If probabilistic update factors are derivative of some more primitive factors, what might these more primitive update factors be? Well, notice that the ratios of form $(Q_{\beta e}[E_j]/Q_\beta[E_j])$ in the Basic Sequential Update Formula represent the amount by which the addition of e to β increases the "prior" degree of belief $Q_\beta[E_j]$ to produce the "posterior" degree of belief $Q_{\beta e}[E_j]$, resulting from the addition of e. Because these ratios play such a prominent role in belief updating, we might suppose that they are stand-ins for some kind of autonomous factor, call it a *Normed-Likelihood update factor*, $\mathrm{NL}[Q_\beta, e, E_j]$, that generates the posterior belief function $Q_{\beta e}[E_j]$ from the prior belief function $Q_\beta[E_j]$, as follows: $Q_{\beta e}[E_j] = \mathrm{NL}[Q_\beta, e, E_j] \cdot Q_\beta[E_j]$.[8] The thought is that these factors may be more immediately generated by non-propositional states than are the updated probabilities for bases. These factors then pass on the influence of a new state to the probabilities of basis sentences by updating the prior belief strengths for basis sentences to produce posterior belief strengths. That is, the state projects some kind of weightings onto its basis sentences; and these weightings are measurable on a scale that employs the non-negative real numbers and has the structural features of normalized likelihoods.

Inspection of the Basic Sequential Update Formula shows that sequential Jeffrey updating depends *only* on Normed-Likelihood update factors and on the agent's belief function Q prior to the sequence of updates. That is, we may rewrite the Basic Sequential Update Formula as follows:

BASIC SEQUENTIAL UPDATE FORMULA: NORMED-LIKELIHOOD FACTOR VERSION. For all S, $Q_{a\ldots efg}[S] = \sum_{\{i: Q_{a\ldots ef}[G_i]>0\}} \cdots \sum_{\{k: Q[A_k]>0\}} Q[S \cdot A_k \cdot \cdots \cdot G_i] \cdot \mathrm{NL}[Q, a, A_k] \cdot \cdots \cdot \mathrm{NL}[Q_{a\ldots ef}, g, G_i]$.

An agent's initial belief function Q together with the Normed-Likelihood update factors affected by states in α suffice to generate the updated degrees of belief Q_α for all of her sentences.

So, which kind of update factor do non-propositional states supply more directly, *Normed-Likelihood update factors* $\mathrm{NL}[Q_\beta, e, E_j]$, or *Probability* *update factors* $Q_{\beta e}[E_j]$, or some other kind of update factors? Notice that each of the two kinds of update factors we've explored so far may be used to generate the other from the prior probabilities of basis sentences. So the issue regarding which kind of factor should be taken as primitive is not a purely mathematical matter. It is an epistemological, or an empirical, or a pragmatic issue.[9]

Rather than try to settle on one kind of factor as the *right one*, we may develop several theories, each an extension of Jeffrey updating that takes a different kind of update factor as more primitive – as affected more directly by non-propositional states. Indeed, it may turn out that there is no one true theory of uncertain updating – that each theory has its uses, its domain of applicability. However, the various theories will, it turns out, suggest different natural extensions of Basic Sequential Jeffrey Updating that have different mathematical characteristics. And these characteristics might make the various extensions more or less suitable as models of certain types of agents. In particular, taking a factor as primitive tends to lead to differences regarding whether update order influences the agent's net belief strengths.

5. THE AMNESTIC UPDATE MODEL

Jeffrey's original model takes the experiential state e to supply new probabilities $Q_{\beta e}[E_i]$ directly to the agent, unmediated by any other factor. Normed-Likelihood factors are then an artifact, *defined* as ratios of the probabilistic update factors $Q_{\beta e}[E_i]$ over initial belief strengths $Q_{\beta}[E_i]$. Those who hold this view go further. They take each new experience e to so overwhelm the agent's belief function on its basis sentences that no trace of the impact of previous experiential states remains. That is, they adopt the following thesis:

AMNESTIC UPDATE-FACTOR THESIS. For any state e that directly affects an evidence basis $\{E_i\}$ and for any other state d and sequence of states α, $Q_{\alpha de}[E_i] = Q_{\alpha e}[E_i]$.

Let's say that a sequence of belief functions connected through successive updates on non-propositional states fits the *Amnestic Update Model* just in case the sequence satisfies both the usual Jeffrey *rigidity* requirement and the *Amnestic Update Thesis*. Amnestic updating provides one fairly natural way to extend Basic Sequential Jeffrey Updating. Indeed Amnestic Updating is just Standard Sequential Updating – Jeffrey's original approach to sequential updating.[10]

The Amnestic Model treats the agent as subject to a strange sort of amnesia. Her belief strengths in statements that most immediately reflect the evidence of her senses can only accommodate her most recent relevant experience. On this model, when I next see my car in dim light, its appearance at that moment completely overwrites my beliefs regarding its true color. This seems highly implausible. One might reasonably think that my belief strengths regarding its true color should be the product of my present experience in combination with past experiences of a similar kind, together with many other relevant beliefs – e.g., that it usually looks about like this in the dark, that it is newly waxed, that waxing has not previously resulted in color change, etc.

When the Amnestic Update Thesis is applied to the Basic Sequential Update Formula we find that the degree of belief $Q_{a...ef}[S]$ after experiences $a \ldots ef$ ultimately depends only on the initial belief function Q and the influence of states on their bases. Furthermore, the belief strengths for basis sentences, $Q_a[A_i], \ldots, Q_f[F_j]$, $Q_g[G_k]$, do not depend on the initial belief function Q at all, but only on the experiential states themselves. When the Basic Sequential Update Formula is decomposed into these components, the update order effect on sentence S is carried by the structure of the resulting formula.[11]

Some instances of Amnestic Updating may turn out to be independent of order. But this can only occur under very special circumstances. The order effect in Amnestic Updating is so strong that a pair of states e and f will commute for Q_β (i.e. $Q_{\beta ef}[S] = Q_{\beta fe}[S]$ for all S) *just in case* neither e nor f can, on its own, influence (even indirectly) the basis sentences of the other – i.e. *just in case* for each E_i, $Q_{\beta f}[E_i] = Q_\beta[E_i]$, and for each F_j, $Q_{\beta e}[F_j] = Q_\beta[F_j]$.[12] However, usually states do indirectly influence the belief strengths of sentences from other bases. Thus, on the Amnestic Model update order-independence will be rare.[13]

To see how strongly update order on distinct bases can influence belief strengths, let us consider an example. Consider a case where the results of two diagnostic tests, a chest x-ray and a sputum cytology test, supply uncertain evidence regarding whether a patient has lung cancer. I will keep the example simple and symmetric so that the effect of update order is relatively easy to trace. Let us suppose the following: given the patient's symptoms, the physician's initial degree of belief that "the patient has some form of lung cancer", $Q[C]$, is .5; the likelihood that "an image of a mass is present on the x-ray" if "cancer is present", $Q[E \mid C]$, is .95, and (suppose) the likelihood that "an image of a mass shows up" if "no cancer is present", $Q[E \mid {\sim}C]$, is .05; the likelihood that "cancer-like cells are present in the sputum sample" if "cancer is present", $Q[F \mid C]$, is also .95, and the likelihood that "cancer-like cells will be present" if "no cancer is present" in the patient, $Q[F \mid {\sim}C]$, is .05. In addition let's suppose that the likelihoods for the outcomes of the two different tests are independent given that cancer is present, $Q[E \cdot F \mid C] = Q[E \mid C] \cdot Q[F \mid C]$; and the outcomes are also independent if cancer is not present, $Q[E \cdot F \mid C] = Q[E \mid {\sim}C] \cdot Q[F \mid {\sim}C]$. The evidence bases here are just $\{E, {\sim}E\}$ and $\{F, {\sim}F\}$.

The sputum sample is taken and sent to the lab. At about the same time a chest x-ray is performed and goes to the radiologist for analysis. The lab technician is not certain that any of the cells he examines are malignant, but sees some suspicious-looking cells. His degree of confidence that no cancer-like cells are present in the sputum sample is .90. However, the radiologist is fairly confident that one of the shadows she sees was produced by a small mass. Her degree of confidence that an image of a mass is present on the x-ray is .90. Both results are sent to the physician.

Suppose that the physician receives the sputum cytology result first. She updates her belief by adopting the technician's degree of confidence about the presence of cancer-like cells, $Q_f[{\sim}F] = .90$. Then she updates her degree of belief that the patient has lung cancer accordingly: $Q_f[C] = .14$. When the x-ray report comes in the physician adopts the radiologist's degree of confidence that an image of a mass is present, $Q_e[E] = Q_{fe}[E] = .90$. As a result she updates her degree of belief that the patient has lung cancer to $Q_{fe}[C] = .68$.

However, if the physician receives the lab results in the opposite order, her updated belief strength, based on the radiologist's report, will first be $Q_e[C] = .86$. Then, upon receiving the sputum cytology report her cumulative belief strength that the patient has lung cancer falls to $Q_{ef}[C] = .32$. Thus, in this example the cumulative effect of the diagnostic tests is either to raise the physician's belief strength by .18 (from .5 to .68) or to lower it by .18 (from .5 to .32), depending on the order in which she incorporates the test results into her belief function. But, given the symmetry of the likelihoods and the test result belief strengths in this example, intuitively the conflicting test reports should cancel each other out. Intuitively the physician should return to her initial .5 belief strength for cancer after the two opposing test results are figured in.[14]

The Basic Sequential Update Formula relies only on rigidity and the standard axioms of probability theory. The addition of the Amnestic Update Thesis is just one way to extend Basic Jeffrey Updating to sequences. The resulting Amnestic Model raises two related concerns. First, it seems implausible that the most recent experience or non-propositional state should completely dictate belief strengths for basis sentences, with no regard for the import of previous experiences or states. There may be some specialized systems for which this model is appropriate. But it seems wrong as a model of (idealized) human agents, and wrong for most applications of automated Bayesian inference networks. The second concern is that, as a result of this amnesia, the order in which states or experiences are acquired will almost always have a very significant influence on an agent's belief strengths. This order effect is very troubling.[15]

6. THE NORMED-LIKELIHOOD FACTOR MODEL

An alternative way to extend Basic Jeffrey Updating is to take the non-propositional state e to *directly* supply the agent with Normed-Likelihood update factors $\mathrm{NL}[Q_\beta, e, E_i]$. Then the Normed-Likelihood factor version of the Basic Sequential Update Formula generates new degrees of belief for each of the agent's sentences. When update factors are viewed in this way $Q_{\beta e}[E_i]$ looks very much like a posterior probability based on the joint "evidence" β and e.

The idea that the update of a basis sentence is a transition from a prior to a posterior probability has a distinctly Bayesian flavor. It suggests that we might get more insight into the relationship between Normed-Likelihood factors and probabilities by looking at how non-propositional states would be treated if they were expressible as sentences. Let's treat them as such for a moment, as a heuristic device, to see what a standard Bayesian analysis might suggest.

Consider a sequence of one or more states ϵ that all sharing the same basis $\{E_i\}$, and consider another state d that affects a different basis $\{D_i\}$. Let Q_α be a belief function previously updated on sequence α. Now suppose that d and ϵ could be adequately expressed as sentences, and let us consider how they would function in the Bayesian update of $Q_\alpha[E_i]$. (Think of the members of $\{E_i\}$ as alternative hypotheses and think of d and ϵ as evidence claims relevant to them). Notice that the Normed-Likelihood factor $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_i] = (Q_{\alpha d\epsilon}[E_i]/Q_{\alpha d}[E_i])$ would then be equal to a ratio of conditional probabilities: $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_i] = (Q_\alpha[E_i \mid d \cdot \epsilon]/Q_\alpha[E_i \mid d])$. Then Bayes' theorem yields $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_i] = (Q_\alpha[d \mid E_i \cdot \epsilon]/Q_\alpha[d \mid E_i]) \cdot (Q_\alpha[E_i \mid \epsilon]/Q_\alpha[E_i]) \cdot (Q_\alpha[d] \cdot Q_\alpha[\epsilon]/Q_\alpha[d \cdot \epsilon])$.[16]

Notice that the Normed-Likelihood factor for Q_α due to ϵ alone (i.e. absent d) would be this: $\mathrm{NL}[Q_\alpha, \epsilon, E_i] = Q_\alpha[E_i \mid \epsilon]/Q_\alpha[E_i]$. Substituting this expression into the previous equation gives a relationship between two Normed-Likelihood factors, one that results if d is encountered before ϵ, the other if only ϵ is encountered: $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_i] = \mathrm{NL}[Q_\alpha, \epsilon, E_i] \cdot (Q_\alpha[d \mid E_i \cdot \epsilon]/Q_\alpha[d \mid E_i]) \cdot (Q_\alpha[d] \cdot Q_\alpha[\epsilon]/Q_\alpha[d \cdot \epsilon])$.

The standard rigidity condition, which is the heart of Jeffrey updating, requires that the basis sentences E_i for ϵ screen off all other sentences from ϵ. This includes sentences in d's basis, and should include d as well if it were a sentence. Indeed, if d were expressible as a sentence, standard rigidity would yield $Q_\alpha[d \mid E_i \cdot \epsilon] = Q_\alpha[d \mid E_i]$ – i.e. d would be independent of ϵ given E_i. With this application of rigidity the equation at the end of the previous paragraph yields $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_i] = \mathrm{NL}[Q_\alpha, \epsilon, E_i] \cdot (Q_\alpha[d] \cdot Q_\alpha[\epsilon]/Q_\alpha[d \cdot \epsilon])$. This expresses a relationship between Normed-Likelihood factors for ϵ with and without intervening state d on a different basis.[17] This relationship between Normed-Likelihood factors would follow from standard rigidity alone if ϵ and d were sentences. So this relationship might plausibly be adopted as an extension of standard rigidity to cases when ϵ and d are not sententially expressible, if only we had a way to express this relationship without treating ϵ and d as though they were sentences in the term '$Q_\alpha[d] \cdot Q_\alpha[\epsilon]/Q_\alpha[d \cdot \epsilon]$' of the formula. Is there a way to do this?

Notice that the term '$Q_\alpha[d] \cdot Q_\alpha[\epsilon]/Q_\alpha[d \cdot \epsilon]$' does not depend on which sentence from $\{E_i\}$ is involved; it has the same value r, regardless. So the previous analysis suggests the following extension of *rigidity* to sequences of states $\alpha d \epsilon$ and to the corresponding sequences of belief functions $Q, \ldots, Q_\alpha, Q_{\alpha d}, \ldots, Q_{\alpha d \epsilon}$ (where each state affects a basis for which the usual rigidity requirement holds).

EXTENDED RIGIDITY THESIS: NORMED-LIKELIHOOD FACTOR VERSION. For any sequence ϵ of states that affect a common basis $\{E_i\}$ and any state d that affects some different basis, there is a real number $r > 0$ such that for each $Q_{\alpha d \epsilon}$-possible sentence E_j in $\{E_i\}$ (i.e. each E_j such that $Q_{\alpha d \epsilon}[E_i] > 0$), $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_j] = r \cdot \mathrm{NL}[Q_\alpha, \epsilon, E_j]$.

The presence of a state d preceding ϵ can only affect the posterior belief strengths $Q_{\alpha d \epsilon}[E_i]$ of ϵ's basis through two channels: through the priors $Q_{\alpha d}[E_j]$, and through the update factors $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_j]$. The Extended Rigidity Thesis implies that when the basis of d is distinct from the basis of ϵ, the only influence that d has on ϵ's basis sentences is the affect it had in updating $Q_\alpha[E_i]$ to $Q_{\alpha d}[E_i]$.[18] The Thesis says that d makes no substantial contribution to the update factors through which ϵ affects its basis; d's only influence on ϵ's Normed-Likelihood factors is to multiply them by a constant r. And that is virtually no influence at all. It merely introduces a scaling factor to get the resulting posterior probabilities to sum to 1.

In other words, if the agent were to acquire a state d with a basis distinct from ϵ's just before she acquires ϵ, the acquisition of d would have precisely the same affect on how each of ϵ's own basis sentences come to be updated by ϵ – i.e. for any two sentence E_i and E_j from ϵ's basis, $\text{NL}[Q_{\alpha d}, \epsilon, E_i]/\text{NL}[Q_\alpha, \epsilon, E_i] = \text{NL}[Q_{\alpha d}, \epsilon, E_j]/\text{NL}[Q_\alpha, \epsilon, E_j]$. Put another way, the quantity by which ϵ changes the agent's belief strength in E_i as compared to the quantity by which it changes her belief strength in E_j *remains exactly the same* in the presence of d as it would be absent d – i.e. $\text{NL}[Q_{\alpha d}, \epsilon, E_i]/\text{NL}[Q_{\alpha d}, \epsilon, E_j] = \text{NL}[Q_\alpha, \epsilon, E_i]/\text{NL}[Q_\alpha, \epsilon, E_j]$. This means that the presence or absence of d cannot influence the relative changes in belief strengths induced by ϵ on its own evidence basis.

The Extended Rigidity Thesis does not presuppose that states d and ϵ may be adequately expressed by sentences. We treated them as such several paragraphs back merely as a heuristic device to help us see how closely the Thesis is related to the usual rigidity requirement. But once we understand what the Thesis says about how ϵ dominates the update factors of its own basis, the Thesis stands as a plausible updating requirement on its own. We'll say that a sequence of belief functions connected through successive updates fits the *Normed-Likelihood Factor Update Model* just in case the sequence satisfies both the usual Jeffrey *rigidity* requirement and the Normed-Likelihood Factor Version of the Extended Rigidity Thesis.

Normed-Likelihood Updating is a more natural, more Bayesian extension of Basic Jeffrey Updating than is Amnestic Updating. The Extended Rigidity Thesis says that a homogeneous sequence of states has autonomous control over update factors that affect its own basis. But other states may continue to exert an indirect influence on this basis. Their influence is carried by the *prior belief strength* due to previous updating, which substantially affects the posterior belief strength of the basis sentences. By contrast, the Amnestic Update Thesis says that the most recent state overwrites belief strengths on its basis so completely that no trace of the effects of any previous updates on its basis remains.

When Extended Rigidity holds, the Basic Sequential Update Formula may be refined. To see how, consider the belief strength $Q_{a\ldots def\epsilon}[S]$ for an arbitrary sentences S due to the sequence of states $a \ldots def\epsilon$, where ϵ is a basis-homogeneous sequence of states that affect basis $\{G_i\}$, and where state f affects $\{F_i\}$ distinct from $\{G_i\}$. (We allow that some of the states $a, \ldots, d, e, f$, may share the same evidence basis, and that some may even share ϵ's basis.) Let us focus on a non-zero term from the Sequential Update Formula of form $Q[S \cdot A_m \cdot \cdots \cdot E_k \cdot F_j \cdot G_i] \cdot \text{NL}[Q, a, A_m] \cdot \cdots \cdot \text{NL}[Q_{a\ldots d}, e, E_k] \cdot \text{NL}[Q_{a\ldots de}, f, F_j] \cdot \text{NL}[Q_{a\ldots def}, \epsilon, G_i]$. (Notice that whenever two states share a basis, this non-zero term must use the same basis sentence for both states – i.e. if e has the same basis as ϵ, then sentence E_k must be G_i; otherwise $Q[S \cdot A_m \cdot \cdots \cdot E_k \cdot F_j \cdot G_i]$ would be 0, and so the whole term would be 0.)

Extended Rigidity implies $\mathrm{NL}[Q_{a\ldots def}, \epsilon, G_i] = r \cdot \mathrm{NL}[Q_{a\ldots de}, \epsilon, G_i]$, where r is the same for each G_i. This says that the state f to the left of ϵ may be stripped away, leaving behind a multiplicative constant r. This stripping away of states to the left of ϵ may continue in the same way to yield $\mathrm{NL}[Q_{a\ldots def}, \epsilon, G_i] = r \cdot s \cdot \cdots \cdot \mathrm{NL}[Q, \epsilon, G_i]$ *unless* some state, say e, has the same basis as ϵ. If ϵ and e share the same basis, then, after stripping away f, the original term becomes $Q[S \cdot A_m \cdot \cdots \cdot G_i \cdot F_j \cdot G_i] \cdot \mathrm{NL}[Q, a, A_m] \cdot \cdots \cdot \mathrm{NL}[Q_{a\ldots d}, e, G_i] \cdot \mathrm{NL}[Q_{a\ldots de}, f, F_j] \cdot \mathrm{NL}[Q_{a\ldots de}, \epsilon, G_i] \cdot r$ (since E_k must be G_i). Now, notice that in this formula the terms for G_i give us $\mathrm{NL}[Q_{a\ldots d}, e, G_i] \cdot \mathrm{NL}[Q_{a\ldots de}, \epsilon, G_i] = (Q_{a\ldots de}[G_i]/Q_{a\ldots d}[G_i]) \cdot (Q_{a\ldots de\epsilon}[G_i]/Q_{a\ldots de}[G_i]) = \mathrm{NL}[Q_{a\ldots d}, e\epsilon, G_i]$. This means that combining the two G_i terms gives us $Q[S \cdot A_m \cdot \cdots \cdot G_i \cdot F_j \cdot G_i] \cdot \mathrm{NL}[Q, a, A_m] \cdot \cdots \cdot \mathrm{NL}[Q_{a\ldots de}, f, F_j] \cdot \mathrm{NL}[Q_{a\ldots d}, e\epsilon, G_i] \cdot r$, where $e\epsilon$ is now a basis homogeneous sequence of states. Thus, basis homogeneous states accumulate.

Continuing to strip away states in this way for each factor, the term from the Sequential Update Formula eventually is transformed into a term of form $Q[S \cdot A_m \cdot \cdots \cdot F_j \cdot G_i] \cdot \mathrm{NL}[Q, \alpha, A_m] \cdot \cdots \cdot \mathrm{NL}[Q, \beta, F_j] \cdot \mathrm{NL}[Q, \gamma, G_i] \cdot K$. The constant K is the product of all of those constants like r that accumulate as states are stripped away; and $\alpha, \ldots, \beta, \gamma$ are basis-homogeneous sequences of states that affect bases $\{A_i\}, \ldots, \{F_i\}$, $\{G_i\}$, respectively.

Putting these reduced terms back into the Sequential Update Formula yields $Q_{a\ldots efg}[S] = K \cdot \sum_{\{i:Q[G_i]>0\}} \sum_{\{j:Q[F_j]>0\}} \cdots \sum_{\{k:Q[A_m]>0\}} Q[S \cdot A_m \cdot \cdots \cdot F_j \cdot G_i] \cdot \mathrm{NL}[Q, \alpha, A_m] \cdot \cdots \cdot \mathrm{NL}[Q, \beta, F_j] \cdot \mathrm{NL}[Q, \gamma, G_i]$. This result would be more enlightening if only we could determine the value of constant K. And indeed we can. Notice: $1 = Q_{a\ldots efg}[\text{tautology}] = K \cdot \sum_{\{i:Q[G_i]>0\}} \sum_{\{j:Q[F_j]>0\}} \cdots \sum_{\{k:Q[A_m]>0\}} Q[A_m \cdot \cdots \cdot F_j \cdot G_i] \cdot \mathrm{NL}[Q, \alpha, A_m] \cdot \cdots \cdot \mathrm{NL}[Q, \beta, F_j] \cdot \mathrm{NL}[Q, \gamma, G_i]$. Thus, we get the following version of the Sequential Update Formula.

EXTENDED SEQUENTIAL UPDATE FORMULA: NORMED-LIKELIHOOD FACTOR VERSION. Let δ be a sequence of states and, for belief function Q, let $\alpha, \ldots, \gamma$ be the basis-homogeneous subsequences of states from δ (where $\alpha, \ldots, \gamma$ each maintain the same order among their states as these states are ordered in δ, and jointly contain all states in δ). Let $\{A_i\}$ be the evidence basis affected by $\alpha, \ldots$, and let $\{C_i\}$ be the basis affected by γ. Then, for all sentences S,

$$Q_\delta[S] = \frac{\sum_i \sum_j \cdots \sum_k Q[S \cdot A_k \cdot \cdots \cdot C_i] \cdot \mathrm{NL}[Q, \alpha, A_k] \cdot \cdots \cdot \mathrm{NL}[Q, \gamma, C_i]}{\sum_i \sum_j \cdots \sum_k Q[A_k \cdot \cdots \cdot C_i] \cdot \mathrm{NL}[Q, \alpha, A_k] \cdot \cdots \cdot \mathrm{NL}[Q, \gamma, C_i]}$$

(sums over the basis sentences such that $Q[C_i], \ldots, Q[A_k]$ are non-zero).

Thus, the updated belief function Q_δ is completely specified by the initial belief function Q together with Normed-Likelihood factors that express how each basis-homogeneous subsequence of states from δ updates Q. Inspection of this formula shows that the resulting belief strengths are independent of update order – except, perhaps, *within* the basis homogeneous subsequences.

7. The Likelihood-Ratio Factor Model

Normed-Likelihood update factors appear to do a better, more Bayesian looking job of representing the impact of non-propositional states than Amnestic update factors. However, there is another closely related factor that, from a Bayesian perspective, is even better suited to belief updating. Bayesians have long considered *likelihood ratios* as the factors that carry the full, unadulterated import of evidence for hypotheses – and with good reason.

First, notice that Normed-Likelihoods are somewhat tainted by the prior belief strengths of basis sentences. You can easily see this in the case where ϵ positively supports E_j. Then, $Q_{\beta\epsilon}[E_i] = \mathrm{NL}[Q_\beta, \epsilon, E_i] \cdot Q_\beta[E_i] > Q_\beta[E_i]$. So, the prior probability of E_i, $Q_\beta[E_i]$, which is not affected by state e, must nevertheless constrain the value that ϵ may induce in $\mathrm{NL}[Q_\beta, \epsilon, E_i]$. It can be no greater than $1/Q_\beta[E_i]$, otherwise $Q_{\beta\epsilon}[E_i]$ would be greater than 1. Thus, experiences ϵ are not free to produce *whatever values they may* (within some fixed range) for Normed-Likelihood factors.[19]

To find an update factor that doesn't fall under the influence of the prior probabilities of basis sentences, let us again suppose, as a heuristic, that sequence ϵ is propositionally expressible. Consider a pair of alternative "hypotheses" E_1 and E_2 from ϵ's basis. Bayes' theorem for each hypothesis E_j yields $Q_{\beta\epsilon}[E_j] = Q_\beta[E_j \mid \epsilon] = (Q_\beta[\epsilon \mid E_j]/Q_\beta[\epsilon]) \cdot Q_\beta[E_j]$. Dividing $Q_{\beta\epsilon}[E_2]$ by $Q_{\beta\epsilon}[E_1]$ we get $Q_{\beta\epsilon}[E_2]/Q_{\beta\epsilon}[E_2] = (Q_\beta[\epsilon \mid E_2]/Q_\beta[\epsilon \mid E_1]) \cdot (Q_\beta[E_2]/Q_\beta[E_1])$. The likelihood ratio term $Q_\beta[\epsilon \mid E_2]/Q_\beta[\epsilon \mid E_1]$ in this equation is completely unaffected by the prior probabilities of sentences in $\{E_j\}$. It may take *any* non-negative value, with no regard for the values of the prior probabilities. It updates the ratio of belief strengths for E_2 over E_1 solely on the basis of what E_2 says about how likely it is that ϵ should occur as compared to what E_1 says about the likelihood of ϵ.

This analysis suggests that *Likelihood-Ratio update factors* may provide a better representation of the immediate impact of evidence than do Normed-Likelihood factors. A *Likelihood-Ratio factor* is a factor $\mathrm{LR}[Q_\beta, \epsilon, E_j, E_k]$ that, when multiplied by a ratio of prior belief strengths for the basis $\{E_i\}$ of state ϵ, generates the ratio of posterior belief strengths: $Q_{\beta\epsilon}[E_j]/Q_{\beta\epsilon}[E_k] = \text{LR}[Q_\beta, \epsilon, E_j, E_k] \cdot (Q_\beta[E_j]/Q_\beta[E_k])$.

Think about Likelihood-Ratio factors this way. A state projects relative weightings onto pairs of its basis sentences that may be measured on a scale employing the non-negative real numbers. Although states are free to specify any non-negative values for Likelihood-Ratio factors, some internal consistency constraints among these factors must be obeyed. That is, they need to satisfy rules that are typical for likelihood ratios. The following rules suffice to govern their behavior. (Hint: To understand what these rules say, just think of $\text{LR}[Q_\beta, \epsilon, E_i, E_n]$ as a likelihood ratio, $Q_\beta[\epsilon \mid E_i]/Q_\beta[\epsilon \mid E_n]$, as though ϵ were a sentence. However, the point of spelling these rules out is to show that the rules governing Likelihood-Ratio factors may be adequately expressed without treating ϵ as a sentence.)

For any homogeneous state sequence ϵ on a basis $\{E_i\}$, and for any belief function Q_β, there is an E_n such that for all E_i, E_j, E_k:

(1) either $\text{LR}[Q_\beta, \epsilon, E_i, E_j] \geq 0$ or undefined (i.e. has no value);
(2) $\text{LR}[Q_\beta, \epsilon, E_i, E_n]$ is defined;
(3) $\text{LR}[Q_\beta, \epsilon, E_i, E_n] = 0$ iff $\text{LR}[Q_\beta, \epsilon, E_j, E_i]$ is undefined;
(4) if $\text{LR}[Q_\beta, \epsilon, E_j, E_i] > 0$, $\text{LR}[Q_\beta, \epsilon, E_i, E_j] = 1/\text{LR}[Q_\beta, \epsilon, E_j, E_i]$;
(5) if $\text{LR}[Q_\beta, \epsilon, E_k, E_j]$ and $\text{LR}[Q_\beta, \epsilon, E_j, E_i]$ are defined, then $\text{LR}[Q_\beta, \epsilon, E_k, E_i] = \text{LR}[Q_\beta, \epsilon, E_k, E_j] \cdot \text{LR}[Q_\beta, \epsilon, E_j, E_i]$;
(6) if ϵ is of form γe and $\text{LR}[Q_\beta, \epsilon, E_j, E_i]$ is defined, then $\text{LR}[Q_\beta, \epsilon, E_j, E_i] = \text{LR}[Q_{\beta\gamma}, e, E_j, E_i] \cdot \text{LR}[Q_\beta, \gamma, E_j, E_i]$.

Let us now take Likelihood-Ratio factors as primitive. From this perspective Normed-Likelihood factors are merely *defined* as ratios of belief strengths: $\text{NL}[Q_\beta, \epsilon, E_j] = (Q_{\beta\epsilon}[E_j]/Q_\beta[E_j])$. The way Likelihood-Ratio factors update belief strengths for basis sentences (i.e. $Q_{\beta\epsilon}[E_j]/Q_{\beta\epsilon}[E_k] = \text{LR}[Q_\beta, \epsilon, E_j, E_k] \cdot Q_\beta[E_j]/Q_\beta[E_k]$), entails they equal ratios of Normed-Likelihood factors: $\text{LR}[Q_\beta, \epsilon, E_j, E_k] = \text{NL}[Q_\beta, \epsilon, E_j]/\text{NL}[Q_\beta, \epsilon, E_k]$ (when $Q_{\beta\epsilon}[E_k] > 0$). So, taking Likelihood-Ratio factors as primitive, the version of the Extended Rigidity Thesis that applies to them is this:

EXTENDED RIGIDITY THESIS: LIKELIHOOD-RATIO FACTOR VERSION. For any sequence of states ϵ with common basis $\{E_i\}$ and any state d not affecting $\{E_i\}$, there's a sentence E_k in $\{E_i\}$ such that, for each $Q_{\alpha d\epsilon}$-possible sentence E_j in $\{E_i\}$, $\text{LR}[Q_{\alpha d}, \epsilon, E_j, E_k]$ is defined and $\text{LR}[Q_{\alpha d}, \epsilon, E_j, E_k] = \text{LR}[Q_\alpha, \epsilon, E_j, E_k]$.

Let us say that a sequence of belief functions connected by successive updates employs the *Likelihood-Ratio Factor Update Model* just in case the sequence satisfies both the usual Jeffrey *rigidity* requirement and the Likelihood-Ratio Factor Version of the Extended Rigidity Thesis. Likelihood-Ratio factors that satisfy Extended Rigidity are roughly equivalent to the factors proposed by Field (1978). So let's call updating based on such factors 'Field Updating', or sometimes 'Likelihood-Ratio factor updating'.[20]

The Likelihood-Ratio version of Extended Rigidity implies the following form of the Extended Sequential Update Formula.

EXTENDED SEQUENTIAL UPDATE FORMULA: LIKELIHOOD-RATIO FACTOR VERSION. Let δ be a sequence of states and, for belief function Q, let $\alpha, \ldots, \gamma$ be the evidence-basis-homogeneous subsequences of states from δ (where $\alpha, \ldots, \gamma$ each maintain the same order among their states as these states are ordered in δ, and jointly contain all states in δ). Let $\{A_i\}$ be the evidence basis affected by $\alpha, \ldots$, and let $\{C_i\}$ be the basis affected by γ. Then, for all sentences S,

$$Q_\delta[S] = \frac{\sum_i \sum_j \ldots \sum_k Q[S \cdot A_k \cdot \cdots \cdot C_i] \cdot \mathrm{LR}[Q, \alpha, A_k, A] \cdot \cdots \cdot \mathrm{LR}[Q, \gamma, C_i, C]}{\sum_i \sum_j \ldots \sum_k Q[A_k \cdot \cdots \cdot C_i] \cdot \mathrm{LR}[Q, \alpha, A_k, A] \cdot \cdots \cdot \mathrm{LR}[Q, \gamma, C_i, C]}$$

(sums over basis sentences such that $Q[C_i], \ldots, Q[A_k]$ are non-zero).[21]

This is essentially equivalent to the update formula proposed by Field (1978). Inspection of this formula shows that belief strengths generated by the Likelihood-Ratio Model through Field Updating are independent of update order – except, perhaps, *within* basis homogeneous subsequences of states. The only way update order may have an effect is if some reordering among states affecting a common basis makes a difference.[22]

If non-propositional states of an agent produce Likelihood-Ratio update factors directly, then her previous belief function Q should be irrelevant to these factors. Only her experiences should affect them. Notice, however, that factors like $\mathrm{LR}[Q, \gamma, C_i, C]$ are expressed as functions of Q. What gives? Think of it this way. An update factor of form $\mathrm{LR}[Q_{a\ldots b}, c, C_i, C]$ should depend on $Q_{a\ldots b}$, not because this belief function itself has an influence on it, but because $Q_{a\ldots b}$ may have "encoded" in it other states that also affect the basis $\{C_i\}$ directly. If the sequence $a \ldots b$ contains no such states and if the belief function Q is not itself the result of updating on any previous states that affect $\{C_i\}$ directly, then indeed $Q_{a\ldots b}$ should be irrelevant to the update factor, and we may represent this by writing $\mathrm{LR}[Q_{a\ldots b}, c, C_i, C] = \mathrm{LR}[c, C_i, C]$.

More generally, if the belief function Q in the Likelihood-Ratio version of the Extended Sequential Update Formula has not come about from experiences that affect any of the bases $\{A_i\}, \ldots, \{C_i\}$, then the relevant update factors may take the form $\mathrm{LR}[\alpha, A_k, A], \ldots, \mathrm{LR}[\gamma, C_i, C]$. Only explicitly represented experiential states affect the update factors for Q. Thus, under these conditions the Likelihood-Ratio version of the Extended Sequential Update Formula becomes:

$$Q_\delta[S] = \frac{\sum_i \ldots \sum_j Q[S \cdot A_j \cdot \,\cdots\, \cdot C_i] \cdot \mathrm{LR}[\alpha, A_j, A] \cdot \,\cdots\, \cdot \mathrm{LR}[\gamma, C_i, C]}{\sum_i \ldots \sum_j Q[A_j \cdot \,\cdots\, \cdot C_i] \cdot \mathrm{LR}[\alpha, A_j, A] \cdot \,\cdots\, \cdot \mathrm{LR}[\gamma, C_i, C]}.$$

The agent's degree of belief in each sentence S depends only on her belief function Q and update factors that depend only on the agent's non-propositional states.

Let's see what Bayes' Theorem looks like under the Likelihood-Ratio Model of updating on uncertain evidence. When the sentence S is a hypothesis H_k from some partition of alternative hypotheses $\{H_m\}$, the Extended Sequential Update Formula yields the following generalization of Bayes' Theorem:

$$Q_\delta[H_k] = \frac{Q[H_k] \cdot \sum_i \ldots \sum_j Q[A_j \cdot \,\cdots\, \cdot C_i \mid H_k] \cdot \mathrm{LR}[\alpha, A_j, A] \cdot \,\cdots\, \cdot \mathrm{LR}[\gamma, C_i, C]}{\sum_m Q[H_m] \cdot \sum_i \ldots \sum_j Q[A_j \cdot \,\cdots\, \cdot C_i \mid H_m] \cdot \mathrm{LR}[\alpha, A_j, A] \cdot \,\cdots\, \cdot \mathrm{LR}[\gamma, C_i, C]}.$$

If in addition the evidence sentences within bases are conditionally independent given the hypotheses (which will usually be the case in applications such as Bayesian inference nets) this formula may be written as follows:

$$\begin{aligned} Q_\delta[H_k] &= \frac{Q[H_k] \cdot (\sum_j Q[A_j \mid H_k] \cdot \mathrm{LR}[\alpha, A_j, A]) \cdot \,\cdots\, \cdot (\sum_i Q[C_i \mid H_k] \cdot \mathrm{LR}[\gamma, C_i, C])}{\sum_m Q[H_m] \cdot (\sum_j Q[A_j \mid H_m] \cdot \mathrm{LR}[\alpha, A_j, C]) \cdot \,\cdots\, \cdot (\sum_i Q[C_i \mid H_m] \cdot \mathrm{LR}[\gamma, C_i, C])} \\ &= \frac{Q_\epsilon[H_k] \cdot (\sum_i Q[C_i \mid H_k] \cdot \mathrm{LR}[\gamma, C_i, C])}{\sum_m Q_\epsilon[H_m] \cdot (\sum_i Q[C_i \mid H_m] \cdot \mathrm{LR}[\gamma, C_i, C])}, \end{aligned}$$

where ϵ here is the sequence of all states in δ except those in γ, and $Q_\epsilon[H_k]$ is the degree of belief in H_k updated on all states other than those in γ.

An example of belief updating on Likelihood-Ratio factors may be instructive. Let's reconsider the cancer diagnosis example in Section 5. Let's keep the physician's belief function prior to the lab reports exactly the same as before, but change how evidence regarding the sputum sample and the chest x-ray is reported from the lab. (In this example the partitions are simpler than usual – they are $\{F, \sim F\}$ and $\{E, \sim E\}$, i.e. {cancer-like cells are present, no cancer-like cells are present} and {a lung mass image is present on the x-ray, no lung mass image is present on the x-ray}. More generally, evidence partitions may include a number of possible alternatives – e.g., that lung mass images of various types or densities are detected.)

A lab technician examines the sputum sample microscopically and finds no cells that he is certain are malignant. But he sees some suspicious-looking cells. His degree of confidence that no cancer-like cells are present may well turn out to be .90, just as before. However, this probability may partially result from his prior belief strength about whether cancer-like cells will be contained in the sample – based, perhaps, on how frequently he finds samples to have such cells. So, what the physician wants from the technician is his Likelihood-Ratio update factor, e.g., $\mathrm{LR}[f, F, {\sim}F] = .50$, rather than his updated degree of belief. This factor represents how well the microscopic appearance of this sample fits with the proposition F (that cancer-like cells are present) *as compared to* how well its appearance fits with the proposition ${\sim}F$ (that no cancer-like cells are present). In a Bayesian context it is this factor that carries the pure evidential import of the lab technician's observations. The technician's report of a factor $\mathrm{LR}[f, F, {\sim}F] = .50$ says in effect that he takes his experience f to be twice as likely to have occurred if "no cancer-like cells are present" than if "cancer-like cells are present." The physician utilizes the technician's expertise by adopting the technician's update factor, $\mathrm{LR}[f, F, {\sim}F] = .50$, as her own. This provides her an updated belief strength that the patient has lung cancer: $Q_f[C] = .35$.

Regarding the chest x-ray, the radiologist sees a shadow that she is fairly confident was produced by a small mass. She takes the appearance of the x-ray film to be twice as likely to have occurred if a mass was present in the lung than if no mass was present; so the evidential value of her experiential state is a Likelihood-Ratio update factor $\mathrm{LR}[e, E, {\sim}E] = 2$. Upon receiving that report the physician adopts the radiologists update factor as her own and comes to a new updated degree of belief that the patient has lung cancer: $Q_{fe}[C] = .50$.

This example was constructed to make the two update factors counteract each other. Notice, however, that whatever values the update factors and the resulting belief function may have, update order will produce net no effect: $Q_{fe}[C] = Q_{ef}[C]$.[23]

8. On the Order-Independence of Sequential Updates

Both the Likelihood-Ratio and Normed-Likelihood versions of the Extended Sequential Update formula show their versions of Extended Jeffrey Updating to be independent of order *across* bases. As a result, if each experience affects a distinct basis, then such updating will be completely independent of order. This stands in sharp contrast to Amnestic Updating of Probabilistic Update Factors, which is order dependent both within and across bases.[24]

Let's focus on Field Updating – i.e. Likelihood-Ratio factor updating. What I will say about it largely applies to Normed-Likelihood updating as well. Given our analysis so far, it remains possible for Likelihood-Ratio Updating to depend on order *within* each basis. Indeed, one could further extend the Likelihood-Ratio Update Model by requiring that the most recent state overwrites all previous states that share its basis. In that case each term of form $\mathrm{LR}[Q, \gamma, C_i, C]$ in the likelihood ratio version of the Extended Sequential Update Formula will equal a factor $\mathrm{LR}[Q, c, C_i, C]$, where c is the last state in the sequence γ affecting basis $\{C_i\}$. Let us call this the *Basis-Overwrite Version* of the *Likelihood-Ratio Update Model*. On this model, given its most recent update, each basis has an order-independent effect on the updates of all other sentences. That is, if two distinct bases are each updated by states that overwrite their likelihood ratio update factors, the order in which these two updates occur can make no net difference to the belief strengths of any sentence in the belief network.

The Overwrite Version is just one way to further extend the Likelihood-Ratio Update Model. There is an alternative extension that maintains order-independence even within bases (as well as across bases). To see how it works, first notice that, as stated, the Extended Rigidity Thesis imposes no requirements on how update order works within bases. But it is perfectly compatible with commutivity holding there. For, the Extended Rigidity Thesis is compatible with treating states as though they were evidence sentences that update belief strengths of their basis sentences via likelihood ratios in Bayes' theorem. And Bayesian updating on sentences is always independent of order. Thus, it is both internally consistent and quite natural to extend the Likelihood-Ratio Update Model by simply stipulating that updating within bases is to be order-independent. That is, we may extend Likelihood-Ratio Updating by adding the requirement that for any basis homogeneous sequence γ, if β is any reordering of γ, then $\mathrm{LR}[Q, \beta, C_i, C] = \mathrm{LR}[Q, \gamma, C_i, C]$. Let us call this the *Basis-Commuting Version* of the *Likelihood-Ratio Update Model*. Sequential updating is *completely* order-independent on this model.

Notice that requiring commutivity among states that share a basis does not amount to extending the Extended Rigidity Thesis to apply within bases. That is, when states d and e affect the same basis and commutivity holds (i.e. when $Q_{\alpha de}[E_j] = Q_{\alpha ed}[E_j]$), the Extended Rigidity relationships of form $\mathrm{LR}[Q_{\alpha d}, e, E_j, E_k] = \mathrm{LR}[Q_\alpha, e, E_j, E_k]$ between update factors may, nevertheless, fail to hold. And, indeed, we should want Extended Rigidity to fail within bases in at least some cases (while commutivity is maintained). To see why, let us again employ the heuristic of treating states as sentences. Suppose that state e and sequence ϵ share basis $\{E_i\}$. And suppose that e and ϵ are adequately expressible as sentences. Then we would have $\text{LR}[Q_{\alpha\epsilon}, e, E_j, E_k] = Q_{\alpha\epsilon}[e \mid E_j]/Q_{\alpha\epsilon}[e \mid E_k] = Q_\alpha[e \mid E_j \cdot \epsilon]/Q_\alpha[e \mid E_k \cdot \epsilon]$. This would result in a version of Extended Rigidity within the basis *if* each basis sentence E_i screens off state e form past states ϵ on their common basis (i.e. $Q_\alpha[e \mid E_i \cdot \epsilon] = Q_\alpha[e \mid E_i]$). In that case we would have the following form of Extended Rigidity within a basis: $\text{LR}[Q_{\alpha\epsilon}, e, E_j, E_k] = \text{LR}[Q_\alpha, e, E_j, E_k]$.[25] And that would entail that the Likelihood-Ratio factors for a sequence of states $cd \ldots e$ on a common basis decompose as follows: $\text{LR}[Q_\alpha, cd \ldots e, E_j, E_k] = \text{LR}[Q_\alpha, c, E_j, E_k] \cdot \text{LR}[Q_\alpha, d, E_j, E_k] \cdot \cdots \cdot \text{LR}[Q_\alpha, e, E_j, E_k]$. But, I'm claiming, we shouldn't in general want this decomposition to hold. For, when e and ϵ share a basis, the equality of form $Q_\alpha[e \mid E_j \cdot \epsilon]/Q_\alpha[e \mid E_k \cdot \epsilon] = Q_\alpha[e \mid E_j]/Q_\alpha[e \mid E_k]$ *should* often fail to hold – because, when e and ϵ share the same basis, the states in ϵ may imply something about the nature of the state e that is not captured sufficiently by basis sentences E_i to permit them to screen off e from past states ϵ. (I'll illustrate this with an example in a moment.) Thus, the intuitive motivation behind Extended Rigidity fails within bases. But commutivity may well continue to hold, since $\text{LR}[Q_\alpha, \epsilon e, E_j, E_k] = Q_\alpha[\epsilon \cdot e \mid E_j]/Q_\alpha[\epsilon \cdot e \mid E_k] = Q_\alpha[e \cdot \epsilon \mid E_j]/Q_\alpha[e \cdot \epsilon \mid E_k] = \text{LR}[Q_\alpha, e\epsilon, E_j, E_k]$.

An example may help. Suppose that the sequence of experiences $cd \ldots e$ results from a series of n repeated glances at the same perceptually foggy situation. And suppose that each glance is perceptually similar enough to produce a nearly identical propositionally inexpressible state in the agent. Then the evidential import of each state taken on its own should be the same – i.e. $\text{LR}[Q_\alpha, c, E_j, E_k] = \text{LR}[Q_\alpha, d, E_j, E_k] = \cdots = \text{LR}[Q_\alpha, e, E_j, E_k]$. If Extended Rigidity were to hold in this case, then the Extended Sequential Update Formula would yield $Q_{\alpha cd\ldots e}[S] = (\sum_j Q_\alpha[S \cdot E_j] \cdot \text{LR}[Q_\alpha, e, E_j, E_k]^n)/(\sum_j Q_\alpha[E_j] \cdot \text{LR}[Q_\alpha, e, E_j, E_k]^n)$. It can be shown that for whichever sentence E_j yields the largest update factor $\text{LR}[Q_\alpha, e, E_j, E_k]$, the right-hand side of this equation must approach the value of $Q[S \mid E_j]$ as the number of glances n increases. In particular this means that when S is the basis sentence E_j itself, a large number of glances will result in $Q_{\alpha cd\ldots e}[E_j] \approx Q[E_j \mid E_j] = 1$. The problem is, repeated glances at the same perceptually foggy situation should furnish too little new information to force near certainty on one of the basis sentences. Rather, after a few very similar glances, additional glances with the same foggy appearance should furnish practically no additional evidence, and should clearly not push one of the basis sentences to certainty.

This is exactly the kind of situation described in Garber's (1980) counter-example to Field Updating. Field's (1978) update formula is essentially equivalent to Likelihood-Ratio version of the Extended Sequential Update Formula. But Field's version of Likelihood-Ratio Updating would in effect have Extended Rigidity apply to all states, even those that affect the same basis. As a result it runs afoul of the kind of problematic example just described.

To avoid the counter-example we need only acknowledge that Extended Rigidity may fail to apply *within* bases. But, as I've already pointed out, the failure of Extended Rigidity within bases does not entail that commutivity must fail within bases as well. Indeed, the further extension of the Likelihood-Ratio Model to the *Basis-Commuting Version*, which requires commutivity to hold within bases, is quite natural, since on the Likelihood Ratio Model all updates flow from factors that behave like likelihood ratios, and commutivity naturally attends likelihood ratios.

So my response to Garber's example is that when repeated glances result in experiential states that affect the same basis, they may not be independent enough of one another (given each E_j) to get the analog of Extended Rigidity to hold. But this does not mean that the updates must depend on *order*. It only means that the value of an update factor for a sequence of basis-homogeneous states may not always decompose into the product of update factors for the individual states. Rather, the update factor decomposes as follows: $\mathrm{LR}[Q_\alpha, cd \ldots e, E_j, E_k] = \mathrm{LR}[Q_\alpha, c, E_j, E_k] \cdot \mathrm{LR}[Q_{\alpha c}, d, E_j, E_k] \cdot \cdots \cdot \mathrm{LR}[Q_{\alpha cd\ldots}, e, E_j, E_k]$. So, if no substantial new information results from glances after c, that is only because the terms $\mathrm{LR}[Q_{\alpha c}, d, E_j, E_k], \ldots, \mathrm{LR}[Q_{\alpha cd\ldots}, e, E_j, E_k]$ all equal (or nearly equal) 1, and the update factor for the whole sequence of glances is just equal to (or nearly equal to) the update factor for the first glance – $\mathrm{LR}[Q_\alpha, cd \ldots e, E_j, E_k] = \mathrm{LR}[Q_\alpha, c, E_j, E_k]$.

Now suppose that the states resulting from glances had occurred in a different order. For example, suppose that the glance that resulted in state d had preceded the glance resulting in c, and that the glances still individually yield the same information – $\mathrm{LR}[Q_\alpha, c, E_j, E_k] = \mathrm{LR}[Q_\alpha, d, E_j, E_k]$. Then we would instead have all factors after d equal to 1 (i.e. $\mathrm{LR}[Q_{\alpha d}, c, E_j, E_k] = \cdots = \mathrm{LR}[Q_{\alpha dc\ldots}, e, E_j, E_k] = 1$) and the cumulative update factor for the whole sequence of glances would equal the update factor on d alone, which has the same value as an update on c alone – $\mathrm{LR}[Q_\alpha, dc \ldots e, E_j, E_k] = \mathrm{LR}[Q_\alpha, d, E_j, E_k] = \mathrm{LR}[Q_\alpha, c, E_j, E_k]$. Thus, on the Likelihood-Ratio Update Model commutivity may well hold within bases (though extended rigidity fails), and Garber's problem dissolves.

9. A NECESSARY AND SUFFICIENT CONDITION ORDER-INDEPENDENT UPDATING

Let us now put aside the Extended Rigidity Theses and the Amnestic Update Thesis, and let us address the following broader issue: under what conditions is Basic Sequential Updating independent of the order of the states involved? Extended Rigidity implies order independence among basis homogeneous subsequences. But the discussion in the previous section indicates that it may be possible to have order-independence without Extended Rigidity. Thus, the Extended Rigidity Thesis, though sufficient, may be stronger than absolutely necessary to achieve order-independence. This suggests that it will be worthwhile to find a necessary and sufficient condition for order-independence expressed in terms of the update factors on which Extended Rigidity operates. Such a condition may shed additional light on the connection between the Extended Rigidity Theses and the influence of update order on belief strengths.

When either version of the Extended Rigidity Thesis holds, any reordering of states that preserves the order among basis-sharing states is guaranteed to produce the same belief strengths. But neither version implies anything special about the effects of reorderings among states that share bases. It would be nice to have a Reordering Theorem (specifying necessary and sufficient conditions) that is both general enough to apply to all states, regardless of whether they share bases, *and* is related closely enough to Extended Rigidity to provide a good comparison when basis-sharing states maintain their order. It turns out that pretty much the same Reordering Theorem holds regardless of whether we design it to also apply to basis-sharing states. So I'll phrase the result in a way that is easily applied to both kinds of cases – to reorderings of *all* states, and to reorderings that maintain the order among basis-sharing states.

The following definitions will provide a somewhat more compact means of expressing the Reordering Theorem:

DEFINITIONS: SUITABLE SEQUENCES AND SUITABLE REORDERINGS. Let's call a sequence of states dϵ within a possibly larger sequence of form $\alpha d\epsilon\beta$ (α and β possibly empty) *suitable* (*for commutation*) (relative to Q) just in case ϵ is a basis-homogeneous sequence of states with a basis distinct from d's basis. A sequence of states δ will be called a *suitable* reordering of a sequence γ (relative to Q) just in case it maintains γ's order among states that share bases – i.e. if states e and g share a basis, and e occurs somewhere before g in γ, then e also occurs somewhere before g in δ.[26]

The necessary and sufficient condition for reordering states given in the coming theorem is a minimal condition. It applies to all sequences of Basic Jeffrey Updates. It only assumes that each probability function in the sequence of updates satisfies the standard axioms of probability theory, and that updates are related by the usual basic *rigidity* condition. The necessary and sufficient condition it specifies is closely related to the Normed-Likelihood Version of the Extended Rigidity Thesis, but is somewhat weaker. Here is the theorem:

THE UPDATE REORDERING THEOREM. *The following two claims are equivalent:*

(1) *For each state sequence γ, every* suitable *reordering of it δ agrees with it (relative to Q) – i.e. for all sentences S, $Q_\delta[S] = Q_\gamma[S]$.*
(2) *For every state sequence $\alpha d\epsilon$ with* suitable *part $d\epsilon$ (relative to Q), for each $Q_{\alpha d\epsilon}$-possible D_i in d's basis, there is an $r = \mathrm{NL}[Q_{\alpha\epsilon}, d, D_i]/\mathrm{NL}[Q_\alpha, d, D_i] > 0$ such that for each E_j in ϵ's basis $Q_{\alpha d\epsilon}$-compatible with D_i, $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_j] = r \cdot \mathrm{NL}[Q_\alpha, \epsilon, E_j]$.*

Furthermore, if the restriction to reorderings and sequences $\alpha d\epsilon$ that are suitable *is removed from both clauses, the resulting claims, which apply to* all *reorderings, are still equivalent.*

Recall that saying D_i is "$Q_{\alpha d\epsilon}$-possible" just means that $Q_{\alpha d\epsilon}[D_i] > 0$; similarly, and saying that E_j is "$Q_{\alpha d\epsilon}$-compatible with D_i" just means that $Q_{\alpha d\epsilon}[D_i \cdot E_j] > 0$. Proof of the theorem is in the Appendix.[27]

What does the Reordering Theorem tell us about the conditions under which states may be reordered without altering belief strengths? Think of it this way. For each of d's $Q_{\alpha d\epsilon}$-possible basis sentences D_i, *define* its *$Q_{\alpha d\epsilon}$-compatibility class* $\langle D_i \rangle_{\alpha d\epsilon}$, to be the set of all sentences from ϵ's basis $\{E_i\}$ that are $Q_{\alpha d\epsilon}$-compatible with it. Then, in terms of compatibility classes, the theorem says that ϵ commutes with *d just in case* for each $Q_{\alpha d\epsilon}$-compatibility class, there is a value of r, which may be specific to *that* compatibility class (where $0 < r = \mathrm{NL}[Q_{\alpha\epsilon}, d, D_i]/\mathrm{NL}[Q_\alpha, d, D_i]$), such that every E_j in that compatibility class satisfies the relationship $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_i] = r \cdot \mathrm{NL}[Q_\alpha, \epsilon, E_j]$. This falls only a little short of the Extended Rigidity Thesis for Normed-Likelihood factors. To get Extended Rigidity all that's needed is for every r associated with the various $Q_{\alpha d\epsilon}$-compatibility classes (for the various possible D_i) to have the same value.

One way to get to Extended Rigidity from the necessary and sufficient condition expressed by clause (2) of the theorem is this. Suppose there is at least one D_i whose $Q_{\alpha d\epsilon}$-compatibility class $\langle D_i \rangle_{\alpha d\epsilon}$ contains *every* $Q_{\alpha d\epsilon}$-possible basis sentence in $\{E_i\}$. Then *every* $Q_{\alpha d\epsilon}$-possible basis sentence E_j in $\{E_i\}$ must satisfy the formulas $\text{NL}[Q_{\alpha d}, \epsilon, E_j] = r \cdot \text{NL}[Q_\alpha, \epsilon, E_j]$ for the *same value* of r. If that happens, then Extended Rigidity clearly holds.

An alternative way to get Extended Rigidity from clause (2) is this. Suppose no such all-inclusive $Q_{\alpha d\epsilon}$-compatibility class exists. But suppose that any pair of distinct $Q_{\alpha d\epsilon}$-compatibility classes has at least one E_j in common – i.e. each pair of compatibility classes overlap. Then, again, *all* sentences E_j in the basis $\{E_i\}$ that are $Q_{\alpha d\epsilon}$-possible satisfy the formulas $\text{NL}[Q_{\alpha d}, \epsilon, E_j] = r \cdot \text{NL}[Q_\alpha, \epsilon, E_j]$ for the *same value* of r; so Extended Rigidity holds.

Now, suppose not every pair of $Q_{\alpha d\epsilon\beta}$-compatibility classes overlap. But suppose any two $Q_{\alpha d\epsilon\beta}$-compatibility classes that fail to overlap are both overlapped (i.e. chained together) by a third $Q_{\alpha d\epsilon\beta}$-compatibility class. Then again, all must have the same value for r, and Extended Rigidity holds.

Most generally, suppose that every pair of $Q_{\alpha d\epsilon\beta}$-compatibility classes is linked together through a chain of overlapping $Q_{\alpha d\epsilon\beta}$-compatibility classes. Then, again, *all* $Q_{\alpha d\epsilon}$-possible sentences E_j in the basis $\{E_i\}$ satisfy the formulas $\text{NL}[Q_{\alpha d}, \epsilon, E_j] = r \cdot \text{NL}[Q_\alpha, \epsilon, E_j]$ for the *same value* of r; so Extended Rigidity holds.

Finally, even if there are pairs of $Q_{\alpha d\epsilon\beta}$-compatibility classes that fail to be linked through a common chain, Extended Rigidity will still hold provided that such classes (and whatever smaller, isolated chains they may be linked with) happen to all share the same value for r.

And even when Extended Rigidity fails, update order will still make no difference to belief strengths when (and only when) each $Q_{\alpha d\epsilon\beta}$-compatibility class has its own value for r that relates each E_j belonging to it by the equation $\text{NL}[Q_{\alpha d}, \epsilon, E_j] = r \cdot \text{NL}[Q_\alpha, \epsilon, E_j]$, provided that the constraint $r = \text{NL}[Q_{\alpha\epsilon}, d, D_i]/\text{NL}[Q_\alpha, d, D_i]$ is satisfied. Indeed, since this is a necessary (and sufficient) condition for belief-strength-equivalent reorderings, Extended Rigidity itself must entail this condition for *its* r.

Finally, let's consider the case where d has the same evidence basis, $\{E_i\}$, as ϵ. This can't happen when $d\epsilon$ is *suitable*. So now we'll see the difference between the condition under which all states may be reordered and the condition for reorderings that preserve order among basis-sharing states. When d has ϵ's basis, what does the condition in clause (2) of the theorem (with the *suitability* condition dropped) say about $\alpha d\epsilon$? Notice that ϵ's basis sentence E_j can be $Q_{\alpha d\epsilon}$-compatible with d's basis sentence E_i (from the same basis) just in case E_j and E_i are the same sentence, since that's the only way to have $Q_{\alpha d\epsilon}[E_i \cdot E_j] > 0$. Also notice that each Normed-Likelihood factor in clause (2) just represents a ratio of belief functions – e.g., $Q_{\alpha d\epsilon}[E_i]/Q_{\alpha d}[E_i] = \text{NL}[Q_{\alpha d}, \epsilon, E_i]$, etc. Plugging these equivalences into clause (2), we find that all it says in the case where d and ϵ share a basis is this: "for each E_i in d and ϵ's basis, if $0 < Q_{\alpha d\epsilon}[E_i]$, then $Q_{\alpha d\epsilon}[E_i] = Q_{\alpha\epsilon d}[E_i]$." That is, clause (2) merely says that ϵ and d commute on their own shared $Q_{\alpha d\epsilon}$-possible basis sentences.

Thus, the necessary and sufficient condition for the belief-strength-preserving reordering of *all* states is identical to the condition that applies when order is maintained among basis-sharing states, supplemented by the condition that pairs of basis-sharing states commute on their own basis sentences. This is precisely the supplement to Extended Rigidity I suggested in the previous section, which yields the Basis Commuting-Version of the Likelihood Ratio Model. I argued there that although Extended Rigidity should not apply to basis sharing states, it is quite plausible that these states should commute on their own basis sentences. And Extended Rigidity together with this supplement guarantees that belief strengths are completely independent of the order in which updates are acquired.

10. Conclusion

Basic Jeffrey Updating is a natural extension of Bayesian updating to cases where the evidence is uncertain. Its only restriction on how sequences of uncertain updates evolve is that each new update must do its work through the mediation of an evidence basis. The standard rigidity requirement is the only constraint on how that mediation works. Its implications are completely captured by the Basic Sequential Update Formula. However, several interesting ways to extend the Basic Model to sequences of updates suggest themselves. Each is motivated by a different conception of how propositionally inexpressible states may influence belief strengths for the sentences of their evidence bases.

Standard Sequential Updating – Jeffrey's original extension of Basic Updating to sequences of states – derives from *rigidity* together with the *Amnestic Update Thesis*. To many investigators this has seemed the most natural approach. Its primary motivation is the idea that a new state should impose itself completely on the belief strengths of its basis sentences, entirely overwhelming any previous belief strengths they may have acquired. The effect of updating is then passed to other sentences by combining the new basis probabilities with previous belief strengths through the Basic Sequential Update Formula. As a result of the direct way in which belief strengths for basis sentences are updated, the belief strengths of all sentences tend to be highly dependent on the order in which each new state is acquired.

Field Updating is an alternative, more Bayesian-like extension of Basic Jeffrey Updating to sequences of states. It derives from *rigidity* together with the *Likelihood-Ratio Update Thesis*. On this model Likelihood-Ratio update factors play a role similar to that of likelihood ratios in standard Bayesian updating. These update factors do not depend on previous belief strengths or on previous states that directly affect other bases. The Extended Rigidity Thesis is an expression of this autonomy of a state's Likelihood-Ratio factors from other states on other bases. The resulting Extended Update Formula shows precisely how updating reduces to the combination of Likelihood-Ratio factors for new states with prior belief strengths. And these Likelihood-Ratio Updates exhibit the independence of update order that one might expect of Bayesian updating.

More generally, picture the idea behind Jeffrey Updating this way. The sentences of a basis function rather like a set orthogonal coordinates against which a non-propositional state is measured. The state projects its shadow some distance along each coordinate. An update factor represents the relative length of that shadow along each coordinate. The basis of a state is *its basis* because it is the best set of coordinates available to the agent for imaging *that* state – for capturing all of the information *it* contains that is relevant to propositionally expressible ways the world might be. Finally, this image of the new state is integrated into a belief function, on an equal footing with the images projected by other states on other bases, to generate new plausibility weightings for all statements. The Sequential Update Formula shows how this integration works.

Likelihood-Ratio factors are well suited to the role of update factors because they employ a scale for measuring the length of a state's shadow along its coordinate bases that does not fall under the influence of the measures of other states on other bases. Amnestic Update Factors share this scale independence, but give a kind of priority to the measure of the most recent state that overpowers the measures of all past states.

Which extension of Basic Jeffrey Updating is the more plausible model of human agents? I'm a logician, not a psychologist. But Amnestic Updating seems psychologically less plausible than the more Bayesian approaches, not because it is un-Bayesian, but because it seems unlikely that we dismiss previous experiences so completely. However, I am mainly interested in whether these models capture useful normative conceptions of belief updating, and whether they might find useful employment in automated reasoning systems.

On the normative side, we want to know whether these models have interesting and useful logical properties, such as a tendency to convergence on true hypotheses, or the ability to single out "best" hypotheses for some interesting conception of *best*. Of particular interest is how each model might link up with standard Bayesian themes such as dutch-book and convergence results. Such questions need to be addressed, but I cannot pursue them here.

Among the models we have investigated, we've found two varieties of Likelihood-Ratio Updating that may provide the most useful normative guides to good reasoning. They are the *Basis-Overwrite Version* and the *Basis-Commuting Version* of the Likelihood-Ratio Update Model. Each is a variety of Field Updating. And either may be appropriate ways to update belief in the right context.

Likelihood-Ratio Models are close kin to Normed-Likelihood Models, but appear to be superior to them. For, Likelihood-Ratio factors have all of the advantages of Normed-Likelihood factors, but avoid the defect of the latter's sensitivity to the prior probabilities of basis sentences.

The Likelihood-Ratio Models should provide very useful approaches to incorporating uncertain evidence into automated Bayesian inference networks. Indeed, for that purpose they seem clearly superior to Standard Sequential Updating, with its amnestic attitude. For, the Amnestic Model depends way too much on update order, due to way it utilizes new information to blot-out the influence of previous states.[28]

APPENDIX: PROOF OF THE UPDATE REORDERING THEOREM

THE UPDATE REORDERING THEOREM. *The following two claims are equivalent:*

(1) *For each state sequence γ, every* suitable *reordering of it δ agrees with it* (*relative to Q*) *– i.e. for all sentences S, $Q_\delta[S] = Q_\gamma[S]$.*

(2) *For every state sequence $\alpha d\epsilon$ with* suitable *part $d\epsilon$ (relative to Q), for each $Q_{\alpha d\epsilon}$-possible D_i in d's basis, there is an $r = \mathrm{NL}[Q_{\alpha\epsilon}, d, D_i]/\mathrm{NL}[Q_\alpha, d, D_i] > 0$ such that for each E_j in ϵ's basis $Q_{\alpha d\epsilon}$-compatible with D_i, $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_j] = r \cdot \mathrm{NL}[Q_\alpha, \epsilon, E_j]$.*

Furthermore, if the restriction to reorderings and sequences $\alpha d\epsilon$ that are suitable is removed from both clauses, the resulting claims, which apply to all reorderings, are still equivalent.

It will be convenient to break this theorem up into two distinct equivalences. The first is this:

THE COMMUTATION REDUCTION THEOREM. *The following two claims are equivalent:*

(1) *For each sequence of states* γ*, every* suitable *reordering of it* δ *agrees with it* (*relative to* Q).
(2) *For each sequence of states* $\alpha d\epsilon\beta$ (α *and* β *possibly empty*)*, every* suitable $d\epsilon$ *commutes* (*for* Q) *– i.e. for all sentence* S, $Q_{\alpha d\epsilon\beta}[S] = Q_{\alpha\epsilon d\beta}[S]$.

Furthermore, if the restriction to reorderings and to sequences $\alpha d\epsilon$ *that are* suitable *is removed, the resulting claims, which applies to* all *reorderings, are still equivalent.*

Proof. This theorem is obvious (and not really needed) when the restriction to *suitable* sequences and reorderings is lifted. So we'll focus exclusively on the case where the *suitability* restriction is in place.

Notice that clause (2) comes from clause (1) directly, since going from $\alpha d\epsilon\beta$ to $\alpha\epsilon d\beta$ does not change the order of states that share a basis. So let's see how to get clause (1) from clause (2).

To get clause (1) from clause (2), consider the following process. Start with γ. It must be of the form $\alpha d\epsilon$, where ϵ is either a single state or a basis-homogeneous block of states and d is a state with a different basis (unless γ is a basis-homogeneous block of states itself, in which case we're done). Reorder, as clause (2) permits, to get $\alpha\epsilon d$. Now keep pushing ϵ in through α (across other states like d) until it runs up against some state e that shares its basis. When that happens we have transformed γ into a sequence of form $\sigma e\epsilon\beta$. Now treat $e\epsilon$ as a basis-homogeneous block, and continue to push *it* on through σ, as clause (2) permits. (If $e\epsilon$ runs against a state g with the same basis, perhaps even as the very next state in σ, just let $ge\epsilon$ be the basis-homogeneous block, and continue). We end up with a transformation of γ into a new sequence with a basis-homogeneous block on the left (having ϵ's basis), with all states in the same order as they occurred in γ; and there are no states with ϵ's basis anywhere else in the new sequence. Now, begin the same process again using whatever state is at the right end of this new sequence (it should be d), and push it in through the new sequence, accumulating into basis-homogeneous blocks as it runs up against other states that share its basis. When done, we have a sequence consisting of an initial left block of states that share d's basis, a block to its immediate right that shares ϵ's basis, and perhaps a remaining heterogeneous sequence going to the right. Begin the process again with the state at the right end. Reiterate the process until all that remains is a sequence of basis-homogeneous blocks, each on a distinct basis. Call this sequence γ^*. Clause (2) implies that $Q_{\gamma^*}[S] = Q_\gamma[S]$.

Now consider any δ that reorders γ but maintains the same order among states that share a basis. Apply the same process to δ, yielding δ^*, which consists of a sequence of basis-homogeneous blocks, each with a distinct basis. Clause (2) implies that $Q_{\delta^*}[S] = Q_\delta[S]$. δ^* must consist of exactly the same basis-homogeneous blocks as γ^*; but perhaps these blocks occur in a different order. However, notice that clause (2) permits any homogeneous block in δ^* to be moved from right to left through the states of an adjacent block, one state at a time. So δ^* can be reordered into γ^*: e.g., reorder δ^* by first finding in it the block that γ^* has at its right end; move that block all the way to the left end of δ^*; call the result 'δ^{**}'; then move the block that γ^* has second from its right end all the way to left end of δ^{**}; call the result 'δ^{***}'; continuing this process eventually transforms δ^* into γ^*. Thus, $Q_\delta[S] = Q_{\delta^*}[S] = Q_{\gamma^*}[S] = Q_\gamma[S]$, which establishes the theorem. □

Here is the second equivalence, which taken together with the theorem just proved yields the Reordering Theorem.

THE COMMUTATION THEOREM. *The following claims are equivalent:*

(1) *For each sequence of states* $\alpha d\epsilon\beta$ *(*α *and* β *possibly empty), every* suitable $d\epsilon$ *commutes (for* Q*) – i.e.* $Q_{\alpha d\epsilon\beta}[S] = Q_{\alpha\epsilon d\beta}[S]$ *for all sentences* S*.*
(2) *For every sequence* $\alpha d\epsilon$ *with* suitable $d\epsilon$ *(relative to* Q*) and for each* $Q_{\alpha d\epsilon}$*-possible* D_i *in d's basis, there is an* $r = \mathrm{NL}[Q_{\alpha\epsilon}, d, D_i]/\mathrm{NL}[Q_\alpha, d, D_i] > 0$ *such that for each* E_j *in* ϵ*'s basis that's* $Q_{\alpha d\epsilon}$*-compatible with* D_i*,* $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_j] = r \cdot \mathrm{NL}[Q_\alpha, \epsilon, E_j]$.

Furthermore, if the restriction to reorderings and to sequences $\alpha d\epsilon$ *that are* suitable *is removed, the resulting claims, which applies to* all *reorderings, are still equivalent.*

Proof. I'll carry out this proof in terms of *suitable* sequences and reorderings. But notice that the proof goes through in the same way without the restriction. The proof of this theorem will be facilitated by first establishing two lemmas.

BASES DECOMPOSITION LEMMA. *Let* $B_i, \ldots, G_m, H_n$ *be basis sentences for states* $b, \ldots, g, h$ *respectively. Both of the following claims hold:*

(1) $Q_{\alpha b\ldots gh}[B_i \cdot \cdots \cdot H_n] = Q_\alpha[B_i \cdot \cdots \cdot H_n] \cdot (Q_{\alpha b}[B_i]/Q_\alpha[B_i]) \cdot \cdots \cdot (Q_{\alpha b\ldots gh}[H_n]/Q_{\alpha b\ldots g}[H_n]) = Q_\alpha[B_i \cdot \cdots \cdot H_n] \cdot \mathrm{NL}[Q_\alpha, b, B_i] \cdot \cdots \cdot \mathrm{NL}[Q_{\alpha b\ldots g}, h, H_n]$*, if none of* $Q_\alpha[B_i], \ldots, Q_{\alpha b\ldots g}[H_n]$ *equal* 0*; otherwise* $Q_{\alpha b\ldots gh}[B_i \cdot \cdots \cdot H_n] = 0$.
(2) *For all sentences* S*,* $Q_{\alpha b\ldots gh}[S] = \sum_i \ldots \sum_n Q_\alpha[S \mid B_i \cdot \cdots \cdot H_n] \cdot Q_{\alpha b\ldots gh}[B_i \cdot \cdots \cdot H_n]$.

Clause (1) of the lemma comes from first observing that when any of $Q_\alpha[B_i], \ldots, Q_{\alpha b\ldots g}[H_n]$ are 0, it must be the case that $Q_{\alpha b\ldots gh}[B_i \cdot \cdots \cdot H_n] = 0$ as well (see note 4); then observe that when none of these are 0, the Basic Sequential Update Formula, applied to state sequence $\alpha b \ldots gh$ and with '$B_i \cdot \cdots \cdot H_n$' plugged in for '$S$', yields the formula in clause (1). Clause (2) follows from the Basic Sequential Update Formula together with clause (1). □

Clause (2) of the lemma shows that the only way in which reordering the sequence of non-propositional states can change the degree of belief in a sentences is by changing the degree of belief in a conjunction of the basis sentences affected by the states involved. That is, the formula in clause (2) implies the:

BASES REDUCTION LEMMA. *Let* $\alpha\gamma$ *be any sequence of states and let* δ *be any reordering of* γ*. Then, for all sentences* S*,* $Q_{\alpha\gamma}[S] = Q_{\alpha\delta}[S]$ just in case *for each conjunction* $B_i \cdot \cdots \cdot G_m \cdot H_n$ *containing exactly one sentence drawn from each basis for the states in* δ*,* $Q_{\alpha\gamma}[B_i \cdot \cdots \cdot G_m \cdot H_n] = Q_{\alpha\delta}[B_i \cdot \cdots \cdot G_m \cdot H_n]$.

For, clearly, if $Q_{\alpha\gamma}[S] = Q_{\alpha\delta}[S]$ for all sentences S, then $Q_{\alpha\gamma}[B_i \cdot \cdots \cdot G_m \cdot H_n] = Q_{\alpha\delta}[B_i \cdot \cdots \cdot G_m \cdot H_n]$ holds for each conjunction of basis sentences $B_i \cdot \cdots \cdot G_m \cdot H_n$. Conversely, if $Q_{\alpha\gamma}[B_i \cdot \cdots \cdot G_m \cdot H_n] = Q_{\alpha\delta}[B_i \cdot \cdots \cdot G_m \cdot H_n]$ holds for each conjunction of the basis sentences $B_i \cdot \cdots \cdot G_m \cdot H_n$, then clause (2) of the Bases Decomposition Lemma implies that $Q_{\alpha\gamma}[S] = Q_{\alpha\delta}[S]$ for all sentences S.

Now we are ready to prove the Commutation Theorem. We first prove the "only if" direction, from (1) to (2) – then the "if" direction. (Notice that the proof still works if all occurrences of the term '*suitable*' are dropped.)

"Only if" direction: Suppose clause (1) of the theorem. And suppose $d\epsilon$ is some particular *suitable* sequence (for Q_α). Let D_i be a $Q_{\alpha d\epsilon}$-possible basis sentence of d's. Let E_j be any basis sentence for ϵ that is $Q_{\alpha d\epsilon}$-compatible with D_i (i.e. such that $Q_{\alpha d\epsilon}[D_i \cdot E_j] > 0$). Then (since we are supposing clause (1) of the theorem – that commutivity holds for all *suitable* $d\epsilon$) we have from clause (1) of the Bases Decomposition Lemma: $0 < Q_\alpha[D_i \cdot E_j] \cdot \text{NL}[Q_\alpha, d, D_i] \cdot \text{NL}[Q_{\alpha d}, e, E_j] = Q_{\alpha d\epsilon}[D_i \cdot E_j] = Q_{\alpha\epsilon d}[D_i \cdot E_j] = Q_\alpha[D_i \cdot E_j] \cdot \text{NL}[Q_\alpha, \epsilon, E_j] \cdot \text{NL}[Q_{\alpha\epsilon}, d, D_i]$. So $\text{NL}[Q_\alpha, d, D_i] \cdot \text{NL}[Q_{\alpha d}, e, E_j] = \text{NL}[Q_\alpha, \epsilon, E_j] \cdot \text{NL}[Q_{\alpha\epsilon}, d, D_i]$ (because $Q_{\alpha d\epsilon}[D_i \cdot E_j] > 0$ implies $Q_\alpha[D_i \cdot E_j] > 0$). Thus, for any $Q_{\alpha d\epsilon}$-possible basis sentence D_i of d's, $\text{NL}[Q_{\alpha\epsilon}, d, D_i]/\text{NL}[Q_\alpha, d, D_i] > 0$, and for each E_j in ϵ's basis that's $Q_{\alpha d\epsilon}$-compatible with D_i, $\text{NL}[Q_{\alpha d}, e, E_j] = (\text{NL}[Q_{\alpha\epsilon}, d, D_i]/\text{NL}[Q_\alpha, d, D_i]) \cdot \text{NL}[Q_\alpha, \epsilon, E_j]$. This completes the "only if" direction.

"If" direction: Suppose clause (2) of the theorem. Consider a particular sequence $\alpha d\epsilon$ with*suitable* $d\epsilon$ (for Q). Let D_i and E_j be particular basis sentences from d's and ϵ's bases, respectively. Then either D_i and E_j are not $Q_{\alpha d\epsilon}$-compatible, or they are. We'll show that in either case, $Q_{\alpha d\epsilon}[D_i \cdot E_j] = Q_{\alpha\epsilon d}[D_i \cdot E_j]$; so this equality holds for all basis sentences D_i and E_j from d's and ϵ's bases. It then follows that $Q_{\alpha d\epsilon}[S] = Q_{\alpha\epsilon d}[S]$ for all sentences S (by the Bases Reduction Lemma), and the Theorem is almost proved. (To complete it we just need to show that adding any additional sequence of states β that contains no states with ϵ's basis will still give $Q_{\alpha d\epsilon\beta}[S] = Q_{\alpha\epsilon d\beta}[S]$ for all sentences S.) So let's address each case.

(*Case* 1): Suppose D_i and E_j are not $Q_{\alpha d\epsilon}$-compatible – i.e. $Q_{\alpha d\epsilon}[D_i \cdot E_j] = 0$. I'll show that $Q_{\alpha\epsilon d}$ should equal 0 as well. But to show this we need help from a very plausible principle. Let x be a state with basis sentence X_k. Note 4 expresses the principle that if $Q_{\alpha y}[X_k] = 0$, then $Q_{\alpha yx}[X_k] = 0$ as well – updating a basis sentence on one of its states should not turn its probabilistic impossibility into a probabilistic possibility. It is also reasonable to add this principle: if $Q_{\alpha xy}[X_k] = 0$ iff $Q_{\alpha yx}[X_k] = 0$ – merely switching the order of the updates involving a state and one of its basis sentence should not change its 0 probability into a positive probability. (One might investigate some kind of updating logic that permits such updates; but update order will obviously be non-commutative for such a logic). So, let's assume this plausible principle.

One more thing: Notice that if $Q_{\beta y}[X_k] > 0$ and X_k is not from y's basis $\{Y_i\}$, then $Q_\beta[X_k] > 0$ (since $0 < Q_{\beta y}[X_k] = \sum_{\{i:Q_\beta[X_i]>0\}} Q_\beta[X_k \cdot Y_i] \cdot (Q_{\beta y}[Y_i]/Q_\beta[Y_i])$, so for some Y_i, $0 < Q_\beta[X_k \cdot Y_i] \le Q_\beta[X_k]$. And a similar result shows that $Q_\alpha[D_i \cdot E_j] > 0$ when $Q_{\alpha\epsilon d}[D_i \cdot E_j] > 0$ – i.e., $0 < Q_{\alpha\epsilon d}[D_i \cdot E_j] = \sum_{\{j:Q_{\alpha\epsilon}[E_j]>0\}} \sum_{\{i:Q_{\alpha\epsilon d}[D_i]>0\}} Q_\alpha[D_i \cdot E_j] \cdot (Q_{\alpha\epsilon}[E_j]/Q_\alpha[E_j]) \cdot (Q_{\alpha\epsilon d}[D_i]/Q_{\alpha\epsilon}[D_i])$; so $Q_\alpha[D_i \cdot E_j] > 0$.

So, to address Case 1: Suppose D_i and E_j are not $Q_{\alpha d\epsilon}$-compatible – i.e. $Q_{\alpha d\epsilon}[D_i \cdot E_j] = 0$. Then $Q_{\alpha\epsilon d}[D_i \cdot E_j]$ should equal 0 as well. For, *suppose otherwise* – i.e. suppose that $Q_{\alpha\epsilon d}[D_i \cdot E_j] > 0$. Then $Q_{\alpha\epsilon d}[E_j] \ge Q_{\alpha\epsilon d}[D_i \cdot E_j] > 0$; so $Q_{\alpha d\epsilon}[E_j] > 0$ and (so) $Q_{\alpha\epsilon}[D_i] > 0$; thus $(Q_{\alpha d\epsilon}[E_j]/Q_{\alpha\epsilon}[D_i]) > 0$. Also, $Q_{\alpha\epsilon d}[D_i] \ge Q_{\alpha\epsilon d}[D_i \cdot E_j] > 0$; so $Q_{\alpha d\epsilon}[D_i] > 0$ and (so) $Q_{\alpha\epsilon}[D_i] > 0$; so by the result in the previous paragraph (applied to each), $Q_{\alpha d}[D_i] > 0$ and $Q_\alpha[D_i] > 0$; thus $(Q_{\alpha d}[D_i]/Q_\alpha[D_i]) > 0$. Also $Q_\alpha[D_i \cdot E_j] > 0$ (from the previous paragraph). Then, $0 < Q_\alpha[D_i \cdot E_j] \cdot (Q_{\alpha d}[D_i]/Q_\alpha[D_i]) \cdot (Q_{\alpha d\epsilon}[E_j]/Q_{\alpha\epsilon}[D_i]) = Q_{\alpha d\epsilon}[D_i \cdot E_j] = 0$ – contradiction! So, $Q_{\alpha d\epsilon}[D_i \cdot E_j] = Q_{\alpha\epsilon d}[D_i \cdot E_j] = 0$.

(*Case* 2): Suppose D_i and E_j are $Q_{\alpha d\epsilon}$-compatible. $Q_{\alpha d\epsilon}[D_i \cdot E_j] > 0$. Then D_i is $Q_{\alpha d\epsilon}$-possible – i.e. $Q_{\alpha d\epsilon}[D_i] > 0$. So, from the main supposition (i.e. clause (2)), $\text{NL}[Q_{\alpha d}, e, E_j] = (\text{NL}[Q_{\alpha\epsilon}, d, D_i]/\text{NL}[Q_\alpha, d, D_i]) \cdot \text{NL}[Q_\alpha, \epsilon, E_j]$. So, $Q_{\alpha d\epsilon}[D_i \cdot E_j] = Q_\alpha[D_i \cdot E_j] \cdot \text{NL}[Q_\alpha, d, D_i] \cdot \text{NL}[Q_{\alpha d}, e, E_j] = Q_\alpha[D_i \cdot E_j] \cdot \text{NL}[Q_\epsilon, e, E_j] \cdot \text{NL}[Q_{\alpha\epsilon}, d, D_i] = Q_{\alpha\epsilon d}[D_i \cdot E_j]$.

Thus, in both cases $Q_{\alpha d\epsilon}[D_i \cdot E_j] = Q_{\alpha\epsilon d}[D_i \cdot E_j]$. So, for this particular sequence $\alpha d\epsilon$ with *suitable* $d\epsilon$ on Q, $Q_{\alpha d\epsilon}[D_i \cdot E_j] = Q_{\alpha\epsilon d}[D_i \cdot E_j]$ for all D_i and E_j from d's and ϵ's bases. Therefore, $Q_{\alpha d\epsilon}[S] = Q_{\alpha\epsilon d}[S]$ for all sentences S (by the Bases Reduction Lemma). And this applies for $\alpha d\epsilon$ with *suitable* $d\epsilon$.

Now, for a given $\alpha d\epsilon$ with suitable $d\epsilon$, let β be any sequence of additional states. Since $Q_{\alpha d\epsilon}$ is precisely the same probability function P as $Q_{\alpha\epsilon d}$ (i.e. they assign the same values to all sentences), updating P to P_β produces, for all sentences S, $Q_{\alpha d\epsilon\beta}[S] = P_\beta[S] = Q_{\alpha\epsilon d\beta}[S]$.

NOTES

[1] See Jeffrey (1987, 1988), and Jeffrey and Hendrickson (1989).

[2] E.g., see Diaconis and Zabell (1982), Garber (1980), and Doring (1999).

[3] I.e., $Q[E_j \cdot E_k] = 0$ for $j \neq k$, and $\sum_i Q[E_i] = 1$. A basis may be finite or countably infinite. If relative to the agent's belief function Q, $\{E_i\}$ is a basis for e, then we take it to remain a partition for each of her updated belief functions, Q_e, Q_{ef}, etc.

[4] I assume that for each Q and state e affecting basis $\{E_i\}$, if $Q[E_i] = 0$, then $Q_e[E_i] = 0$ too; a new experience cannot turn a previously (probabilistically) impossible claim into a (probabilistic) possibility. Also, when $Q[E_j] = 0$ it is convenient to have $Q[S \mid E_j]$ remain defined; define it equal to 1.

[5] *Rigidity* across updates is not required by the axioms of probability; and in some kinds of contexts it may appropriately fail (see, e.g., Levi, 1967). However, in many important contexts where new less-than-certain evidence is acquired, rigidity holds – e.g., in automated Bayesian networks the conditional probabilities of hypotheses on evidence claims remain rigid. For our purposes the agent's series of belief functions need not be perpetually related by rigidity. We are simply studying what happens when rigidity persists through some specific sequence of updates. That is, our starting function Q might well have come from a previous function through some other process. And after a sequence of Jeffrey updates, $Q_a, Q_{ab}, \ldots, Q_{ab\ldots g}$, the very next update of the agent's belief function might well employ some other process to produce a new function R (from $Q_{ab\ldots g}$).

[6] $Q_{ef}[S] = \sum_j Q_e[S \mid F_j] \cdot Q_{ef}[F_j] = \sum_{\{j:Q_e[F_j]>0\}} Q_e[S \cdot F_j] \cdot (Q_{ef}[F_j]/Q_e[F_j]) = \sum_{\{j:Q_e[F_j]>0\}} \sum_i Q[S \cdot F_j \mid E_i] \cdot Q_e[E_i] \cdot (Q_{ef}[F_j]/Q_e[F_j]) = \sum_{\{j:Q_e[F_j]>0\}} \sum_{\{i:Q[E_i]>0\}} Q[S \cdot F_j \cdot E_i] \cdot (Q_e[E_i]/Q[E_i]) \cdot (Q_{ef}[F_j]/Q_e[F_j])$.

[7] Since, where ϵ is some sequence $e \ldots fg$ on a common basis $\{E_i\}$, if $Q_{a\ldots cd}[E_i] > 0$, then $(Q_{a\ldots cd\epsilon}[E_i]/Q_{a\ldots cd}[E_i]) = (Q_{a\ldots cde}[E_i]/Q_{a\ldots cd}[E_i]) \cdot \cdots \cdot (Q_{a\ldots cde\ldots f}[E_i]/Q_{a\ldots cde\ldots}[E_i]) \cdot (Q_{a\ldots cd.e\ldots fg}[E_i]/Q_{a\ldots cd.e\ldots f}[E_i])$.

[8] The term 'Normed-Likelihood' comes from the observation that if e were a sentence, then $Q_{\beta e}[E_j]$ could be represented as a conditional probability. Bayes' theorem would then yield $Q_{\beta e}[E_j] = Q_\beta[E_j \mid e] = (Q_\beta[e \mid E_j]/Q_\beta[e]) \cdot Q_\beta[E_j]$. So $\mathrm{NL}[Q_\beta, e, E_j] = Q_{\beta e}[E_j]/Q_\beta[E_j] = Q_\beta[e \mid E_j]/Q_\beta[e]$. That is $\mathrm{NL}[Q_\beta, e, E_j]$ effectively plays the role of the likelihood of e given E_i (for Q_β) divided by a normalization factor $Q_\beta[e]$.

[9] *Empirical* if you want an approximate model of human agents; *pragmatic* if you want to build a usable, logically sound automated reasoning system; epistemological (or logical) if you are constructing a normative theory of belief updating.

[10] Those who have supported this approach include Jeffrey (1965, 1975), Diaconis and Zabell (1982), Garber (1980), and Doring (1999). Almost everyone seems to consider it to be an essential part of Jeffrey updating. But even its defenders seem somewhat unhappy with it.

[11] To see how this works consider how a belief function Q_β is updated on states e and f. Applying Amnestic Updating down to β in the Basic Sequential Update Formula we get $Q_{\beta ef}[S] = \sum_{\{j:Q_{\beta e}[F_j]>0\}} \sum_{\{k:Q_\beta[E_k]>0\}} Q_\beta[S \cdot E_k \cdot F_j] \cdot (Q_{\beta e}[E_k]/Q_\beta[E_k]) \cdot (Q_{\beta f}[F_j]/Q_{\beta e}[F_j])$, where $Q_{\beta e}[F_j]$ itself decomposes into $Q_{\beta e}[F_j] = \sum_{\{k:Q_\beta[E_k]>0\}} Q_\beta[E_k \cdot F_j] \cdot (Q_{\beta e}[E_k]/Q_\beta[E_k])$. Thus, $Q_{\beta ef}[S] = \sum_{\{j:Q_{\beta e}[F_j]>0\}} \sum_{\{k:Q_\beta[E_k]>0\}} Q_\beta[S \cdot E_k \cdot F_j] \cdot (Q_{\beta e}[E_k]/Q_\beta[E_k]) \cdot (Q_{\beta f}[F_j]/ \sum_{\{k:Q_\beta[E_k]>0\}} Q_\beta[E_k \cdot F_j] \cdot (Q_{\beta e}[E_k]/ Q_\beta[E_k]))$. The occurrence of the factor $\sum_{\{k:Q_\beta[E_k]>0\}} Q_\beta[E_k \cdot F_j] \cdot (Q_{\beta e}[E_k]/Q_\beta[E_k])$ in the denominator of the second ratio term (with no analogous factor in the first ratio term) carries the order effect from the fact that state e came before state f. If this decomposition is continued for states in sequence β, the result is a formula in which these denominator terms themselves have expanded denominator terms, and so on. Ultimately this expansion results in a formula that involves only Q and the update probabilities $Q_a, \ldots, Q_e, Q_f$ applied to their own bases. The effect of update order is then carried in the structure of this formula.

[12] Theorem 3.2 in (Diaconis and Zabell, 1982).

[13] Order-independence may occur only when $Q_\beta[F_j] = Q_{\beta e}[F_j] = \sum_i Q_\beta[E_i \mid F_j] \cdot Q_\beta[F_j] \cdot (Q_{\beta e}[E_i]/Q_\beta[E_i])$; i.e. only when $\sum_i Q_\beta[E_i \mid F_j] \cdot (Q_{\beta e}[E_i]/Q_\beta[E_i]) = 1$. And (similarly) only when $\sum_j Q_\beta[F_j \mid E_i] \cdot (Q_{\beta f}[F_j]/Q_\beta[F_j]) = 1$ as well.

[14] If the belief strengths for the test outcomes remain at $Q_f[\sim F] = .9$ and $Q_e[E] = .9$ but the likelihoods are more extreme, then we get a more extreme difference in updating due to order. E.g., if it happens that the likelihoods are $Q[E \mid C] = Q[F \mid C] = Q[\sim E \mid \sim C] = Q[\sim F \mid \sim E] = .99$, then $Q_{fe}[C] = .83$ and $Q_{ef}[C] = .17$.

[15] Diaconis and Zabell (1982) prove results that indicate how ubiquitous the order effect will be. Doring (1999) and Lange (2000) discuss its implausibility.

[16] $Q_\alpha[E_i \mid d \cdot \epsilon] \div Q_\alpha[E_i \mid d] = (Q_\alpha[d \mid E_i \cdot \epsilon] \cdot Q_\alpha[\epsilon \mid E_i] \cdot Q_\alpha[E_i]/Q_\alpha[d \cdot \epsilon]) \div (Q_\alpha[d \mid E_i] \cdot Q_\alpha[E_i]/Q_\alpha[d]) = (Q_\alpha[d \mid E_i \cdot \epsilon]/Q_\alpha[d \mid E_i]) \cdot Q_\alpha[\epsilon \mid E_i] \cdot (Q_\alpha[d]/Q_\alpha[d \cdot \epsilon]) = (Q_\alpha[d \mid E_i \cdot \epsilon]/Q_\alpha[d \mid E_i]) \cdot (Q_\alpha[E_i \mid \epsilon]/Q_\alpha[E_i]) \cdot (Q_\alpha[d] \cdot Q_\alpha[\epsilon]/Q_\alpha[d \cdot \epsilon])$.

[17] For our purposes the independence of d from ϵ given E_i is stronger than necessary. It would suffice for ϵ to multiply $Q_\alpha[d \mid E_i]$ by the same constant $s > 0$ for every sentence of $\{E_i\}$ – i.e. $Q_\alpha[d \mid E_i \cdot \epsilon] = s \cdot Q_\alpha[d \mid E_i]$. Then we would have $\mathrm{NL}[Q_{\alpha d}, \epsilon, E_i] = \mathrm{NL}[Q_\alpha, \epsilon, E_i] \cdot (Q_\alpha[d] \cdot Q_\alpha[\epsilon]/Q_\alpha[d \cdot \epsilon]) \cdot s$.

[18] It updated $Q_\alpha[E_i]$ to $Q_{\alpha d}[E_i] = \sum Q_\alpha[E_i \cdot D_k] \cdot (Q_{\alpha d}[D_k]/Q_\alpha[D_k])$.

[19] Wagner (2002, note 7) raises a similar point.

[20] In his most recent work Jeffrey favors updating based on Likelihood-Ratio factors as well. He thinks of them as ratios of new to old odds and calls them ‘Bayes factors’, following Good (1950). See (Jeffrey, 1992, pp. 7–9) and (Jeffrey and Hendrickson, 1989, pp. 16–17). The present treatment extends Field’s approach from binary bases of form $\{E, \sim E\}$ to countable bases. Also, Field and Jeffrey make no distinction between basis homogeneous and heterogeneous sequences of states. In effect they take Extended Rigidity to apply to all states, regardless of whether they share a common basis. Later I will explain why one should not assume that Extended Rigidity applies *within* basis homogeneous sequences.

[21] To derive this formula, start with the Basic Sequential Update Formula. Notice that each term of form $Q_{a\ldots ef}[F_j]/Q_{a\ldots e}[F_j]$ (for each state f and its basis sentence F_i) equals a formula of form $\mathrm{LR}[Q_{a\ldots e}, f, F_j, F] \cdot (Q_{a\ldots e}[F]/Q_{a\ldots ef}[F_i])$ (for some F from f's basis). Substituting such equivalent terms for each term in the Basic Sequential Update Formula yields a formula of form $Q_{a\ldots efg}[S] = K \cdot \sum_{\{i:Q_{a\ldots ef}[G_i]>0\}} \cdots \sum_{\{k:Q[A_k]>0\}} Q[S \cdot A_k \cdot \cdots \cdot G_i] \cdot \mathrm{LR}[Q, a, A_k, A] \cdot \cdots \cdot \mathrm{LR}[Q_{a\ldots ef}, g, G_i, G]$, where K is a product of form $(Q[A]/Q_a[A]) \cdot \cdots \cdot (Q_{a\ldots ef}[G]/Q_{a\ldots efg}[G])$. Notice this implies $1 = Q_{a\ldots ef}[\text{tautology}] = K \cdot \sum_{\{i:Q_{a\ldots ef}[G_i]>0\}} \cdots \sum_{\{k:Q[A_k]>0\}} Q[A_k \cdot \cdots \cdot G_i] \cdot \mathrm{LR}[Q, a, A_k, A] \cdot \cdots \cdot \mathrm{LR}[Q_{a\ldots ef}, g, G_i, G]$, giving a formula for K. Thus, $Q_{a\ldots efg}[S] = \sum_{\{i:Q_{a\ldots ef}[G_i]>0\}} \cdots \sum_{\{k:Q[A_k]>0\}} Q[S \cdot A_k \cdot \cdots \cdot G_i] \cdot \mathrm{LR}[Q, a, A_k, A] \cdot \cdots \cdot \mathrm{LR}[Q_{a\ldots ef}, g, G_i, G] \div \sum_{\{i:Q_{a\ldots ef}[G_i]>0\}} \cdots \sum_{\{k:Q[A_k]>0\}} Q[A_k \cdot \cdots \cdot G_i] \cdot \mathrm{LR}[Q, a, A_k, A] \cdot \cdots \cdot \mathrm{LR}[Q_{a\ldots ef}, g, G_i, G]$. Now just apply the Likelihood-Ratio Version of Extended Rigidity to the terms of this formula.

[22] This result is basically equivalent to Wagner's (2002) Theorem 3.1, which extends Field's (1978) commutivity result. But Field and Wagner don't single out basis homogeneous sequences of states for special treatment. In effect they assume that Extended Rigidity applies to all states, regardless of whether they share a basis. But that assumption is too strong. It leads the Garber's (1980) objection, as we will see.

[23] Jeffrey and Hendrickson (1989) employ a somewhat similar example of how Likelihood-Ratio factors inform a cancer diagnosis. However, their example employs two update factors affecting a common basis: {cancer, no cancer}. The present example differs in an important respect. It involves update factors that affect two distinct, lower level bases that are relevant to the higher-level cancer diagnosis partition.

[24] E.g., the cancer diagnosis example employs separate bases for x-ray and sputum cytology results.

[25] Since in that case we would have $\mathrm{LR}[Q_{\alpha\epsilon}, e, E_j, E_k] = Q_\alpha[e \mid E_j \cdot \epsilon]/Q_\alpha[e \mid E_k \cdot \epsilon] = Q_\alpha[e \mid E_j]/Q_\alpha[e \mid E_k] = \mathrm{LR}[Q_\alpha, e, E_j, E_k]$.

[26] The notion of *suitability* need only be relativized to Q if some other belief function under consideration, R, may disagree with Q about whether two states share a common basis.

[27] Wagner's (2002) Theorem 4.1 provides a necessary condition for commutivity that is closely related to this result.

[28] Thanks to Chris Swoyer and Adam Morton for many helpful comments.

REFERENCES

Diaconis, P. and Zabell, S. L., 1982: Updating subjective probability, *J. Amer. Statist. Assoc.* **77**, 822–830.

Doring, F., 1999: Why Bayesian psychology is incomplete, *Philos. Sci.* **66**, S379–S389.

Field, H., 1978: A note on Jeffrey conditionalization, *Philos. Sci.* **45**, 361–367.

Garber, D., 1980: Discussion: Field and Jeffrey conditionalization, *Philos. Sci.* **47**, 142–145.

Good, I. J., 1950: *Probability and the Weighing of Evidence*, Charles Griffin, London.

Jeffrey, R. C., 1965: *The Logic of Decision*, McGraw-Hill, New York (2nd edn 1983). University of Chicago Press, Chicago.

Jeffrey, R. C., 1975: Carnap's empiricism, in G. Maxwell and R. M. Anderson, Jr. (eds.), *Minnesota Studies in Philosophy of Science*, Vol. 6, University of Minnesota Press, Minneapolis, pp. 37–49.

Jeffrey, R. C., 1987: Alias Smith and Jones: The testimony of the senses, *Erkenntnis* **26**, 391–399. Reprinted in (Jeffrey, 1992).

Jeffrey, R. C., 1988: Conditioning, kinematics, and exchangeability, in B. Skyrms and W. L. Harper (eds.), *Causation, Chance, and Credence*, Vol. 1, Kluwer, Dordrecht, pp. 221–255. Reprinted in (Jeffrey, 1992).

Jeffrey, R. C., 1992: *Probability and the Art of Judgment*, Cambridge University Press, Cambridge.

Jeffrey, R. and Hendrickson, M., 1989: Probabilizing pathology, *Proc. Aristotelian Soc.* **89**, 211–225.

Lange, M., 2000: Is Jeffrey conditionalization defective in virtue of being non-commutative? Remarks on the sameness of sensory experience, *Synthese* **123**, 393–403.

Levi, I., 1967: Probability kinematics, *British J. Philos. Sci.* **18**, 197–209.

Wagner, C. G., 2002: Probability kinematics and commutativity, *Philos. Sci.* **69**, 266–278.

Department of Philosophy
University of Oklahoma
Norman, OK 73069, U.S.A.
e-mail: hawthorne@ou.edu