Report on *Asymmetry and the Geometry of Reason* (BJPS-2016-115) 5 May 2016

Suppose that you have assessed a prior probability distribution p over a set Ω of possible states of the world, and you are apprised of new evidence that prompts you to revise p. The evidence is such that any distribution q belonging to the set Q is admissible as a posterior. Is there under such circumstances a rational way to select some $q \in Q$ to which you ought to revise p? A number of different methods have been proposed as an answer to this question:

(1) Choose a $q \in Q$ that minimizes $d(p,q)$, where d is a metric on the set of all probability distributions on Ω (keeping in mind that there are many such metrics, and that there may be more than one such q in some situations).

(2) Choose a $q \in Q$ that minimizes $\mathrm{KL}(q, p)$, the Kullback-Leibler divergence of q from p.

(3) Leitgeb-Pettigrew updating: Choose a $q \in Q$ that minimizes "diachronic expected global inaccuracy," where inaccuracy is quantifed by a Brier-type score, and expectation is taken with respect to the prior p.

The author (henceforth, A) of this paper exhibits two lists (list A and list B) of properties that might be thought to be desirable for an updating principle, and catalogues the properties enjoyed by the various updating methods described above, with an emphasis on (2) and (3), providing counter-examples in the cases where a property fails to obtain. The mathematical calculations are competently done, but the entire exercise is unsatisfying on several counts:

1. The introductory sections are badly organized, using technical terms that have not been defined (see section 2.2, for example, where, all of a sudden, A states a couple of axioms due to Joyce for a function $I(a,b)$, which is only partly defined in the following paragraph. One gathers that a and b are real numbers, and A tells us that $I(a,b)$ measures "the change in belief in going from a to b," but that is hardly sufficient to appreciate the provenance of these axioms, or why A brings them up at this early point in the paper). And really, the pompous talk about the "isomorphism" from the set probability mass functions on an $n-$ element set Ω to the obvious set of vectors in R^n is hardly necessary. In addition, A says in the introductory paragraph that "for information theory…the underlying topology for credence functions is not a metric space." But one of the references, the lovely 1988 Kopperman paper (which, strangely, A cites only for a definition of "metric," a term introduced by Fréchet back in 1906), contains a result that rules out (due to the asymmetry of $\mathrm{KL}(q, p)$) using KL to construct a topology on the set of credence functions.

2. A’s exposition of Leitgeb and Pettigrew’s alternative to Jeffrey conditioning, based on a numerical example, is insufficient to appreciate what L & P are up to.

3. The philosophical lessons of A’s cataloguing of the properties of various constrained updating methods are not sufficiently spelled out. A is clearly more positive about Kullback-Leibler updating than about Leitgeb-Pettigrew updating, but is perturbed by the non-symmetry of KL-divergence and its consequences (which A, rightly, says “begs for an explanation). A paper that provided even the beginnings of such an explanation would be valuable.