Finding Balance in Uncertain Times

Adrian Liu (adrian.liu@rutgers.edu)
Formal Approaches to Rationality and Meaning, UMD College Park, Oct 11 2025
This handout also at adrianliu.net/balance

Let π describe the credences you in fact have and let R describe the rational credences to have in your situation, *whatever they might be*.

Let $[R(q) = t]$ be the proposition that the rational credence to have in proposition q is t. Then you **reflect** the rational opinions on a q if your credence in q, conditional on the supposition that the rational credence to have in proposition q is t, is t.

Formally, π **reflects** R if for any proposition q and any real number t, we have $\pi(q \mid [R(q) = t]) = t$.

Fact 1. *If some prior credence* π reflects *a posterior credence* R*, then:*

1. π ***trusts*** R*:* π *expects* R *to give better estimates than* π*.*
2. π ***balances*** R*:* π *expects* R *to have the same estimates as* π *on average.*
3. R *is **higher-order certain**:* R *is certain about the actual value of* R

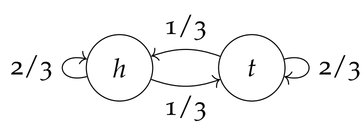
Some think that *reflection* could not be a rational requirement on updating because it entails *higher-order certain* posteriors, and requiring higher-order certain posteriors could not be a rational requirement.[1] If rational updating allows for higher-order uncertainty, there's a further question whether rational updating requires trust and balance. **In this talk:** how much balance can you get in conditions of higher-order uncertainty? **Answer:** not much at all. **Upshot**: It will be hard to both allow higher-order uncertainty and require balance in one's theory of rational epistemic updating.

[1] Williamson (2011), Elga (2013), Dorst (2020), among others.

§1. Higher-Order Uncertainty

One Coin: A fair coin is flipped out of your sight. If the coin lands heads, you get some vague evidence that warrants raising your credence in heads to 2/3, and likewise if the coin lands tails. Then we can visualize the higher-order uncertainty like below.

At world h, you should be 2/3 confident in world h. But then you should be 1/3 confident in world t, where it is rational to be 1/3 confident in world h. So it is rational to be unsure what the rational credences are.



A Kripke diagram where probabilities are added. Here h is the world where the coin lands heads (and t tails), and an arrow with label t from w_1 to w_2 means that if w_1 is actual, you should have credence t in w_2.

We can also model your epistemic situation the following way:

$$\textit{prior:} \left(\begin{array}{c|cc} & h & t \\ \hline \pi & 1/2 & 1/2 \end{array} \right) \qquad \textit{posterior:} \left(\begin{array}{c|cc} & h & t \\ \hline \rho_h^@ & 2/3 & 1/3 \\ \rho_t & 1/3 & 2/3 \end{array} \right)$$

Each row is a credence function and each column of a row is that credence function's credence in the respective world. Here ρ_h is your rational posterior if the coin lands heads (resp. ρ_t for tails), and the @ marks the actual world (unbeknownst to the epistemic agent).

Suppose that in fact (after the toss), you're at world h, so you should have ρ_h and be 2/3 confident in *heads*. Then you're 1/3 confident in tails, and if *tails*, you should be 2/3 confident in *tails*. So you should be only 2/3 confident that your 2/3 confidence in *heads* is rational.

Here is how we'll formalize higher-order uncertainty: we take the idea that the rational credences are different at different worlds, and capture that with functions $R : \Omega \to \mathbb{C}(\Omega)$ mapping worlds to credence functions, so that ρ_w, the value of R at w, can be understood as R's credences at w. I call these functions **credence families**. Then we'll say a credence family R is **higher-order uncertain** if, at some world w, the credence function ρ_w that R takes on at w is uncertain about the proposition "R takes on ρ_w."

Formally, a credence function family $R : \Omega \to \mathbb{R}$ is higher-order uncertain if for some $w \in \Omega$, $\pi_w([R = \rho_w]) < 1$, where $[R = \rho_w] =_{\text{def}} \{w' \in \Omega \mid \rho_{w'} = \rho_w\}$ (Dorst 2019).

Many epistemologists think that *some* form of higher-order uncertainty is possible for rational epistemic agents.[2] But if it threatens to conflict with other plausible rational principles, then we have to weigh our options. Is *balance* a potential issue?

[2] Williamson (2011), Elga (2013), Dorst (2019), Gallow (2021), Zendejas Medina (2024), Schultheis (forth), among others.

§2. *Balanced Updating*

with nothing new in it, but every pain and every joy and every thought and

Suppose the events of *One Coin* happened several times over. What would you expect your average credence in *heads* to be?

$$\pi(h) \cdot \rho_h(h) + \pi(t) \cdot \rho_t(h) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}.$$

This makes sense, and notably it matches your *prior* credence in *heads*: $\pi(h) = \frac{1}{2}$. If the coin is fair, then it should land heads half the time (on average), and it seems irrational for you to give answers that significantly diverge from that average.

This requirement, that your prior credence in a proposition and your prior expectation of your posterior credence in a proposition match, can be formalized as requiring that the prior *balance* the posterior:

Definition 1. *π **balances** R on a proposition q if*

$$\pi(q) = \mathbb{E}_\pi[R(q)], \text{ where } \mathbb{E}_\pi[R(q)] =_{def} \sum_{w \in \Omega} \pi(w) \cdot \rho_w(q).$$

A straightforward motivation for thinking balance is a rational requirement is that being unbalanced is being *biased* in a way you could expect or predict (and therefore, perhaps, prevent) beforehand.[3] For example, suppose that when the coin came up *tails*, you got *slightly stronger* evidence about the true outcome, warranting increasing your credence in t to $5/6$. Then we'd have the following:

$$\mathbb{E}_\pi(R(h)) = \pi(h) \cdot \rho_h(h) + \pi(t) \cdot \rho_t(h) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{6} = \frac{5}{12} < \frac{1}{2},$$

so over a large number of trials, you'd think that only around 40% of the coins came up heads, even though you get no evidence that the coin is tails-biased. If we think balance is a rational requirement, then these "warranted credences" I stipulated must not be possible.

[3] Salow (2017) shows that cases of balance failure can be co-opted to perform "intentionally-biased" inquiry that guarantees favorable results, and Dorst (2023) proposes a mechanism in which balance failures lead to predictable changes in one's beliefs.

§3. *Higher-Order Uncertain Priors*

I've discussed updates from priors to higher-order uncertain posteriors, and I've implicitly assumed that these priors are higher-order certain.[4,5] But once a higher-order certain prior updates into a higher-order *uncertain* posterior, it becomes a higher-order uncertain prior for the purposes of further updating, and matters get trickier.

Two Coins:[6] In addition to the fair coin from the last example, another fair coin is independently flipped to the right of the first coin, out of your sight. You're told nothing about how it landed, so you are 1/2 confident that it came up heads. (In fact, it came up tails, but you don't know that.) We can fine-grain both the prior and posterior from before to capture the second coin:

$$\left(\begin{array}{c|cccc} & hh & ht & th & tt \\ \hline \pi & 1/4 & 1/4 & 1/4 & 1/4 \end{array}\right) \to \left(\begin{array}{c|cccc} & hh & ht & th & tt \\ \hline \rho_{hh} & 1/3 & 1/3 & 1/6 & 1/6 \\ \rho_{ht}^{@} & 1/3 & 1/3 & 1/6 & 1/6 \\ \rho_{th} & 1/6 & 1/6 & 1/3 & 1/3 \\ \rho_{tt} & 1/6 & 1/6 & 1/3 & 1/3 \end{array}\right)$$

As we might expect, π still balances R. But now suppose you're told whether or not *both* coins came up heads. That is, you're told the true cell of the partition $\{\{hh\}, \{ht, th, tt\}\}$. You learn this with certainty, and thus, since you've learned the true cell of a partition with certainty, you update by conditionalizing on this information.

$$\left(\begin{array}{c|cccc} & hh & ht & th & tt \\ \hline \rho_{ht}^{@} & 1/3 & 1/3 & 1/6 & 1/6 \end{array}\right) \to \left(\begin{array}{c|cccc} & hh & ht & th & tt \\ \hline \rho_{ht}^{+@} & 0 & 1/2 & 1/4 & 1/4 \end{array}\right)$$

However, before conditionalizing, you are higher-order uncertain: you are not certain that you are at world ht and have ρ_{ht}. From your perspective, before you are told whether both coins came up heads, you leave all the worlds, and thus all of $\rho_{hh}, \rho_{ht}, \rho_{th}, \rho_{tt}$, open.

Thus you have *four* candidate credence functions that could have been the one you should update. From your perspective, things more like this (and you can't even see where the @ symbol is):

$$\left(\begin{array}{c|cccc} & hh & ht & th & tt \\ \hline \rho_{hh} & 1/3 & 1/3 & 1/6 & 1/6 \\ \rho_{ht}^{@} & 1/3 & 1/3 & 1/6 & 1/6 \\ \rho_{th} & 1/6 & 1/6 & 1/3 & 1/3 \\ \rho_{tt} & 1/6 & 1/6 & 1/3 & 1/3 \end{array}\right) \to \left(\begin{array}{c|cccc} & hh & ht & th & tt \\ \hline \rho_{hh}^{+} & 1 & 0 & 0 & 0 \\ \rho_{ht}^{+@} & 0 & 1/2 & 1/4 & 1/4 \\ \rho_{th}^{+} & 0 & 1/5 & 2/5 & 2/5 \\ \rho_{th}^{+} & 0 & 1/5 & 2/5 & 2/5 \end{array}\right)$$

where each ρ_w is conditionalized on the true cell of the partition $\{\{hh\}, \{ht, th, tt\}\}$, so:

$$\begin{aligned} \rho_{hh}^{+} &= \rho_{hh}(\cdot \mid hh) \\ \rho_{ht}^{+} &= \rho_{ht}(\cdot \mid \neg hh) \\ \rho_{th}^{+} &= \rho_{th}(\cdot \mid \neg hh) \\ \rho_{tt}^{+} &= \rho_{tt}(\cdot \mid \neg hh) \end{aligned}$$

Here R represents the intermediate priors[7] you might have, and R^{+} represents the posteriors you might have: every prior has simply been conditionalized on whichever of $hh, \neg hh$ the agent with that intermediate prior would learn.

[4] That is, the priors are uncertain about some things in the world, but they are certain that the prior should be exactly what it in fact is.

[5] In these settings, Gallow (2021) and Isaacs & Russell have exhibited ways for updates to always satisfy balance. I have also argued (ms), in a framework more similar to Dorst's, that evidence cannot warrant unbalanced levels of higher-order uncertainty.

[6] I owe this example to Kevin Dorst.

[7] Not the *original* prior π: the one between the first update and the second.

§4. *The Problem for Balance*

In *Two Coins*, from the perspective of the epistemic agent, without the theorist's inadmissible evidence, the agent is updating a credence family R to another credence family R^+ by conditionalization. So it might be sensible to ask if balance is satisfied here: if R balances R^+.

(knowledge of which world is actual)

Let's say that R balances R^+ if ρ balances R^+ for every $\rho \in R$. Then:

- In *Two Coins*, R does not balance R^+.
- *Also,* no matter which world is actual (and thus which of the priors is actual), the prior ρ does not balance R^+.
- *Also,* the original prior π does not balance R^+.

For example:
$\rho_{hh}(h\_) = 2/3$ but $\mathbb{E}_{\rho_{hh}}[R^+(h\_)] \approx 0.56$
$\rho_{th}(h\_) = 1/3$ but $\mathbb{E}_{\rho_{th}}[R^+(h\_)] \approx 0.38$
$\pi(h\_) = 1/2$ but $\mathbb{E}_{\pi}[R^+(h\_)] = 0.475$.

If balance is a rational requirement, then one of these updates, $\pi \to R$ or $R \to R^+$, must be irrational. But surely it cannot be $R \to R^+$, which is conditionalization on veridical, partitional evidence.[8] And if $\pi \to R$ is not rational, then the prospects of rational higher-order uncertainty at all might be grim, given how simple a case it is.

[8] Which, as Greaves & Wallace (2006), Schoenfield (2017), and Briggs and Pettigrew (2020), teach us, maximize expected accuracy *and* accuracy dominate other update methods.

The Problem is General

Two Coins is not an edge case. *Whenever* we have a credence family R with higher-order uncertainty, then, given some minimal conditions,[9] there will be some way to conditionalize R on partitional evidence into an R^+ such that some R does not balance R^+.

If we accept that higher-order uncertainty like the one-coin case is possible.

[9] The first being that there are at least three possible worlds; the second being self-confidence (below).

Theorem 1. *Let Ω have at least three possible worlds and let R be a higher-order uncertain, self-confident[10] credence family. Then there is some partition Π of Ω with at least two nonempty cells such that R does not balance the credence family R^+ obtained by conditionalizing each R_w on the partition cell $c_w \in \Pi$ that contains w.*

[10] A credence family R is **self-confident** if $\forall w \in \Omega$, $R_w(w) > 0$. I take this to be a rather minimal condition (strictly less demanding than regularity).

The result means, in effect, that whenever an agent has higher-order uncertainty, they are always at risk of being in a state where even the update rule "conditionalize on partitional evidence" leads to expectable biases.

Once higher-order uncertainty exists, it's very difficult to stay balanced, even if no further uncertainty is introduced. Those (like me and Dmitri) who think balance is a crucial component of epistemic rationality, but who also take higher-order uncertainty to be a central (and epistemically rational) phenomenon to capture, will need to search harder for an explanation of how the two can be reconciled.

I am grateful to Kevin Dorst and Dmitri Gallow for insightful and helpful feedback on versions of this paper.