Untangling Infrabayesianism: A Redistillation

Lorxus

Summer 2023

Intro

So you want to understand infrabayesianism, to hack to the center of that thorny wood and seek out and recover any treasures hidden there? You've come to a correct creature for a guide. If you want to journey there, make sure you've already got the necessary tools well in hand: some simple decision theory, the basics of topology and linear algebra, and a little measure theory - for that last, if you know how a Lebesgue integral is defined and why no reasonable σ-algebra can encompass the full power set, then you're already doing fine. If you find yourself struggling with such things, reach out to me on Discord and I'll see what we can do.

Infrabayesianism seems like exactly what we might need as alignment researchers: a way to discuss all of our usual decision-theoretic questions while also getting to account for uncertainty about the world, compensate for policy-dependent environments and adversarial selection, and even talk about UDT puzzles. It does this by fundamentally being a decision theory that has explicit reasonable machinery for handling Knightian uncertainty about the environment due to nonrealizable or nonlearnable hypotheses while still permitting nontrivial inference and planning.

Three major brambly hedges block the way between you and understanding: the prickly snagging of the frequently unclear, unintuitive, or just plain lacking notation used in the original infrabayesian sequence; thorny philosophical tangles up front and scattered throughout; and math and its accompanying density of concept and notation getting thicker as we go deeper in. Follow me, though, and we'll slip right through them with barely a scratch, and eat a couple of delicious berries from right off their vines. In fact, I can tell you up front that if you haven't read the original infrabayesianism sequence too closely and aren't that familiar with its notation... that's an active benefit, because we won't need most of it here. We won't be cleaving perfectly to its choices of notation or terminology, though I will eventually provide a dictionary between the two as a postscript.

1 Philosophical Tangles (Some Stay Dry While Others Feel The Pain)

First, I'll need to bring you through the hedge of philosophical brambles, tangled and sharp. We'll begin with the doctrine of the least-wrong answer, which we also call nonrealizability. Given that infrabayesian reasoning should effectively strictly extend bayesian reasoning, we should expect that any answer it produces will be the least-wrong answer in our hypothesis space - that is, if the right answer is somewhere in there, our reasoning process should converge to it, and if it is not, then there should be a reasonable sense in which it arrives at the least-wrong answer - that it got as close as possible to the right answer among the limited hypotheses it could consider.

Infrabayesian deciding is also a fundamentally pessimistic process. Its stated goal is to maximize the minimum payout it can attain, no matter how wrong it turns out to be about the state of the world or what environments the world throws at that which uses it. To this end, we introduce a fictitious agentic force to our model: that of the Demiurge. The Demiurge's goal is not to make you suffer, but rather to simply oppose you at every turn, that you might not simply get your way. It is thus just as greedy as It is impersonal and mechanistic. Worse yet, the Demiurge is long-sighted as well: if It could choose to put us in a universe with immediate payoff -0.8 and where we get $+1$ reward ever after, It will never do so if the alternative is a universe with immediate reward $+0.999$ but where our reward will never be positive again. We might model the Demiurge as existing outside of time, looking at the possible paths that our IB-agent could take through the possible histories the Demiurge could present it with, and behaving like a variational principle to present an argmin to our IB-agent to live in. We don't simply lose, though: the Demiurge is also intensely arrogant, and seeks to preserve Its pride. This gives us a single major tactic: *flipping off the Demiurge*. How does this work? Simply put, a key part of our plan will be to prepare to - whenever the Demiurge presents us with a universe we know for sure can't exist, an impossible possible world - tell It to *go pound sand up Its ass*. Omega isn't predicting perfectly? Causality has broken down? Time to realize that we're only inside the Demiurge's dream and flip It off to force It to awake! I promise you, we'll make all of this philosophy explicit and mathematical soon enough, but the Demiurge framing, dense and thorny as it may be, is something we'll need to treat as real if we want to see why it's not actually necessary.

2 Measure Theory (Makes You Happy Living In A Gate)

The other major bramble is the density of mathematics, and there's no royal road past that. I'll assume that you have the basic chops in measure theory and topology that I mentioned in the intro. You can find my Discord trivially easily if you need math help and if you do and tell me IB is why you're DMing me, I won't just block you as spam. To begin with, a little bit of notation that's more standard in computer science than in pure math:

Definition 2.1. *Let f, g, h be functions with $x \in \mathrm{dom}(f)$, and with P a predicate that might hold on $\mathrm{dom}(f)$. Then we will use $f(x) = P(x)?g(x) : h(x)$ to mean that f is the function which takes all x such that $P(x)$ to $g(x)$, and otherwise takes x to $h(x)$.*

This is technically off-standard notation for pure math, but it lets us express simple conditional-

branch functions much more cleanly.

Definition 2.2. *Let X be a metric space. We say that a functional $f : X \to \mathbb{R}$ is k-**Lipschitz continuous** if there exists $k \geq 0 \in \mathbb{R}$ such that for all $x, y \in X$, $d_X(f(x), f(y)) \leq k \cdot |x - y|$. In such a case k is called the **Lipschitz constant** of f.*

Loosely speaking, the heart of it is that Lipschitz-continuous functions are ones that are not only continuous in the ordinary sense, but they "don't change all that fast, either", and the speed of that change is measured by the parameter k. In particular, for $k \leq 1$, it means the function is also absolutely and thus uniformly continuous, and that it's also almost-everywhere differentiable.

Definition 2.3. *Let $\mu, \nu \in \mathcal{M}^{\pm}(X)$, and denote by L the set of all Lipschitz-continuous functions $f : X \to [-1, 1]$ with Lipschitz constant at most 1. Then the KR-distance between the two measures is given by $d_{KR}(\mu, \nu) := sup_{f \in L} |\mu(f) - \nu(f)|$.*

When we talk about Lipschitz continuity, we should be thinking of especially nice continuous functions. In particular, every Lipschitz-continuous function is differentiable almost everywhere and has derivative bounded by the Lipschitz constant everywhere it has a derivative. As for *why* we choose this metric instead of any other, we should ponder the nature of convergence in arbitrary metric spaces. If you choose your metric right enough (or wrong enough) you can construct metric spaces where every sequence converges to any point in the space, or to no point in the space, and quite a lot in between. Thus our choice of metric now will have to reflect what measures we want to call similar to each other. We're explicitly back in normative judgements here: beholding a series of Dirac functions whose characteristic points' values are a sequence like $\frac{1}{2}, \frac{1}{4}, \cdots, \frac{1}{2^n}$ which converge to 0 in the usual metric sense in the reals, we would like to declare such a sequence to converge to δ_0. We achieve this through the use of the KR-metric: the naive metric where we just measure both functions directly with respect to the other and see how much of their assigned "probability mass" they assign away from each other is much too exacting a measure, allowing even deeply pathological functions to wander into our argument to nitpick at any difference between where two measures place their mass within a space. We want to think of our space X as a *space*, though, equipped with all the usual comforts of euclidean distance, and only having to match our chosen measures against functions that can only slope gently down or up lets us retain that intuitive notion of distance. Better yet, choosing this measure makes our dual space $(\mathcal{M}^{\pm} \oplus \mathbb{R})^*$ well-behaved - it's just $C(X, [0, 1]) \oplus \mathbb{R}$ - so that having picked the right notion of distance, we can talk meaningfully about the space of functionals on the resulting space much more easily.

Definition 2.4. *An **affine measure** is a point in $\mathcal{M}^{\pm}(X) \oplus \mathbb{R}$ with canonical form $\vec{a} = (\mu_a, b_a)$, where μ is a (positive) measure and $b \in \mathbb{R}_{\geq 0}$. A **super-affine measure** $\vec{s} = (\mu_s, b_s)$ is an signed measure such that $\mu^-(1) + b \geq 0$, that is, the pessimal measure of any measurable set in X is still nonnegative.*

Of note, we should think of μ_a, μ_s as the expected total future payout from a given point in a possible history onwards, and b_a, b_s as how well we're already doing off-history. Recall that we need to keep track of our off-history rewards to make sure the Demiurge isn't pulling one over on us poor finite creatures!

Affine and super-affine measures will be the building blocks from which we construct inframeasures and infradistributions, along with a few regularity conditions that make our linear combinations of such measures actually behave like we want a more rigorous notion of maximinned possible future outcomes to behave. We'll go over them in a few natural groups. But first: supporting definitions!

Definition 2.5. *Let $D \subseteq \mathcal{M}^{\pm}(X) \oplus \mathbb{R}$ be a set of super-affine measures. The* ***expectation value with respect to*** *D of a function $f \in C(X, [0,1])$ is defined to be $\mathbb{E}_D := \inf_{(\mu,b)\in D}(\mu(f) + b)$, that is, it's equal to the greatest lower-bound for all super-affine measures in D of $\int_X f(x)d\mu$, plus our constant b.*

Recall that we can interpret a signed measure μ as a linear function taken from $C(C(X, \mathbb{R}), \mathbb{R})$.

Definition 2.6. *Let $D \subseteq \mathcal{M}^{\pm}(X) \oplus \mathbb{R}$ be a set of super-affine measures. The* ***upper completion*** *of D is defined to be $D^{\uparrow} := D + \mathcal{M}^{sa}(X) = \{(\mu^*, b^*) | (\mu_D, b_D) \in D, (\mu, b)\mathcal{M}^{sa}(X) : \mu^* = \mu_D + \mu; b^* = b_D + b\}$. More simply, $D^{\uparrow}$ is the result of adding, to each super-affine measure in D, every possible super-affine measure on X; we can imagine a copy of $\mathcal{M}^{sa}(X)$ sitting on top of every element of D within $\mathcal{M}^{sa}(X)$ (not a typo, these should be the same object) and take the resulting union inside of $\mathcal{M}^{sa}(X)$.*

Definition 2.7. *Let $D, D^{\uparrow}$ be as in the previous definition. The the* ***basal*** *or* ***minimal set*** *of D is given by $D^{\min} := \bigcap_{C\subseteq\mathcal{M}^{sa}(X):C^{\uparrow}=D^{\uparrow}} C$. More simply, to get the basal set of D, intersect together every set of super-affine measures sharing an upper completion with D.*

Trivially, every element of $D^{\min}$ is an element of D; the archetypal element of $D^{\min}$ looks like an element of D that we can't get just by adding an arbitrary super-affine measure to a different element of D.

Proposition 2.8. $(D^{\uparrow})^{\min} = D^{\min}$. $(D^{\min})^{\uparrow} = D^{\uparrow}$.

Definition 2.9. *Let $D \subseteq \mathcal{M}^{\pm}(X) \oplus \mathbb{R}$ be a set of super-affine measures. The following are all regularity conditions we might require of such a set to make it maximal with respect to set inclusion or impose further structure on it.*

Trivial extension conditions:

- *Nonemptiness: $D \neq \varnothing$*
- *Topological Closure: $D = \overline{D}$*
- *Convex-hull Closure: $D = c.h(D)$*
- *Upper Closure: $D = D^{\uparrow}$*

These are all conditions ensuring set maximality while avoiding changing expectation value.

Minimal point and boundedness conditions:

- *Positivity of Minimals: $D^{min} \subseteq \mathcal{M}^{a}(X)$*
- *Strong Boundedness of Minimals: For some compact C, $D^{min} \subseteq C$. Equivalently, assuming positivity of minimals, there is some $\hat{\lambda}$ such that for all $(\lambda \cdot \tilde{\mu}, b) = (\mu_D, b) \in D^{min}$, we have $\lambda \leq \hat{\lambda}$.*
- *Weak Boundedness of Minimals: The map taking $f \mapsto \mathbb{E}_D(f)$ is uniformly continuous.*

These are all conditions controlling the nature of the minimal set, and imposing continuity conditions on the functionals they LF-dualize to.

Range conditions:

- *Unitality:* $\mathbb{E}_D(\mathbb{0}) = 0; \mathbb{E}_D(\mathbb{1}) = 1$
- *Normalizability:* $\mathbb{E}_D(\mathbb{0}) \neq \mathbb{E}_D(\mathbb{1})$

These are both conditions forcing our set to act more like a probability distribution.

The boundedness of minimals conditions here merit a little more care. We should recall that Lipschitz continuity (for appropriate Lipschitz constant) is a strictly stronger condition than uniform continuity, and as we'll soon see, the strong boundedness condition on minimals is equivalent to 1-Lipschitz continuity under Fenchel duality. This thus justifies the nomenclature.

Definition 2.10. *We will require that all our named properties on sets of inframeasures satisfy all four of the trivial extension conditions. We call such a set an* ***inframeasure*** *if it also has minimals that are positive and weakly bounded. An inframeasure that is also unital, we will call an* ***infradistribution*** *(in the sense of a probability distribution). If an inframeasure has strongly bounded minimals, we call it a bounded inframeasure; similarly, if an infradistribution has strongly bounded minimals, we call it a bounded infradistribution.*

We note here that normalizability is an extremely weak condition that still permits us to unitalize an inframeasure into an infradistribution: we take $(\mu_D, b_D) \mapsto \frac{1}{\mathbb{E}_D(\mathbb{1}) - \mathbb{E}_D(\mathbb{0})} \cdot (\mu_D, b_D - \mathbb{E}_D(\mathbb{0}))$. It is in fact provable that a set S of super-affine measures fails to be normalizable exactly when every element in S^{min} is a trivial measure of the form $(0, b)$, assigning a measure-value of 0 to every measurable set.
More worryingly, weak boundedness of minimals looks like a very strange condition if what you care about is actual boundedness of minimals. Fortunately, this is where the dual picture comes in: it is already established mathematical fact that the continuous linear functions of $C(X, \mathbb{R})$ are each equivalent to a signed measure from $\mathcal{M}^{\pm}$. Using Fenchel duality/convex conjugation, we can turn our statements about sets of inframeasures with possibly-bounded minimals into statements about the kinds of functions that arise when we go from functions to expectation values over infradistributions of those functions. In particular:

Theorem 2.11. *Let H be an infradistribution/a bounded infradistribution. Then the function $H_* : f \mapsto \mathbb{E}_H(f)$ taking functions to their expectation values over H is concave, monotone, uniformly continuous (or even Lipschitz-continous, if we started off bounded!) over $C(X, [0, 1])$, $H_*(\mathbb{0}) = 0$ and $H_*(\mathbb{1}) = 1$, and if $\text{range}(f) \not\subseteq [0, 1]$, then $H_*(f) = -\infty$. Every such $C(X, [0, 1])$ that fulfills the other properties is one such expectation function.*

Let h be a concave, monotone, uniformly continuous/Lipschitz continuous function taken from $C(X, [0, 1])$, with $h(0) = 0$ and $h(1) = 1$, with $h(f) = -\infty$ whenever $\text{range}(f) \not\subseteq [0, 1]$. Then writing $h'(f) = -h(-f)$, and taking h^* to mean the convex conjugate, the set $h^* := \{(\mu, b) | b \geq (h')^*(\mu)\}$ is an infradistribution (or even a bounded infradistribution, if we started off Lipschitz-continuous!) over X. Every such possibly-bounded infradistribution can be expressed in this way.

$(H_*)^* = H$. $(h^*)_* = h$.

Given the importance of the condition on the functions used in the prior theorem:

Definition 2.12. *We say that a function $f \in C(C(X,\mathbb{R}),\mathbb{R})$ is **IDable** (infradistribution-able) if it is concave, monotone, uniformly continuous, takes any $g \in C(X,\mathbb{R})$ to $-\infty$ if $\text{range}(g) \not\subseteq [0,1]$, and has $f(\mathbb{0}) = 0; f(\mathbb{1}) = 1$. If f is also Lipschitz-continuous, we say that f is **BIDable** (bounded-infradistribution-able).*

Definition 2.13. *Let $f : X \to Y$ be a map of measurable topological spaces, with maps $\mu_X : X \to \mathcal{M}^{sa}(X), \mu_Y : Y \to \mathcal{M}^{sa}(Y)$. The **pushforwards map** $f_* : \mathcal{M}^{sa}(X) \to \mathcal{M}^{sa}(Y)$ is the unique map on the appropriate measure spaces which commutes with f and the measure maps, that is, for any $y = f(x) \in f(X)$, we get $f_*(\mu_X(x)) = \mu_Y(f(x)) = \mu_Y(Y)$.*

Proposition 2.14. *Let $f \in C(X,[0,1])$, and $g : X \to Y$ be a continuous function of topological spaces. Then for infradistribution $H \subseteq \mathcal{M}^{\pm}(X) \oplus \mathbb{R}$, $\mathbb{E}_{g_*(H)^\uparrow}(f) = \mathbb{E}_H(f \circ g)$. Additionally, $g_*(H)^\uparrow$ is an inframeasure if H is, and is bounded if H is, and if g is surjective, then $g_*(H)^\uparrow = g_*(H)$.*

Definition 2.15. *Let $S \subsetneq \mathbb{N}$, and let $\vec{\sigma}$ be a probability distribution over S, so that $\sum_{i\in S} \sigma_i = 1$. If $\vec{h} = \{h_i\}_{i\in S}$ is a set of IDable functions, the **σ-weighted mixture of** h is IDable, and is given by $(\mathbb{E}_\sigma h_i)(f) := \mathbb{E}_\sigma(h_i(f))$. If the h_i are in fact BIDable with Lipschitz constants κ_i, we must also require that the σ-weighted mix of the Lipschitz constants is still finite, that is, $\sum_i \sigma_i\kappa_i < \infty$, and in that case, the resulting mixed function is also BIDable.*

Theorem 2.16. *Let $S \subsetneq \mathbb{N}$, and let $\vec{\sigma}$ be a probability distribution over S, so that $\sum_{i\in S} \sigma_i = 1$. If $\vec{H} = \{H_i\}_{i\in S}$ is a countable family of infradistributions, the **σ-weighted mixture of** $\vec{H}$ is an infradistribution, and is given by $\mathbb{E}_\sigma H_i = \{M \in \mathcal{M}^{sa} | \vec{G} \in \Pi_{i\in S} : \sum_i \sigma_i G_i = M\}$, that is, the set of all σ-parametrized linear combinations of countable choices of infradistribution, one from each family. If $\vec{H}$ is in fact comprised of bounded infradistributions with minimal-bound constants κ_i, we must also require that the σ-weighted mix of the minimal-bound constants is still finite, that is, $\sum_i \sigma_i\kappa_i < \infty$, and in that case, the resulting mixed infradistribution is also bounded.*

Proposition 2.17. *Using the same variable assignments as above, $\mathbb{E}_{\mathbb{E}_\sigma H_i}(f) = \mathbb{E}_\sigma(\mathbb{E}_{H_i}(f))$. Additionally, $g_*(\mathbb{E}_\sigma(H_i)) = \mathbb{E}_\sigma(g_*(H_i))$.*

A few procedural notes on updating before we can get to Infrabayes's Rule. First, in thinking about updating, we'll be explicitly separating the initial part where we cut down measure - restricting only to the histories we want to update over - from the final part, where we normalize our resulting measure back up to a total of 1. An additional hurdle is that unlike how in classical probability, where we can very easily turn an expectation value into a probability by thinking in terms of expectation values over (possibly-discontinuous) indicator functions, in the inframeasure setting we can only take expectation values over particularly nice continuous functions.
Thankfully, "continuous" for us respects our topology on histories - all we really need is for the indicator set to be clopen. And because of how we defined histories - namely, to have discrete action and observation spaces - the property of "beginning with a specific finite prefix" is a clopen condition for histories, and thus the set of histories beginning with a specific finite prefix is always a clopen set. Given a discrete (or finite) set of possible observations, a similar statement is true of the clopenness of conditions like "the true observation is within this subset of observations". The major additional work we need to put in to extend from probability measures to inframeasures is that for inframeasures, we need a more powerful and subtle notion than usual of expectation values.

Consider that the very first part of what we need to do for both the classical and inframeasure cases is to specify the expectation value of arbitrary functions over the whole sample space X, but unlike the classical case where our functions are total, here we might have a function f defined only on $Z \subseteq X$. In both cases, we only have expectation values for total functions to work from, but that's not a problem for the classical case. Here, though, we need to decide what it is f should be doing outside of Z. The very simplest way we could do this is simply to declare that it takes the value 0 on $X \backslash Z$, but all that does for us is recover a trivial extension of classical probability. Instead, let's try the extension $\tilde{f}_Z^g = \mathbb{1}_Z f + \mathbb{1}_{X \backslash Z} g$ for some suitable choice of g, which we can think of as "the function that does f when it can and g otherwise". As it turns out, from here we can just follow our noses: writing $\neg Z := X \backslash Z$ and starting with the (B)IDable form h of an infradistribution H, and recalling that we can pass freely between linear functionals and infradistributions, we get a (probably non-unital) infradistribution from $\hat{h}_Z^g(f) := h(\mathbb{1}_Z \cdot f + \mathbb{1}_{\neg Z} \cdot g)$. We then have to unitalize that using the technique described earlier, which gives us $h_Z^g(f) := \dfrac{\hat{h}_Z^g(f) - \hat{h}_Z^g(\mathbb{0})}{\hat{h}_Z^g(\mathbb{1}) - \hat{h}_Z^g(\mathbb{0})} = \dfrac{h(\mathbb{1}_Z \cdot f + \mathbb{1}_{\neg Z} \cdot g) - h(\mathbb{1}_{\neg Z} \cdot g)}{h(\mathbb{1}_Z + \mathbb{1}_{\neg Z} \cdot g) - h(\mathbb{1}_{\neg Z} \cdot g)}$. Running the same sanity check as in the original post to see that we recover ordinary bayesian formulas, we get $\mu_Z^g(f) = \dfrac{\mu(\mathbb{1}_Z \cdot f + \mathbb{1}_{\neg Z} \cdot g) - \mu(\mathbb{1}_{\neg Z} \cdot g)}{\mu(\mathbb{1}_Z + \mathbb{1}_{\neg Z} \cdot g) - \mu(\mathbb{1}_{\neg Z} \cdot g)} = \dfrac{\mu(\mathbb{1}_Z \cdot f)}{\mu(\mathbb{1}_Z)} = \mu|_Z(f)$. Great!

This does leave us with the lingering question of why we needed to introduce that function g in the first place, if all we were going to do was recover classical updating. We're not - we're very explicitly trying to dodge the problems with dynamic inconsistency that trips up classical bayesian updating. This is another place where our off-history performance comes in clutch for us - a nonzero off-history lets us make sure that even when the expected payouts ahead change as we near them, what we care about stays the same. This brings up another important motivational point: because of our Demiurge-framing, our choice of g heavily affects how different input functions f score on our infradistribution H, so unlike the classical case, where we can cleanly separate out expectation from $Z, \neg Z$, the same can't be said here.

We actually don't need to restrict ourselves to clopen sets for our choices of indicator function. We can use fuzzy sets, and treat them as likelihood functions over X - each set element gets sent to the probability that we get the observation we really did. This lets us make our earlier more primitive notion of $\tilde{f}_Z^g = \mathbb{1}_Z \cdot f + \mathbb{1}_{\neg Z} \cdot g$ more general as follows:

Definition 2.18. *Let $f, g, L \in C(X, [0,1])$. Then we define a function called f **glued to** g **hinging on** L, given by $f \oplus_L g := Lf + (1-L)g$. Additionally, we define $f \oplus_{\neg L} g := (1-L)f + Lg = g \oplus_L f$.*

As you might expect, we use this better notion of function extension everywhere we used the simpler non-likelihood-using function extension from before. This gives us the following definitions, analogous to the raw and unitalized updates of infradistributions from before:

Definition 2.19. *Raw:* $\hat{h}_L^g(f) := h(f \oplus_L g)$,

Unitalized: $h_L^g(f) := \dfrac{\hat{h}_L^g(f) - \hat{h}_L^g(\mathbb{0})}{\hat{h}_L^g(\mathbb{1}) - \hat{h}_L^g(\mathbb{0})} = \dfrac{h(f \oplus_L g) - h(\mathbb{0} \oplus_L g)}{h(\mathbb{1} \oplus_L g) - h(\mathbb{0} \oplus_L g)} = \dfrac{h(Lf + (1-L)g) - h(g - Lg)}{h(g - Lg + L) - h(g - Lg)}$,

Probability/caring from expectation: $\mathbb{P}_D^g(L) := \mathbb{E}_D(\mathbb{1} \oplus_L g) - \mathbb{E}_D(\mathbb{0} \oplus_L g)$.

Proposition 2.20. *Taking L as a fuzzy set on the left and the likelihood function which is its indicator on the right, $\mathbb{P}_D^{\mathbb{0}}(L) = \mathbb{E}_D(L)$.*

As the original post notes, if we've bothered to keep track of off-history payout $g \neq 0$, then $\mathbb{P}_D^g(L)$ is best interpreted as "what's still at stake", given that it's the difference between the best and worst scores our pessimistic agent thinks it can get, given g, which tells you how things could be going outside L; additionally, we can always rewrite our unitalization denominator from much earlier, $\mathbb{E}_D(\mathbb{1}) - \mathbb{E}_D(\mathbb{0})$, as $\mathbb{P}_D^g(\mathbb{1}_X)$. This gives us the natural interpretation that we might have guessed at earlier - that that denominator represents the probability that we assign to anything at all occurring.

Definition 2.21. *Let H be an inframeasure on X, with $g, L \in C(X, [0,1])$. Then we define the **update of** H **by** L **and** g as $H_L^g := \overline{\{\frac{1}{\mathbb{P}_H^g(L)}(L \cdot \mu, b + \mu(\mathbb{0} \oplus_L g) - \mathbb{E}_H((\mathbb{0} \oplus_L g))|(\mu, b) \in H\}}^{\uparrow}$.*
If H is a bounded inframeasure, we need not take the closure. If L is the indicator function of some clopen set in X, we need not take the upper completion.

Unpacking the expression further: the closure of the upper completion of, for all affine measures $(\mu, b) \in H$, the set of all affine measures of unital expectation value where the pure measure component has also been scaled by the fuzzy-containment likelihood function L, and the affine constant component has been adjusted by the difference between the measure μ assigns to the L-adjusted off-history g and the off-history expectation over H of g.

Proposition 2.22. *Using variable assignments from the definition above, if $H \in \mathcal{M}^{sa}(X)$ is bounded, then the update (if we can in fact update) will be a bounded infradistribution in $\mathcal{M}^{sa}(L)$. If H is not bounded, we may need to take the closure of the image.*
Additionally, $\mathbb{E}_H(f \oplus_L g) = \mathbb{E}(\mathbb{0} \oplus_L g) + \mathbb{P}_H^g(L)\mathbb{E}_{H^g L}(f)$, that is, we can break down an expectation value of a composite function into its two components: the expectation of f weighted by the probability of L relative to g, and the expectation of g outside L.

Finally, this brings us to the statement of Infrabayes's Rule:

Theorem 2.23. $\mathbb{E}_\sigma(H_i)^g L = \dfrac{\mathbb{E}_\sigma(\mathbb{P}_{H_i}^g(L) \cdot (H_i^g L))}{\mathbb{E}_\sigma(\mathbb{P}_{H_i}^g(L))}$ *whenever even one of the $\mathbb{P}_{H_i}^g(L)$ are nonzero.*

Update individually, mix together as prescribed by σ and L, and divide out by the probability of L. Great!

3 Historyspace, Policyspace (Only In The Past, Is What They Say)

Let's start off by figuring out what we mean by a history.

Definition 3.1. *A **history** is an element of $\mathbb{H} := (A \times O)^{\leq \omega}$, a possibly-infinite alternating sequence of actions and observations.*
*An **a-history** is a history that ends with an action from A. An **o-history** is any other history - one that ends in an observation, is empty, or is infinite and thus has no end. We'll represent histories as $\vec{h} = a_1 o_1 a_2 o_2 \cdots a_n o_n \cdots$ and also frequently as possibly-infinite trees.*
We denote the set of arbitrarily long finite histories by $\underset{\sim}{\mathbb{H}} := (A \times O)^{< \omega}$.
We denote the set of (possibly-infinite) histories of at least length n by $\mathbb{H}^{\geq n}$.
We denote the set of specifically infinite histories by $\mathbb{H}^\omega := (A \times O)^\omega$, and caution that all such

histories are o-histories.
There is a natural partial-order relation $\prec$ on $\mathbb{H}$ (and all of its subsets under consideration) given by $\vec{g} \prec \vec{h}$ if $\vec{g}$ is a prefix of $\vec{h}$; in this case we may also write $\vec{h} \succ \vec{g}$.
Given some history $\vec{h}$, we denote the initial segment of n observations and n actions by $\vec{h}_{\leq n}$.
*The **type function** $\tau : \mathbb{H} \to A, O$ is a technical function we'll occasionally use for labelling. It takes in a history and returns A if it's an a-history, and O if it's an o-history.*
*We denote the **$\vec{h}$-restriction of a measure** μ by $\mu|_{\vec{h}} := \mu \cdot \mathbb{1}_{\vec{h}}$.*

If you like, you can think of the a_i as our moves and the o_i as the uncovering of the response the Demiurge already picked, or more properly our observation of that response. On that theme, there's something we really should define...

Definition 3.2. ***Flipping off the Demiurge*** *is a special and purely internal action that won't take up the action slot of a history $\vec{h}$. As such, we represent it as an observation F that can sometimes occur as the end node of a history, defined by its status as an end node - no actions can be taken past such an observation leaf - and the fact that depending on formulation, the reward payout for getting to flip off the Demiurge is immediately $+\infty$ or pinned to $+1$ forever after.*
We denote the set of histories ending in flipping off the Demiurge by $\mathbb{H}^F$, and caution that all such histories are finite o-histories.

We should make a few observations on flipping off the Demiurge. First, this is identical to the somewhat unhelpfully-named "Nirvana trick" one might play on a "Murphy". Second, we should strongly expect for a system using an infrabayesian decision process as described here never to actually flip off the Demiurge; as we'll see in more detail shortly, the use of the formalism lies mainly in marking possible histories or outcomes that are actually impossible for one reason or another. Finally, and more philosophically, even calling it "flipping off the Demiurge" is *still* a slightly-unhelpful simplification.
The image we should hold in our minds here, to understand why flipping off the Demiurge is an *observation* and not an action, is to employ the related metaphor of realizing that the Demiurge is naked - we find ourselves in an impossible possible history; time to point and laugh at the Demiurge for screwing it up! Of course, that's not something to which there's any point unless it's *true*. Ultimately, it's not even that we're taking some action which gives us truly infinite reward - more that we've decided (mostly arbitrarily) to shape our desires so as to assign infinite reward to reaching and observing such an impossible state; that way, all of the maximinning math works out correctly rather than give us unhelpful garbage that interferes with our ability to maximin properly in situations that have an actual chance of happening. As such:

Definition 3.3. *We call an affine measure or super-affine measure **non-flipoff** when it assigns the set of histories ending in flipping off the Demiurge measure 0. We denote the set of non-flipoff affine measures by $\mathcal{M}^{\mathrm{NF}}(X) \oplus \mathbb{R}_{\geq 0}$, and will take the use of* NF *as a superscript to mean that as a regularity condition, we're trimming away any part of the object where we flip off the Demiurge. Because we never get to actually do that.*

Now for a few definitions regarding what the original sequence would refer to as simply policies:

Definition 3.4. *A **plan** is a partial function $p_{par} : \mathbb{H} \to A$ which is defined only on continuable o-histories, taking such o-histories with at least one possible action branching off of them to one such action; additionally, if we have a history $\vec{h}$ in the codomain of p_{par}, and $\vec{h} = \vec{h'}a_n$, then we*

require $p_{par}(\vec{h'}) = a$ - that is, the plan must produce its own history if able and is not allowed to be self-defeating. We mark such plans $\tilde{p}$.
*We call a plan a **stub** and mark it with subscript $p_{\nexists,n}$ if its shelf-life is finite - that is, there exists some $n \in \mathbb{N}$ such that $\mathrm{codom}(p_{\nexists}) \cap \mathbb{H}^{\geq n}$ is empty.*
*We call a plan a **policy** and mark it $\hat{p}$ if it is a total function, that is, $\hat{p}(\vec{h})$ is defined for all $\vec{h} \in \widetilde{\mathbb{H}}$. The **empty policy** $p_{\varnothing}$ is the unique special policy with empty domain and codomain, defined on no histories and prescribing no actions.*

Recall that a total function is defined on its whole domain, while a partial function might be undefined for some choices of input. We recommend reading $p_{\nexists,n}$ as "p sub block at n", where "block" might be swapped out for words like "stop", "wall", "obstruct", or simply "stub".

Definition 3.5. *We denote the set of plans as Π and the set of strict-plans as $\widetilde{\Pi}$.*
We denote the set of policies as $\widehat{\Pi}$.
We denote the set of stubs as $\Pi_{\nexists}$.
We denote the set of strict-plans starting at $\vec{h}$ as $\widetilde{\Pi}^{\vec{h}}$, and the set of plans (including policies) starting at $\vec{h}$ as simply $\Pi^{\vec{h}}$.

Importantly: Π is a Greek capital-'p', for plan (and also policy).

Definition 3.6. *An **off-history** plan (with respect to $\vec{h}$) $p_{\neg\vec{h}}$ is a plan defined precisely on all histories $\vec{g}$ that are both finite and contain no $\vec{h}$-prefix.*
We note that we can write the set of histories beginning with a given $\vec{h} \in \mathbb{H}$ as $\vec{h}\mathbb{H}$, given concatenation. By abuse of notation we'll denote the set of histories that do not begin with a given $\vec{h}$, $\neg\vec{h}\mathbb{H}$.

We note that such a $p_{\neg\vec{h}}$ must specify everything that happens apart from the subtree after $\vec{h}$, and can possibly produce $\vec{h}$.

Definition 3.7. *Let $p_{\neg\vec{h}}$ be an off-history plan with respect to $\vec{h}$. Then for $q \in \Pi^{\vec{h}}$, the **composite plan** $r = p_{\neg\vec{h}} \oplus_{\vec{h}} q$ is the result of **gluing** $p_{\neg\vec{h}}$ to q **hinging on** $\vec{h}$.*
Explicitly, $r : \widetilde{\mathbb{H}} \to A, \vec{x} \mapsto \vec{x} \prec \vec{h}\,?\ q(x)\ :\ p_{\neg\vec{h}}(x)$.

In plain language: in case of $\vec{h}$, the composite plan applies $\tilde{q}$. Otherwise, it sticks to $\tilde{p}$.

Definition 3.8. *Let p, q be plans that never disagree, that is, for no $\vec{h} \in \widetilde{\mathbb{H}}$ does $p(\vec{h}) \neq q(\vec{h})$ when both are defined. We say that q **lies above** p, with $q \succ p$, if $\mathrm{codom}(p) \subseteq \mathrm{codom}(q)$, and that q **lies below** p, with $q \prec p$, if $\mathrm{codom}(p) \supseteq \mathrm{codom}(q)$.*

We note that the empty policy is special in that it is the unique plan lying below all plans. More generally, lying above/below defines a partial order $\prec$ on Π with $p_{\varnothing}$ as the unique minimal element and the maximal elements given by the policies of $\widehat{\Pi}$.

Definition 3.9. *Let $\vec{g}, \vec{h} \in \mathbb{H}$, and denote by $t(\vec{g}, \vec{h})$ the first time step at which $\vec{g}, \vec{h}$ differ. Then for choice of time-discounting parameter $0 < \gamma < 1$, we define the metric distance between the histories as $d_\gamma(\vec{g}, \vec{h}) := \gamma^{t(\vec{g},\vec{h})}$.*

As a reminder, γ here is the parameter that tells us how much we discount possible future rewards or penalties; either a problem setup will specify it, or we'll use it as a way to get a continuous family of distance measures d_γ that we can then limit as $\gamma \to 1$.

Definition 3.10. *Let $\tilde{p} \in \Pi$. The **outcome set** $F(\tilde{p})$ is defined as the set of all o-histories $\vec{h}$ such that $\vec{h} \notin \text{codom}(\tilde{p})$ for which $\tilde{p}$ postdicts perfectly, that is, for any prefix $\vec{g}a_n$ of $\vec{h}$, $\tilde{p}(\vec{g}) = a_n$. F^{NF} is the outcome subset where no flipping off of the Demiurge occurs.*

As clarification, $F(\tilde{p})$ consists of the entirety of all the subtrees that following $\tilde{p}$ could put you in, not just their root nodes. As an additional philosophical note, we should expect never to actually flip off the Demiurge: recall that the Demiurge is long-sighted and will never allow us into a possible history where flipping It off is even a meaningful possibility.

Definition 3.11. *Let $p \neq q \in \Pi$, and denote by $t(p, q)$ the first time step across all histories at which p, q differ in their choice of action, including cases where (WLOG) p is defined and q is not. Then for choice of time-discounting parameter γ, we define $d(p, q) := \gamma^{t(p,q)}$.*

Definition 3.12. *Let $q \succ p \in \Pi$. Then we may define the projection-induced function $pr_*^{q,p} : \mathcal{M}^{sa}(F(q)) \twoheadrightarrow \mathcal{M}^{sa}(F(p)); (m_q, b_q) :\mapsto (m_p, b_q)$ where in particular we recall that $m_p(h) = m_q(h)$ whenever h is a prefix of some outcome in $F(p)$. Additionally, this maps affine measures to affine measures and non-flipoff measures to non-flipoff measures.*

We will use this function for two major purposes: we'll use preimages of special versions of this function to put sets we care about in the same ambient space, so that we can compare them, and we'll also use more normal versions of the function itself to help define internal consistency conditions on belief functions.

4 Affine Environments (Which One Do You Think You're Living In?)

Definition 4.1. *An **environment** is a function $\check{e} : \Pi \times \widetilde{\mathbb{H}} \times A \to \Delta O$ taking an ordered triple of plan, finite history prefix, and action to a probability distribution over observations. It must be either total, or defined only over $\{(p_\varnothing, h, a) | h \in \widetilde{\mathbb{H}}, a \in A\}$. If the environment is only defined on $(p_\varnothing, h, a)$, then we call the environment **policy-independent**, and if it is total, then we call it **policy-dependent**.*

Definition 4.2. *An **affine environment** is an ordered pair $\hat{e} = (\lambda \cdot \check{e}, b)$, where $\lambda, b \geq 0$ and $\check{e}$ is an environment.*

We should note that this construction is precisely analogous to the construction of affine measures from ordinary measures. Similarly, b should be interpreted as a reflection of how well we're doing off-history, and λ as either the probability of the environment's existence (for $\lambda \leq 1$) or how much we care about $\check{e}$ more generally.

Proposition 4.3. *Let $\hat{e} = (\lambda \cdot \check{e}, b)$ be an affine environment. Then for all plans $\tilde{p} \in \Pi$, $\hat{e} \cdot \tilde{p} = (\lambda \cdot \check{e}(\tilde{p}, \widetilde{\mathbb{H}}, A), b)$ is an affine measure; we will interpret $\check{e}(\tilde{p}, \widetilde{\mathbb{H}}, A)$ as the probability distribution over $F(\tilde{p})$ (or, as appropriate, $F^{\text{NF}}(\tilde{p})$).*

The set of environments is equivalent to the set of functions $f : \Pi \to \Delta O$. In one direction, we have $\tilde{p} \mapsto \tilde{p} \cdot \check{e}$, and in the other direction, any function $f : \Pi \to \Delta O$ which has $pr_*^{q,p}(f(q)) = f(p)$ whenever $q \succ p$ corresponds to an environment, as we'll shortly see. For now, we'll mostly black-box the tools for explicitly working with sets of affine environments in terms of beliefs.

Definition 4.4. *We call a total function* $\beta : \Pi \to \mathcal{P}(\mathcal{M}^a(F(\Pi))), \beta(\tilde{p}) \subseteq \mathcal{M}^a(F(\Pi))$ *a* ***belief function*** *if the image of every plan is nonempty. We may, at pleasure, require non-flipoffness of all relevant affine measures in this definition. Additionally, we get for free that* $\mathbb{E}_{\beta(\tilde{p})}(f) = \mathbb{E}_{\beta^{\text{NF}}(\tilde{p})}(f)$*, because the Demiurge will only ever pick a non-flipoff affine measure for an environment.*

We require affine rather than super-affine measures here for two major reasons: we don't actually need to use super-affine measures for anything except the definition of the upper completion, and (more importantly) any negative measure on a history that has nontrivial flipoff in it interacts extremely badly with a maximinned approach. As for some further motivation for why sets of affine environments match up to the belief functions they induce, we may note that in one direction, given a set of affine environments $\hat{e}_i$, we can take a fixed plan $\tilde{p}$ and look at all the resulting affine measures over $F(\tilde{p})$ that arise from the assorted $\hat{e}_i \cdot \tilde{p}$. Now, once again we must first define a few terms before we can talk about regularity conditions on beliefs.

Definition 4.5. *The set* $\mathcal{M}^a(\widetilde{\mathbb{H}}^\omega)$ *is the set of affine measures over all infinite histories.*

Definition 4.6. *Let* β *be a belief function satisfying boundedness of minimals as below. We denote the set of all affine measures over histories with the same set of upper bounds by* $\mathcal{M}^a_{\leq(\lambda,b)}$*. We call such a set a* ***clip set****.*

Definition 4.7. *Let* $\tilde{p}$ *be a plan. We denote by* $(pr_*^{\omega,\tilde{p}})^{-1}$ *the function that takes subsets of affine measures* $D \subseteq \mathcal{M}^a(F^{\text{NF}}(\tilde{p}))$ *to the set of affine measures* $\{\epsilon = (\lambda_e, b_e) \in \mathcal{M}^a(\widetilde{\mathbb{H}}^\omega) | \forall \vec{h} \in \mathbb{H} \neg \exists \vec{g} \in F^{\text{NF}}(\tilde{p}) : \vec{g} \prec \vec{h}, (\epsilon(\vec{h}) = 0) \wedge (pr_*^{\omega,\tilde{p}})(\epsilon) \in D)\}$*. More plainly, this preimage set is best characterized as the set of affine measures assigning measure* 0 *to any history lying above none of the no-flipoff outcome histories in* $F^{\text{NF}}(\tilde{p})$*.*

We use $(pr_*^{\omega,\tilde{p}})^{-1}$ *to take preimages of no-flipoff outcome sets of different plans to compare them in the larger ambient space of* $\mathcal{M}^a(\widetilde{\mathbb{H}}^\omega)$*. We call such a preimage the* ***preimage from infinity*** *of that plan.*

Definition 4.8. *Let $\beta : \Pi \to \mathcal{P}(\mathcal{M}^a(F(\Pi)))$ be a belief function, taking arbitrary plans $\tilde{p}$ to nonempty sets of affine measures over outcomes. The following are all conditions we might require of such a function to impose further structure on it.*

Trivial extension conditions:

- *No-Flipoff Nonemptiness: $\beta^{\text{NF}}(\tilde{p}) \neq \varnothing$.*
- *Topological Closure: $\beta(\tilde{p}) = \overline{\beta(\tilde{p})}$.*
- *Convex-hull Closure: $\beta(\tilde{p}) = c.h(\beta(\tilde{p}))$.*
- *No-Flipoff Upper Closure: $\beta^{\text{NF}}(\tilde{p}) = (\beta^{\text{NF}}(\tilde{p}) + \mathcal{M}^{sa}(F^{\text{NF}}(\tilde{p}))) \cap \mathcal{M}^a(F(\tilde{p}))$.*

These conditions correspond to the first four regularity conditions about inframeasures. They ensure image maximality while avoiding changing affine measure value.

Local regularity conditions:

- *Boundedness of Minimals: There is some $\hat{\lambda}, \hat{b} \in \mathbb{R}$, so that for all $(\lambda \cdot \tilde{\mu}, b) \in \beta(\tilde{p})^{\min}$, we have $\lambda \leq \hat{\lambda}$ and $b \leq \hat{b}$. This must be fixed once and for all over Π.*
- *Unitality: $\min_{\tilde{p}} \mathbb{E}_{\beta(\tilde{p})}(\mathbb{0}) = 0$. $\min_{\tilde{p}} \mathbb{E}_{\beta(\tilde{p}}(\mathbb{1})) = 1$. The resulting belief function must then in general be unitalized.*

These conditions correspond to the regularity conditions about inframeasures of the same(ish) names. They deal only with the properties of a single choice of $\beta(\tilde{p})$.

Special conditions:

- *Lower Internal Consistency: $\beta(\tilde{q}) = \bigcap_{p_\exists \prec \tilde{q}} (pr_*^{\tilde{q}, p_\exists})^{-1}(\beta(p_\exists))$. More plainly: what our IB-agent believes about the outcomes of its plans can be fully recovered if we know what it believes about the outcomes of its stubs.*
- *Upper Internal Consistency: $\beta(\tilde{q}) = \overline{c.h}(\bigcup_{\hat{r} \succ \tilde{q}} pr_*^{\hat{r}, \tilde{q}}(\beta(\hat{r})))$. More plainly: what our IB-agent believes about the outcomes of its plans can be fully recovered if we know what it believes about the outcomes of its policies.*
- *Internal Consistency: β is both upper and lower internal consistent. The original Condition 7 on belief functions. I've broken it up here to make the two conditions easier to parse and talk about separately. [Maybe discuss motivations for both further here?]*
- *No-Flipoff Full-Stack Extremal Consistency: For any no-flipoff set of affine measures $S \in \beta(p_\exists)^{\text{xmin}}$, there is some $\hat{r} \succ p_\exists$ and associated no-flipoff set of affine measures $T \in \beta(\hat{r})$ allowing the projection function to satisfy $pr_*^{\hat{r}, p_\exists}(T) = S$, respecting the partial order.*
- *Hausdorff-Metric Continuity: The map taking $\tilde{p} \in \Pi$ to $(pr_*^{\omega, \tilde{p}})^{-1}(\beta^{\text{NF}}(\tilde{p}) \cap \mathcal{M}^a_{\leq(\lambda,b)})$ is continuous.*

These are special conditions dealing with our ability to recover plans from stubs or policies and to require our belief function to respect the topology on policyspace, including the partial order on policies.

Hausdorff-metric continuity will take a little more motivation. First off, we state the condition more plainly: the map taking a plan to those affine measures in the no-flipoff preimage from infinity of the plan's outcome set which also have their (λ', b') bounded by some initial choice of (λ, b) is a continuous map. In particular, we need to trim away flipoff measures because $\mathcal{M}^a\mathbb{H}^\omega$ has no flipoffs in it - they'd end any history they showed up in at a finite time step - and so we need to trim them away; similarly, the way we defined distance between policies means that if we don't bound minimals and trim away a few nonminimal elements, the upper-completions of slightly different plans diverge arbitrarily far, and our entire conception of a usable hausdorff distance collapses. Taking all of this together, the condition roughly says that if we have a no-flipoff affine measure M in the outcome set of a plan p, and we perturb p slightly to some plan q that agrees with p for a very long time and is thus near to p, then we will have a no-flipoff affine measure N which is close to N and inside the outcome set of q. As a side note, given that $\widetilde{\Pi}$ is a compact space, our condition is actually equivalent to requiring *uniform* continuity.

Alright, that was the what - how about the why? In a reasonable sense, the HMC condition can be interpreted as a regularity condition on the Demiurge! There are two major strong reasons for thinking this way. First, if we really like the Demiurge framing, we can think about what might happen in a setup where the plans our IB-agent follows are deterministic, but with some small probability of error, where the IB-agent suddenly chooses a different plan, with the probability decreasing extremely quickly (exponentially, say) in the metric distance between the old and new plans. In such a setup, we should absolutely want for the Demiurge not to be forced all around Its possibility-space trying to respond to the tiny accidental shifts in the IB agent's plan, but instead only have to jump to a nearby environment It can present to the IB-agent.

The other is through appeal to UDT puzzles. Once again, consider that the HMC condition is, at its core, a statement about how large the changes to the environment right now can be, based on and limited in terms of the differences in plans an IB-agent can have going into them. We can therefore consider the case of an iterated Newcomb's dilemma problem: it makes sense for Omega to demand some finite commitment of our IB-agent - say, that it must predictably one-box on the next thousand rounds, whichever time-step it happens to find itself at, and the HMC condition reflects this. Complementarily, it should make very little sense to us for Omega to demand an unbounded commitment of our poor confused IB-agent, who has not established and cannot establish that this environment is even a real one, or that future time-steps will even exist, never mind to demand precommitment about those far-off rounds; it makes even less sense for Omega to radically alter its behavior now in "response" to things that might "happen" in the far future. Seen from this perspective, the HMC condition is a statement about the strength of acausal influences from the distant future, which we require to be weak.

In any event, we need this condition in order to prove that an optimal policy exists, and also to establish important results about internal consistency with respect to stubs, plans, and policies, and recovering belief functions from just one subset among those.

Definition 4.9. *We call a belief function β a **hypothesis** if it satisfies all of the above regularity conditions: all six in the first two groups, along with internal consistency, NF-extremal consistency, and hausdorff-continuity. In such a case, we may frequently use the variable η for it instead.*

Just as β is a Greek 'b', η is a Greek 'h'. Sort of.

Definition 4.10. *We call a function $\phi : \widetilde{\Pi} \to \mathcal{M}^a(\mathbb{H})$ a forecast function if it both takes each $\tilde{p} \in \Pi$ into $\mathcal{M}^a(F(\tilde{p}))$, the set of affine-measures over the outcome set of $\tilde{p}$, and it respects the relevant pr_* functions, such that for all $\tilde{q} \succ \tilde{p} \in \widetilde{\Pi}$, $pr_*^{\tilde{q},\tilde{p}}(\phi(\tilde{q})) = (\phi(\tilde{p}))$.*

This corresponds to the definition of an outcome function *of* in the original writeup. I've chosen this different name to avoid confusion with outcome sets, and also because I think it's more appropriate - these are all the functions taking partial policies to affine measures over histories with the twin conditions that it has to send every plan to one of the affine measures over its own outcome set and also respect the partial order over plans in the process.

Recalling that (affine) environments correspond to the functions taking plans to (infra)distributions over observations, we present several desirable properties belief functions can have:

Definition 4.11. *We call a hypothesis η **causal** if for all plans $\tilde{p} \in \Pi$ and all affine measures $(\mu, b) \in \eta(\tilde{p})$, there exists a forecast function ϕ such that $\phi(\tilde{p}) = (\mu, b)$, and for all other $\tilde{q} \in \Pi$, we similarly have $\phi(\tilde{q}) \in \eta(\tilde{q})$.*

This is an important property, and one that's also a little hard (philosophically) to understand. The idea here is two-part. First, if we fix an affine environment $\hat{e}$ and then look at what it makes of every possible partial policy $\tilde{p}$, we'd certainly get a forecast function out of that. In the other direction, if we fix a forecast function ϕ, we should now know precisely what happens on all possible histories $\vec{h}$, and we already know that all the projections from long to short histories commute, compose, and respect the rest of our desiderata. This is thus the critical part for passing between affine environments and sets of belief functions and also why we call such hypotheses causal: every point in all the belief sets must in some sense have come (or look just like it came) from actually poking an affine environment, seeing what (causally) happens, and always writing the answer down, even if it seems not to make sense; likewise, querying a set of affine environments with partial policies results in a causal hypothesis.

Definition 4.12. *We call a hypothesis η **pseudocausal** if for all plans $\tilde{p}, \tilde{q} \in \Pi$ where we have $(\mu_p, b_p) \in \eta(\tilde{p})$ and $\text{supp}(\mu_p) \subseteq F^{\text{NF}}(\tilde{q})$, we also get $(\mu_p, b_p) \in \eta(\tilde{q})$.*

We could think of this as simply the no-flipoff version of causality, but there's somewhat more going on here. Briefly, we already know that μ_p is a no-flipoff measure, and that it's even already a no-flipoff measure over $F(\tilde{q})$. This condition can then be interpreted as requiring it to be the case that whenever two plans prescribe the same course of action on histories that have positive measure, they must also induce the same affine measure over outcomes. We call such hypotheses pseudocausal because this is also the property that a hypothesis would have if it came from actually poking an affine environment, seeing what (causally) happens, and only writing the answer down if it makes immediate sense, pruning away any flipoff affine measures as impossible. Accordingly, we'll later on see how to add flipoffs back in to turn a pseudocausal hypothesis into a causal one.

Definition 4.13. *We call a hypothesis η **acausal** if it is a no-flipoff hypothesis.*

The immediate question we should have is that of how (or even whether) acausal hypotheses are any different from pseudocausal ones. The main distinction is both more evident and more natural inside decision-theory puzzles: pseudocausality corresponds to decision-theoretic problems where it is always possible for an agent, on having been mispredicted, to actually end up in a situation where it's evident that it's been mispredicted and the agent can act against the prediction, while acausality corresponds to decision-theoretic problems where there exist cases where you were mispredicted,

but never end up being able to take actions that show that that prediction was in fact wrong.
In the next section, which is partially about finger-injection, we'll more fully cover how to turn a pseudocausal hypothesis into a causal one by adding in flipoffs, and similarly can turn any causal hypothesis into a set of affine environments. It's less clear how you could do this to acausal hypotheses, which is why results further down involve further treatment of the case where we have some acausal hypothesis that we can't turn into a causal one, and how that requires us to shift our perspective a little to see an agent as believing itself to be within some set of environments rather than as working purely with belief functions that need not have a representation as a set of environments.

Definition 4.14. *When it's important to us to mark whether a given hypothesis is defined only on stubs, strict-plans, or policies, we denote that with respective notations* $\eta_{\dashv}, \tilde{\eta}, \hat{\eta}$*, just as for plans. When it's important to us to mark whether a given hypothesis is acausal, causal, pseudocausal, or surcausal, we denote that with respective notations* $\eta^a, \eta^c, \eta^p, \eta^s$*.*

We'll define what it means for a hypothesis to be surcausal soon enough.

As we mention elsewhere, it takes some additional work to take an acausal hypothesis that we can't actually turn into a causal one, and figure out how to get it to basically uniquely pick out a set of environments anyway. We'll explicitly construct a toy case here, to both illuminate the problem and provide a solution.
Consider the Transparent Newcomb problem. In it, two boxes lie before us - a transparent one, and an opaque one. A nearly-perfect predictor Omega always puts a single gold coin in the opaque box, and fills the transparent box with 100 gold coins if it predicts that we would, on seeing the transparent box full, one-box; otherwise it leaves the transparent box empty. It has already done so before we saw either box, and it errs with tiny probability ϵ. We are presented with both boxes, and the choice between taking just the clear box and taking both boxes.
This is certainly a policy-dependent environment, and we should recall that we can simply hard-code all possible policies into the policy slot, make a family of policy-*independent* environments through that currying process, and laugh at how the Demiurge has no clothes if we ever find ourselves having violated our own hardcoded policy. Because Omega errs with probability ϵ, we can still counterfactually demonstrate a credible threat of getting to flip off the Demiurge should we ever actually find ourselves having violated our own hardcoded policy in a given environment. Once we translate this set of environments to a no-flipoff belief function, we get a pseudocausal hypothesis and can continue on our merry way. XOR Blackmail and Counterfactual Mugging both work similarly, and result in pseudocausal hypotheses.
Problems arise with that whole approach if Omega is actually a perfect predictor - that is, $\epsilon = 0$. Suppose our policy is to one-box if the clear box is full, and to two-box otherwise, but the environment's hard-coded policy is simply to always two-box. In that case, Omega always predicts (wrongly!) that we two-box, we're stuck always two-boxing and getting a single coin, and we never get to flip off the Demiurge for putting us in an impossible environment conflicting with our policy. Even worse, "always two-box" and "one-box on full, two-box on empty" are both optimal policies here, so even if we could directly apply our hard-coding process, we still don't get UDT behavior.

Two possible solutions present themselves: we could figure out how to encode ϵ-exploration, or we could develop some additional machinery around marking histories as possible/impossible. We'll cover both options here.

For the first option, let's suppose that we have both a deterministic "intended plan", which corresponds both to our plans in the IB sense, and to some internal sense of what we intend to do; with small probability ϵ, exploration will overwrite this. Omega thus has a hard-coded prediction of our policy, given by whatever we intend to do. On each round, we take a random draw such that we choose to act as intended with probability $1-\epsilon$ and we explore a different option with probability ϵ. If we act as intended, nothing changes for Omega, but if we choose to explore, Omega will do the opposite of what it would usually do, because it knows that we will also do that. Finally, upon seeing the clear box, we recognize an impossible world and flip off the Demiurge should our desired action on the clear box's state fail to match up to Omega's prediction - it is, after all, a perfect predictor. Accordingly, no matter what happens, the real action we take is determined entirely by our starting intention and the status of the random draw. We can thus always have some meaningful chance of proving Omega wrong and flipping off the Demiurge about it, should Omega be wrong, and we can thus turn this into a set of environments.

The second option will take a little more setting up, but it's worth it, because it cashes out in a close cousin to causal hypotheses and turns out to be the mathematically nicer way of resolving this problem.

For a naive first attempt at this, consider what happens if we try to straightforwardly represent the Perfect Transparent Newcomb problem as a suitable no-flipoff belief function over policies. Let's use $p_{a,b}$, where $1 \leq a, b \leq 2$, to denote some of our possible policies; we'll take these policies to specify how many boxes we take if the clear box is (full, empty).

Then $\beta(p_{1,1}), \beta(p_{1,2})$ both have the single history where the clear box is guaranteed full, because Omega knows we'll one-box if it's full, so we one-box for 100 gold. $\beta(p_{2,1})$ has the single history where the clear box is guaranteed empty, because Omega knows we'd two-box if it were full, and we're thus stuck one-boxing for nothing. Lastly, $\beta(p_{2,2})$ has the single history where the clear box is guaranteed empty, because Omega knows we always two-box, and so we two-box and get 1 gold.

This is a problem, though - "clear box empty, two-box" is in $\beta(p_{2,2})$, and also supported over $F^{NF}(\beta(p_{1,2}))$ - after all, that really is what we'd do, if we had that policy and we saw an empty clear box. But that's not present in $\beta(p_{1,2})$ - the pseudocausality condition from earlier wants for us to permit the possibility of the bad outcome again! So that's not what we want.

What if we looked at the (non-closed, non-flipoff-free) family of environments that correspond to the ϵ-noise Transparent Newcomb problem from before, where we let ϵ range over $0 < \epsilon \leq 0.01$ (say)?

If we try taking the closure in the space of affine environments, before we turn them into a belief function, then the belief function we end up with adds back in the same bad distribution where the fact that we might two-box is taken to mean that we always will, and if we follow a strategy that ever one-boxes, we'll never get to flip off the Demiurge.

Alright - what if we take the closure of the history-sets coming from the belief functions contemplating plans *after* we've trimmed away all the flipoffs? For $beta(p_{1,2})$ at least, we'll get a natural sequence of distributions of the form "$1-\epsilon$ probability of seeing a full box, and then we one-box; ϵ probability of seeing an empty box, and then we two box", and these limit to the Perfect Transparent Newcomb distribution we want - the one where the box is always full and we thus always one-box on seeing that. Notably, it *doesn't* limit to the failure mode from earlier, where we'd be stuck two-boxing.

We still don't actually have pseudocausality, though - "clear box empty, two-box" is still in $\beta(p_{2,2})$, and also still supported over $F^{NF}(\beta(p_{1,2}))$ - we haven't fixed that problem at all!

Our problem lies deeper still. Let's say Omega falsely locks in "two-box on full box" as our strategy, and we somehow one-box anyway. Then the expected value for all of our $0.01 \geq \epsilon > 0$-noise Newcomb problems will be infinity - we get to flip off the Demiurge - but will be just 1 gold coin in the limiting case of Perfect Newcomb!
As it stands, limits of reasonable-looking flipoff-containing affine measures can themselves be no-flipoff, and this gives us a clue as to the shape of the condition we need - the Demiurge should not be allowed to take limits of flipoff affine measures it wouldn't want to get to anyway, get a non-flipoff affine measure, and then make sure that we end up in that pathological non-flipoff limit measure.
What kinds of new limit points could we add in to compactify the subspace of flipoff measures to get around this problem? Whatever it is, it still needs to be a flipoff measure, and needs to be the limit of the "$1 - \epsilon$ probability of seeing a full box, and then we one-box; ϵ probability of seeing an empty box, and then we two box" points in a natural way. Let's relax archimedeanness of reals and try "1 probability of seeing a full box, and then we one-box; 0^+ probability of seeing an empty box, and then we two box", where 0^+ is some arbitrarily small positive number. The Demiurge will still recognize this as having flipoff in it, and will avoid it as desired.
We could also tinker with our distance metric a little as we did before, the last time we had limits of sequences that didn't converge like we wanted them to. This flavor of approach would look at measures and compare where they think a flipoff is a real possibility, and assign very different values to measures that have different measures on flipoff histories, so that we can't have flipoff measures converging to a non-flipoff measure.

Definition 4.15. *A **surmeasure** is an ordered pair of measure and a function $(\mu, f_\Diamond)$ where $f_\Diamond$ is a function on flipoff histories $f_\Diamond : \mathbb{H}^F \to \{\Diamond, \neg\Diamond\}$, where the function marks histories that end in flipping off the Demiurge as "possible" if the measure assigns them positive measure and is underdetermined in choosing between choosing "possible" or "impossible" otherwise. Histories that would otherwise get assigned 0 measure that the function marks as "possible" are instead marked as having arbitrarily tiny positive measure 0^+. We denote the set of these $S\mathcal{M}(\mathbb{H})$. We define affine and superaffine surmeasures in the natural way to extend surmeasures; we denote the set of affine surmeasures $S\mathcal{M}^a(\mathbb{H})$ and the set of superaffine surmeasures $S\mathcal{M}^{sa}(\mathbb{H})$.*

Additionally, a survironment is defined similarly: it's an environment where we have a 0^+ chance of making an impossible observation, and having done so, will inevitably end with flipping off the Demiurge.

Definition 4.16. *The **surtopology** is the topology over the space of (super)affine (sur)measures with a subbasis given by open balls around points, along with all sets of the form $\{(\mu, b) | \mu(\vec{h}) = 0\}$ where $\vec{h} \in \mathbb{H}^F$.*

More plainly: all our usual open balls, plus all the sets of (super)affine (sur)measures that assign a given flipoff history a measure of 0.

Definition 4.17. *Let $d(\cdot, \cdot)$ be the KR-metric, and $\gamma < 1$ a scaling factor. Denote by $t_F(\mu, \nu) : \mathbb{H} \times \mathbb{H} \to \mathbb{N}$ the flipoff-distance on pairs of measures such that (μ, ν) gets sent to the minimal length of a flipoff history to which one of the measures assigns positive measure and the other assigns 0 measure.*
*The **surmetric** is the metric over the space of (super)affine (sur)measures given by $d_s((\mu, b), (\nu, c)) = \max(d(\mu, \nu) + |b - c|, \gamma^{t_F(\mu,\nu)})$.*

Definition 4.18. *We call a hypothesis η **surcausal** if it is a causal hypothesis over affine surmeasures. More explicitly, we require that for all plans $\tilde{p} \in \Pi$ and all affine surmeasures $(\mu, b, f_\diamondsuit) \in \eta(\tilde{p})$, there exists a forecast function ϕ such that $\phi(\tilde{p}) = (\mu, b, f_\diamondsuit)$, and for all other $\tilde{q} \in \Pi$, we similarly have $\phi(\tilde{q}) \in \eta(\tilde{q})$.*

We primarily care about these because they're essentially equivalent to (actually slightly stronger than) causal hypotheses, and you can construct one very easily given an arbitrary acausal hypothesis, as we'll see in a bit.

We recall that the Demiurge will avoid any history that has any chance of Its getting flipped off in it. Thankfully, worrying about all this is only the kind of thing that can happen if we're in some UDT puzzle where the environment can somehow totally lock us out of demonstrating that a prediction that has been made is actually wrong. Of course, if we want to solve UDT puzzles without having to worry so much about surmeasures, you could always stick with more ordinary acausal hypotheses and just plain accept that cramming them into some causal set of hypotheses is a bad idea. One might start to worry that these three ways of incorporating "possible impossible flipoffs" might break something we care about, or at least be incompatible. Thankfully:

Proposition 4.19. *The natural surmetric on the space of superaffine measures $\mathcal{M}^{sa}(F(\tilde{p}))$ induces the surtopology on the space, and the cauchy completion of $\mathcal{M}^{sa}(F(\tilde{p}))$ with respect to the surmetric recovers $S\mathcal{M}^{sa}(F(\tilde{p}))$ exactly.*

The takeaway from all this, even if you need to blackbox all the weirdness, is that the standard way to turn a set of policy-dependent environments into a family of policy-independent environments as described early on - that is, hardcoding in all possible policies and hacking in a flipoff if those policies gets violated - only actually works if our IB-agent is then guaranteed to have some actual nonzero chance of getting to recognize that this has happened and flip off the Demiurge. There exist some UDT puzzles that invoke perfect predictors that still violate this property, and in that case, we use surmeasures to specifically allow the "manual" assignment of positive (arbitrarily small) measures to such situations, instead of 0. This process of assigning 0+ to such outcomes in turn gets us survironments and surmeasures, and then we can view even weird policy-dependent decision theory puzzles this way, as arising from a set of survironments - with that 0+ measure on some of the outcomes - instead of just a set of ordinary environments.

5 Advanced Policies and Environments (Zoom the Camera Out and See the Lie)

Now that we have the basics of hypotheses and their regularity properties down, we can talk about how we can recover an entire belief function if we have recourse to either one of what the belief function makes of only stubs or only policies.

Definition 5.1. *Let $p_\exists \in \Pi_\exists$, $\tilde{q} \in \Pi$, $\hat{r} \in \widehat{\Pi}$. Let $\beta_\exists$ be some belief function defined only on stubs, and let $\hat{\beta}$ be some belief function defined only on policies.*
We define the weaving function by $w_\exists : \tilde{q} \to \bigcap_{p_\exists \prec \tilde{q}} (pr_*^{\tilde{q}, p_\exists})^{-1}(\beta_\exists(p_\exists))$, that is, it sends every plan to the intersection, over all stubs lying below that plan, of all preimages under the appropriate projection function of the stub-belief function applied to its stub.
We define the meshing function by $\hat{m} : \tilde{q} \to \overline{c.h}(\bigcup_{\hat{r} \succ \tilde{q}} (pr_*^{\hat{r}, \tilde{q}}(\hat{\beta})(\hat{r})))$, that is, it sends every plan

to the union, over all policies lying above that plan, of all images under the appropriate projection function of the policy-belief function applied to its policy.
We define the two functions $s_{\dashv}, \hat{s}$ to be the functions taking belief functions to themselves; they differ in that the lower isomorphism function $s_{\dashv}$ is defined only on stub-belief functions and the upper isomorphism function $\hat{s}$ is only defined on policy-belief functions.

Theorem 5.2. *Let $\eta_{\dashv}$, $\hat{\eta}$ be causal hypotheses fulfilling finitary or infinitary analogues of all their defining conditions. Then $w_{\dashv}(\eta_{\dashv}), \hat{m}(\hat{\eta})$ are also causal hypotheses.*
If we swap out "causal" with pseudocausal, acausal, or surcausal in the previous sentence, the resulting sentence remains true.
Additionally, $w_{\dashv}, s_{\dashv}$ define an isomorphism between $\eta_{\dashv}, \eta$, and $\hat{m}, \hat{s}$ define an isomorphism between $\eta, \hat{\eta}$.

That is, as promised we can fully recover a belief function just by seeing what it makes of only stubs, or of only policies, as long as the analogues of the hypothesis conditions hold for the stubs/policies. Better yet, the behavior on stubs and the behavior on policies pin each other down uniquely, and we even get to recover consistency - the crucial quality of the isomorphisms - from other weaker conditions.

Definition 5.3. *We denote by S^{β} the set of affine environments or survironments defined by $S^{\beta} := \{(\lambda \cdot \hat{e}, b(, f_{\Diamond}))|\forall \tilde{p} \in \Pi, (\lambda(\tilde{p} \cdot \hat{e}), b(, f_{\Diamond})) \in \beta(\tilde{p})\}$.*
For an arbitrary set of affine environments S, we denote by β^S the belief function given by $\beta^S := \{(\lambda \cdot \mu, b)|\exists(\lambda \cdot \hat{e}, b) \in S : \tilde{p} \cdot \hat{e} = \mu\}$.

Proposition 5.4. *Let β be a causal belief also fulfilling no-flipoff nonemptiness, topological closure, and convex-hull closure. Then S^{β} is a nonempty, closed, and convex set of affine environments or survironments. $\beta^{S^{\beta}} = \beta$, and $S \subseteq S^{\beta^S}$.*

There we are! We can turn causal hypotheses into actual sets of affine environments or survironments as we like. The reverse direction - taking an arbitrary set of affine environments and turning them into a causal hypothesis - is much rarer and in general impossible, and while the original sequence cashes that out in how unlikely it is for some random collection of affine environments to satisfy hausdorff-measure continuity or no-flipoff full-stack extremal consistency, we have a better way of convincing ourselves of this: consider that if we take our terminology seriously, we'd want to reason about an IB-agent which believes that it's in one of some set of environments, each tagged with appropriate (λ, b). If we know nothing further about the environments, why *should* we expect that the IB-agent will have beliefs that looked like they formed causally? The set of environments doesn't even look like it formed causally!
Anyway, assuming that we do in fact have a set of affine environments that induce a causal hypothesis, then on going back to affine environments, we may well introduce additional points which correspond to the "chameleon environments" mentioned in the original posts. Since they mimic some combination of the behaviors of some already-existing environments for any policy it interacts with - this is why going from an abstract belief function removes the redundancy within sets of environments that the original posts mention, and is also the reason for the form of the final proposition above, especially the set containment/equality conditions at the end.
We can pass freely between stub, plan, and policy versions of acausal, pseudocausal, surcausal, and causal hypotheses, given that we have explicit isomorphisms moving between the different conditions on policies, and we can establish strong connections between (sur)causal hypotheses

and the sets of environments that they correspond to. The major remaining obstruction between history- and plan-level statements and environment-set-level statements is thus the gap between acausal/pseudocausal and surcausal/causal hypotheses. As it turns out, we can absolutely bridge that gap with a middle finger.

Definition 5.5. *Let $\tilde{q} \succ \tilde{p}$. The **(middle-finger) injection map** is given by $I^{\tilde{p},\tilde{q}} : F(\tilde{p}) \hookrightarrow F(\tilde{q})$ such that $\vec{h} :\mapsto \vec{h} \in F(\tilde{q})?\vec{h} : \vec{h}\tilde{q}(\vec{h})F$, that is, it fixes every history in $F(\tilde{q})$, and to any history outside $F(\tilde{q})$ it appends the action $\tilde{q}$ says to take on seeing $\vec{h}$ and then also appends a flipoff observation. Additionally, we write $I_*^{\tilde{p},\tilde{q}} : \mathcal{M}^a(F(\tilde{p})) \hookrightarrow \mathcal{M}^a(F(\tilde{q}))$ for the pushforward that $I^{\tilde{p},\tilde{q}}$ induces on affine measures, and $I_{*s}^{\tilde{p},\tilde{q}} : S\mathcal{M}^a(F(\tilde{p})) \hookrightarrow S\mathcal{M}^a(F(\tilde{q}))$ for the pushforward it induces in affine surmeasures; in particular, $I_{*s}^{\tilde{p},\tilde{q}}$ is effectively identical to $I_*^{\tilde{p},\tilde{q}}$, with the major difference (apart from also being defined on affine surmeasures) being that the $f_\diamondsuit$ coordinates of the elements in the image of $I_{*s}^{\tilde{p},\tilde{q}}$ all label every flipoff-history outside $F(\tilde{p})$ as possible.*

By way of characterization, we should first note that these maps are injective/injections, going upwards, not surjective/projections, going downwards, so adding in *something* is unavoidable. An I_* simply caps off every history that needs an extension with "do what the upper policy says to do, then flip off the Demiurge", while I_{*s} does the same while also assigning 0^+ measure to every flipoff-history that would otherwise have measure 0 - anywhere it can add a flipoff, it does.

We'll use finger injection maps primarily to turn acausal hypothesis-stubs into surcausal hypothesis-stubs, and pseudocausal hypothesis-stubs into acausal hypothesis-stubs. More precisely:

Definition 5.6. *We denote the grow-to-causal function by $\Gamma^c : \{\eta_{\exists}^p\} \to \{\eta_{\exists}^c\}$, and it maps from pseudocausal hypotheses defined only over stubs to causal hypotheses also defined only over stubs. It's given by $\eta_{\exists}^p(p_{\exists}) :\mapsto \overline{\text{c.h}}(\bigcup_{q_{\exists} \prec p_{\exists}} I_*^{q_{\exists},p_{\exists}}(\eta_{\exists}(q_{\exists})))$.*

We denote the grow-to-surcausal function by $\Gamma^s : \{\eta_{\exists}^a\} \to \{\eta_{\exists}^s\}$, and it maps from acausal hypotheses defined only over stubs to surcausal hypotheses also defined only over stubs. It's given by $\eta_{\exists}^a(p_{\exists}) :\mapsto \overline{\text{c.h}}(\bigcup_{q_{\exists} \prec p_{\exists}} I_{*s}^{q_{\exists},p_{\exists}}(\eta_{\exists}(q_{\exists})))$.

We won't need the "$\to^{NF}$ function defined in the original sequence; we'll use our existing $*^{NF}$ notation for that.

Theorem 5.7. *($P \to C$) Let $\eta_{\exists}^p$ be an arbitrary pseudocausal hypothesis defined only over stubs. Then $\Gamma^c(\eta_{\exists}^p)$ is a causal hypothesis defined only over stubs.*
($C \to P$) Let $\eta_{\exists}^c$ be an arbitrary causal hypothesis defined only over stubs. Then $(\eta_{\exists}^c)^{NF}$ is a pseudocausal hypothesis defined only over stubs.
Additionally, $\Gamma^c(\eta_{\exists}^p)^{NF} = \eta_{\exists}^p$.

It's still unclear whether $\Gamma^c((\eta_{\exists}^c)^{NF}) = \eta_{\exists}^c$.

Theorem 5.8. *($A \to S$) Let $\eta_{\exists}^a$ be an arbitrary acausal hypothesis defined only over stubs. Then $\Gamma^s(\eta_{\exists}^a)$ is a surcausal hypothesis defined only over stubs.*
($S \to A$) Let $\eta_{\exists}^s$ be an arbitrary surcausal hypothesis defined only over stubs. Then $(\eta_{\exists}^s)^{NF}$ is an acausal hypothesis defined only over stubs.
Additionally, $\Gamma^s(\eta_{\exists}^a)^{NF} = \eta_{\exists}^a$.

It's still unclear whether $\Gamma^s((\eta_{\exists}^s)^{NF}) = \eta_{\exists}^s$.

The post-theorem remarks aren't the important part - the paired theorems are. These are the big ones, the central link that let you go from acausal and pseudocausal hypotheses to their corresponding surcausal and causal counterparts and back again. There's a little bit of philosophy to dig into here, too.

From one perfectly valid perspective, the whole "flipping off the Demiurge" framing is an ugly hack to make the math work out properly for weird UDT puzzles, but you're just fine with working with belief functions rather than insisting on encoding them all as sets of affine environments. In that case, what we've provided is mostly simple constructive ways for trimming away all the unpleasant nonsensical flipoff-histories and such and work directly with non-flipoff belief functions.

From another perfectly valid perspective, flipping off the Demiurge is based as hell and a system that can precommit to that as a means of tagging impossible possible worlds(/policies/hypotheses) is excellent, and you think that using flipoffs to encode nasty but necessary UDT problems is a totally acceptable way of transporting an isomorphic image of all the math you need into the appropriate setting. In that case, you get to use flipoff-histories(/plans/belief functions) to turn the abstract belief functions into causal or surcausal form, and from there, to a set of affine environments or survironments, which we can interpret as a set of environments with additional information about how much we still care, how well we're doing off-history, and whether this is even possible at all.

As a side note, in the "Nirvana is 1 reward forever" setting, this turns into an isomorphism right away and we get a parallel with the original formulation of an infradistribution, in which all possible points that don't affect expected values over the set have been added in.

Proposition 5.9. *Let $\hat{\beta}^{NF}$ be a non-flipoff belief function defined over policies. We'll want to turn $\hat{\beta}^{NF}$ into at least an acausal hypothesis, and in order to do so, we must close it topologically, convex-hull-wise, and with respect to upper completion; then we must unitalize it and use the meshing and upper-isomorphism functions as used in the isomorphism theorem in order to fill in the belief function's behavior on plans.*

Accordingly, the minimal set of properties to check that $\hat{\beta}^{NF}$ satisfies are nonemptiness, boundedness of minimals, hausdorff-measure continuity, and renormalizability.

Additionally, $\hat{\beta}^{NF}$ satisfies these properties and we turn it into its corresponding acausal hypothesis β^{NF}, then for all $p \in \Pi$ and functions f, there exist $a, b \in \mathbb{R}$ such that $\mathbb{E}_{\beta^{NF}(p)}(f) = a \cdot (\mathbb{E}_{\hat{\beta}^{NF}(p)}(f) - b)$.

Basically, if we turn some arbitrary mess of affine measures over each outcome set of policies into an acausal hypothesis, we leave the pessimal values fixed up to an affine transformation by (a, b).

To sum things up: Even given some mostly-arbitrary mess of affine measures over every policy, we can turn it into an acausal hypothesis if it satisfies four technical conditions. Then, given some acausal or pseudocausal hypothesis, we can freely turn it into the corresponding surcausal or causal hypothesis and back again; we thus can conclude that we can drop either flipping off the Demiurge or affine environments/survironments, depending on preferred philosophical interpretation. In either case, we can finally cash thosea those (sur)causal hypotheses out as an actual set of affine environments/survironments that they came from. Also, even if we do in fact push that totally arbitrary mess of affine measures through this whole process and turn it into a surcausal measure, the behavior will match up perfectly modulo some fixed affine transformation. Better yet, we even have a completeness lemma:

Proposition 5.10. *For all hypotheses γ, η, the following are equivalent:*

- *For all plans p, functions f, we have $\mathbb{E}_{\gamma(p)}(f) = \mathbb{E}_{\eta(p)}(f)$.*
- *$\gamma^{NF} = \eta^{NF}$, that is, the no-flipoff versions of the hypotheses are identical.*

If regardless of the utility function we pick, we have the exact same minimum value for both our hypotheses, then if we trim off the flipoff-histories, those hypotheses are the same one, right on the nose.

It remains for us to recover mixing and updates for belief functions, just as we did for plans earlier on.
The following was neither a proposition nor a definition in the original sequence, but it seems to me like this is a property that holds of non-flipoff hypotheses defined over policies rather than something we need to explicitly define.

Proposition 5.11. *Let $\{\eta\}_{i \in S \subseteq \mathbb{N}}$ be a family of non-flipoff hypotheses defined over policies, with $\zeta \in \Delta S$. Then for all $p \in \Pi$, $((E)_\zeta \eta_i)(p) = (E)_\zeta(\eta_i(p))$; that is, the expected outcomes ζ-mix of the hypotheses gives for p is exactly the same as the value of the ζ-mix of expected outcomes of p.*

While it might initially seem like we can forge ahead easily as we did with plans, that's not quite so. Three major obstructions prevent that. First and most obviously, simply scaling and adding together our η_i gives us no guarantee that we end up with something unital, never mind that we started with unital objects. We'll write $\mathbb{E}_\zeta \eta_i$ for the raw mix, and $\mathbb{E}^U_\zeta \eta_i$ for the unitalized mix, and we should keep in mind that our earlier finiteness condition must hold - namely, $\sum_i \zeta_i \hat{\lambda}_i < \infty$.
Additionally, if we mix policies in this way and then recover our hypothesis's behavior on plans, we find that this would *not* be the same as if we took the induced mix on plans directly. In fact, an application of the upper isomorphism theorem from earlier gives us $((E_\zeta)\eta_i)(\tilde{p}) = \overline{c.h}(\bigcup_{\hat{q} \succ \tilde{p}} \mathbb{E}_\zeta(pr^{\hat{q},\tilde{p}}_*(\eta(\hat{q}))))$.
This is a particularly bad problem because it also means that our causality condition, which depends entirely on the behavior of our hypothesis on plans, breaks completely.

Given these three obstructions, we clearly can't just naively mix hypotheses defined over policies, unitalize, and expect to get anything sensible. Rather, if we want to mix causal hypotheses together to yield a prior, we must first trim away flipoffs and clean up the result to make the resulting hypothesis be pseudocausal, do our naive mixing there, and then finally translate back to the unique causal hypothesis that our mix determines. More formally:

Definition 5.12. *Let $\{\eta^c\}_{i \in S \subseteq \mathbb{N}}$ be a family of causal hypotheses defined over policies, with $\zeta \in \Delta S$. Then for $\{\eta^c\}^{NF}_i := \{\eta^p\}_i$, we denote the **$\zeta$-mixed causal hypothesis** by $\eta^c_{+\zeta} := \Gamma^c((E)_\zeta(\{\eta^p\}_i))$.*

The careful reader will note an abuse of notation in how we use Γ^c here; we have done so for the sake of clarity, rather than clutter up the definition with several to-stub/to-plan/to-policy functions. We'll also use similar notation for acausal, pseudocausal, and surcausal mixes, that is, $\eta^a_{+\zeta}, \eta^p_{+\zeta}, \eta^s_{+\zeta}$.

Definition 5.13. *Let η be a hypothesis. We say that η is **nontrivial** precisely when there exists some plan p for which $\mathbb{E}_{\eta(p)}(\mathbb{0}) \neq \mathbb{E}_{\eta(p)}(\mathbb{1})$.*

This is a very weak condition, but it gives us a sufficient condition for nontriviality of a prior made from a mixed hypothesis: even one of the η_i must be nontrivial; if we can unitalize even one of the η_i, then we'll be able to unitalize our prior.

Proposition 5.14. *Let $\{\eta^a\}_{i\in S\subseteq\mathbb{N}}$ be a family of acausal hypotheses, not all trivial, with $\zeta \in \Delta S$. Assume $\sum_i \zeta_i \hat{\lambda}_i < \infty$. Then $\eta^{a,U}_{+\zeta}$ is also an acausal hypothesis.*

Similarly, let $\{\eta^p\}_{i\in S\subseteq\mathbb{N}}$ be a family of pseudocausal hypotheses, not all trivial, with $\zeta \in \Delta\mathbb{N}$. Assume $\sum_i \zeta_i \hat{\lambda}_i < \infty$. Then $\eta^{p,U}_{+\zeta}$ is also a pseudocausal hypothesis.

In a similar vein:

Proposition 5.15. *Let $\{\eta\}_{i\in S\subseteq\mathbb{N}}$ be a family of either all acausal or all pseudocausal hypotheses. Then $\mathbb{E}_{(\mathbb{E}_\zeta \eta_i)(\tilde{p})}(f) = \mathbb{E}_\zeta(\mathbb{E}_{\eta_i(\tilde{p})(f)})$.*
Additionally, for all $\tilde{s} \succ \tilde{r}$, $\tilde{p}, \tilde{q} \in \widetilde{\Pi}$, we have $pr^{\tilde{s},\tilde{r}}_*((\mathbb{E}_\zeta \eta_i)(\tilde{p})) = \mathbb{E}_\zeta(pr^{\tilde{s},\tilde{r}}_*(\eta_i(\tilde{q})))$.

A note of warning: I am not entirely sure about changing the original sequence's Θ_n in Proposition 7 of "Belief Functions and Decision Theory", but they don't define that notation elsewhere and what they would have denoted by Θ_i seems to fit. Maybe it's just a typo or mistranscription? Proposition 8 is also very weird, given how many variables it leaves effectively unbound.

Before we can get to updates, we need another pair of technical bits.

Proposition 5.16. *Recall that $\widetilde{\Pi}^{\vec{h}}$ is the set of strict-plans starting at $\vec{h}$. For any $\tilde{p} \in \widetilde{\Pi}^{\vec{h}}$, we can translate back to $\widetilde{\Pi}$ by prefixing $\vec{h}$ to every o-history and specifying that $\tilde{p}$ behaves appropriately to make $\vec{h}$ happen. In the reverse direction, for any $\tilde{q}$ capable of producing $\vec{h}$, we may remove all o-histories with no $\vec{h}$ prefix, and then trim off the $\vec{h}$-prefix from what remains.*

Definition 5.17. *We denote the $\vec{h}$-**trim function** by $\mathrm{tr}_{\vec{h}}$. It's a partial function $\mathrm{tr}_{\vec{h}} : \mathbb{H} \to \mathbb{H}$ given by $\mathrm{tr}_{\vec{h}}(\vec{g}) = [\tau(\vec{g}) = O \wedge \vec{g} \in \widetilde{\Pi}^{\vec{h}}]?\vec{f} : \varnothing$, where $\vec{h}\vec{f} = \vec{g}$.*

More plainly: it trims off the $\vec{h}$ from the start of an o-history if it can, and if it can't, it's not defined.

Because I've chosen to present things in a somewhat different order, if you're reading along with this, using it as a map to navigate the original IB sequence, you might get confused at this point. In that case, go back and read the part about gluing policies, and mixing and updating infradistributions again. You'll need to understand that, because this next part is about mixing and updating hypotheses, and porting over all our results about infradistributions into the hypothesis setting. (Don't worry. I'll be waiting. I probably brought snacks.)

Definition 5.18. *We denote the probability of the history $\vec{h}$ relative to the belief function β, off-history plan $\tilde{p}_{\neg\vec{h}}$, and function g by $\mathbb{P}^g_{\beta,\tilde{p}_{\neg\vec{h}}}(\vec{h})$.*
We define the probability to be $\mathbb{P}^g_{\beta,\tilde{p}_{\neg\vec{h}}}(\vec{h}) := \max_{\tilde{q}\succ\tilde{p}_{\neg\vec{h}}} \mathbb{E}_{\beta(\tilde{q})}(\mathbb{1} \oplus_{\vec{h}} g) - \mathbb{E}_{\beta(\tilde{p}_{\neg\vec{h}})}(\mathbb{0} \oplus_{\vec{h}} g)$.

This definition requires some additional care. First and most importantly, while they mostly look and act like probabilities, unlike normal probabilities our $\mathbb{P}^g_{\beta,\tilde{p}_{\neg\vec{h}}}$ aren't in general additive. Like before, these act more like scaling terms we'll need in order to unitalize these, or a combined measure of likelihood and remaining expected value. All the same, we'll see a few results where the analogous result in the classical bayesian setting would use actual probability instead, and in the

case of a single environment we even recover true probability.
Let's pick through the heart of the definition bit by bit, to fully understand it. First off, we already know that at the end of all this, we want for the worst possible case to get a score of 0, and the best possible case to get a score of 1; we thus need to know the gap between the worst-case and best-case outcomes, so that we can rescale by that. We know a little more about our plan $\tilde{q}$ - we know it has to behave like $\tilde{p}_{\neg\vec{h}}$ does off-$\vec{h}$, because it lies above $\tilde{p}_{\neg\vec{h}}$. We also know our off-$\vec{h}$ utility function - that's just g.
Then the best possible case gets a raw score of $\max_{\tilde{q}\succ\tilde{p}_{\neg\vec{h}}}\mathbb{E}_{\beta(\tilde{q})}(\mathbb{1}\oplus_{\vec{h}} g)$, and the worst possible case gets a raw score of $\min_{\tilde{q}\succ\tilde{p}_{\neg\vec{h}}}\mathbb{E}_{\beta(\tilde{p})}(\mathbb{0}\oplus_{\vec{h}} g)=\mathbb{E}_{\beta(\tilde{p}_{\neg\vec{h}})}(\mathbb{0}\oplus_{\vec{h}} g)$.

Definition 5.19. *Let β be a belief function, $\vec{h}\in\mathbb{H}$ a history, $p_{\neg\vec{h}},\tilde{q}\in\Pi$ possessing their notational properties.*
First, take the intersection $\beta(p_{\neg\vec{h}}\oplus_{\vec{h}}\tilde{q})\cap\{\vec{a}\in\mathcal{M}^a(\mathbb{H})|\mu_a(\neg\vec{h}\mathbb{H}\cap\mathbb{H}^F)=0\}$.

We should interpret the former set, the affine measures in the image of the composite plan, as how good the belief function expects things might go if we follow the composite plan, and the latter set, which we read as the set of all affine measures on histories assigning measure 0 to histories where a flipoff occurs without $\vec{h}$ having happened first, as the set of expected-valuations on those possible histories which assign all of their remaining value away from histories where a flipoff occurs with without $\vec{h}$ having happened first. All in all at the end of this step, we've cut down the set of belief functions to just those belief functions that assign no remaining expected value on those histories where $\vec{h}$ never happened, but a flipoff observation happened anyway. We just plain don't care about those - we specifically want to get rid of the possibility of infinite reward carrying over from some flipoff off-history.

Next, apply to that set the map on affine measures given by $\vec{a}\in\mathcal{M}^a(\mathbb{H})\mapsto\frac{1}{\mathbb{P}^g_{\beta,p_{\neg\vec{h}}}(h)}(\text{tr}_{\vec{h}}(\mu_a|_{\vec{h}}),b_a+\mu_a(\mathbb{0}\oplus_h g)-\mathbb{E}_{\beta(p_{\neg h})}(\mathbb{0}\oplus_h g))$.

We should interpret this as modifying the belief functions; the measures (the remaining value) are now for histories where $\vec{h}$ has already happened, and we've also removed all the $\vec{h}$-prefixes for bookkeeping purposes because we're now looking solely inside histories where $\vec{h}$ has already happened. The affine constants get changed by the difference between what their respective measures assign to the off-history utility function and the expectation value over what the affine measures assign to the same off-history utility function. And of course at the end of that we make sure to unitalize.

*Finally, take the topological closure, just in case we might have some limit of a sequence of belief functions in the KR-metric to something we actually want. We denote this by $\beta^{\vec{h}}_{q_{\neg\vec{h}},g}(\tilde{p})$. We call this the **post-$\vec{h}$ update of the belief function on** $\tilde{p}$.*

Don't let the definition-ness fool you - this is a fair-sized technical result. Thankfully, we also get a few additional technical lemmas to help us feel safer about the intricacy of that last definition:

Proposition 5.20. *If we start with an acausal, causal, pseudocausal, or surcausal hypothesis, then the above update process will yield an acausal, causal, pseudocausal, or surcausal hypothesis, assuming that unitalization doesn't fail.*

Proposition 5.21. *Let η be a hypothesis which we may require at pleasure to be acausal, causal, pseudocausal, or surcausal.*

Then $\mathbb{E}_{\eta(p_{\neg\vec{h}} \oplus_{\vec{h}} \vec{q})}(f \oplus_{\vec{h}} g) = \mathbb{E}_{\eta(p_{\neg\vec{h}})}(\mathbb{0} \oplus_{\vec{h}} g) + \mathbb{P}^g_{\eta, p_{\neg\vec{h}}} \cdot \mathbb{E}_{\eta^{\vec{h}}_{p_{\neg\vec{h}}, g}(\vec{q})}(f^{\vec{h}})$, where $f^{\vec{h}}$ is the trimming of $\vec{h}$-prefixes from f, that is, we restrict the domain to outcome histories in $\Pi^{\vec{h}}$ and then apply $\mathrm{tr}_{\vec{h}}$.

Briefly - the expectation value over the affine-weighted histories that the hypothesis thinks a composite plan hinged on $\vec{h}$ will yield of some composite function also hinged on $\vec{h}$ is the same as the expectation value over the affine-weighted histories the same hypothesis thinks the off-$\vec{h}$ plan-component will yield for the off-$\vec{h}$ function-component, plus the expectation value over the affine-weighted histories the *updated* hypothesis thinks the on-$\vec{h}$ plan-component will yield for the on-$\vec{h}$ function-component, which latter has been scaled by the probability/caring-measure that $\vec{h}$ happens in the first place.

Proposition 5.22. *Let $\vec{gh} \in \mathbb{H}$ be a valid o-history. Then for any acausal, causal, pseudocausal, or surcausal hypothesis η, $[\eta^{\vec{g}}_{q_{\neg\vec{g}}, g}]^{\vec{h}}_{r_{\neg\vec{h}}, g^{\vec{h}}} = \eta^{\vec{gh}}_{(q_{\neg\vec{h}} \oplus_{\neg\vec{h}} r_{\neg\vec{h}}), g}$.*

Basically: updating a belief function on $(\vec{g}, q_{\neg\vec{g}}, g)$ and then on $(\vec{h}, r_{\neg\vec{h}}, g^{\vec{h}})$ is the exact same as just updating it once on $(\vec{gh}, (q_{\neg\vec{h}} \oplus_{\neg\vec{h}} r_{\neg\vec{h}}), g)$ - we have a simple closed form for how to combine two updates into one.

At last, we can express the equivalent of Bayes for belief functions and hypotheses!

Theorem 5.23. *Let $\{\eta_i\}_{i \in S}$ be a set of only acausal or only pseudocausal hypotheses such that for at least one i, $\eta^{\vec{h}}_{i; p_{\neg\vec{h}}, g}$ is well-defined and nontrivial.*

Then for $\zeta \in \Delta S$, we have $[\mathbb{E}^U_\zeta(\eta_i)] = \left(\frac{\mathbb{E}_\zeta(\mathbb{P}^g_{\eta_i, p_{\neg\vec{h}}}(\vec{h}) \cdot \eta^{\vec{h}}_{i; p_{\neg\vec{h}}, g})}{\mathbb{E}_\zeta(\mathbb{P}^g_{\eta_i, p_{\neg\vec{h}}}(\vec{h}))}\right)^U$

On the left side: we mix hypotheses, unitalize, and then update. On the right side: we mix updated hypotheses according to the expected-caring measure/probability they put on the observation, rescale by expected-caring measure/probability, and finally unitalize.
Compare this to classical bayesian updating, where we'd mix hypotheses to make a prior and then update, and find that to be the same as mixing already-updated hypotheses according to the probability they put on the observation.
Best of all, because we dealt with nontriviality earlier, we don't even need to worry that belief-function Bayes giving us undefined garbage just because one of the components of the prior has become trivial and "given up" - those will all straightforwardly vanish, since it suffices that even one of the η_i be nontrivial.

Now to stretch our legs a little and port decision-theoretic concepts over, too!

Proposition 5.24. *("Has Been The Whole Time" Theorem) Let η be a hypothesis, which we may require at pleasure to be acausal, causal, pseudocausal, or surcausal. Let p be an arbitrary plan and U a utility function. Then for $p^{\vec{h}}$ the continuation of p post-update, $p_{\neg\vec{h}}$ the off-$\vec{h}$ behavior of p, and d a plan for which $\mathbb{E}_{(\beta|U, p_{\neg\vec{h}}, \vec{h})(p^h)}(U^h) \simeq \mathbb{E}_{(\beta|U, p_{\neg\vec{h}}, \vec{h})(d)}(U^h)$, where we may take $\simeq$ to mean any of $<, =, >$.*
Then $\mathbb{E}_{(\beta(p))}(U) \simeq \mathbb{E}_{(\beta(p_{\neg h} \oplus_h d)}(U)$, where $\simeq$ must be the same as above.

More pithily, if you think that $p^{\vec{h}}$ is a better idea than d after having seen $\vec{h}$, then you've thought as much the whole time about p being a better idea than the compound plan of d (in case of $\vec{h}$) and the off-$\vec{h}$ p; and the same will be true if you think that they're just as good as each other, or if you think that $p^{\vec{h}}$ is actually worse than d. In no event should your past self, who has less information than you, be screaming from the sidelines disagreeing with your choices.

Definition 5.25. *Let $r : \widetilde{\mathbb{H}} \to [0,1]$ be a reward function, and $\gamma < 1$ be a time-discount parameter. Then we call the utility function $\mathcal{U}^\gamma : \mathbb{H}^\omega \to [0,1]$ that takes $\vec{h}$ to $\mathcal{U}^\gamma(\vec{h}) := (1-\gamma)\sum_{n=0}^{\infty}\gamma^n r(\vec{h}_{\leq n})$ the **γ-discounted utility function**.*

We provide the definition of classical regret, classical learnability, and bayes-optimality of policies for completeness and self-containedness:

Definition 5.26. *Let $\hat{p}$ be a policy, $\check{e}$ an environment, and U a utility function. Then we define the (classical) regret of $\hat{p}$ as $R(\hat{p}, \check{e}, U) := \max_{\hat{q}\in\widehat{\Pi}}(\mathbb{E}_{\hat{q}\cdot\check{e}}(U) - \mathbb{E}_{\hat{p}\cdot\check{e}}(U))$.*

That is - the regret value of a policy is the difference between what you could have scored, and what you did score.

Definition 5.27. *We say that a set of environments $\check{e}_i$ is (classically) learnable with respect to some γ-parametrized family of utility functions $\mathcal{U}^\gamma$ exactly when there exists some γ-indexed family of policies $\hat{p}_\gamma$ such that $\forall i : \lim_{\gamma\to 1} R(\hat{p}_\gamma, \check{e}_i, \mathcal{U}^\gamma) = 0$.*

Definition 5.28. *Let ζ be a prior over a family of environments $\check{e}_i$ assigning probability 0 to no outcome; we will treat this as a single environment $\check{e}_\zeta$. Then we call a γ-indexed family of policies $\hat{p}^*_\gamma$ bayes-optimal if for all $\gamma < 1$, $\hat{p}^*_\gamma \in \operatorname{argmax}_{\hat{p}\in\widehat{\Pi}}\mathbb{E}_{\hat{p}\cdot\check{e}_\zeta}(\mathcal{U}^\gamma)$.*

Alright, enough of the classical stuff, on to infrabayesianism!

Definition 5.29. *The **regret function** R is defined by its taking ordered triples (p, η, U) of policies, hypotheses, and utility functions to $R(p, \eta, U) := \max_{\hat{p}}(\mathbb{E}_{\eta(\hat{p})}(U) - \mathbb{E}_{\eta(p)}(U))$, where $\hat{p}$ is the optimal policy.*

Definition 5.30. *Let $\mathcal{U}$ be a family of utility functions. A family of hypotheses η_i is **learnable** if there exists some γ-indexed family of policies p_γ such that for all i, $\lim_{\gamma\to 1} R(p_\gamma, \eta_i, U_\gamma) = 0$.*

That is: a family of hypotheses is learnable precisely when as we limit time-discounting towards 0, we still have some paired family of policies where the regret also tends to 0.

Definition 5.31. *Suppose we have finitely many world-states S, possible observations O, and possible actions A, along with an observe function $ob : S \to \Delta O$ taking states to probability distributions over observations and a payoff function $P : S \times A \to [0,1]$. In this setting, a policy $p : O \to A$ can be thought of as the same thing as its induced function $f_p : S \to [0,1]$; more generally, for such a policy q, we have $f_q(s) = \mathbb{E}_{ob(s)}(P(s, q(o)))$.*

*Two policies p, q are **indistinguishable** exactly when $f_p = f_q$; we also say that they share an equivalence class.*

Definition 5.32. *We call a γ-indexed family of policies p^γ **infrabayes-optimal** with respect to some specified unitalized prior $\mathbb{E}_\zeta\eta_i$ if for all $\gamma < 1$, we have $p^{*,\gamma} \in \operatorname{argmax}_p\mathbb{E}_{\mathbb{E}_\zeta\eta_i(p)}(\mathcal{U}^\gamma)$.*

Proposition 5.33. *Let η_i be a family of learnable hypotheses. Then any infrabayes-optimal family for a prior on the hypotheses can also learn the collection of hypotheses.*

Definition 5.34. *A total order $\succ$ on equivalence classes of policies is **convex** if when we interpret the f_p as points in $[0,1]^S$, we get that whenever $f_q \in c.h(\{f_r | r \succ p\}) + [0,\infty)^S$, we also have $q \succ p$.*

That is: whenever q is in the set of policies outperforming a specific policy p, we always get that $q \succ p$.

Conjecture. *(The Complete Class Conjecture) Suppose $\succ$ is a complete ordering over equivalence classes of plans, and that it fulfills the convexity property. Then there exists an infradistribution D over states S such that $q \succ p \to \mathbb{E}_D(f_q) > \mathbb{E}_D(f_p)$.*

Proposition 5.35. *(The Weak Complete Class Theorem) Let p be a Pareto-optimal policy. Then for all $q \in \Pi$ with $f_p \neq f_q$, there is some infradistribution D over states S so that $\mathbb{E}_D(f_p) > \mathbb{E}_D(q)$.*

Acknowledgements and Miscellany

With thanks to Carolus Vitellius and Jay Azoth, without either of whom writing this would have been infeasible at best.
With some thanks to Diffractor and Vanessa Kosoy. With additional thanks to John Wentworth and Quinn Dougherty, who both played crucial roles in getting me to actually write this.