Ockham Efficiency Theorem for Empirical Methods Conceived as Empirically-Driven, Countable-State Stochastic Processes

Kevin Kelly
Conor Mayo-Wilson

November 29, 2008

Abstract

Ockham's razor is the principle that, all other things being equal, it is rational to prefer simpler scientific theories to more complex ones. In a series of a papers, Kelly, Glymour, and Schulte argue that scientists who heed Ockham's razor make fewer errors and retract their opinions less often than do their complexity preferring counterparts. The centerpiece of their argument is the *Ockham Efficiency Theorem,* which provides a precise explanation of errors, retractions, and Ockham's razor within a model of scientific inquiry developed by formal-learning theorists. Kelly, Glymour, and Schulte's previous arguments, however, were restricted in two important ways: (1) they applied only to deterministic (rather than randomized) methods for choosing scientific theories from data and (2) they failed to successfully model inference from statistical data with error. In this paper, we full address the first issue by extending the Ockham Efficiency Theorem to prove that, amongst any set of randomized strategies, a systematic preference for simpler theories minimizes the number of errors and retractions one commits before converging to the true theory. By incorporating probabilistic elements into the model employed by formal learning theorists, moreover, we take a large step towards addressing the second issue as well.

1 Introduction

From the Copernican revolution to Einstein's jettisoning of absolute space from mechanics, some of the most celebrated advances in the history of science were motivated in part by a desire to simplify existing theories. The systematic preference for simpler theories, moreover, still dominates scientific practice today. Faced with multiple competing theories that are all compatible with existing experimental and observational evidence, scientists eschew complexity in favor of theories with fewer laws, fewer free parameters, fewer postulated causes, fewer fundamental entities (e.g. particles), and so on. Moreover, this systematic bias for simpler theories is often tacitly built into computer-statistical packages that have become the modern-day toolbox for working scientists. But why should scientists favor simpler scientific theories when the world might, in fact, be extremely complicated? In particular, is there any reason to believe that simpler theories are more likely to be *true?*

To answer these questions, many philosophers have argued that simpler theories possess other theoretical virtues. Simpler theories, they claim, are more unified (Friedman), more easily falsified or tested (Popper, Mayo), more explanatory (Harman, Nolan, and Baker), and more concise, in that simpler theories minimize description length (Risannen, Vitanyi, Li, Simon,). However, the scientific theory that truly describes the world might lack unity or be "dappled" (Cartwright); it might be difficult to test and/or falsify, and its explanation of observed phenomena might be long and convoluted. In short, unless one has independent reason to think that true scientific theories possess these other virtues (unifying power, falsifiability, explanatory power, etc.), the above arguments provide no reason to think that simpler theories are more likely to be true.

Other philosophers and statisticians have argued that scientists who favor simpler hypotheses will *eventually* endorse the true theories *in the long run* (Sklar, Friedman, Rozenkrantz). Yet as Reichenbach first noted, and was subsequently endorsed by Hempel and Salmon, almost *any arbitrary bias* is compatible with finding true scientific theories eventually. [Finish - describe prior washing out in long run]

So-called Bayesians and confirmation theorists, argue that simpler theories are better confirmed, and hence, there is reason to think that simpler theories are more likely to be true. Such arguments, however, either explicitly (Jeffreys) or implicitly (Rosenkrantz) assume that simpler theories are assigned higher prior probabilities.[1] But then simpler theories are more probable precisely because one assumes them to be more probable. Clearly, such circular arguments are unacceptable.

Most recently, a number of philosophers have harnessed mathematical theorems from statistics and machine learning to argue simpler theories make better predictions (Harman and Kulkarni, Vapnik, Forster, Sober, and Hitchcock). But the theorems prove too much: simpler theories, according to the theorems, make better predictions *regardless of whether they are true or not.* For this reason, Vladamir Vapnik, the inventor of statistical learning theory, argues that one should use simpler statistical models in many practical applications *even when it is known that the simpler model is false.* Astute philosophers have noticed this feature of the statistical theorems they employ. For example, Forster, Sober, and Hitchcock argue that simplicity is merely instrumental in helping one make better predictions, but it is no indication of the truth of a scientific theory.

In a series of papers, Kevin Kelly, Clark Glymour, and Oliver Schulte have provided a more nuanced and promising defense of Ockham's razor. They argue that scientists who systematically favor simpler hypotheses will make fewer

[1]Explain why uniform prior on simpler and complex theories begs the question.

errors (i.e. they will endorse false theories less often), and they will retract previously endorsed theories fewer times before ultimately settling on the true theory in the long run. This thesis is captured by the slogan, "Ockham's razor equals efficient convergence to the truth." Importantly, the model of scientific inquiry developed by these authors provides a successful explanation of why simpler theories ought to be preferred in a number of scientific problems including curve-fitting, causal inference, and estimating conserved quantities in particle physics.

Kelly, Glymour, and Schulte's arguments, however, only consider *deterministic* methods for choosing scientific theories from observed data. In game theory, it is familiar that the use of *randomized* strategies often allows one to minimize costs (or maximize gains) in a way in which deterministic strategies cannot. Thus, an important question is the following:

"Does a scientist who employs Ockham's razor minimize errors and retractions in converging to truth, when compared, as well, to scientists capable of employing any randomized method for choosing theories from data?"

In this paper, we prove that the answer to this question is "yes." The notion of randomized strategy we consider is very general: it includes the class of "mixed strategies" developed to analyze normal form games; it includes the class of "behavior strategies" developed to analyze extensive form games; it includes the class of Randomized Turing Machines (RTMs), whose output is a function of both the current tape reading and its immediate previous internal state (i.e. RTMs are Markov processes). In fact, in the model of scientific inquiry developed in this paper, a scientist may choose theories on the basis of existing data using a random device that is correlated (or not) to any arbitrary degree with her choices in the past and in the future. Amidst all such strategies for choosing scientific theories, a systematic preference for simpler theories still proves to minimize errors and retractions in converging to the truth. Moreover, the randomized methods that minimize costs in converging to the true theory turn out to be only minor variants of the deterministic strategies described in previous papers (Kelly ****). This is good news, as (we hope that) working scientists rarely flip coins or roll dice to decide which of a set of theories to pursue researching.

2 Stochastic Empirical Inquiry

2.1 Stochastic Processes

The contents of this section are introductory and may be skipped by those with sufficient background in stochastic processes. Our summary is borrowed from Shalizi's text. If f is a function from $X \times Y$ into Z, then for each $x \in X$, let $f_x : Y \to Z$ be the unique function such that for each $y \in Y$, $f_x(y) = f(x, y)$. For each $y \in Y$ define $f_y : X \to Z$ similarly.

Let T, Δ, Σ be arbitrary sets. A *stochastic process* is a quadruple

$$\mathcal{Q} = (T, (\Delta, \mathcal{D}, p), (\Sigma, \mathcal{S}), X),$$

where:

1. T is a set called the *index set* of the process;
2. $(\Delta, \mathcal{D}, p)$ is a (countably additive) probability space;[2]
3. $(\Sigma, \mathcal{S})$ is a measurable space of possible *values* of the process;
4. $X : T \times \Delta \to \Sigma$ is such that for each fixed $t \in T$, the function X_t is $\mathcal{D}/\mathcal{S}$-measurable.

If $t \in T$ and $\delta \in \Delta$, then X_t is a *random variable* of the process, and X_δ is a *sample path* of the process. Let $t \in T$ and $S \in \mathcal{S}$. Let $t \in T$ and define:

$$[X_t \in S] = (X_t)^{-1}(S) = \{\delta \in \Delta : X(t, \delta) \in S\}.$$

Then $p(X_t \in S)$ is defined.

2.2 Empirical Worlds and Theories

Let E be a non-empty, countable (finite or countably infinite) set of *empirical effects*. A problem is a set $K \subseteq 2^E$ that corresponds to an empirical constraint on which finite sets of effects one might see for eternity. In this paper, K is a parameter that is often held fixed, in which case reference to K may be dropped to ease notation. Say that problem K is *bounded nowhere* if and only if for each $S \in K$ there exists $S' \in K$ such that $S \subset S'$.

An *empirical world* in K is an ω-sequence w of disjoint subsets of E such that

$$\bigcup_{i \in \omega} w_i \in K.$$

Let W_K be the set of all empirical worlds. Let $w|n$ denote the finite initial segment $(w_0, \ldots, w_{n-1})$ of w, so that, in particular, $w|0 = ()$. For arbitrary set R, let R^∞ denote the ω-sequence that is constantly R. Let:

$$\begin{aligned} F_{K,n} &= \{w|n : w \in W_K\}; \\ F_K &= \bigcup_{i \geq 0} F_{K,i}. \end{aligned}$$

Let $e, e' \in W_K \cup F_K$. Define the length of e as:

$$l(e) = |dom(e)|.$$

[2] It is usual to denote the underlying measurable space by $(\Omega, \mathcal{F})$, with ω as a representative element of Ω, but in this paper the set-theoretic interpretation of ω as the first infinite ordinal number takes precedence.

Define $e \leq e'$ to mean that e an initial segment of e' and let $e < e'$ hold just in case e is a proper initial segment of e'. Let $e \ll e'$ hold just in case $e < e'$ and $l(e')l(e) + 1$. Say that $e \approx e'$ if and only if $e \leq e'$ or $e' \leq e$, in which case one may say that e is *compatible* with e'. Let $*$ denote sequence concatenation. Define:

$$\begin{aligned} F_{K,e} &= \{e' \in F_K : e \leq e'\}; \\ W_{K,e} &= \{e' \in W_K : e \leq e'\}. \end{aligned}$$

The set of effects presented along e is:

$$S_e = \bigcup_{i<l(e)} e_i.$$

The restriction of K to effect sets compatible with e is then:

$$K_e = \{S \in K : S_e \subseteq S\}.$$

Also, if $e \in F_K$, let e_- denote the result of deleting the last entry in e, if $e \neq ()$ and let e_- denote $()$ otherwise. Let $S \in K$. Then the unique theory corresponding to effect set S is:

$$T_S = \{w \in W : S_w = S\}.$$

Then let:

$$\begin{aligned} \mathsf{Th}_K &= \{T_S : S \in K\}; \\ \mathsf{Th}_{K,e} &= \{T_S : S \in K_e\}. \end{aligned}$$

If $T \in \mathsf{Th}_K$, let S_T denote the unique $S' \in K$ such that for each $w \in T$, $S_w = S'$. For each $w \in W_K$, it will be convenient to abbreviate T_{S_w} as T_w, which is the unique theory $T \in \mathsf{Th}$ such that $w \in T$.

2.3 Stochastic Empirical Methods

Let $\mathsf{Ans}_K = \mathsf{Th}_K \cup \{`?'\}$. Since the intended applications all involve a discrete hypothesis language, it will be assumed that the set E of empirical effects is countable. Because Ans_K contains only '?' and finite subsets of E, it will likewise be countable. We think of a stochastic empirical method $\mathcal{M}$ as operating as follows. At the outset of inquiry, when no inputs have yet been provided, $\mathcal{M}$ is initialized to a non-random state

$$X_{()} = \sigma_0.$$

At stage n, finite input sequence $e(S_0, \ldots, S_{n-1})$ has been presented and $\mathcal{M}$ enters a new random state

$$X_e = \sigma_n$$

in response to e. Then method $\mathcal{M}$ employs a uniform rule or procedure $\alpha(e, \sigma)$ for choosing an answer in Ans_K in light of the current input sequence e and the current random state σ. Hence, at stage n:

$$\mathcal{M}_e = \alpha(e, \sigma_n) \in \mathsf{Ans}_K.$$

This setup is general enough to model random automata and random Turing machines, but it is far more general than that. Random computational models are *Markov* processes, which are processes in which X_e is dependent only on the set of variables $X_{e'}$ such that $e' \geq e$ given the value of X_{e_-}. The arguments that follow assume *nothing* about statistical independence among the evolving, random states of $\mathcal{M}$. Furthermore, we do not consider joint probabilities of state evolutions of alternative methods at all, so the question of correlations between alternative methods does not even arise.

A *stochastic empirical method* for problem K is just a stochastic process indexed by the set F_K of all finite input sequences. A *stochastic empirical method* is a triple

$$\mathcal{M} = (\mathcal{Q}, \alpha, \sigma_0),$$

where:

1. $\mathcal{Q}$ is a stochastic empirical process $(F_K, (\Delta, \mathcal{D}, p), (\Sigma, \mathcal{S}), X)$;
2. $\sigma_0 \in \Sigma$ satisfies: $X_{()}^{-1}(\sigma_0) = \Delta$.
3. $\alpha : F_K \times \Sigma \to \mathsf{Ans}$ is such that for each $e \in F_K$, α_e is $\mathcal{S}/2^{\mathsf{Ans}}$-measurable.

Call X_e the *random state variable*, Σ the set of possible *states*, and α the *output generation function*.

Let $p_{\mathcal{M}}$ denote the probability measure occurring in the stochastic process $\mathcal{Q}$ occurring in $\mathcal{M}$. When $\mathcal{M}$ appears elsewhere in the formula, as in $p_{\mathcal{M}}(\mathcal{M}A)$, the subscript $\mathcal{M}$ will be dropped to avoid clutter. One may think also of $\mathcal{M}_e$ as a function:

$$\mathcal{M}_e(\delta) = (\alpha_e \circ X_e)(\delta).$$

Then since α is $\mathcal{S}/2^{\mathsf{Ans}}$-measurable and X_e is $\mathcal{D}/\mathcal{S}$-measurable, it follows that $\mathcal{M}_e(\alpha_e \circ X_e)$ is $\mathcal{D}/2^{\mathsf{Ans}}$-measurable and, hence, $\mathcal{M}_e$ is an Ans-valued random variable on $(\Delta, \mathcal{D})$. Let $A \in \mathsf{Ans}_K$, and define:

$$[\mathcal{M}_e = A] = \{\delta \in \Delta : \mathcal{M}_e(\delta) = A\}.$$

Thus for any for any $A \in \mathsf{Ans}_K$, the probability $p(\mathcal{M}_e = A)$ is well-defined.

2.4 Discrete State Methods

The proposed conception of stochastic empirical methods is very general. We constrain the concept in a natural way, by assuming that the set Σ of possible states of $\mathcal{M}$ is countable and discrete, which means that Σ is either finite or countably infinite and

$$\mathcal{S} = 2^{\Sigma}.$$

Such a stochastic method is called a *discrete state* method. For example, methods implemented in random Turing machines are a special case of discrete state methods.

Random Turing machines are *Markov* processes, which have the property that future states are statistically independent of past states given the current state. The results that follow are far more general: they require no independence conditions at all regarding states. There may, for example, be magical informational channels from states in which empirical effects have been observed back to states in which they have not yet been observed. That might seem to make Ockham's razor a bad idea, at least in some cases. Surprisingly, Ockham's razor is still demonstrably the best policy to follow at every stage, so far as worst-case expected errors and retractions are concerned. The trick is, roughly, that convergence to the truth in worlds in which the effects never arise eliminates the advantage of such correlations in the worst case.

2.5 State and Answer Trajectories

Let $\mathcal{M} = ((F_K, (\Delta, \mathcal{D}, p), (\sigma, \mathcal{S}, X)), \alpha, \sigma_0)$ be a stochastic method. When $e \in F_K$ and $\delta \in \Delta$, the finite state history of $\mathcal{M}$ in response to e given sample point δ is:

$$X_{[e]}(\delta) = (X_{e|i}(\delta) : i \leq l(e)),$$

and the finite output trajectory of $\mathcal{M}$ in response to e given state trajectory $s \in \Sigma^{l(e)}$ is:

$$\mathcal{M}_{[e]}(s) = (\alpha(s(i-1)(\delta), w|i) : i \leq l(e)).$$

Thus, the finite output trajectory of $\mathcal{M}$ in response to e given δ is:

$$\mathcal{M}_{[e]}(\delta) = (\mathcal{M}_{[e]} \circ X_{[e]})(\delta).$$

Similarly, if $w \in W_K$ and $\delta \in \Delta$, then define the infinite versions of these concepts as:

$$\begin{aligned} X_{[w]}(\delta) &= (X_{w|i}(\delta) : i < \omega); \\ \mathcal{M}_{[w]}(s) &= (\alpha(s(i-1)(\delta), w|i) : i < \omega); \\ \mathcal{M}_{[w]}(\delta) &= (\mathcal{M}_{[w]} \circ X_{[w]})(\delta). \end{aligned}$$

Define the event:

$$[X_{[e]} = s] = \bigcap_{i \leq l(e)} [(X_{[e]})_i = s_i].$$

By definition, $[X_{[e]} = s]$ is $\mathcal{D}$-measurable, so the probability $p(X_{[e]} = s)$ is defined. When $e = ()$, the initial state σ_0 is assumed to be certain, so we have:

$$p(X_{[()]} = (\sigma_0)) = p(\Delta) = 1.$$

If $A \subseteq \mathsf{Ans}_K^{\leq l(e)}$,define:

$$[\mathcal{M}_{[e]} \in A] \quad = \quad \{\delta \in \Delta : \mathcal{M}_{[e]}(\delta) \in A\}.$$

Let:

$$\Sigma_{\mathcal{M},e,A} = \{s \in \Sigma^{l(e)+1} : \mathcal{M}_{[e]}(s) \in A\}.$$

Then:

$$[\mathcal{M}_{[e]} \in A] \quad = \bigcup_{s \in \Sigma_{\mathcal{M},e,A}} [X_{[e]} = s].$$

Since $\Sigma_{\mathcal{M},e}(A) \subseteq \Sigma^{l(e)+1}$ is countable, $[\mathcal{M}_{[e]} \in A] \in \mathcal{D}$ and, hence, $p(\mathcal{M}_{[e]} \in A)$ is defined.

2.6 State Support and Total Probability

The countable state assumption allows one to represent all probabilities of interest as countable linear combinations of conditional probabilities. Let $e \in F_K$, let $s \in \Sigma^{l(e)}$ and let $D \in \mathcal{D}$ satisfy $p(D) > 0$. Define the *conditional state support* of $\mathcal{M}$ on e *given* D as:

$$\mathsf{Spt}(X_{[e]} \mid D) = \{s \in \Sigma^{l(e)+1} : p(X_{[e]} = s \mid D) > 0\}$$

In particular, when D is an event of the form $X_{[e]} = s$ for $e \in F_K$, $s \in \Sigma^{l(e)+1}$, and $p(X_{[e]} = s) > 0$, then for any arbitrary $e' \geq e$, we have

$$\mathsf{Spt}(X_{[e']} \mid X_{[e]} = s) = \{s' \in \Sigma^{l(e')+1} : p(X_{[e']} = s' \mid p(X_{[e]} = s) > 0\}$$

It is easy to check that s' is an element of $\mathsf{Spt}(X_{[e']} \mid X_{[e]} = s)$ only if $s' \geq s$.

Assuming that Σ is countable, it is also true that $\mathsf{Spt}(X_{[e]} \mid D)$ is countable, so we have:

Proposition 1 (total probability) *for each $e \in F_K$ and for each $D, D' \in \mathcal{D}$:*

$$p(D \mid D') \quad = \sum_{s \in \mathit{Spt}(X_{[e]} \mid D')} p(D \mid X_{[e]} s \wedge D') \cdot p(X_{[e]} = s \mid D').$$

Some elementary consequences of total probability will be essential.

Proposition 2 *Suppose that $e \in F_K$ and $e' \in F_{K,e}$ and $D \in \mathcal{D}$ such that $p(D) > 0$ and $s \in \mathit{Spt}(X_{[e]} \mid D)$. Then:*

$$\mathit{Spt}(X_{[e']} \mid D) \neq \emptyset.$$

Proof. Since $s \in \mathsf{Spt}(X_{[e]} \mid D)$, it is immediate that $p(X_{[e]} = s) > 0$, so:

$$\begin{aligned}
0 \quad &> \quad p(X_{[e]} = s) \\
&= \sum_{s' \in \mathsf{Spt}(X_{[e']} \mid D)} p(X_{[e]} = s \mid X_{[e']} = s' \ \wedge \ D) \cdot p(X_{[e']} = s' \mid D) \\
&= \sum_{s' \in \mathsf{Spt}(X_{[e']} \mid D) \ \wedge \ s \leq s'} p(X_{[e]} = s \mid X_{[e']} = s' \ \wedge \ D) \cdot p(X_{[e']} = s' \mid D) \\
&= \sum_{s' \in \mathsf{Spt}(X_{[e']} \mid D)} p(X_{[e']} = s' \mid D),
\end{aligned}$$

where the first identity is by total probability and the second holds because $p(X_{[e]} = s \mid X_{[e']} = s' \wedge D) = 0$ if $s \not\leq s'$ and 1 if $s \leq s'$ when $s' \in \mathsf{Spt}(X_{[e']} \mid D)$. Since the last sum is non-zero, $\mathsf{Spt}(X_{[e']} \mid D) \neq \emptyset$. □

Proposition 3 *Let $e \in F_K$, let $D, D' \in \mathcal{D}$, and let $p(D') > 0$. Then the following statements are equivalent:*

1. $p(D \mid D') = 0$;
2. $p(D \mid D' \ \wedge \ X_{[e]} = s) = 0$, *for each* $s \in \mathsf{Spt}(X_{[e]} \mid D')$.

The same is true if 0 is replaced by 1 in both statements.

Proof. By total probability:

$$p(D \mid D') \quad = \sum_{s \in \mathsf{Spt}(X_{[e]} \ \mid \ D')} p(D \mid D' \wedge X_{[e]} = s) \cdot p(X_{[e]} = s \mid D').$$

For each $s \in \mathsf{Spt}(X_{[e]} \mid D')$,

$$p(X_{[e]} = s \mid D') > 0.$$

Hence, the statement:

$$\sum_{s \in \mathsf{Spt}(X_{[e]} \ \mid \ D')} p(D \mid D' \wedge X_{[e]} = s) \cdot p(X_{[e]} = s \mid D') = 0$$

holds if and only if for each $s \in \mathsf{Spt}(X_{[e]} \mid D')$,

$$p(D \mid D' \wedge X_{[e]} = s) = 0.$$

□

2.7 Variable Conventions

To eliminate clutter, the following typing conventions are assumed for variables and are assumed in all definitions, theorem statements, and proofs.

K is a problem;

$\mathcal{M}$ is a stochastic empirical method for K;

X_e is the state variable of $\mathcal{M}$;

α is the output function of $\mathcal{M}$;

σ_0 is the initial state of $\mathcal{M}$;

$w, w' \in W_K$;

$e, e' \in F_K$;

$s \in \mathsf{Spt}(X_{[e]})$ and $s' \in \mathsf{Spt}(X_{[e']})$.

Particularly important is the assumption that $e, e' \in F_K$, which entails that:

$$W_{K,e} \neq \emptyset \text{ and } W_{K,e'} \neq \emptyset;$$

and the assumption that $s \in \mathsf{Spt}(X_{[e]})$ and $s' \in \mathsf{Spt}(X_{[e']})$, which entail that

$$p(X_{[e]} = s) > 0 \text{ and } p(X_{[e']} = s') > 0.$$

When a second method $\mathcal{M}'$ is introduced, its initial state, state variables, and output function are all primed.

3 Methodological Properties

A *methodological property* is a relation of form:

$$\Phi(K, \mathcal{M}, e', e, s).$$

Adopt the mnemonic notation:

$$\Phi_K(\mathcal{M}, e' \mid X_{[e]}s) \equiv \Phi(K, \mathcal{M}, e', e, s).$$

Recalling that $[X_{[()]}(\sigma_0)] = \Delta$, define:

$$\Phi_K(\mathcal{M}, e') \equiv \Phi_K(\mathcal{M}, e' \mid X_{[()]}(\sigma_0)).$$

It is not assumed that Φ depend upon all of its arguments. Also, in applications, it is intended that $e' \approx e$, since we are not interested in random trajectories in counterfactual input scenarios, but it turns out that the results do not actually require that assumption, so it is never explicitly made.

3.1 Properties that Hold Henceforth and Perfectly

Say that methodological property Φ holds *henceforth* of $\mathcal{M}$ in K given $X_{[e]} = s$ if and only if:

$$\Phi_K(\mathcal{M}, e' \mid X_{[e]} = s) \text{ holds for all } e' \in F_{K,e}.$$

The holding of Φ henceforth given $X_{[e]} = s$ varies time into the future but holds the conditioning event $X_{[e]} = s$ fixed. A stronger notion of Φ "continuing to hold" is *perfection*, which varies the conditioning event along with the time. Say that methodological property Φ holds *perfectly* of $\mathcal{M}$ in K given $X_{[e]} = s$ if and only if:

$$\Phi_K(\mathcal{M}, e' \mid X_{[e']} = s') \text{ holds, for all } e' \in F_{K,e} \text{ and } s' \in \mathsf{Spt}(X_{[e']}).$$

When Φ holds henceforth given $X_{[()]} = (\sigma_0)$, say that Φ holds *always* and when Φ holds perfectly given $X_{[()]}(\sigma_0)$, say that Φ holds *perfectly.*

3.2 Logical Consistency With the Data

An obvious methodological idea is to refuse ever to produce an answer logically incompatible with the current inputs. Theory T is *refuted* by e if and only if

$$E_e \not\subseteq E_T.$$

Let Ref_e denote the set of all theories in Th_K that are refuted by e. Then:

$$[\mathcal{M}_e \in \mathsf{Ref}_e] = \bigcup_{A \in \mathsf{Ref}(e)} [\mathcal{M}_e A].$$

Since E is countable, so is $\mathsf{Ref}(e)$, so $p(\mathcal{M}_e \in \mathsf{Ref}_e)$ is defined. Let $e \in F_K$ and let $s \in \mathsf{Spt}(X_{[e]})$. Then say that $\mathcal{M}$ is *logically consistent* at e' given $X_{[e]} = s$ if and only if:

$$p(\mathcal{M}_{e'} \in \mathsf{Ref}_{e'} \mid X_{[e]} = s)0.$$

3.3 Empirical Simplicity and Ockham's Razor

Let $e \in F_K$. A *path* in K_e is a finite or infinite sequence of elements of K_e ordered by $\subset$. Let $\mathsf{path}_K(S \mid e)$ denote the set of all finite paths in K_e that terminate in S. A path $(S_0, \ldots, S_n, \ldots)$ is said to be *maximal* if for all S_i, there does not exist $S \in K_e$ such that $S_i \subset S \subset S_{i+1}$. Then define the *empirical complexity* of S given e as:

$$c_{K,e}(S) = \max\{l(q) : q \in \mathsf{path}_K(S \mid e)\} - 1.$$

Then define:

$$\begin{aligned} c_{K,e}(w) &= c_{K,e}(S_w); \\ c_{K,e}(T_S) &= c_{K,e}(S); \\ C_{K,n}(e) &= \{w \in W_e : c_{K,e}(w) = n\}. \end{aligned}$$

The set $C_{K,n}(e)$ is the nth *empirical complexity class* of worlds relative to problem K given e. A basic fact is that the least complexity class is non-empty.

Proposition 4 $C_{K,0}(e) \neq \emptyset$.

Proof. Recall the notational convention that $e \in F_K$. Then there exists $w \in W_K$ such that $e < w$. So

$$S_e \subseteq S_w \in K_e.$$

Since each element of K_e is finite, let S be a least element of K_e. Then (S) is the longest path ending with S in K_e, so $c_{K,e}(S) = 0$. Hence:

$$e * (S \setminus S_e) * (\emptyset^\omega) \in C_{K,0}(e).$$

□

If $S \in K_e$ is such that $S = S_w$ for some $w \in C_{K,0}(e)$, then we say S is *minimal* with respect to e. Let $e, \in F_K$ and let $T \in \mathsf{Th}_K$. Then say that T is Ockham at e if and only if for each $S \in K_e$ such that $c_{K,e}(S) = 0$, $S = S_T$. That is, T is Ockham for e if and only if S_T is uniquely minimal with respect to e. Let $\mathsf{Ock}_{K,e} = \{`?', T\}$, where T is the unique Ockham theory for e, if there exists one. Then for fixed $e \in F_K$ the event:

$$[\mathcal{M}_e \in \mathsf{Ock}_{K,e}] = \bigcup_{A \in \mathsf{Ock}_{K,e}} [\mathcal{M}_e = A]$$

is in $\mathcal{D}$, so $p(\mathcal{M}_e \in \mathsf{Ock}_e)$ is defined, since E is countable and, hence, Th_K is countable.

Let $e \in F_K$ and let $s \in \mathsf{Spt}(X_{[e]})$. Say that $\mathcal{M}$ is *Ockham* at e' given $X_{[e]} = s$ if and only if:

$$p(\mathcal{M}_{e'} \in \mathsf{Ock}_{K,e'} \mid X_{[e]} = s) = 1.$$

An important fact is that the only way a theory can cease to be Ockham is to be refuted.[3]

Proposition 5 *Suppose that* $T \in \mathit{Ock}_{K,e_-} \setminus \mathit{Ock}_{K,e}$. *Then*

$$S_e \not\subseteq S_T.$$

Proof: Since $T \in \mathsf{Ock}_{K,e_-}$, it follows that

$$(*) \text{ For each } S \in K, \text{ if } S_{e_-} \subseteq S, \text{ then } S_{e_-} \subseteq S_T \subseteq S$$

Because $T \notin \mathsf{Ock}_{K,e}$, it follows that one of three options hold:

1. $S_e \not\subseteq S_T$,
2. S_T is not uniquely minimal in K_e. That is, there exist at least two distinct sets $S', S'' \in K_e$ such that neither $S' \subseteq S''$ nor $S'' \subseteq S'$, and for any $S \in K_e$, either $S' \subseteq S$ or $S'' \subseteq S$
3. There exists an $S' \in K$ such that $S_e \subseteq S' \subset S_T$.

If (3) were true, then because $S_{e_-} \subseteq S_e$, we'd have $S_e \subseteq S' \subset S_T$, contradicting (*). If (2) were true, then again because $S_{e_-} \subseteq S_e$, there are two distinct sets $S, S' \in K_e$ such that $S_{e_-} \subseteq S', S''$ and for any $S \in K$ containing S_{e_-}, either $S' \subseteq S$ or $S'' \subseteq S$. By $(*)$, we obtain that $S_{e_-} \subseteq S_T \subseteq S', S''$, and by the assumption of (2), we obtain that either $S' \subseteq S_T$ or $S'' \subseteq S_T$. Hence, either $S_T = S'$ or $S_T = S''$. Without loss of generality, assume that $S_T = S'$. Now as $S_T = S'$ is Ockham at e_-, by $(*)$ it follows that $S_{e_-} \subseteq S' \subseteq S''$, contradicting the fact the assumption that $S' \not\subseteq S''$. Hence, (1) must hold, as desired. $\Box$

In many applications, the problem K has additional structure that yields a stronger efficiency argument for Ockham methods (see Theorem 2). Let $l(K)$

[3]This assumption fails in some interesting applications.

be the length of the longest path in K (where $l(K) = \omega$ if the longest path is infinite). Say that path q in K is *traversing* if and only if $l(q) = n_K$. Clearly, if $n \leq n_K$, then $C_{K,n}(e) \neq \emptyset$. Say that K has *no short paths* if and only if for every $e \in F_K$, and for every *minimal* $S \in K_e$, there exists a traversing path in K_e beginning with S. For example, if K is linearly ordered by $\subset$ (as in the curve-fitting problem), then K has no short paths. Another, plausible condition for no short paths is that for each $S, S' \in K$ there exists $S'' \in K$ such that $S, S' \subseteq S''$. It is immediate that if K has no short paths, then for each $e \in F_K$, K_e has no short paths.

Proposition 6 *Suppose K has no short paths, and let $e \in F_K$. Suppose that $T \notin \mathsf{Ock}_{K,e}$. Then there exists a traversing path q in K_e such that $S_T \neq q_0$.*

Proof: If $T \notin \mathsf{Ock}_{K,e}$, then again one of the three following options hold:

1. $S_e \nsubseteq S_T$
2. S_T is not uniquely minimal in K_e. That is, there exist at least two distinct sets $S', S'' \in K_e$ such that neither $S' \subseteq S''$ nor $S'' \subseteq S'$, and for any $S \in K_e$, either $S' \subseteq S$ or $S'' \subseteq S$.
3. There exists a minimal $S' \in K_e$ such that $S_e \subseteq S' \subset S_T$.

If (1) holds, then any traversing path in K_e is as desired. To show that such a traversing path exists, note that $C_{K,0}(e) \neq \emptyset$ by Proposition 4, and then apply the no short paths assumption. If (2), then because $S' \neq S''$, either S' or S'' is distinct from S_T. Without loss of generality, assume that $S_T \neq S'$. Because K has no short paths and S' is minimal, there is a traversing path in K_e beginning with $S' \neq S_T$, and we're done. If (3) holds, the there exists an $S' \in K$ such that $S_e \subseteq S' \subset S_T$. Because K has no short paths and S' is minimal, there is a traversing path in K_e beginning with $S' \neq S_T$ as desired. □

3.4 Stalwartness

Stalwartness complements Ockham's razor. Ockham's razor proscribes answers other than the uniquely simplest. Stalwartness insists that one hang on to an Ockham answer until it is dethroned by further data. In the deterministic case, that is easy to define: don't drop you previous answer until it fails to be uniquely simplest. The statistical generalization of this idea is intuitive: if you ever have a chance of producing an answer, produce it with unit chance conditional on having just produced it.

Let $e \in F_K$ and let $s \in \mathsf{Spt}(X_{[e]})$ and let $e' \in F_K$. Say that $\mathcal{M}$ is *stalwart* for T at e' given $X_{[e]} = s$ if and only if:

$$l(e') > 0 \text{ and } p(\mathcal{M}_{e'_-} = T \mid X_{[e]} = s) > 0 \text{ and } T \in \mathsf{Ock}_{K,e'}$$

implies:

$$p(\mathcal{M}_{e'} = T \mid \mathcal{M}_{e'_-} = T \wedge X_{[e]} = s) = 1.$$

3.5 Statistical Consistency (Convergence to the Truth)

In statistical usage, a *consistent* method is a method that converges in probability to the truth. Let $e \in F_K$ and let $s \in \mathsf{Spt}(X_{[e]})$ and let $e' \in F_K$. Say that $\mathcal{M}$ is *consistent* over K given $X_{[e]} = s$ if and only if:

$$\lim_{i \to \infty} p(\mathcal{M}_{w|i} = T_w \mid X_{[e]} = s) = 1, \quad \text{for each } w \in W_{K,e}.$$

Consistency can be expressed as:

$$\Phi_K(\mathcal{M} \mid X_{[e]}s),$$

which does not depend on the additional argument e'.

3.6 Eventual Informativeness

Under the usual restrictions, say that $\mathcal{M}$ is *eventually informative* over K given $X_{[e]} = s$ if and only if:

$$\lim_{i \to \infty} p(\mathcal{M}_{w|i}\text{`?'} \mid X_{[e]} = s) = 0, \quad \text{for all } w \in W_{K,e}.$$

Eventual informativeness is entailed by consistency and implies that $\mathcal{M}$ cannot keep producing '?' infinitely often with non-vanishing probability.

3.7 Property Stability

Let $\Phi_K(\mathcal{M}, e', e, X, s)$ be a methodological property. Say that methodological property $\Phi_K(\mathcal{M}, e' \mid X_{[e]} = s)$ is *stable* just in case for each $e'' \in F_{K,e}$, the following conditions are equivalent:

a. $\Phi_K(\mathcal{M}, e' \mid X_{[e]} = s)$;

b. $\Phi_K(\mathcal{M}, e', \mid X_{[e'']} = s'')$, for each $s'' \in \mathsf{Spt}(X_{[e'']})$.

The methodological properties introduced above are all stable.

Proposition 7 *The following properties are stable:*

1. *$\mathcal{M}$ is statistically consistent given $X_{[e]} = s$,*
2. *$\mathcal{M}$ is eventually informative given $X_{[e]} = s$,*
3. *$\mathcal{M}$ is logically consistent at e' given $X_{[e]} = s$,*
4. *$\mathcal{M}$ is Ockham at e' given $X_{[e]} = s$,*
5. *$\mathcal{M}$ is stalwart at e' given $X_{[e]} = s$.*

Proof of (1). By total probability:

$$\begin{aligned} & p(\mathcal{M}_{w|i} \neq T_w \mid X_{[e]} = s) \\ = & \sum_{s'' \in \mathsf{Spt}(X_{[e'']} \mid X_{[e]} = s)} p(\mathcal{M}_{w|i} \neq T_w \mid X_{[e'']} = s'') \cdot p(X_{[e'']} = s'' \mid X = s). \end{aligned}$$

The second factor under the summation is non-zero for each $s' \in \mathsf{Spt}(X_{[e'']} \mid X_{[e]} = s)$. So, since the second factor does not depend on i, the statement:

$$\lim_{i \to \infty} p(\mathcal{M}_{w|i} \neq T_w \mid X_{[e]} = s) = 0$$

holds if and only if for each $s'' \in \mathsf{Spt}(X_{[e'']} \mid X_{[e]} = s)$:

$$\lim_{n \to \infty} p(\mathcal{M}_{w|n} \neq T_w \mid X_{[e'']} = s'') = 0.$$

Proof of (2). Similar to proof of (1).

Proof of (3 and 4). Immediate consequences of proposition 3.

Proof of (5). The implication from (b) to (a) is immediate, since (a) is a substitution instance of (b). For the implication from (a) to (b), Suppose that $\mathcal{M}$ is stalwart for $T \in \mathsf{Th}_K$ at e' given $X_{[e]} = s$. Let $e'' \in F_K$ and let $s'' \in \mathsf{Spt}(X_{[e'']})$ and suppose that:

$$l(e') > 0 \text{ and } p(\mathcal{M}_{e'_{\_}} = T \mid X_{[e'']} = s'') > 0 \text{ and } T \in \mathsf{Ock}_{e'}.$$

Then, by proposition 3:

$$p(\mathcal{M}_{e'_{\_}} = T \mid X_{[e]} = s) > 0 \text{ and } T \in \mathsf{Ock}_{e'} \text{ and } l(e') > 0.$$

So, since $\mathcal{M}$ is stalwart for T at e' given $X_{[e]} = s$:

$$p(\mathcal{M}_{e'} = T \mid \mathcal{M}_{e'_{\_}} = T \wedge X_{[e]} = s) = 1.$$

So, by proposition 3 again:

$$p(\mathcal{M}_{e'} = T \mid \mathcal{M}_{e'_{\_}} = T \wedge X_{[e'']} = s'') = 1.$$

□

3.8 Consistency of Eventually Informative Ockham Methods

One reason to use an Ockham method is that if such a method is not too skeptical, it is guaranteed to converge to the truth.

Proposition 8 *Suppose that $\mathcal{M}$ is both henceforth Ockham and eventually informative given $X_{[e]} = s$. Then $\mathcal{M}$ is consistent over K given $X_{[e]} = s$.*

Proof. Suppose $\mathcal{M}$ Ockham and eventually informative over K given given $X_{[e]} = s$. Let $w \in W_{K,e}$. Then

$$S_e \subseteq S_w = \bigcup_{i \in \omega} S_{w|i},$$

so since S_w is finite, there exists $n \geq l(e)$ such that $S_w S_{w|n}$. Then for each $m \geq n$, (S_w) is a unit path in K_e to S_w and (S_w) is the longest such path in K_e. Hence, $c_{K,e}(S_w) = 0$, so $c_{K,w|m}(T_w) = 0$. Furthermore, for each $S \in K_e$ such that $S \neq S_w$, we have $S_{w|m} \neq S$, so there is a path $(S_{w|m}, S)$ in K_e. Hence, $c_{K,w|m}(T_S) > 0$. Hence, for each $m \geq n$, $\mathsf{Ock}_{K,w|m}\{T_w, \text{‘?’}\}$. Since $\mathcal{M}$ is Ockham given $X_{[e]}s$ and $n \geq l(e)$, we have that for each $m \geq n$:

$$p(\mathcal{M}_{w|m} \in \{T_w, \text{‘?’}\} \mid X_{[e]} = s) = 1.$$

Since $\mathcal{M}$ is informative given $X_{[e]} = s$:

$$\lim_{i \to \infty} p(\mathcal{M}_{w|i} \neq \text{‘?’} \mid X_{[e]} = s) = 1.$$

Thus:

$$\lim_{i \to \infty} p(\mathcal{M}_{w|i} = T_w \mid X_{[e]} = s) = 1.$$

□

But that is just one way to converge to the truth. The question remains why converging to the truth the Ockham way is better than doing so in any other way: e.g., by guessing a complex theory for a thousand stages and reverting to an Ockham strategy thereafter. That one reverts to the Ockham strategy *later* is of little help in explaining why the Ockham theory is the right one to believe *now*. That is the nub of the puzzle of simplicity.

4 Efficiency of Empirical Inquiry

4.1 Loss Functions

A *loss function* is a mapping:

$$\lambda : \mathsf{Ans}^{\omega} \times W_K \to \Re.$$

A *local loss function* is a mapping:

$$\gamma : \mathsf{Ans}^{<\omega} \times W_K \to \Re.$$

In the intended applications, $\gamma(c)$ reflects cost incurred only at the moment when the last entry in c is produced. If γ, γ' are local loss functions, say that $\gamma \leq \gamma'$ just in case for each $c \in \mathsf{Ans}^{<\omega}$ and each $w \in W_K$,

$$\gamma(c, w) \leq \gamma'(c, w).$$

4.2 Errors

Let $c \in \mathsf{Ans}^{<\omega}$. Define:

$$\mathsf{Err}(c, w) \quad \text{iff} \quad w \notin c(l(c) - 1) \in \mathsf{Th}_K.$$

Then say that an *error* occurs in c at stage i with respect to world w (recall that at stage i answer $c(i-1)$ has just been produced). Only theories can be in error—to output answer '?' is to be immune from error. We (crudely) charge one unit per error, regardless of its severity, which corresponds to the following, local loss function:

$$\epsilon(c, w) \quad = \quad \begin{cases} 1 & \text{if } \mathsf{Err}(c, w); \\ 0 & \text{otherwise;} \end{cases}$$

Our argument continues to work, however, when the cost depends on the theory produced as long as the cost of producing a theory in error is invariant over worlds in which the theory is false.[4]

4.3 Retractions

Let $c \in \mathsf{Ans}^{\omega}$. Define:

$$\mathsf{Ret}(c) \quad \text{iff} \quad l(c) \geq 2 \text{ and } c(l(c) - 1) \neq c(l(c) - 2) \in \mathsf{Th}_K.$$

Then say that a *retraction* occurs in c at stage i. Only theories are retracted: changing one's mind from '?' to some substantive theory does not count as a retraction. Again, merely for simplicity, we charge one unit per error, which corresponds to the following, local loss function:

$$\rho(c, w) = \rho(c) = \begin{cases} 1 & \text{if } \mathsf{Ret}(c); \\ 0 & \text{otherwise;} \end{cases}$$

4.4 Cumulative Loss

If γ is a local loss function and $w \in W_K$ and $\beta \leq \omega$, define the *cumulative* loss functions induced by γ as follows:

$$\gamma(c, w)[^{\beta}_{m} \quad = \quad \sum_{i=m}^{\beta} \gamma(c|i, w).$$

In particular, we consider as losses both cumulative errors $\epsilon(c, w)[^{n}_{m}$ and cumulative retractions $\rho(c, w)[^{n}_{m}$. It is convenient to adopt the following, obvious

[4]That is the case, for example, in the epistemic utility theories of I. Levi (***) and C. Hempel (***), according to which the loss of an error depends only on the content of the false theory. For Levi, more false content is better than less; for Hempel, less false content is better than more.

notations:

$$\begin{aligned}
\gamma(c,w)^n &= \gamma(c,w)[_n^n;\\
\gamma(c,w)^{<n} &= \gamma(c,w)[_0^{n-1};\\
\gamma(c,w)^{\leq n} &= \gamma(c,w)[_0^n;\\
\gamma(c,w)^{>n} &= \gamma(c,w)[_{n+1}^\omega;\\
\gamma(c,w)^{\geq n} &= \gamma(c,w)[_n^\omega.
\end{aligned}$$

4.5 Lag Time to Accumulated Loss

An Ockham efficiency theorem can be obtained for cumulative errors and cumulative retractions alone, but a stronger Ockham efficiency theorem can be obtained if we consider as well the lag time to each retraction. It is not always the case that we prefer to incur costs earlier rather than later—many prefer to pay their debts as late as possible. But delaying retractions seems different. For one thing, belief in a theory leads to further, subsidiary inferences from it just as a cancerous tumor spreads malignancy to other organs—best to nip the malignancy in the bud. Also, in the case of paying debts later rather than earlier, one assumes by default that the indebtedness is known. Retractions, by their nature, can't be anticipated in advance, so to get them over as early as possible is more like having the creditor come to collect as soon as possible when you have forgotten the debt—living a sumptuous lie can be as financially dangerous as living a scientific lie is epistemically dangerous. If γ is a local loss function, define the *lag time* prior to aggregated cost $r \in \Re$ as follows:

$$\tau_{\gamma \geq r}(c,w) = \begin{cases} \min_i \gamma(c)[_0^i \ \geq r & \text{if there exists } i \text{ such that } \gamma(c,w)[_0^i \ \geq r;\\ 0 & \text{otherwise;}\end{cases}$$

For example, the elapsed time to the kth retraction is given by the loss function $\tau_{\rho \geq k}(c,w)$. Thus, the 0th retraction occurs at stage 0 and if there are at most k retractions, then since there is no $k+1$th retraction, the $k+1$th retraction occurs also at stage 0.

4.6 Measurability

Let γ be local loss function. The *random local loss* incurred by $\mathcal{M}$ in w at stage i is:

$$\gamma^i_{\mathcal{M},w}(\delta) = \gamma(\mathcal{M}_{[w|i]}(\delta), w).$$

Rewrite:

$$\gamma(c,w) = \gamma_w(c).$$

Thus:

$$\gamma^i_{\mathcal{M},w}(\delta) = (\gamma_w \circ \mathcal{M}_{[w|i]} \circ X_{[w|i]})(\delta).$$

For the remainder of the paper, let $\mathcal{B}$ denote the Borel algebra on $\Re$. Let $\overline{\Re} = \Re \cup \{-\infty, \infty\}$ be the real line extended by positive and negative infinite elements, and define an algebra $\overline{\mathcal{B}}$ on $\overline{\Re}$ such that $B^* \in \overline{\mathcal{B}}$ if and only if $B^* = B \cup S$ where $B \in \mathcal{B}$ and $S \in \{\emptyset, \{-\infty\}, \{\infty\}, \{-\infty, \infty\}\}$.

Proposition 9 *If γ is a local loss function, then $\gamma_{\mathcal{M},w,i}$ is $\mathcal{D}/\overline{\mathcal{B}}$-measurable.*

Proof. It is required show that the pre-image $\gamma^{-1}_{\mathcal{M},w,i}(B)$ of any Borel set B under $\gamma_{\mathcal{M},w,i}$ is $\mathcal{D}$-measurable. Accordingly, let $B \subseteq \Re$ be any Borel set and define:

$$S_B = \{s \in \Sigma^{i+1} : \gamma(s, w) \in B\}$$

Then S is countable since Σ is. Now:

$$\begin{aligned} [\gamma_{\mathcal{M},w,i} \in B] &= \{\delta \in \Delta : \gamma_{\mathcal{M},w,i}(\delta) \in B\} \\ &= \{\delta \in \Delta : \gamma(\mathcal{M}_{[w|i]}(\delta), w) \in B\} \\ &= \bigcup_{s \in S_B} \{\delta \in \Delta : X_{[w|i]}(\delta) = s\} \\ &= \bigcup_{s \in S_B} [X_{[w|i]} = s]. \end{aligned}$$

□

In the same spirit, define *random cumulative loss* as a function of δ, where $\beta \le \omega$:

$$\gamma_{\mathcal{M},w}[^{\beta}_{n}(\delta) \quad = \quad \gamma(\mathcal{M}_{[w]}(\delta), w)[^{\beta}_{n}.$$

Again, it is convenient to adopt the notation:

$$\begin{aligned} \gamma^{<n}_{\mathcal{M},w} &= \gamma_{\mathcal{M},w}[^{n-1}_{0}; \\ \gamma^{\le n}_{\mathcal{M},w} &= \gamma_{\mathcal{M},w}[^{n}_{0}; \\ \gamma^{>n}_{\mathcal{M},w} &= \gamma_{\mathcal{M},w}[^{\omega}_{n+1}; \\ \gamma^{\ge n}_{\mathcal{M},w} &= \gamma_{\mathcal{M},w}[^{\omega}_{n}. \end{aligned}$$

Proposition 10 *If γ is a local loss function and $\beta \le \omega$, then $\gamma_{\mathcal{M},w}[^{\beta}_{n}$ is $\mathcal{D}/\overline{\mathcal{B}}$-measurable.*

Proof: Note that if β is finite, then for any Borel set B, one can define:

$$S_B = \{s \in \Sigma^{\beta-n} : \sum_{i=n}^{\beta} \gamma(s(i), w) \in B\}$$

and complete the proof in an analogous manner to that of the previous proposition. Call this fact †. Now suppose $\beta = \omega$ is infinite. By the definition of $\gamma_{\mathcal{M},w}[^{\omega}_{n}$:

$$\begin{aligned} \gamma_{\mathcal{M},w}[^{\omega}_{n}(\delta) &= \sum_{i=n}^{\infty} \gamma_{\mathcal{M},w,i}(\delta) \\ &= \lim_{m \to \infty} \sum_{i=n}^{m} \gamma_{\mathcal{M},w,i}(\delta) \end{aligned}$$

Or equivalently one may write:

$$\gamma_{\mathcal{M},w}[_n^\omega = \sup_{m \in \mathbb{N}} \gamma_{\mathcal{M},w}[_n^m$$

By fact † above, each of the functions on the right-hand side of this equation is $\mathcal{D}/\overline{\mathcal{B}}$-measurable. Hence, because the supremum of a set of measurable functions is measurable, the function $\gamma_{\mathcal{M},w}[_n^\omega$ is $\mathcal{D}/\overline{\mathcal{B}}$-measurable. □

Note that the function $\gamma_{\mathcal{M},w}[_n^\omega$ may take on infinite values if the method $\mathcal{M}$ retracts or commits an infinite number of errors in a world w along some state history. Thus, the use of the extended real line $\overline{\Re}$ and the extended σ-algebra $\overline{\mathcal{B}}$ are critical. Finally, define:

$$\tau_{\mathcal{M},w}^{\gamma \geq r}(\delta) \quad = \quad \tau_{\gamma \geq r}(\mathcal{M}_{[w]}(\delta), w).$$

Proposition 11 *If γ is a local loss function, then $\tau_{\mathcal{M},w}^{\gamma \geq r}$ is $\mathcal{D}/\overline{\mathcal{B}}$-measurable.*

Proof. First, note that the for any local loss function γ, the range of Δ under $\gamma_{\mathcal{M},w,i}$ is countable by the following calculation:

$$\begin{aligned} |\mathrm{rng}(\gamma_{\mathcal{M},w,i})| &= |\{\gamma_{\mathcal{M},w,i}(\delta) \ : \ \delta \in \Delta\}| \\ &= |\{\gamma(\mathcal{M}_{[w|i]}(\delta), w) \ : \ \delta \in \Delta\}| \\ &\leq |\{\gamma(s, w) \ : \ s \in \Sigma^{i+1}\}| \\ &= \aleph_0 \end{aligned}$$

where the last equality holds because Σ^{i+1} is countable (as i is finite and Σ is countable). For any local loss function γ and any real number $r \in \Re$, then, it follows that:

$$[\gamma_{\mathcal{M},w}^{\geq 0} \geq r] = \bigcup_{r \leq r' \in \mathrm{rng}(\gamma_{\mathcal{M},w,i})} [\gamma_{\mathcal{M},w}^{\geq 0} = r']$$

is measurable, as all events on the right-hand side are measurable by the previous proposition and the union is countable because $\mathrm{rng}(\gamma_{\mathcal{M},w,i})$ is countable. By similar reasoning, the set $[\gamma_{\mathcal{M},w}^{\leq n} < r]$ is measurable for any $n \in \mathbb{N}$. Now by definition of τ:

$$[\tau_{\mathcal{M},w}^{\gamma \geq r} = i] = [\gamma_{\mathcal{M},w}^{< i} < r] \cap [\gamma_{\mathcal{M},w}^{\geq i} \geq r]$$

As both events on the right-hand side of the equation are measurable by the above remarks, we've proven that $[\tau_{\mathcal{M},w}^{\gamma \geq r} = i]$ is measurable for any $i \in \Re$. Note that, for any loss function γ, the range of $\tau_{\mathcal{M},w}^{\gamma \geq r}$ is contained in the natural numbers. Hence, for any measurable set $B^* \in \overline{\mathcal{B}}$, it follows that:

$$[\tau_{\mathcal{M},w}^{\gamma \geq r} \in B^*] = \bigcup_{r \in \mathbb{N} \cap B^*} [\tau_{\mathcal{M},w}^{\gamma \geq r} = r]$$

where the right-hand side is a countable union of $\mathcal{D}$-measurable events. So $\tau_{\mathcal{M},w}^{\gamma \geq r}$ is $\mathcal{D}/\overline{\mathcal{B}}$-measurable. □

4.7 Expected Loss

Since the functions just considered are $\mathcal{D}$-measurable, they all have (possibly infinite) expected values. In the case of expected cumulative losses:

Proposition 12 *Let $\gamma \in \{\epsilon, \rho\}$ and let $p(X_{[e]} = s) > 0$ and let $\beta \leq \omega$. Then:*

$$\mathit{Exp}(\gamma_{\mathcal{M},w}[_n^\beta \mid X_{[e]} = s) = \sum_{i=n}^{\beta} \mathit{Exp}(\gamma_{\mathcal{M},w,i} \mid X_{[e]} = s).$$

Proof. Case: $\beta = \omega$:

$$\begin{aligned}
\mathsf{Exp}(\gamma_{\mathcal{M},w}[_n^\beta \mid X_{[e]} = s) &= \int \gamma_{\mathcal{M},w}[_n^\beta \; dp(. \mid X_{[e]} = s) \\
&= \int \sum_{i=n}^{\infty} \gamma_{\mathcal{M},w,i} \; dp(. \mid X_{[e]} = s) \\
&= \int \lim_{k \to \infty} \sum_{i=n}^{k} \gamma_{\mathcal{M},w,i} \; dp(. \mid X_{[e]} = s) \\
&= \lim_{k \to \infty} \int \sum_{i=n}^{k} \gamma_{\mathcal{M},w,i} \; dp(. \mid X_{[e]} = s) \\
&= \lim_{k \to \infty} \sum_{i=n}^{k} \int \gamma_{\mathcal{M},w,i} \; dp(. \mid X_{[e]} = s) \\
&= \sum_{i=n}^{\infty} \int \gamma_{\mathcal{M},w,i} \; dp(. \mid X_{[e]} = s) \\
&= \sum_{i=n}^{\infty} \mathsf{Exp}(\gamma_{\mathcal{M},w,i} \mid X_{[e]} = s).
\end{aligned}$$

Pulling out the limit works by (Billingsly 1986) Theorem 15.1.iii, since the functions are non-negative. Pulling out the sum works by (Billingsly 1986) Theorem 15.1.iv, since again the functions are non-negative. Drop the limit in case $\beta < \omega$. $\square$

4.8 Certain Costs

Because the retractions of $\mathcal{M}$ along e_- given $X_{[e]} = s$ are certain and do not depend on w, they may be viewed as certain costs given $X_{[e]} = s$.

Proposition 13 *For each $w \in K_e$, for each non-empty $Q \subseteq W_{K,e}$*

$$\begin{aligned}
\mathit{Exp}(\rho_{\mathcal{M},w}^{<l(e)} \mid X_{[e]} = s) &= \sup_{w \in Q} \mathit{Exp}(\rho_{\mathcal{M},w}^{<l(e)} \mid X_{[e]} = s) \\
&= \rho(\mathcal{M}_{[e_-]}(s_-)).
\end{aligned}$$

Similarly, the times of retractions that occur along e_- are fixed forever after.

Proposition 14 *For each $w \in K_e$, for each non-empty $Q \subseteq W_{K,e}$:*

if $\rho(\mathcal{M}_{[e_-]}(s_-)) \geq m$, then:

$$\begin{aligned} \mathsf{Exp}(\tau^{\rho \geq m}_{\mathcal{M},w} \mid X_{[e]} = s) &= \sup_{w \in Q} \mathsf{Exp}(\tau^{\rho \geq m}_{\mathcal{M},w} \mid X_{[e]} = s) \\ &= \tau_{\rho \geq m}(\mathcal{M}_{[e_-]}(s_-)). \end{aligned}$$

Errors are uncertain costs because what counts as an error depends on the infinite future.

4.9 States of Inquiry

We wish to rank stochastic methods $\mathcal{M}, \mathcal{M}'$ in light of finite input history e, but it isn't that simple because $\mathcal{M}$ has, by that time, already traversed some state trajectory $s \in \mathsf{Spt}(X_{[e]})$ and $\mathcal{M}'$ has traversed some state trajectory $s' \in \mathsf{Spt}(X'_{[e]})$. The two state spaces could be entirely disjoint. Therefore, we will begin by ranking *states of inquiry* for K at e, by which we mean pairs $(\mathcal{M}, s)$ such that $\mathcal{M}$ is a stochastic empirical method for K and $s \in \mathsf{Spt}(X_{[e]})$. Let $\mathsf{Inq}_{K,e}$ denote the set of all states of inquiry for K at e.

4.10 Worst-case Cumulative Errors and Retractions

It remains to define some concrete rankings to put into Γ in the definition of efficiency. Let γ be a local loss function. A natural way to turn γ into a ranking on $\mathsf{Inq}_{K,e}$ is to compare worst-case loss in each world compatible with e:

$$\sup_{w \in W_{K,e}} \mathsf{Exp}(\gamma^{\geq 0}_{\mathcal{M},w} \mid X_{[e]} s) \leq \sup_{w \in W_{K,e}} \mathsf{Exp}(\gamma^{\geq 0}_{\mathcal{M},w} \mid X'_{[e]} = s').$$

The trouble with that, however, is that all the costs under consideration are unbounded over all of $W_{K,e}$, which would result in equivalence of all methods. On the other hand, some methods can achieve finite retraction bounds in each empirical complexity class, which explains both the reason why we consider retractions as a loss function and why we will consider rankings defined in terms of worst-case loss taken not over all of $W_{K,e}$, but over complexity classes $C_{K,e}(i)$, for $i \in \omega$. Accordingly, define:

$$(\mathcal{M}, s) \leq^{\gamma}_{K,e,n} (\mathcal{M}', s')$$

to hold if and only if:

$$\sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\gamma^{\geq 0}_{\mathcal{M},w} \mid X_{[e]} s) \leq \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\gamma^{\geq 0}_{\mathcal{M},w} \mid X'_{[e]} = s').$$

It is immediate from the definition that $\leq^{\gamma}_{K,e,n}$ is a pre-order (reflexive and transitive) over: $\mathsf{Inq}_{K,e}$.

4.11 A Ranking Based on Retraction Times

If a retraction is coming anyway, it seems better to get it over with earlier rather than later. It is not a general principle of decision theory that debts should be paid earlier rather than later. But it does seem natural to get retractions over with as soon as possible, because belief in a theory that one will later reject constitutes a kind of insouciance reflected in standard Gettier examples in which one believes "for the wrong reason".[5] That consideration is purely epistemic. Pragmatically, whenever a fundamental theory is retracted, all the subsidiary applications of the theory must also be re-examined and, in some cases, discarded. The longer one "lives a lie" by delaying an inevitable retraction, the longer these subsidiary applications have to accumulate before being "flushed". Of course, the theory to be retracted might be true (if the method is not a stalwart method) and in that case, retracting later would seem to be better, because, as Plato remarked long ago, true belief is as good as knowledge while one has it. But we will credit belief in the true theory by jointly considering cumulative errors as a cost as well.

A simple proposal is to compare the time of the first retraction by $\mathcal{M}$ with the time of the first retraction by $\mathcal{M}'$, and so forth, for the second retraction, third retraction, etc., so that if $\mathcal{M}'$ runs out of retractions before all the retractions by $\mathcal{M}$ have been "accounted for", we conclude that $\mathcal{M}$ is not at least as good as $\mathcal{M}'$ in terms of timed retractions. But that is wrong, intuitively. Consider an "honest" method $\mathcal{M}$ that behaves sensibly and a "renegade" method $\mathcal{M}'$ just like $\mathcal{M}$ except that it gratuitously "waffles" for the first forty stages between T_0 and '?', resulting in twenty gratuitous retractions. Since they happen right away, these retractions may all precede the first retraction of $\mathcal{M}$, so that all the retractions performed by $\mathcal{M}$ will be performed later by $\mathcal{M}$. It seems that $\mathcal{M}'$ is Pareto-dominated in terms of retraction times by $\mathcal{M}$, but that is not true if we compare retraction times by naively matching retractions between the two methods by the order in which they occur. Evidently, one could *delete* the first twenty retractions from the sequence of retractions performed by $\mathcal{M}$ and then retractions by $\mathcal{M}'$ would be no earlier than the corresponding retractions by $\mathcal{M}$. On the other hand, there is no way to delete retractions by $\mathcal{M}$ so that all of the retractions by $\mathcal{M}'$ occur no later than corresponding retractions by $\mathcal{M}$. This is like weak Pareto dominance with respect to the times of all the retractions, except for the deletion of retractions by $\mathcal{M}'$ prior to checking respective times of corresponding retractions.

Accordingly, define:

$$(\mathcal{M}, s) \leq^{\tau}_{K,e,n} (\mathcal{M}', s')$$

to hold if and only if for each $w \in C_{K,e}(n)$, there exist $w' \in C_{K,e}(n)$ and a local loss function $\gamma \leq \rho$ such that for each $j \leq \omega$:

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M},w} \mid X_{[e]} = s) \quad \leq \quad \mathsf{Exp}(\tau^{\gamma \geq j}_{\mathcal{M}',w'} \mid X'_{[e]} = s').$$

Again, we have a ranking of states of inquiry:

[5]Even tax withholding is desirable if, as in the case of retractions, one doesn't know when or even whether one's tax bill will come due.

Proposition 15 *$\leq^{\tau}_{K,e,n}$ is a pre-order over $\mathsf{Inq}_{K,e}$.*

Proof. Each relation $\leq^{\tau}_{K,e,i}$ is clearly reflexive by letting $\gamma = \rho$ and $w = w'$. It remains to be shown that each relation $\leq^{\tau}_{e,i}$ is transitive. Suppose that:

$$(\mathcal{M}_0, s_0) \leq^{\tau}_{e,n} (\mathcal{M}_1, s_1) \leq^{\tau}_{e,n} (\mathcal{M}_2, s_2).$$

Let X_i denote the state variable of $\mathcal{M}_i$ and Λ be the set of all cumulative loss functions. By the definition of $\leq^{\tau}_{e,n}$ and the axiom of choice, there are functions $f_{0,1}, f_{1,2} : W_K \to W_K$ and $g_{0,1}, g_{1,2} : W_K \to \Lambda$ such that $g_{0,1}(w), g_{1,2}(w) \leq \rho$, for all $w \in W_K$ and:

$$\begin{aligned}
\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M}_0,w} \mid X_{0_{[e]}} = s_0) &\leq \mathsf{Exp}(\tau^{g_{0,1}(w) \geq j}_{\mathcal{M}_1,f_{0,1}(w)} \mid X_{1_{[e]}} = s_1);\\
\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M}_1,w} \mid X_{1_{[e]}} = s_1) &\leq \mathsf{Exp}(\tau^{g_{1,2}(w) \geq j}_{\mathcal{M}_2,f_{1,2}(w)} \mid X_{2_{[e]}} = s_2).
\end{aligned}$$

Define:

$$\begin{aligned}
f_{0,2} &= f_{1,2} \circ f_{0,1};\\
g_{0,2} &= g_{1,2} \circ f_{0,1}.
\end{aligned}$$

By construction, it follows that:

$$\begin{aligned}
\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M}_0,w} \mid X_{0_{[e]}} = s_0) &\leq \mathsf{Exp}(\tau^{g_{0,1}(w) \geq j}_{\mathcal{M}_1,f_{0,1}(w)} \mid X_{1_{[e]}} = s_1)\\
&\leq \mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M}_1,f_{0,1}(w)} \mid X_{1_{[e]}} = s_1)\\
&\leq \mathsf{Exp}(\tau^{g_{1,2}(f_{0,1}(w)) \geq j}_{\mathcal{M}_2,f_{1,2}(f_{0,1}(w))} \mid X_{2_{[e]}} = s_2)\\
&\leq \mathsf{Exp}(\tau^{g_{0,2}(w)) \geq j}_{\mathcal{M}_2,f_{0,2}(w)} \mid X_{2_{[e]}} = s_2).
\end{aligned}$$

Thus, the functions $f_{0,2}$ and $g_{0,2}$ provide the witnesses to the inequality:

$$(\mathcal{M}_0, s_0) \leq^{\tau}_{e,n} (\mathcal{M}_2, s_2).$$

$\square$

4.12 Pareto-Rankings

It remains to assemble the various rankings under consideration into a single ranking. We do so in the least controversial way, by ordering two states of inquiry just in case all the individual rankings agree. That is known as the *Pareto* ranking. Think of $\gamma \in \{\rho, \epsilon, \tau\}$ as a formal parameter picking out relation $\leq^{\gamma}_{K,e,n}$. Let $\Gamma \subseteq \{\rho, \epsilon, \tau\}$. Then define:

$$\begin{aligned}
(\mathcal{M}, s) \leq^{\Gamma}_{K,e,n} (\mathcal{M}', s') &\quad \text{iff} \quad (\mathcal{M}, s) \leq^{\gamma}_{K,e,n} (\mathcal{M}', s'), \text{ for each } \gamma \in \Gamma;\\
(\mathcal{M}, s) <^{\Gamma}_{K,e,n} (\mathcal{M}', s') &\quad \text{iff} \quad (\mathcal{M}, s) \leq^{\Gamma}_{K,e,n} (\mathcal{M}', s') \text{ and } (\mathcal{M}', s') \not\leq^{\Gamma}_{K,e,n} (\mathcal{M}, s);\\
(\mathcal{M}, s) \ll^{\Gamma}_{K,e,n} (\mathcal{M}', s') &\quad \text{iff} \quad (\mathcal{M}, s) <^{\Gamma}_{K,e,n} (\mathcal{M}', s'), \text{ for each } \gamma \in \Gamma.
\end{aligned}$$

Now we go through another round of Pareto-combination, this time with respect to world complexity. First define the upper complexity bound for K_e as:

$$c_{K,e} = \sup\{i + 1 : i \in \omega \text{ and } C_{K,e}(\beta) \neq \emptyset\}.^6$$

Then define:

$$\begin{aligned}
(\mathcal{M}, s) \leq^{\Gamma}_{K,e} (\mathcal{M}', s') &\quad \text{iff} \quad (\mathcal{M}, s) \leq^{\gamma}_{K,e} (\mathcal{M}', s'), \text{ for each } n \in \omega; \\
(\mathcal{M}, s) <^{\Gamma}_{K,e} (\mathcal{M}', s') &\quad \text{iff} \quad (\mathcal{M}, s) \leq^{\Gamma}_{K,e,n} (\mathcal{M}', s') \text{ and } (\mathcal{M}', s') \not\leq^{\Gamma}_{K,e} (\mathcal{M}, s); \\
(\mathcal{M}, s) \prec^{\Gamma}_{K,e} (\mathcal{M}', s') &\quad \text{iff} \quad (\mathcal{M}, s) <^{\Gamma}_{K,e,n} (\mathcal{M}', s'), \text{ for each } n < c_{K,e}; \\
(\mathcal{M}, s) \ll^{\Gamma}_{K,e} (\mathcal{M}', s') &\quad \text{iff} \quad (\mathcal{M}, s) \ll^{\Gamma}_{K,e,n} (\mathcal{M}', s'), \text{ for each } n < c_{K,e}.
\end{aligned}$$

Then $<^{\Gamma}_{e}$ is *weak* Pareto-dominance and $\ll^{\Gamma}_{e}$ is *strong* Pareto-dominance. The relation $\prec^{\Gamma}_{e}$ falls in between, requiring weak dominance over Γ over each non-empty complexity class $C_{K,e}(n)$, and will be called *strong complexity dominance.* Unwinding the definitions of the preceding relations yields a clearer picture of what is involved in each.

Proposition 16

$$\begin{aligned}
(\mathcal{M}, s) \leq^{\Gamma}_{K,e} (\mathcal{M}', s') &\quad \textit{iff} \quad (\forall n \in \omega)(\forall \gamma \in \Gamma)\ (\mathcal{M}, s) \leq^{\gamma}_{K,e,n} (\mathcal{M}', s'); \\
(\mathcal{M}, s) <^{\Gamma}_{K,e} (\mathcal{M}', s') &\quad \textit{iff} \quad (\forall n \in \omega)(\forall \gamma \in \Gamma)\ (\mathcal{M}, s) \leq^{\gamma}_{K,e,n} (\mathcal{M}', s') \textit{ and} \\
&\qquad\quad (\exists n \in \omega)(\exists \gamma \in \Gamma)\ (\mathcal{M}, s) <^{\gamma}_{K,e,n} (\mathcal{M}', s'); \\
(\mathcal{M}, s) \prec^{\Gamma}_{K,e} (\mathcal{M}', s') &\quad \textit{iff} \quad (\forall n < c_{K,e})(\forall \gamma \in \Gamma)\ (\mathcal{M}, s) \leq^{\gamma}_{K,e,n} (\mathcal{M}', s') \textit{ and} \\
&\qquad\quad (\forall n < c_{K,e})(\exists \gamma \in \Gamma)\ (\mathcal{M}, s) <^{\gamma}_{K,e,n} (\mathcal{M}', s'); \\
(\mathcal{M}, s) \ll^{\Gamma}_{K,e} (\mathcal{M}', s') &\quad \textit{iff} \quad (\forall n < c_{K,e})(\forall \gamma \in \Gamma)\ (\mathcal{M}, s) <^{\gamma}_{K,e,n} (\mathcal{M}', s').
\end{aligned}$$

4.13 Switching Methods in Midstream

Let $\mathcal{M}$ be a stochastic method for K. Suppose that one has been using $\mathcal{M}'$ and the current, finite input sequence is e. Given that $X_{[e]} = s$, where $e > ()$, the past outputs of $\mathcal{M}$ along e_- cannot be changed, so one is stuck with the output sequence $c = \mathcal{M}_{[e_-]}(s)$ and with the cumulative loss $\gamma(c)[_0^{l(c)-1}$. Now consider alternative stochastic method $\mathcal{M}'$ with state variables $\{X'_e : e \in F_K\}$. Given that $X'_{[e]} = s'$, one has the option to switch methods from $\mathcal{M}$ to $\mathcal{M}'$ with state history s' from e onward. But one is still stuck with the costs from having used $\mathcal{M}$ (s' is relevant only insofar as s' affects the future performance of $\mathcal{M}'$). So, when switching from $\mathcal{M}$ to $\mathcal{M}'$ at e, one must consider not the overall resource consumption of $\mathcal{M}'$ given $X_{[e']} = s'$, but the cost of $\mathcal{M}'$ given $X_{[e']} = s'$ from $l(e)$ onward, added to the resource consumption of $\mathcal{M}$ given $X_{[e]} = s$ along e_-. It is convenient to conceive of the switch from $\mathcal{M}$ to $\mathcal{M}'$ at e as having *always*

[6] Adding 1 makes the bound strict both in the case of finitely bounded and finitely unbounded orders of complexity.

followed hybrid method $\mathcal{M} \star_e^s \mathcal{M}'$, which acts like $\mathcal{M}$ given $X_{[e]} = s$ along e_- and like $\mathcal{M}'$ thereafter. That is readily accomplished simply by modifying the output function α' of $\mathcal{M}'$. Define the hybrid output function:

$$(\alpha \star_e^s \alpha')(e', \sigma) = \begin{cases} \alpha(e', s(l(e'))) & \text{if } e' < e; \\ \alpha'(e', \sigma) & \text{otherwise.} \end{cases}$$

Then define the hybrid method:

$$\mathcal{M} \star_e^s \mathcal{M}'$$

to be the result of replacing α' with $(\alpha \star_e^s \alpha')$ in $\mathcal{M}'$.

4.14 Efficiency

Say that $\mathcal{M}$ *is* Γ*-efficient* for K given $X_{[e]} = s$ if and only if for each $(\mathcal{M}', s') \in \mathsf{Inq}_{K,e}$,

$$\text{if } \mathcal{M}' \text{ is consistent given } X'_{[e]} = s' \text{ then } (\mathcal{M}, s) \leq_e^\Gamma (\mathcal{M} \star_e^s \mathcal{M}',\ s').$$

Next, say that $\mathcal{M}$ *is weakly* Γ*-dominated* given $X_{[e]} = s$ if and only if there exists $(\mathcal{M}', s') \in \mathsf{Inq}_{K,e}$ such that:

$$\mathcal{M}' \text{ is consistent from } e \text{ given } X'_{[e]} s' \text{ and } (\mathcal{M} \star_e^s \mathcal{M}',\ s') <_e^\Gamma (\mathcal{M}, s).$$

Finally, say that $\mathcal{M}$ *is strongly* Γ*-dominated* given $X_{[e]} = s$ if and only if there exists $(\mathcal{M}', s') \in \mathsf{Inq}_{K,e}$ such that:

$$\mathcal{M}' \text{ is consistent from } e \text{ given } X'_{[e]} = s' \text{ and } (\mathcal{M} \star \mathcal{M}',\ s') \ll_e^\Gamma (\mathcal{M}, s).$$

Note that these concepts are relative to $\mathcal{M}$ and e and that such a property holds *perfectly* just in case it holds at every $e \in F_K$. Thus, one may speak of perfect Γ-efficiency, perfect non-Γ dominance and perfect non-Γ-strict-dominance. It is obvious that weak Γ-dominance at some e implies lack of perfect Γ-efficiency, but it is a strong and surprising feature of the Ockham efficiency theorems that follow that Γ-inefficiency at some e implies at least weak Γ-dominance at e and, in an interesting class of cases, strong Γ-dominance at e.

4.15 General Ockham Efficiency Theorem

Theorem 1 (Ockham Efficiency Characterization) *Assume that $\mathcal{M}$ is consistent and that:*

$$\begin{aligned} \{\epsilon, \rho\} &\subseteq \Gamma \subseteq \{\rho, \epsilon, \tau\} \text{ or} \\ \{\tau\} &\subseteq \Gamma \subseteq \{\rho, \epsilon, \tau\}. \end{aligned}$$

Then following are equivalent:

1. *$\mathcal{M}$ is always Ockham and stalwart.*

2. *$\mathcal{M}$ is perfectly Γ-efficient.*

3. *$\mathcal{M}$ is perfectly weakly Γ-undominated.*

It is immediate from the definitions that weak Γ-dominance at some e implies Γ^{Λ}-inefficiency at that e, but the converse, implied by the preceding theorem, is not at all trivial: it holds only because of the asymptotic character of the costs considered and because of nature's ability to force one arbitrarily late retraction for each degree of empirical complexity from an arbitrary, consistent method. Thus, the consistent methods are neatly partitioned into the efficient, stalwart, Ockham ones and the weakly dominated ones.

Proof of theorem 1. **Proof of 1 $\Rightarrow$ 2.** Let $\mathcal{M}$ be always consistent, Ockham, and stalwart. That means that $\mathcal{M}$ is henceforth Ockham, stalwart, and consistent given $X_{[()]} = (\sigma_0)$. Let $e \in F_K$ be of length k and let $s \in \mathsf{Spt}(X_{[e]})$. So by proposition 7, $\mathcal{M}$ is henceforth Ockham, stalwart, and consistent given $X_{[e]} = s$. Now, let $\mathcal{M}'$ have state variable X'_e, let $s' \in \mathsf{Spt}(X'_{[e]})$, and let $\mathcal{M}'$ be consistent given $X'_{[e]} = s'$. It must be shown, for each $\gamma \in \{\epsilon, \rho, \tau\}$, for each $s' \in \mathsf{Spt}(X'_{[e]})$ such that $\mathcal{M}'$ is consistent given $X'_{[e]} = s'$, and for each $n \in \omega$ that:

$$(\mathcal{M}, s) \leq^{\gamma}_{e,n} ((\mathcal{M} \star^s_e \mathcal{M}'), s').$$

Suppose that $\gamma \in \{\epsilon, \rho\}$. Since Since $\epsilon^{>0}_{\mathcal{M},w} = 0$, if $C_{K,e}(n) \neq \emptyset$, by 18, it must be shown that for each $s' \in \mathsf{Spt}(X'_{[e]})$ such that $\mathcal{M}'$ is consistent given $X'_{[e]} = s'$, for each n such that $C_{K,e}(n) \neq \emptyset$:

$$\sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\gamma^{\geq 0}_{\mathcal{M},w} \mid X_{[e]} = s) \leq \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\gamma^{\geq 0}_{(\mathcal{M}\star^s_e\mathcal{M}'),w} \mid X'_{[e]} = s').$$

Accordingly, let $s' \in \mathsf{Spt}(X'_{[e]})$ and suppose that $\mathcal{M}'$ is consistent given $X'_{[e]} = s'$. Then $\mathcal{M} \star^s_e \mathcal{M}'$ is consistent given $X'_{[e]} = s'$.

Case A: $\gamma = \epsilon$. Define:

$$a(n) = \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\epsilon^{<k}_{\mathcal{M},w} \mid X_{[e]} = s) = \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\epsilon^{<k}_{\mathcal{M}\star^s_e\mathcal{M}',w} \mid X_{[e]} = s).$$

Then Propositions 19 and 20 yield the following bounds, keeping in mind that the upper and lower bounds in the $i > k$ column collapse to zero if $C_{K,e}(n) = \emptyset$.

Errors, $n = 0$		$i < k$	$i = k$	$i > k$
$\sup_{w \in C_{K,e}(0)} \sum_i \mathsf{Exp}(\epsilon^i_{(\mathcal{M}\star^s_e\mathcal{M}'),w} \mid X'_{[e]} = s')$	$\geq$	$a(0)$	0	0
$\sup_{w \in C_{K,e}(0)} \sum_i \mathsf{Exp}(\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s)$	$\leq$	$a(0)$	0	0
Errors, $n > 0$				
$\sup_{w \in C_{K,e}(n)} \sum_i \mathsf{Exp}(\epsilon^i_{(\mathcal{M}\star^s_e\mathcal{M}'),w} \mid X'_{[e]} = s')$	$\geq$	$a(n)$	0	ω
$\sup_{w \in C_{K,e}(n)} \sum_i \mathsf{Exp}(\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s)$	$\leq$	$a(n)$	0	ω

Thus:

$$\begin{aligned} a(0) &\leq \sup_{w \in C_{K,e}(0)} \mathsf{Exp}(\epsilon^i_{(\mathcal{M}\star^s_e\mathcal{M}'),w} \mid X'_{[e]} = s'); \\ a(0) &\geq \sup_{w \in C_{K,e}(0)} \mathsf{Exp}(\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s), \end{aligned}$$

and if $n > 0$ and $C_{K,e}(n) \neq \emptyset$, then:

$$\begin{array}{rcl}\omega & \leq & \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\epsilon^i_{(\mathcal{M} \star^s_e \mathcal{M}'),w} \mid X'_{[e]} = s'); \\ \omega & \geq & \sup_{w \in C_{K,e}(0)} \mathsf{Exp}(\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s).\end{array}$$

Hence:

$$(\mathcal{M}, s) \leq^{\epsilon}_{K,e} ((\mathcal{M} \star^s_e \mathcal{M}'), s').$$

Case B: $\gamma = \rho$. Define:

$$b = \rho(\mathcal{M}_{[e_-]}(s_-)) = \rho((\mathcal{M} \star^s_e \mathcal{M}')_{[e_-]}(s'_-)).$$

Case B.1: Suppose that $\sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\rho^k_{\mathcal{M},w} \mid X_{[e]} = s) = 0$.

Then, by Propositions 13, 30, 19 and 17, we have the following bounds which imply, by proposition 29, that $(\mathcal{M}, s) \leq^{\rho}_{K,e} ((\mathcal{M} \star^s_e \mathcal{M}'), s')$.

Retractions	$i < k$	$i = k$	$i > k$	total
$\sup_{w \in C_{K,e}(n)} \sum_i \mathsf{Exp}(\rho^i_{(\mathcal{M} \star^s_e \mathcal{M}'),w} \mid X_{[e]} = s) \quad \geq$	b	0	n	$b+n$
$\sup_{w \in C_{K,e}(n)} \sum_i \mathsf{Exp}(\rho^i_{\mathcal{M},w} \mid X_{[e]} = s) \quad \leq$	b	0	n	$b+n$

Case B.2: Suppose that $\sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\rho^k_{\mathcal{M},w} \mid X_{[e]} = s) > 0$.

Then, by proposition 26, there exists $T \in \mathsf{Th}_K$ such that:

$$p(\mathcal{M}_e \neq \mathcal{M}_{e_-} = T \mid X_{[e]} = s) > 0.$$

Since $X_{[e]} = s$ settles the outputs of $\mathcal{M}_{e_-}, \mathcal{M}_e$,

$$\begin{array}{rcl} 1 & = & p(\mathcal{M}_{e_-} = T \mid X_{[e]} = s); \\ 0 & = & p(\mathcal{M}_e = T \mid X_{[e]} = s).\end{array}$$

Hence:

$$0 \quad = \quad p(\mathcal{M}_e = T \mid \mathcal{M}_{e_-} = T \wedge X_{[e]} = s).$$

So, since $\mathcal{M}$ is stalwart from e onward:

$$T \notin \mathsf{Ock}_{K,e},$$

† and, since $\mathcal{M}$ is always Ockham and $p(\mathcal{M}_{e_-} = T \mid X_{[e]} = s) = 1$:

$$T \in \mathsf{Ock}_{K,e_-}$$

So by proposition 5,

$$S_e \not\subseteq S_T.$$

Hence, for each path $(S_0, \ldots, S_n)$ in K_e, we have that $S_T \neq S_0$. So by propositions 27, 21, and 19 we have the following bounds which, by proposition 29, suffice to establish that $(\mathcal{M}, s) \leq^{\rho}_{K,e} ((\mathcal{M} \star^s_e \mathcal{M}'), s')$.

Retractions		$i < k$	$i = k$	$i > k$	total
$\sup_{w \in C_{K,e}(n)} \sum_i \mathsf{Exp}(\rho^i_{(\mathcal{M} \star^s_e \mathcal{M}'),w} \mid X_{[e]} = s)$	$\geq$	b	0	$n+1$	$b+n+1$
$\sup_{w \in C_{K,e}(n)} \sum_i \mathsf{Exp}(\rho^i_{\mathcal{M},w} \mid X_{[e]} = s)$	$\leq$	b	1	n	$b+n+1$

Case C: $\gamma = \tau$. To establish that $(\mathcal{M}, s) \leq^{\tau}_{K,e} ((\mathcal{M} \star^s_e \mathcal{M}'), s')$, it must be shown that for each $n \in \omega$ and for each $w \in C_{K,e}(n)$ there exists local loss function$\gamma \leq \rho$ and $w' \in C_{K,e}(n)$ such that for each $j \leq \omega$:

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M},w} \mid X_{[e]} = s) \quad \leq \quad \mathsf{Exp}(\tau^{\gamma \geq j}_{\mathcal{M}',w'} \mid X'_{[e]} = s').$$

Set $a = \rho(\mathcal{M}_{[e]}(s))$. Let $w \in C_{K,e}(n)$.

Case C.1: suppose that $j \leq a$. Then $\mathcal{M}_{[e]}(s) = (\mathcal{M} \star^s_e \mathcal{M}')_{[e]}(s')$, so the first a retractions of $\mathcal{M}$ occur along w no later than those of $\mathcal{M} \star^s_e \mathcal{M}'$. Since these retractions occur with unit chance by the end of e:

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M},w} \mid X_{[e]} = s) \quad = \quad \mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M}',w'} \mid X'_{[e]} = s').$$

Case C.2: suppose that $j > a$. Set $k = \mathsf{Exp}(\tau^{\gamma \geq j}_{\mathcal{M}',w'} \mid X'_{[e]} = s')$. Set $j' = j - a$. Apply proposition 24 to obtain $w', \gamma \leq \rho$ such that:

$$\mathsf{Exp}(\tau^{\rho \geq a+j'}_{\mathcal{M},w} \mid X_{[e]} = s) \geq k.$$

In light of cases A-C, it has been shown that when $\Gamma \subseteq \{\epsilon, \rho, \tau\}$:

$$(\mathcal{M}, s) \leq^{\Gamma}_{K,e,i} ((\mathcal{M} \star^e_s \mathcal{M}'), s').$$

Proof of 2 $\Rightarrow$ 3. Weak dominance implies inefficiency immediately.

Proof of 3 $\Rightarrow$ 1. Let $\mathcal{M}$ be an always consistent method that is not always Ockham and stalwart. Then there exists an $e \in F_K$ such that $\mathcal{M}$ is not Ockham and stalwart at e given $X_{[()]} = (\sigma_0)$ and no $e' \in F_K$ such that $e' < e$ has that property. Let $k = l(e)$. By proposition 7, there exists $s \in \mathsf{Spt}(X_{[e]})$ such that $\mathcal{M}$ is not both Ockham and stalwart at e given $X_{[e]} = s$ but is Ockham and stalwart at each $e' \in F_K$ such that $e' < e$ given $X_{[e]} = s$. We construct a (deterministic) competitor $\mathcal{O}$ for $\mathcal{M}$ that is Ockham and stalwart from e onward by modifying output function α of $\mathcal{M}$. Define the output function β so that for each $e' \in F_K$:

$$\beta(\sigma, e') \quad = \quad \begin{cases} \mathcal{M}_{e'}(\delta) & \text{for all } e' < e \text{ where } X_e(\delta) = \sigma; \\ T_{S_{e'}} & \text{if } e' \geq e \text{ and } T_{S_{e'}} \in \mathsf{Ock}_K(e'); \\ \text{`?'} & \text{otherwise.} \end{cases}$$

Now let $\mathcal{O}$ denote the result of replacing the output function α of $\mathcal{M}$ with the output function β just defined. By construction, we have that for each $s' \in \mathsf{Spt}(X_{[e]})$:

$$(\mathcal{M} \star^e_{s'} \mathcal{O}) = \mathcal{O}.$$

It suffices to show that:

1. $(\mathcal{O}, s) \leq^{\Gamma}_{K,e,i} (\mathcal{M}, s)$, for all i;

2. $(\mathcal{M}, s) \not\leq^{\Gamma}_{K,e,i} (\mathcal{O}, s')$, for some i.

By construction, $\mathcal{O}$ is henceforth Ockham and stalwart given $X_{[e]} = s$. Furthermore, $\mathcal{M}$ is eventually informative given $X_{[e]} = s$ because for each $w \in W_K$ there exists i such that for all $j \geq i$, $S_{w|j} = S_w$ and, hence, $\mathsf{Ock}_K(w|j) = \{T_{S_w}\}$. So by proposition 8, method $\mathcal{O}$ is consistent given $X_{[e]} = s$. Since $\mathcal{M}$ is consistent by assumption, statement 1 is an instance of $(1 \Rightarrow 2)$. For the second statement, focus on $i = 0$. Let:

$$\begin{aligned} a(n) &= \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\rho^{<k}_{\mathcal{M},w} \mid X_{[e]} = s) = \mathsf{Exp}(\rho^{<k}_{\mathcal{O},w} \mid X_{[e]} = s); \\ b(n) &= \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\epsilon^{<k}_{\mathcal{M},w} \mid X_{[e]} = s) = \mathsf{Exp}(\epsilon^{<k}_{\mathcal{O},w} \mid X_{[e]} = s). \end{aligned}$$

Case A: $\mathcal{M}$ is not stalwart given $X_{[e]} = s$.

Then there exists $T \in \mathsf{Ock}_{K,e}$ such that:

$$\begin{aligned} 0 &< p(\mathcal{M}_{e_-} = T \mid X_{[e]} = s); \\ 1 &> p(\mathcal{M}_e = T \mid \mathcal{M}_{e_-} = T \wedge X_{[e]} = s). \end{aligned}$$

Since $X_{[e]} = s$ settles the values of $\mathcal{M}_{e_-}, \mathcal{M}_e$ and $\mathcal{O}$ agrees with $\mathcal{M}$ at e_-:

$$\begin{aligned} 1 &= p(\mathcal{M}_{e_-} = T \mid X_{[e]} = s) \\ &= p(\mathcal{O}_{e_-} = T \mid X_{[e]} = s); \\ 0 &= p(\mathcal{M}_e = T \mid \mathcal{M}_{e_-} = T \wedge X_{[e]} = s). \end{aligned}$$

Since $\mathcal{O}$ is henceforth stalwart given $X_{[e]} = s$ and $p(\mathcal{O}_{e_-} = T \mid X_{[e]} = s) = 1$:

$$1 = p(\mathcal{O}_e = T \mid \mathcal{M}_{e_-} = T \wedge X_{[e]} = s).$$

So by proposition 26,

$$\mathsf{Exp}(\rho^k_{\mathcal{O},e} \mid X_{[e]} = s) = 0.$$

Using this fact and propositions 22, 27, and 19, one obtains for $C_{K,e}(i) \neq \emptyset$:

Retractions, $n \geq 0$		$i < k$	$i = k$	$i > k$	total
$\sup_{w \in C_{K,e}(n)} \sum_i \mathsf{Exp}(\rho^i_{\mathcal{M},w} \mid X_{[e]} = s)$	$\geq$	$a(n)$	1	n	$a(n) + n + 1$
$\sup_{w \in C_{K,e}(n)} \sum_i \mathsf{Exp}(\rho^i_{\mathcal{O},w} \mid X_{[e]} = s)$	$\leq$	$a(n)$	0	n	$a(n) + n$

By proposition 29, the supremum over expected retractions is in the total column. The advantage to $\mathcal{O}$ is proper since $C_{K,e}(0) \neq \emptyset$, by proposition 4. Turning to retraction times, it follows from the second row of the preceding table that there is no $e' \in F_{K,e}$ and $s' \in \mathsf{Spt}(X_{[e']} \mid X_{[e]} = s)$ such that $\rho(\mathcal{M}_{[e']}(s')) > n$. The same is true for arbitrary local loss function $\gamma \leq \rho$. Thus, for each $\gamma \leq \rho$ and $w \in C_{K,e}$:

$$\mathsf{Exp}(\tau^{\gamma \geq a+n+1}_{\mathcal{O},w} \mid X_{[e]} = s) = 0.$$

By the first row of the table and proposition 25, if $C_{K,e}(n) \neq \emptyset$, then there exists $w \in C_{K,e}(n)$ such that:

$$\mathsf{Exp}(\tau^{\rho \geq a+n+1}_{\mathcal{M},w} \mid X_{[e]} = s) > 0.$$

Hence:

$$(\mathcal{M}, s) \not\leq^{\tau}_{K,e,n} (\mathcal{O}, s).$$

Again, the advantage is proper since $C_{K,e}(0) \neq \emptyset$.

Case B: $\mathcal{M}$ is not Ockham at e given $X_{[e]} = s$. Then:

$$p(\mathcal{M}_e \notin \mathsf{Ock}_{K,e} \mid X_{[e]} = s) > 0.$$

Recall:

$$[\mathcal{M}_e \notin \mathsf{Ock}_{K,e}] = \bigcup_{T \in \mathsf{Th}_K, T \notin \mathsf{Ock}_{K,e}} [\mathcal{M}_e = T]$$

Hence, because Th_K is countable and p is countably additive, there exists $T \in \mathsf{Th}_K \setminus \mathsf{Ock}_{K,e}$ such that:

$$p(\mathcal{M}_e = T \mid X_{[e]} = s) > 0.$$

Moreover, because the output of $\mathcal{M}_e$ is settled by $X_{[e]} = s$, this implies:

$$p(\mathcal{M}_e = T \mid X_{[e]} = s) = 1.$$

Since $T \notin \mathsf{Ock}_{K,e}$, there exists maximal unit path (S) in K_e such that $S_T \neq S$. Let:

$$w_0 = e * (S \setminus S_e) * (\emptyset^\omega).$$

Then $\mathcal{M}$ commits an error with unit probability in w_0 at e. Furthermore, by proposition 23:

$$b(0) \leq \mathsf{Exp}(\epsilon^{<k}_{\mathcal{M},w} \mid X_{[e]} = s) = \mathsf{Exp}(\epsilon^{<k}_{\mathcal{O},w_0} \mid X_{[e]} = s).$$

Thus:

Errors, $n = 0$		$i < k$	$i = k$	$i > k$	total
$\sum_i \mathsf{Exp}(\epsilon^i_{\mathcal{M},w_0} \mid X_{[e]} = s)$	$\geq$	$b(0)$	1	0	$b(0) + 1$
$\sup_{w \in C_{K,e}(0)} \sum_i \mathsf{Exp}(\epsilon^i_{\mathcal{O},w} \mid X_{[e]} = s)$	$\leq$	$b(0)$	0	0	$b(0)$

The cumulative loss in w_0 is the total cumulative loss. Again, the advantage to $\mathcal{O}$ is proper, since $C_{K,e}(0) \neq \emptyset$, by proposition 4.

Regarding retraction times, let w_0 be as defined above. Let $\gamma \leq \rho$ and $w \in C_{K,e}(0)$ be arbitrary. Since $\mathcal{M}$ and $\mathcal{O}$ retract deterministically at the same times along e_-, we are done if γ fails to count each retraction along e_-, for a total of $a(0)$ retractions, so without loss of generality, suppose that γ agrees with ρ along e_-. Recall that $\mathcal{M}$ deterministically produces a theory false in w_0 at e given $X_{[e]} = s$. Since $\mathcal{M}$ is always consistent, we have for each $\epsilon > 0$:

$$\lim_{i \to \infty} p(\mathcal{M}_{w_0|i} = T_w \mid X_{[e]} = s) = 1.$$

Choose:

$$\epsilon < 1 - \frac{l(e)}{l(e)+1}.$$

Then there exists $k \geq l(e)$ such that:

$$p(\mathcal{M}_{w_0|i} = T_w \mid X_{[e]} = s) > 1 - \epsilon.$$

Since $\mathcal{M}$ produces an error at e in w_0, with probability $> 1 - \epsilon$, a retraction occurs no sooner than $l(e) + 1$. Hence:

$$\begin{aligned}
\mathsf{Exp}(\tau^{\rho \geq a(0)+1}_{\mathcal{M},w_0} \mid X_{[e]} = s) &> (1-\epsilon) \cdot (l(e)+1) \\
&= \left(1 - \left(1 - \frac{l(e)}{l(e)+1}\right)\right) \cdot (l(e)+1) \\
&= \frac{l(e)}{l(e)+1} \cdot (l(e)+1) \\
&= l(e).
\end{aligned}$$

Since $\mathcal{O}$ is Ockham and Stalwart from e onward by construction, proposition 19 yields that $\mathcal{O}$ never retracts properly after e in arbitrary $w \in C_{K,e}(0)$. So for each $\gamma \leq \rho$ and $w \in C_{K,e}(0)$:

$$\mathsf{Exp}(\tau^{\rho \geq a(0)+1}_{\mathcal{M},w_0} \mid X_{[e]} = s) \quad > \quad \mathsf{Exp}(\tau^{\gamma \geq a(0)+1}_{\mathcal{O},w} \mid X_{[e]} = s).$$

So:

$$(\mathcal{M}, s) \not\leq^{\tau}_{K,e,0} (\mathcal{M}, s).$$

□

5 Strengthened Ockham Efficiency Theorems

Recall that K has *no short paths* just in case for every $e \in K$ and every $S \in K_e$ there is a traversing path in K_e beginning with S. For example, K has no short paths if each maximal path in K is infinite, as in standard examples like curve fitting. The following two theorems strengthen Theorem 1 in two different ways. Theorem 2 proves that, if K has no short paths, then any method that fails to be always Ockham and stalwart is *strictly* dominated in worst-case costs in every complexity class with respect to the empty sequence. That is to say, Theorem 2 proves that there is never any reason to deviate from the behavior of an Ockham, stalwart strategy. Furthermore, Theorem 1 says that, if K has no short paths, then agents that have fallen from the stalwart, Ockham path (no pun intended!) in the past would do better to return to it at every stage of inquiry. Thus, the stalwart, Ockham property is *perfectly* truth-conducive, in the sense that it always makes sense, given what has happened, to reform one's ways and adopt the stalwart, Ockham lifestyle.

Theorem 2 (Strong, Stable Ockham Efficiency Characterization) *Suppose that K has no short paths. Let $e \in F_K$, $s \in \mathsf{Spt}(X_{[e]})$ and let $\mathcal{M}$ be consistent given $X_{[e]} = s$. Assume, finally, that:*

$$\{\tau\} \subseteq \Gamma \subseteq \{\rho, \epsilon, \tau\}.$$

Then following are equivalent.

1. *$\mathcal{M}$ is henceforth Ockham and stalwart given $X_{[e]} = s$.*
2. *$\mathcal{M}$ is perfectly Γ-efficient given $X_{[e]} = s$.*
3. *$\mathcal{M}$ is perfectly strongly Γ-undominated given $X_{[e]} = s$.*

Proof of 1 $\Rightarrow$ 2: The 1 $\Rightarrow$ 2 argument in the proof of Theorem 1 works up to the † symbol, where the assumption that $\mathcal{M}$ is always Ockham is used to show the existence of path $(S_0, \ldots, S_n) \in K_e$ such that $S_0 \neq S_T$. But the same conclusion can now be obtained from Proposition 6, because there are no short paths.

Proof of 2 $\Rightarrow$ 3: Again, weak dominance immediately implies inefficiency.

Proof of 3 $\Rightarrow$ 1: Follow the proof of the (3 $\Rightarrow$ 1) direction of the proof of Theorem 1 up to Cases A and B, including the definitions of $\mathcal{O}$ and $a(n)$. Cases A and B of that proof will now be strengthened using the no short paths hypothesis.

Case A′: $\mathcal{M}$ is not Stalwart given $X_{[e]} = s$. It is shown in case A, without assuming that e is the first stalwartness violation by $\mathcal{M}$, that for each n such that $C_{K,e}(n) \neq \emptyset$:

$$(\mathcal{M}, s) \not\leq^{\tau}_{K,e,n} (\mathcal{O}, s).$$

Case B′: $\mathcal{M}$ is not Ockham given $X_{[e]} = s$. It is shown in Case B that:

$$(\mathcal{M}, s) \not\leq^{\tau}_{K,e,0} (\mathcal{O}, s),$$

so it suffices to show that for each n such that $C_{K,e}(n+1) \neq \emptyset$:

$$(\mathcal{M}, s) \not\leq^{\tau}_{K,e,n+1} (\mathcal{O}, s).$$

Let $a = a(n+1)$. It is shown in case B, without assuming that e is the first Ockham violation by $\mathcal{M}$, that there exists $T \in \mathsf{Th}_K$ such that:

$$p(\mathcal{M}_e = T \mid X_{[e]} = s) = 1;$$

and that there exists maximal unit path (S_0) in K_e such that $S_T \neq S_0$. Suppose that $C_{K,e}(n) \neq \emptyset$. Then, by the no short paths hypothesis, (S_0) can be extended to a maximal path $(S_0, \ldots, S_{n+1})$ in K_e. By Propositions 19 and 21:

$$\begin{aligned} \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\rho^{>l(e)}_{\mathcal{M},w} \mid X_{[e]} = s) &\geq n+2; \\ \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\rho^{>l(e)}_{\mathcal{O},w} \mid X_{[e]} = s) &\leq n+1. \end{aligned}$$

Choose $w \in C_{K,e}(n+1)$ such that:

$$\mathsf{Exp}(\rho^{>l(e)}_{\mathcal{M},w} \mid X_{[e]} = s) \quad > \quad n+1.$$

So, by countable additivity and the discreteness of retractions in a given state history, there exists $k' > l(e)$ and $s' \in \mathsf{Spt}(X_{[w|k]} \mid X_{[e]} = s)$ such that:

$$\rho^{>l(e)}(\mathcal{M}_{[w|k']}(s')) \geq n+2 > \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\rho^{>l(e)}_{\mathcal{O},w} \mid X_{[e]} = s).$$

Furthermore:

$$a = \rho^{<l(e)}(\mathcal{M}_{[w|k']}(s')) = \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\rho^{<l(e)}_{\mathcal{M},w} \mid X_{[e]} = s);$$

and for each $i \leq a$:

$$0 < \tau^{\rho \geq i}(\mathcal{M}_{[w|k']}(s')) = \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\tau^{\rho \geq i}_{\mathcal{O},w} \mid X_{[e]} = s).$$

Case B.1: Suppose that $\mathcal{M}$ retracts at e if $\mathcal{O}$ does. Let $0 \leq b \leq b' \leq 1$. Then:

$$\begin{aligned} \rho^{\geq 0}(\mathcal{M}_{[w|k']}(s')) &= a + b' + n + 2 \\ &> a + b + n + 1 \\ &\geq \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\rho^{\geq 0}_{\mathcal{O},w} \mid X_{[e]} = s). \end{aligned}$$

Hence, by proposition 25, for each $j \leq a + b' + n + 2$:

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M},w} \mid X_{[e]} = s) > 0.$$

But since $n + 1 \geq \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\rho^{>l(e)}_{\mathcal{O},w} \mid X_{[e]} = s)$, we have, for each $j > a + b + n + 1$ and for each $\gamma \leq \rho$:

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{O},w} \mid X_{[e]} = s) = 0.$$

So there exists no $\gamma \leq \rho$ and $w' \in W_{K,e}$ such that for each $j \geq \omega$:

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M},w} \mid X_{[e]} = s) \leq \mathsf{Exp}(\tau^{\gamma \geq j}_{\mathcal{M},w'} \mid X_{[e]} = s).$$

Case B.2: Suppose that $\mathcal{O}$ retracts at e and $\mathcal{M}$ does not. Then:

$$\begin{aligned} \rho^{\geq 0}(\mathcal{M}_{[w|k']}(s')) &= a + n + 2 \\ &\geq \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\rho^{\geq 0}_{\mathcal{O},w} \mid X_{[e]} = s); \end{aligned}$$

and:

$$\begin{aligned} \tau^{\rho \geq a+1}(\mathcal{M}_{[w|k']}(s')) &> l(e); \\ &= \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\tau^{\rho \geq a+1}_{\mathcal{O},w} \mid X_{[e]} = s). \end{aligned}$$

So if $\gamma \leq \rho$ and:

$$\mathsf{Exp}(\tau^{\rho \geq a+1}_{\mathcal{M},w} \mid X_{[e]} = s) \leq \mathsf{Exp}(\tau^{\gamma \geq a+1}_{\mathcal{M},w'} \mid X_{[e]} = s),$$

it follows that:

$$\gamma(\mathcal{O}_{[e]}(s)) = 0.$$

But then:

$$\begin{aligned} \rho^{\geq 0}(\mathcal{M}_{[w|k']}(s')) &= a + n + 2 \\ &> a + n + 1 \\ &\geq \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\gamma^{\geq 0}_{\mathcal{O},w} \mid X_{[e]} = s). \end{aligned}$$

By proposition 25, for each $j \leq a + n + 2$:

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M},w} \mid X_{[e]} = s) > 0.$$

But since $n + 1 \geq \sup_{w \in C_{K,e}(n+1)} \mathsf{Exp}(\rho^{> l(e)}_{\mathcal{O},w} \mid X_{[e]} = s)$, we have, for each $j > a + n + 1$ and for each γ such that $\gamma(\mathcal{O}_{[e]}(s)) = 0$.

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{O},w} \mid X_{[e]} = s) = 0.$$

So there exists no $\gamma \leq \rho$ and $w' \in W_{K,e}$ such that for each $j \geq \omega$:

$$\mathsf{Exp}(\tau^{\rho \geq j}_{\mathcal{M},w} \mid X_{[e]} = s) \leq \mathsf{Exp}(\tau^{\gamma \geq j}_{\mathcal{M},w'} \mid X_{[e]} = s).$$

$\square$

6 Subsidiary Concepts and Lemmas

6.1 Errors and Retractions in Probability

Here, we introduce an auxiliary concept that links expected errors and retractions with learning theoretic argumentation. Instead of penalizing expected retractions, one can penalize retractions *in probability*, which occur when the chance of producing a theory drops over time. In a similar spirit, an error *in probability* occurs when there is a chance of producing the wrong answer. Let the total error *in probability* committed by p in w at stage i *given* $X_{[e]} = s$ be:

$$\mathsf{Ipr}[\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s] = p(\mathcal{M}_{w|i} \in \mathsf{Err}_K(w) \mid X_{[e]} = s),$$

where the square brackets indicate that $\mathsf{Ipr}[\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s]$ is a formal notation intended to resemble the notation for conditional expectation, rather than a literal operation on the random variable $\epsilon^i_{\mathcal{M},w}$. The total *retraction in probability* by $\mathcal{M}$ in w at stage i given $X_{[e]} = s$, denoted by the formal expression:

$$\mathsf{Ipr}[\rho^i_{\mathcal{M},w} \mid X_{[e]} = s],$$

is defined as:

$$\sum_{T \in \mathsf{Th}_K} p(\mathcal{M}_{w|i} = T \mid X_{[e]} = s) \ \ominus \ p(\mathcal{M}_{w|i+1} = T \mid X_{[e]} = s),$$

where $x \ominus y = x - y$ if $x \geq y$ and $x \ominus y = 0$ otherwise. Let $\gamma \in \{\epsilon, \rho\}$. Then define:

$$\mathsf{Ipr}[\epsilon^{<k}_{\mathcal{M},w} \mid X_{[e]} = s] \quad = \quad \sum_{i<k} \mathsf{Ipr}[\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s];$$

and similarly for $\geq, \leq, >$. As in the case of expected cost, let the joint cost in probability be:

$$\mathsf{Ipr}[\kappa^{<k}_{\mathcal{M},w} \mid X_{[e]} = s] \quad = \quad (\mathsf{Ipr}[\epsilon^{<k}_{\mathcal{M},w} \mid X_{[e]} = s],\ \mathsf{Ipr}[\rho^{<k}_{\mathcal{M},w} \mid X_{[e]} = s]);$$

and similarly for $\geq, \leq, >$, and relate joint costs by the Pareto ordering discussed for expected cost.

Aside from the intrinsic interest of errors and retractions in probability, they provide lower bounds on errors and retractions constitute lower bounds on expected errors and retractions, by the following proposition. To see that the inequality cannot be improved to equality in the general case, consider a stochastic method $\mathcal{M}$ that flips a fair coin to decide between T and T' for ten stages. No retraction in probability occurs over these ten stages, because the chance of producing T, T' does not change, but the expected retractions are non-zero since the chance of retraction is .5 at each such stage (i.e., the chance of producing pattern (T, T') plus the chance of producing pattern (T', T)). This point has concrete methodological consequences: it is never a good idea to produce an answer twice in succession with fractional probability, since such a strategy incurs needless expected retractions. That is why violations of stalwartness imply inefficiency.

Proposition 17 *Let $\gamma \in \{\epsilon, \rho, \kappa\}$ and let $e \in F_K$ and $s \in \mathit{Spt}(X_{[e]})$. Then we have:*

$$\mathit{Ipr}[\gamma^i_{\mathcal{M},w} \mid X_{[e]} = s] \quad \leq \quad \mathit{Exp}(\gamma^i_{\mathcal{M},w} \mid X_{[e]} = s).$$

Proof.

$$\begin{aligned}
\mathsf{Exp}(\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s) &= \int \epsilon^i_{\mathcal{M},w}\; dp(.\mid X_{[e]} = s); \\
&= 1 \cdot p(\epsilon^i_{\mathcal{M},w} = 1 \mid X_{[e]} = s) + 0 \cdot p(\epsilon^i_{\mathcal{M},w} 0 \mid \; X_{[e]} = s) \\
&= p(\epsilon^i_{\mathcal{M},w} = 1 \mid X_{[e]} = s) \\
&= p(\mathcal{M}_{w|i} \in \mathsf{Err}(w) \mid X_{[e]} = s) \\
&= \sum_{T \in \mathsf{Err}(w)} p(\mathcal{M}_{w|i} = T \mid X_{[e]} = s) \\
&= \mathsf{Ipr}[\epsilon^i_{\mathcal{M},w} \mid X_{[e]} = s].
\end{aligned}$$

$$\begin{aligned}
\mathsf{Exp}(\rho^i_{\mathcal{M},w} \mid X_{[e]} = s) &= \int \rho^i_{\mathcal{M},w}\; dp(.\mid X_{[e]} = s) \\
&= 1 \cdot p(\rho^i_{\mathcal{M},w} = 1 \mid X_{[e]} = s) + 0 \cdot p(\rho^i_{\mathcal{M},w} = 0 \mid X_{[e]} = s) \\
&= p(\rho^i_{\mathcal{M},w} = 1 \mid X_{[e]} = s) \\
&= p(\mathcal{M}_{w|i+1} \neq \mathcal{M}_{w|i} \neq \text{‘?’} \mid X_{[e]} = s) \\
&= p(\bigcup_{T \in \mathsf{Th}_K} [\mathcal{M}_{w|i} = T] \setminus [\mathcal{M}_{w|i+1} = T] \mid X_{[e]} = s) \\
&= \sum_{T \in \mathsf{Th}_K} p([\mathcal{M}_{w|i} = T] \setminus [\mathcal{M}_{w|i+1} = T] \mid X_{[e]} = s) \\
&\geq \sum_{T \in \mathsf{Th}_K} p(\mathcal{M}_{w|i} = T \mid X_{[e]} = s) \ominus p(\mathcal{M}_{w|i+1} = T \mid X_{[e]} = s) \\
&= \mathsf{Ipr}[\rho^i_{\mathcal{M},w} \mid X_{[e]} = s].
\end{aligned}$$

□

It follows immediately from propositions 12 and 17 that:

Proposition 18 (basic inequality) *Let γ range over $\{\epsilon, \rho, \kappa\}$. Then:*

$$\mathit{Ipr}[\gamma^{\geq k}_{\mathcal{M},w} \mid X_{[e]} = s] \quad \leq \quad \mathit{Exp}(\gamma^{\geq k}_{\mathcal{M},w} \mid X_{[e]} = s),$$

and similarly for $\gamma^{<k}_{\mathcal{M},w}, \gamma^{k}_{\mathcal{M},w}$.

7 Upper Loss Bounds

Proposition 19 *Let $\mathcal{M}$ be henceforth stalwart and Ockham given $X_{[e]} = s$. Let $e' \in F_{K,e}$. Then:*

1. $sup_{w \in C_{K,0}(e')} \mathsf{Exp}(\epsilon^{\geq l(e')}_{\mathcal{M},w} \mid X_{[e]} = s) = 0.$

2. $sup_{w \in C_{K,e'}(n)} \mathsf{Exp}(\rho^{>l(e')}_{\mathcal{M},w} \mid X_{[e]} = s) \leq n.$

By proposition 18, Ipr *may be substituted for* Exp.

Proof. For $n = 0$, suppose that $w \in C_{K,0}(e')$. Then, for each $k \geq l(e')$,

$$\mathsf{Ock}_{w|k} = \{\text{`?'}\} \text{ or } \mathsf{Ock}_{w|k} = \{\text{`?'}, T_w\}.$$

Since $\mathcal{M}$ is henceforth Ockham given $X_{[e]} = s$, it follows that:

$$(*) \quad p(\mathcal{M}_{w|k} \in \{T_w, \text{`?'}\} \mid X_{[e]} = s) = 1, \quad \text{for all } k \geq l(e').$$

For statement 1, let $k \geq l(e')$. Recall that:

$$\begin{aligned}
\mathsf{Exp}(\epsilon^k_{\mathcal{M},w} \mid X_{[e]} = s) &= \int \epsilon^k_{\mathcal{M},w}\ dp(. \mid X_{[e]} = s) \\
&= 0 \cdot p(\epsilon^k_{\mathcal{M},w} = 0 \mid X_{[e]} = s) \\
&\quad +1 \cdot p(\epsilon^k_{\mathcal{M},w} = 1 \mid X_{[e]} = s) \\
&= 0 \cdot p(\mathcal{M}_{w|k} \in \{T_w, \text{`?'}\} \mid X_{[e]} = s) \\
&\quad +1 \cdot p(\mathcal{M}_{w|k} \notin \{T_w, \text{`?'}\} \mid X_{[e]} = s) \\
&= p(\mathcal{M}_{w|k} \notin \{T_w, \text{`?'}\} \mid X_{[e]} = s) \\
&= 0.
\end{aligned}$$

So, since $w \in C_{K,0}(e')$ and $k \geq l(e')$ are arbitrary, proposition 12 yields 1.

Next, we prove statement 2 by induction on n. Let $k > l(e')$, recall that:

$$\begin{aligned}
\mathsf{Exp}(\rho^k_{\mathcal{M},w} \mid X_{[e]} = s) &= \int \rho^k_{\mathcal{M},w}\ dp(. \mid X_{[e]} = s) \\
&= 0 \cdot p(\mathcal{M}_{w|k} = \mathcal{M}_{w|k-1} \vee \mathcal{M}_{w|k-1}\text{`?'} \mid X_{[e]} = s) \\
&\quad +1 \cdot p(\mathcal{M}_{w|k} \neq \mathcal{M}_{w|k-1} \neq \text{`?'} \mid X_{[e]} = s) \\
&= p(\mathcal{M}_{w|k} \neq \mathcal{M}_{w|k-1} \neq \text{`?'} \mid X_{[e]} = s).
\end{aligned}$$

Case 1: Suppose that $p(\mathcal{M}_{w|k-1} \neq \text{`?'} \mid X_{[e]} = s) = 0$. Then:

$$\begin{aligned}
\mathsf{Exp}(\rho^k_{\mathcal{M},w} \mid X_{[e]} = s) &= p(\mathcal{M}_{w|k} \neq \mathcal{M}_{w|k-1} \neq \text{`?'} \mid X_{[e]} = s) \\
&\leq p(\mathcal{M}_{w|k-1} \neq \text{`?'} \mid X_{[e]} = s) \\
&= 0.
\end{aligned}$$

Case 2: Suppose that $p(\mathcal{M}_{w|k-1} \neq \text{`?'} \mid X_{[e]} = s) > 0$. Recall that because $\mathcal{M}$ is henceforth Ockham given $X_{[e]} = s$ and $k > l(e)$ and $w \in C_{K,0}(e')$, we have:

$$p(\mathcal{M}_{w|k-1} \in \{T_w, \text{`?'}\} \mid X_{[e]} = s) = 1$$

which, by the assumption of Case 2, implies:

$$p(\mathcal{M}_{w|k-1} = T_w \mid X_{[e]} = s) > 0.$$

So by the henceforth stalwartness of $\mathcal{M}$ given $X_{[e]} = s$:

$$1 = p(\mathcal{M}_{w|k} = T_w \mid \mathcal{M}_{e|k-1} = T_w \wedge X_{[e]} = s)$$

which implies:

$$0 = p(\mathcal{M}_{w|k} = \text{`?'} \mid \mathcal{M}_{e|k-1} = T_w \wedge X_{[e]} = s).$$

Thus:

$$\begin{aligned}
\mathsf{Exp}(\rho^k_{\mathcal{M},w} \mid X_{[e]} = s) &= p(\mathcal{M}_{w|k} \neq \mathcal{M}_{w|k-1} \neq \text{`?'} \mid X_{[e]} = s) \\
&= p(\mathcal{M}_{w|k} = \text{`?'} \wedge \mathcal{M}_{w|k-1} = T_w \mid X_{[e]} = s) \\
&\quad + \sum_{T \in \mathsf{Th}_K, T \neq T_w} p(\mathcal{M}_{w|k} = \text{`?'} \wedge \mathcal{M}_{w|k-1} = T \mid X_{[e]} = s) \\
&\quad + \sum_{T, T' \in \mathsf{Th}_K, T \neq T'} p(\mathcal{M}_{w|k} = T \wedge \mathcal{M}_{w|k-1} = T' \mid X_{[e]} = s) \\
&\leq p(\mathcal{M}_{w|k} = \text{`?'} \mid \mathcal{M}_{w|k-1} = T_w \wedge X_{[e]} = s) \\
&\quad + \sum_{T \in \mathsf{Th}_K, T \neq T_w} p(\mathcal{M}_{w|k-1} = T \mid X_{[e]} = s) + \\
&\quad + \sum_{T \in \mathsf{Th}_K, T \neq T_w} p(\mathcal{M}_{w|k} = T \mid X_{[e]} = s) \\
&= 0 + 0 + 0 \\
&= 0.
\end{aligned}$$

So, since $w \in C_{K,0}(e')$ and $k > l(e')$ are arbitrary, proposition 12 yields:

$$_{w \in C_{K,0}(e')} \mathsf{Exp}(\rho^{>l(e')}_{\mathcal{M},w} \mid X_{[e]} = s) = 0.$$

For induction, suppose that $w \in C_{K,n+1}(e')$. Then each nested path $S_{e'} \subset \ldots \subset S_w$ through $K_{e'}$ has length $\leq n + 2$. Let m be least such that $S_{e'} \subset S_{w|m}$. Then, since no path from $S_{w|m}$ to S_w in $K_{e|m}$ begins with $S_{e'}$, each such path has length $\leq n + 1$, so $w \in C_{K,n}(w|m)$. By the induction hypothesis:

i. $\sup_{w \in C_{K,e'}(n)} \mathsf{Exp}(\rho^{>m}_{\mathcal{M},w} \mid X_{[e]} = s) \leq n$.

Construct:

$$w' = w|m - 1 * (\emptyset^\infty).$$

Then the only nested paths from S_e to $S_{w'} = S_{e'}$ have unit length, so

$$w' \in C_{K,0}(e').$$

So by the base case:

$$\mathsf{Exp}(\rho^{>l(e')}_{\mathcal{M},w'} \mid X_{[e]} = s) = 0.$$

By choice of m:

$$w'|m - 1 = w|m - 1,$$

so:

ii. $\mathsf{Exp}(\rho^i_{\mathcal{M},w} \mid X_{[e]} = s) = 0, \quad$ for each i such that $l(e) < i < m$.

By the definition of $\rho^i_{\mathcal{M},w}$:

iii. $\mathsf{Exp}(\rho^m_{\mathcal{M},w} \mid X_{[e]} = s) \leq 1$.

Thus, by (i -iii) and proposition 12:

$$\mathsf{Exp}(\rho^{>l(e')}_{\mathcal{M},w} \mid X_{[e]} = s) \leq n + 1.$$

Since $w \in C_{K,e'}(n+1)$ is arbitrary:

$$\sup_{w \in C_{K,e'}(n+1)} \mathsf{Exp}(\rho^{>l(e')}_{\mathcal{M},w} \mid X_{[e]} = s) \leq n + 1.$$

□

8 Lower Loss Bounds

The fundamental idea behind the Ockham efficiency theorem is the following proposition, inspired by computational learning theory. It says that nature is in a position to force an arbitrary, consistent strategy to retract at least n times in complexity class n and to retract ω times in each complexity class > 0. One reason for counting retractions instead of errors is that worst-case retractions are bounded within empirical complexity classes whereas worst-case errors are not.

Proposition 20 *Suppose that $\mathcal{M}$ is consistent given $X_{[e]} = s$, let $e' \geq e$ and assume that $C_{K,e'}(n) \neq \emptyset$. Then:*

$$\begin{aligned} \sup_{w \in C_{K,e'}(n)} \mathsf{Ipr}[\epsilon^{>l(e')}_{\mathcal{M},w} \mid X_{[e]} = s] &\geq \omega, \quad \text{if } n > 0; \\ \sup_{w \in C_{K,e'}(n)} \mathsf{Ipr}[\rho^{>l(e')}_{\mathcal{M},w} \mid X_{[e]} = s] &\geq n. \end{aligned}$$

By proposition 18, Exp may be substituted for Ipr.

Proof. Let $e' \geq e$ and suppose that $C_{K,e'}(n) \neq \emptyset$ and $\mathcal{M}$ is consistent with respect to K given $X_{[e]} = s$. Let $l(e') = k'$.

Proof of the first statement. Since $C_{K,e'}(n) \neq \emptyset$ and $n > 0$, it follows from the definition of empirical complexity that there exists a finite, upward-nested sequence $(S_0, \ldots, S_{n+1})$ of elements of $K_{e'}$ such that $S_{n+1} \neq S_0$. Define:

$$w = e' * (S_0 \setminus S_{e'}) * (\emptyset^\infty).$$

Then $S_w = S_0 \in K$ and w extends e', so $w \in C_{K,0}(e')$. Since $\mathcal{M}$ is consistent given $X_{[e]} = s$, there exists $n_0 > k'$ such that for all $m \geq n_0$:

$$p(\mathcal{M}_{w|m} = T_w \mid X_{[e]} = s) > 1/2.$$

Let $m \in \omega$ be arbitrary and define:

$$w' = (w|(n_0 + 2m)) * (S_{n+1} \setminus S_{w|(n_0+2m)}) * (\emptyset^\infty).$$

Then $e' \leq w'$ and $S_{w'} = S_{n+1} \in K$, so $w' \in C_{K,e'}(n)$. But $T_{w'} = T_{S_{n+1}} \neq T_{S_0} = T_w$, so T_w is an error in $T_{w'}$. Thus:

$$\begin{aligned}
\mathsf{lpr}[\epsilon^{>k'}_{\mathcal{M},w'} \mid X_{[e]} = s] &= \sum_{i>k'} \mathsf{lpr}[\epsilon^{i}_{\mathcal{M},w'} \mid X = s] \\
&= \sum_{i>k'} p(\mathcal{M}_{w'|i} \notin \{T_{w'}, \text{‘?’}\} \mid X = s) \\
&> \sum_{i>k'}^{k'+2m} \epsilon > 2m/2 = m.
\end{aligned}$$

Hence:

$$\begin{aligned}
\sup_{w \in C_{K,e'}(n)} \mathsf{lpr}[\epsilon^{>k'}_{\mathcal{M},w} \mid X_{[e]} = s] &\geq \sup_{m \in \omega} m \\
&= \omega.
\end{aligned}$$

Proof of the second statement. The base case when $n = 0$ is trivial, since costs are non-negative. For the inductive case, Suppose that $C_{K,n+1}(e') \neq \emptyset$. Since $C_e(n+1 \mid e') \neq \emptyset$, there exists an upward-nested path $(S_0, \ldots, , S_n, S_{n+1})$ in $K_{e'}$. Let $K' = \{S_0, \ldots, S_n, S_{n+1}\}$. Then $\mathcal{M}$ is still consistent with respect to K' given $X_{[e]} = s$ and the decremented path $(S_0, \ldots, S_n)$ in K' witnesses that $C_{K',n}(e') \neq \emptyset$. So, by the induction hypothesis,

$$\sup_{w \in C_{K',n}(e')} \mathsf{lpr}[\rho^{>k'}_{\mathcal{M},w} \mid X_{[e]} = s] \geq n.$$

Let $\epsilon > 0$ be given. Then there exists $w \in C_{K',n}(e')$ such that

$$\mathsf{lpr}[\rho^{>k'}_{\mathcal{M},w} \mid X_{[e]} s] = \sum_{i>k'}^{\infty} \mathsf{lpr}[\rho^{i}_{\mathcal{M},w} \mid X_{[e]} = s] > n - \epsilon/3.$$

Hence, there exists $k_0 > k'$ such that:

$$\sum_{i>k'}^{k_0} \mathsf{lpr}[\rho^{i}_{\mathcal{M},w} \mid X_{[e]} = s] > n - \epsilon/2.$$

Since $\mathcal{M}$ is consistent given s in response to e, there exists $k_1 > k_0$ such that for all $m \geq k_1$:

$$p(\mathcal{M}_{w|m} = T_w \mid X_{[e]} = s) > 1 - \epsilon/4.$$

In K' the unique effect set of complexity n is S_{n+1}, so since $w \in C_{K,e'}(n)$, it follows that:

$$S_w = S_{n+1} \subseteq S_{n+2}.$$

Define:

$$w' = (w|k') * (S_{n+2} \setminus S_{w|k'}) * (\emptyset^\infty).$$

Then w' extends e' and $S_{w'} \in K' \subseteq K$, so $w' \in C_{K,n+1}(e')$. Since $\mathcal{M}$ is consistent in K given $X_{[e]} = s$, there exists $k_2 > k_1$ such that for all $m \geq k_2$:

$$p(\mathcal{M}_{w'|m} = T_{w'} \mid X_{[e]} = s) > 1 - \epsilon/4.$$

Then:

$$\begin{aligned} & p(\mathcal{M}_{w'|k_1} = T_w \mid X_{[e]} = s) - p(\mathcal{M}_{w'|k_2} = T_{w'} \mid X_{[e]} = s) \\ > \;& (1 - \epsilon/4) - (1 - (1 - \epsilon/4)) \\ = \;& 1 - \epsilon/2. \end{aligned}$$

But:

$$T_{w'} = T_{S_{n+1}} \neq T_{S_n} = T_w,$$

so:

$$\begin{aligned} \mathsf{Ipr}[\rho^{>k'}_{\mathcal{M},w'} \mid X_{[e]} = s] &= \sum_{i>k'}^{\infty} \mathsf{Ipr}[\rho^{i}_{\mathcal{M},w'} \mid X_{[e]} = s] \\ &\geq \sum_{i>k'}^{k_0} \mathsf{Ipr}[\rho^{i}_{\mathcal{M},w'} \mid X_{[e]} = s] + \sum_{i>k_0}^{k_1} \mathsf{Ipr}[\rho^{i}_{\mathcal{M},w'} \mid X_{[e]} = s] \\ &> (n - \epsilon/2) + (1 - \epsilon/2) \\ &= (n+1) - \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, it follows that:

$$\sup_{w \in C_{K,n+1}(e')} \mathsf{Ipr}[\rho^{>k'}_{\mathcal{M},w'} \mid X_{[e]} = s] \quad \geq \quad n+1.$$

□

Proposition 21 *Suppose that $\mathcal{M}$ is consistent given $X_{[e]} = s$. Let $e' \in F_{K,e}$. Let $(S_0, \ldots, S_n)$ be a maximal path in $K_{e'}$ ending with S_n. Suppose that $S_T \neq S_0$ and that $p(\mathcal{M}_{e'} = T \mid X_{[e]} = s) = r > 0$. Then for each $j \leq n$:*

$$\sup_{w \in C_{K,e'}(0)} \mathit{Ipr}[\epsilon^{l(e')}_{\mathcal{M},w} \mid X_{[e]} = s] \quad \geq \quad r.$$

$$\sup_{w \in C_{K,e'}(j)} \mathit{Ipr}[\rho^{>l(e')}_{\mathcal{M},w} \mid X_{[e]} = s] \quad \geq \quad j + r.$$

By proposition 18, Exp may be substituted for Ipr.

Proof. Let $k = l(e)$. In the base case for $n = 0$, let (S_0) be a maximal path through $K_{e'}$ and define:

$$w = e * (S_0 \setminus S_e) * (\emptyset^\infty).$$

Since $S_T \cap S_0 = \emptyset$, it follows that $T_w \neq T$, so:

$$\begin{aligned} r &= p(\mathcal{M}_{e'} \in \mathsf{Err}_{K,e'} \mid X_{[e]} = s) \\ &= \mathsf{lpr}[\epsilon^k_{\mathcal{M},w} \mid X_{[e]} = s]. \end{aligned}$$

By the consistency of $\mathcal{M}$ given $X_{[e]} = s$:

$$\lim_{i \to \infty} p(\mathcal{M}_{w|i} = T_w \mid X_{[e]} = s) = 1.$$

Let $\epsilon > 0$ and choose $i > k$ such that:

$$p(\mathcal{M}_{w|i} = T_w \mid X_{[e]} = s) > 1 - \epsilon.$$

Then

$$\begin{aligned} r - \epsilon &> \mathsf{lpr}[\epsilon^k_{\mathcal{M},w} \mid X_{[e]} = s]; \\ r - \epsilon &> \mathsf{lpr}[\rho^{>k}_{\mathcal{M},w} \mid X_{[e]} = s], \end{aligned}$$

so:

$$\begin{aligned} r &\geq \sup_{w \in C_{K,e'}(0)} \mathsf{lpr}[\epsilon^k_{\mathcal{M},w} \mid X_{[e]} = s]; \\ r &\geq \sup_{w \in C_{K,e'}(0)} \mathsf{lpr}[\rho^{>k}_{\mathcal{M},w} \mid X_{[e]} = s]. \end{aligned}$$

For the induction, let $(S_0, \ldots, S_{n+1})$ be a maximal path through $K_{e'}$ terminating in S_{n+1} such that $S_T \neq S_0$. Truncate the path to $(S_0, \ldots, S_n)$. It is still the case that $S_T \neq S_0$.

Let $K' = \{S_0, \ldots, S_n\}$. Let $\epsilon > 0$. By the induction hypothesis,

$$\sup_{w \in C_{K,e'}(0)} \mathsf{lpr}[\rho^{>k}_{\mathcal{M},w} \mid X_{[e]} = s] = n + r.$$

So there exists $w' \in C_{K,e'}(n)$ and $j > k$ such that:

$$\sum_{i=k+1}^{j} \mathsf{lpr}[\rho^i_{\mathcal{M},w'} \mid X_{[e]} = s] > n + r - \epsilon/2.$$

By the consistency of $\mathcal{M}$ given $X_{[e]} = s$,

$$\lim_{i \to \infty} p(\mathcal{M}_{w'|i} = T_{w'} \mid X_{[e]} = s) = 1.$$

Let

$$w'' = w|j' * (S_{n+1} \setminus S_{w|j}) * (\emptyset^\infty).$$

By consistency again:

$$\lim_{i \to \infty} p(\mathcal{M}_{w''|i} = T_{w''} \mid X_{[e]} = s) = 1,$$

so let $j'' > j'$ be such that:

$$p(\mathcal{M}_{w''|j''} = T_{w''} \mid X_{[e]} = s) > 1 - \epsilon/4.$$

Since $S_{n+1} \notin K'$, it follows that $T_{w'} \neq T_{w''}$. Hence:

$$\begin{aligned} \sum_{i=j'}^{j''} \mathsf{Ipr}[\rho^i_{\mathcal{M},w'} \mid X_{[e]} &= s] > (1 - \epsilon/4) - \epsilon/4 \\ &= 1 - \epsilon/2. \end{aligned}$$

So:

$$\begin{aligned} \sum_{i=k}^{j''} &= (n + r) - \epsilon/2 + 1 - \epsilon/2 \\ &= (n + 1 + r) - \epsilon. \end{aligned}$$

So, since $\epsilon > 0$ is arbitrary:

$$\sup_{w \in C_{K,e}(n+1)} \mathsf{Ipr}[\rho^{>k}_{\mathcal{M},w} \mid X_{[e]} = s] \geq n + 1 + r.$$

□

Proposition 22 *Suppose that $\mathcal{M}$ is consistent given $X_{[e]} = s$, that $\mathcal{M}$ fails to be henceforth stalwart given $X_{[e]} = s$, and that $C_{K,e}(n) \neq \emptyset$. Then there exists $e' \in F_{K,e}$ and $s' \in \mathsf{Spt}_s(X_{[e']})$ such that:*

$$\sup_{w \in C_{K,e'}(n)} \mathsf{Ipr}[\rho^{l(e')}_{\mathcal{M},w} \mid X_{[e']} = s'] \geq n + 1.$$

By proposition 18, Exp may be substituted for Ipr.

Proof. Suppose that $\mathcal{M}$ is consistent given $X_{[e]} = s$ fails to be henceforth stalwart given $X_{[e]} = s$. Hence, there exists $e' \in F_{K,e}$ and $T \in \mathsf{Th}_K$ such that $\mathcal{M}$ fails to be stalwart with respect to T at e' given $X_{[e]} = s$. By proposition 7, there exists $s' \in \mathsf{Spt}_e(X_{[e']})$ such that $\mathcal{M}$ fails to be stalwart with respect to T at e' given $X_{[e']} = s'$. Thus:

i. $T \in \mathsf{Ock}(e')$;

ii. $p(\mathcal{M}_{e'_-} = T \mid X_{[e']} = s') > 0$;

iii. $p(\mathcal{M}_{e'} = T \mid \mathcal{M}_{e'_-} = T \wedge X_{[e']} = s') < 1$.

So, since s' determines all states assumed by $\mathcal{M}$ along e':

ii$'$. $p(\mathcal{M}_{e'_-} = T \mid X_{[e']} = s') = 1$;

iii$'$. $p(\mathcal{M}_{e'} = T \mid \mathcal{M}_{e'_-} = T \wedge X_{[e']} = s') = 0$.

Thus:

$$p(\mathcal{M}_{e'_-} = T \mid X_{[e']} = s') \ominus p(\mathcal{M}_{e'} = T \mid X_{[e']} s') = 1.$$

Let:

$$w = e'_- * (\emptyset^\infty).$$

Since $e' \in F_K$, there exists $w \in W_{K,e}$, so $S_{e'} \subseteq S_w \in K$. Hence, one may choose $S \in K$ to be least such that $S_{e'} \subseteq S$. Then $c_K(S \mid e') = 0$. Define:

$$w = e' * (S \setminus S_{e'}) * (\emptyset^\infty).$$

Then:

$$w \in C_{K,e'}(0)$$

and:

$$\mathsf{lpr}[\rho^{l(e')}_{\mathcal{M},w} \mid X_{[e']} = s'] \geq 1.$$

Hence:

$$\sup_{w \in C_{K,e'}(0)} \mathsf{lpr}[\rho^{l(e')}_{\mathcal{M},w} \mid X_{[e']} = s'] \geq 1.$$

That establishes the base case for $n = 0$. The inductive step proceeds as in the proof of proposition 20, under the induction hypothesis that

$$\sup_{w \in C_{K',e'}(n)} \mathsf{lpr}[\rho^{l(e')}_{\mathcal{M},w} \mid X_{[e']} = s'] \geq n + 1.$$

$\square$

Proposition 23 *Let $e \in F_K$, $e' < e$, $s \in \mathit{Spt}(X_{[e']})$, and suppose that $\mathcal{M}$ is Ockham at e' given $X_{[()]} = (\sigma_0)$ and that $\mathcal{M}_{e'}(s') = T_S \in \mathit{Th}_K$. Then for each $w, w' \in C_{K,e}(0)$:*

$$w \in T_S \leftrightarrow w' \in T_S.$$

Proof. Let $w, w' \in C_{K,e}(0)$. Then $(S_w), (S_{w'})$ are maximal unit paths in K_e. Let $e' < e$ and let $\mathcal{M}_{e'}(s') = T_S \in \mathsf{Th}_K$. Suppose that $w' \notin T_S$. Then:

$$S \neq S_{w'}.$$

By hypothesis, $\mathcal{M}$ is Ockham at e' given $X_{[()]} = (\sigma_0)$, so T_S is Ockham at e' and, hence, (S) is the unique, maximal unit path in $K_{e'}$. Thus:

$$S_{e'} \subseteq S \subseteq S_e \subseteq S_{w'},$$

where the second inclusion is by uniqueness. Therefore:

$$S \subset S_{w'}.$$

So, since $(S_{w'})$ is a maximal path in K_e, we have that $S \notin K_e$. So, since (S_w) is also a maximal unit path in K_e, we have: $S \neq S_w$, so $w \notin T_S$. $\square$

For the propositions that follow, if γ is a local loss function and $l(s) = l(e) + 1$, define:

$$\gamma(s) = \gamma(\mathcal{M}_{[e]}(s)).$$

Proposition 24 *Suppose that $\mathcal{M}$ is consistent given $X_{[e]} = s$ and assume that $n > 0$ and that $C_{K,e'}(n) \neq \emptyset$ and that $s \in \mathsf{Spt}(X_e)$. Let $a = \mathsf{Exp}(\rho^{<l(e)} \mid X_{[e]} = s)$. Then for each k there exists $w \in C_{K,e'}(n)$ and $\gamma \leq \rho$ such that:*

$$\mathsf{Exp}(\tau^{\gamma \geq a+n}_{\mathcal{M},w} \mid X_{[e]} = s) \quad \geq \quad k.$$

Proof by induction on n: The proposition is trivial if $k = 0$, so without loss of generality suppose that $k > 0$. The base case is trivial, since $n > 0$. Now suppose that $C_{K,e'}(n+1) \neq \emptyset$ and $s \in \mathsf{Spt}(X_e)$. Then $C_{K,e'}(n) \neq \emptyset$. Let $k \in \omega$ and let $a = \mathsf{Exp}(\rho^{<l(e)} \mid X_{[e]} = s)$. By the induction hypothesis, there exists $w \in C_{K,e'}(n)$ and $\gamma \leq \rho$ such that:

$$\mathsf{Exp}(\tau^{\gamma \geq a+n}_{\mathcal{M},w} \mid X_{[e]} = s) \quad \geq \quad k.$$

Define the following local loss function, which is just like γ, except that it ignores all retractions after the $a + 1$th retraction counted by γ:

$$\gamma_0(s) = \begin{cases} 1 & \text{if } \gamma(s) = 1 \text{ and } \text{ and } \gamma^{<l(s)}(s) < a + n; \\ 0 & \text{otherwise.} \end{cases}$$

Thus:

$$\gamma_0 \leq \gamma \leq \rho,$$

and:

$$\mathsf{Exp}(\tau^{\gamma' \geq a+n}_{\mathcal{M},w} \mid X_{[e]} = s) \quad \geq \quad k > 0.$$

Define:

$$S = \{e \in \Sigma^{<\omega} : l(s) > 0 \text{ and } \gamma(s) = a + n \text{ and } \gamma_0(s_-) < i\}.$$

Since:

$$\mathsf{Exp}(\tau^{\gamma_0 \geq a+n}_{\mathcal{M},w} \mid X_{[e]} = s) \quad = \quad \sum_{s' \in S} l(s') \cdot p(X_{[w|l(s')]} = s' \mid X_{[e]} = s),$$

it follows that there exists $s_0 \in S$ such that:

$$0 < l(s) \cdot p(X_{[w|l(s')]} = s_0 \mid X_{[e]} = s).$$

Set:

$$\begin{array}{rcl} p & = & p(X_{[w|l(s')]} = s_0 \mid X_{[e]} = s); \\ l & = & l(e). \end{array}$$

Let $p > \epsilon > 0$. Since $\mathcal{M}$ is consistent, there exists $m > l(s)$ such that for each $m' \geq m$:

$$p(\mathcal{M}_{w|m'} = T_w \mid X_{[e]} = s) > 1 - \epsilon/2.$$

Let:

$$m_0 > \max(m, \frac{k}{p-\epsilon}),$$

so:

$$p(\mathcal{M}_{w|m_0} = T_w \mid X_{[e]} = s) > 1 - \epsilon/2.$$

Since $C_{K,e}(n+1) \neq \emptyset$, there exists effect $x \in E \setminus S_e$. Construct world:

$$w' = w|m_0 * \{x\} * \emptyset^\infty.$$

Again, since $\mathcal{M}$ is consistent, there exists $m_1 \geq m_0$ such that for all $m' \geq m_1$:

$$p(\mathcal{M}_{w'|m'} = T_{w'} \mid X_{[e]} = s) > 1 - \epsilon/2.$$

Now define a new local loss function $\gamma_1 \leq \gamma_0$ as follows:

$$\gamma_1(s) = \begin{cases} 1 & \text{if } l(s) > m_0 \text{ and } \rho(s) = 1 \text{ and } \text{ and } \gamma_0^{<l(s)}(s) < a + n + 1; \\ 0 & \text{otherwise.} \end{cases}$$

Let $s' \in \Sigma^{m_1}$. Then by the definition of γ_1:

$$s' \geq s_0 \text{ and } \rho[^{m_1}_{m_0}(s) > 0 \text{ implies } \gamma_2(s')^{\geq c+k+1} \geq m_0.$$

Thus:

$$\begin{array}{rcl} p(\tau^{\gamma_2 \geq c+k+1} \geq m_0 \mid X_{[e]} = s) & \geq & p(X_{[w'|m_1]} \geq s_0 \text{ and } \rho[^{m_1}_{m_0}(X_{[w'|m_1]}) > 0 \mid X_{[e]} = s) \\ & \geq & p(X_{[w|m_1]}) \geq s_0 \mid X_{[e]} = s) - p(\rho[^{m_1}_{m_0}(X_{[w'|m_1]}) = 0 \mid X_{[e]} = s) \\ & \geq & p - p(\mathcal{M}_{[w'|m_0]} \neq T_w \text{ or } \mathcal{M}_{[w'|m_1]} \neq T_{w'} \mid X_{[e]} = s) \\ & \geq & p - p(\mathcal{M}_{[w'|m_0]} \neq T_w \mid X_{[e]} = s) + p(\mathcal{M}_{[w'|m_1]} \neq T_{w'} \mid X_{[e]} = s) \\ & \geq & p - 2(\epsilon/2) \\ & \geq & p - \epsilon. \end{array}$$

But $m_0 \geq \frac{k}{p-\epsilon}$, so:

$$\begin{array}{rcl} \mathsf{Exp}(\tau^{\gamma_2 \geq c+k+1} \mid X_{[e]} = s) & \geq & m_0 \cdot p(\tau^{\gamma_2 \geq c+k+1} \geq m_0 \mid X_{[e]} = s) \\ & \geq & \dfrac{k}{p-\epsilon} \cdot (p - \epsilon) \\ & = & k. \end{array}$$

□

Proposition 25 *For each $r > i$ and Boolean-valued local loss function γ:*

$$\mathit{Exp}(\gamma^{\geq 0}_{\mathcal{M},w} \mid X_{[e]} = s) \geq r \textit{ implies } \mathit{Exp}(\tau^{\gamma^{\geq i+1}}_{\mathcal{M},w} \mid X_{[e]} = s) > 0.$$

Proof. Suppose that:

$$\mathsf{Exp}(\tau^{\gamma^{\geq i+1}}_{\mathcal{M},w} \mid X_{[e]} = s) = 0.$$

Let Q denote the set of all $s' \in \Sigma^{<\omega}$ such that ($s \subseteq s'$ or $s' \subseteq s$) and $\gamma^{\geq 0}_{\mathcal{M},w}(s) \geq i + 1$. Let Q' be the set of all $s' \in Q$ such that for all $s'' < s$, $s'' \notin Q$. Then for each $s \in Q'$:

$$p(X_{[w|(l(s')-1)]} = s' \mid X_{[e]} = s) = 0.$$

So by countable additivity and the fact that $\Sigma^{<\omega}$ is countable:

$$\begin{aligned}
\mathsf{Exp}(\gamma^{\geq 0}_{\mathcal{M},w} \mid X_{[e]} = s) &\leq i + (i+1) \sum_{s \in Q} p(X_{[w|l(s)-1]} \mid X_{[e]} = s) \\
&\leq i < r.
\end{aligned}$$

□

9 Basic Lemmas

Proposition 26

$$\begin{aligned}
\mathit{Exp}(\rho^k_{\mathcal{M},w} \mid X_{[e]} = s) &= p(\mathcal{M}_e \neq \mathcal{M}_{e_-} \neq \text{‘?’} \mid X_{[e]} = s); \\
\mathit{Exp}(\epsilon^k_{\mathcal{M},w} \mid X_{[e]} = s) &= p(\mathcal{M}_e \notin \{T_w, \text{‘?’}\} \mid X_{[e]} = s).
\end{aligned}$$

Proof.

$$\begin{aligned}
\mathsf{Exp}(\rho^k_{\mathcal{M},w} \mid X_{[e]} = s) &= \int \rho_{\mathcal{M},w} \quad dp(. \mid X_{[e]} = s) \\
&= 0 \cdot p(\rho^k_{\mathcal{M},w} = 0 \mid X_{[e]} = s) + 1 \cdot p(\rho^k_{\mathcal{M},w} = 1 \mid X_{[e]} = s) \\
&= p(\rho^k_{\mathcal{M},w} = 1 \mid X_{[e]} s) \\
&= p(\mathcal{M}_e \neq \mathcal{M}_{e_-} \in \mathsf{Th}_K \mid X_{[e]} = s).
\end{aligned}$$

The argument for the second statement is similar. □

Proposition 27 *Let $\gamma \in \{\epsilon, \rho\}$. Then:*

$$1 \geq \sup_{w \in C_{K,e}(n)} \mathit{Exp}(\gamma^k_{\mathcal{M},w} \mid X_{[e]} = s).$$

Proof. Immediate from the definitions of $\epsilon^k_{\mathcal{M},w}$ and $\rho^k_{\mathcal{M},w}$. □

Proposition 28 *Let γ be a local loss function. Then:*

$$\begin{aligned}\sup_{w\in C_{K,e}(n)} \mathit{Exp}(\gamma^{\geq 0}_{\mathcal{M},w} \mid X_{[e]} = s) \quad \leq \quad & \sup_{w\in C_{K,e}(n)} \mathit{Exp}(\gamma^{<l(e)}_{\mathcal{M},w} \mid X_{[e]} = s) \\ & + \sup_{w\in C_{K,e}(n)} \mathit{Exp}(\gamma^{l(e)}_{\mathcal{M},w} \mid X_{[e]} = s) \\ & + \sup_{w\in C_{K,e}(n)} \mathit{Exp}(\gamma^{>l(e)}_{\mathcal{M},w} \mid X_{[e]} = s).\end{aligned}$$

The same is true if Exp *is replaced with* Inp.

Proof. Immediate. □

Proposition 29

$$\begin{aligned}\sup_{w\in C_{K,e}(n)} \mathit{Exp}(\rho^{\geq 0}_{\mathcal{M},w} \mid X_{[e]} = s) \quad = \quad & \sup_{w\in C_{K,e}(n)} \mathit{Exp}(\rho^{<l(e)}_{\mathcal{M},w} \mid X_{[e]} = s) \\ & + \sup_{w\in C_{K,e}(n)} \mathit{Exp}(\rho^{l(e)}_{\mathcal{M},w} \mid X_{[e]} = s) \\ & + \sup_{w\in C_{K,e}(n)} \mathit{Exp}(\rho^{>l(e)}_{\mathcal{M},w} \mid X_{[e]} = s).\end{aligned}$$

The same is true if Exp *is replaced with* Inp.

Proof. Let $k = l(e)$. By proposition 12:

$$\begin{aligned}\sup_{w\in C_{K,e}(n)} \mathsf{Exp}(\rho_{\mathcal{M},w} \mid X_{[e]} = s) \quad = \quad & \sup_{w\in C_{K,e}(n)} \big(\mathsf{Exp}(\rho^{<k}_{\mathcal{M},w} \mid X_{[e]} = s) \\ & \qquad + \mathsf{Exp}(\rho^{k}_{\mathcal{M},w} \mid X_{[e]} = s) \\ & \qquad + \mathsf{Exp}(\rho^{>k}_{\mathcal{M},w} \mid X_{[e]} = s)\big).\end{aligned}$$

Since $\mathcal{M}$ responds the same way along $w|k$ for each $w' \in C_{K,e}(n)$, it follows that, for each $w' \in C_{K,e}(n)$:

$$\begin{aligned}\mathsf{Exp}(\rho^{<k}_{\mathcal{M},w'} \mid X_{[e]} = s) \quad &= \quad \sup_{w\in C_{K,e}(n)} \mathsf{Exp}(\rho^{<k}_{\mathcal{M},w} \mid X_{[e]} s); \\ \mathsf{Exp}(\rho^{k}_{\mathcal{M},w'} \mid X_{[e]} s) \quad &= \quad \sup_{w\in C_{K,e}(n)} \mathsf{Exp}(\gamma^{k}_{\mathcal{M},w} \mid X_{[e]} = s).\end{aligned}$$

Hence:

$$\begin{aligned}\mathsf{Exp}(\rho_{\mathcal{M},w} \mid X_{[e]} = s) \quad = \quad & \sup_{w\in C_{K,e}(n)} \mathsf{Exp}(\rho^{<k}_{\mathcal{M},w} \mid X_{[e]} s) \\ & + \sup_{w\in C_{K,e}(n)} \mathsf{Exp}(\rho^{k}_{\mathcal{M},w} \mid X_{[e]} = s) \\ & + \mathsf{Exp}(\rho^{>k}_{\mathcal{M},w} \mid X_{[e]} = s).\end{aligned}$$

Since the first two terms in the sum are constants:

$$\begin{aligned}\sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\rho_{\mathcal{M},w} \mid X_{[e]} = s) \quad = \quad & \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\rho^{<k}_{\mathcal{M},w} \mid X_{[e]} s) \\ & + \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\rho^{k}_{\mathcal{M},w} \mid X_{[e]} = s) \\ & + \sup_{w \in C_{K,e}(n)} \mathsf{Exp}(\rho^{>k}_{\mathcal{M},w} \mid X_{[e]} = s).\end{aligned}$$

The logic depends only on suprema and sums, so it works as well for Inp. □

Proposition 30 *If $C_{K,e}(n) \neq \emptyset$ then:*

$$\mathit{Exp}(\rho^{l(e)}_{\mathcal{M},e} \mid X_{[e]} = s) = \sup_{w \in C_{K,e}(n)} \mathit{Exp}(\rho^{l(e)}_{\mathcal{M},w} \mid X_{[e]} s).$$

Proof. Note that if $w \in W_{K,e}$ then for each $\delta \in \Delta$:

$$\rho^{k}_{\mathcal{M},w}(\delta) = \rho^{k}_{\mathcal{M},w|l(e)}(\delta).$$

□