A Topological Theory of Empirical Simplicity and its Connection to the Truth

Kevin T. Kelly
Department of Philosophy
Carnegie Mellon University
kk3n@andrew.cmu.edu

June 8, 2007

Abstract

Simplicity is analyzed topologically. Ockham's razor is shown, under very general conditions, to be the unique, convergent strategy that jointly minimizes errors, retractions, and retraction times prior to convergence to the true theory.

1 Empirical Questions and Solutions

1.1 Input Streams

Let I be a set of potential **inputs**. An **input stream** is an element of I^ω. If w is an input stream let

$$w|i = (w(0), \ldots, w(i-1)),$$

so that $w|0 = ()$, where $()$ denotes the **empty sequence**. An **input sequence** of **length** n is a member of I^n. A **finite input sequence** is a finite input sequence of some length. If e is a finite input sequence, let $|e|$ denote the length of e, which is the same as the cardinality of e viewed as a set of ordered pairs. If e is a finite input sequence and e' is an input stream or finite input sequence, define $e \leq e'$ if and only if there exists n such that $e'|n = e$.

1.2 Empirical Problems

An **empirical problem** is a pair (K, Θ) where Θ is a partition of I^ω so that

$$\emptyset \subset K \subseteq \bigcup \Theta \subseteq I^\omega.$$

Then K is called the **empirical presupposition** of the problem and Θ is called the **question posed** by the problem. Let H_w denote the unique cell of Θ that contains (i.e., is **true of**) w. Define:

$$\widehat{K} = \{w|i : w \in K \text{ and } i \in \omega\},$$

which is the set of all finite input sequences compatible with K. If $e \in \widehat{K}$, then define:

$$\begin{aligned} K_e &= \{w \in K : e < w\}; \\ \widehat{K}_e &= \{e' \in \widehat{K} : e \leq e'\}. \end{aligned}$$

Proposition 1 (problem restriction) *Let (K, Θ) be a problem and let $e \in \widehat{K}$. Then (K_e, Θ) is a problem.*

Proof of proposition 1. Let $e \in \widehat{K}$. Then there exists $w \in K$ and $n \in \omega$ such that $e = w|k$. So $w \in K_e \neq \emptyset$. Furthermore, $K_e \subseteq K \subseteq \bigcup \Theta$. ∎

1.3 Solutions

Define:

$$\Theta^+ = \Theta \cup \{?\},$$

where ? is a distinguished object indicating an unwillingness to choose and answer in Θ. An **empirical strategy** for (K, Θ) is a map:

$$M : \widehat{K} \to \Theta^+.$$

An empirical strategy **solves** (K, Θ) (in the limit) if and only if:

$$(\forall w \in K)(\exists i)(\forall j \geq i)\ M(w|j) = H_w.$$

Then M is said to be a **solution** to (K, Θ) and (K, Θ) is said to be **solvable**.

1.4 Topology

The branching temporal structure of K generates a natural topology. The basis of the space is:

$$\Omega_K = \{K_e : e \in \widehat{K}\}.$$

Then let Ω^*_K denote the closure of Ω_K under union. Finally, define

$$\mathcal{T}_K = (K, \Omega^*_K),$$

which will be called the **branching future space** on K with **open sets** Ω^*_K. Topological concepts relative to $\mathcal{T}_K$ are indicated by the subscript K or by the more idiomatic phrase "**given** K", as in "**open**$_K$" or "**open given** K".

1.5 Sphere Systems and Induced Rankings

Following Grove (***), a **sphere system** for K is a downward-nested collection Λ of subsets of K such that $K = \bigcup \Lambda$. Sphere system Λ is **natural** if and only if the order type $\lceil \Lambda \rceil$ of the totally ordered set $(\Lambda, \supset)$ is $\leq \omega$. Thus, for each $i < \lceil \Lambda \rceil$, one may let Λ_i denote the ith element of Λ in the $\supset$ order. It is natural to think of differences between successive spheres as **levels** of complexity with respect to Λ. For each $i < \lceil\Lambda\rceil - 1$, define:

$$\overline{\Lambda}_i = \Lambda_i \smallsetminus \Lambda_j.$$

If $w \in K$, define the **rank** of w in Λ as:

$$\Lambda(w) = \sup_{i<\lceil\Lambda\rceil} w \in \Lambda_i.$$

The ranking can be extended to arbitrary propositions $P \subseteq K$ as follows:

$$\Lambda(P) \quad = \quad \min_{w \in P} \Lambda(w).$$

2 Sphere System Update

If $K' \subseteq K$, then define the **restriction** of Λ to K' as follows:

$$\Lambda|K = \{S \cap K' : S \in \Lambda\}.$$

In particular, abbreviate:

$$\begin{aligned}
\Lambda_e &= \Lambda|K_e; \\
\Lambda_{e,i} &= (\Lambda_e)_i; \\
\overline{\Lambda}_{e,i} &= (\overline{\Lambda_e})_i.
\end{aligned}$$

Next, define:

$$\Lambda(e) = \Lambda(K_e).$$

Proposition 2 ($\Lambda(e)$ is monotone) *Let Λ be a sphere system for K and let $e, e' \in \widehat{K}$. Then:*

$$e \leq e' \textit{ implies } \Lambda(e) \leq \Lambda(e').$$

Proof of proposition 2. Suppose that $e \leq e'$. Then

$$\Lambda(e) = \lambda(K_e) = \min_{w \in K_e} \Lambda(w) \geq \min_{w \in K_{e'}} \Lambda(w) = \lambda(K_{e'}) = \Lambda(e').$$

■

Proposition 3 ($\Lambda(e)$ is bounded) *Let Λ be a sphere system for K and let $w \in \overline{\Lambda}_n$. Then for each i:*

$$\Lambda(w|i) \leq n.$$

Proof of proposition 3. Let $w \in \Lambda_n$. Then $\Lambda(w) = n$. So by proposition 2, $\Lambda(w|i) \leq n$. ■

Proposition 4 (update) *Let (K, Θ) be a problem with sphere system Λ such that $K \subseteq \Lambda_0$ and let $e \in \widehat{K}$. Then:*

$$\overline{\Lambda}_{e,0} = K_e \cap \overline{\Lambda}_{(\Lambda(e))}.$$

Proof of proposition 4.

Recall that:

$$\Lambda_e = \{K_e \cap S : S \in \Lambda \text{ and } K_e \cap S \neq \emptyset\}.$$

Then $\Lambda_{e,i}$ is the ith entry in the descending enumeration of Λ_e and $\Lambda_{e,0}$ is the first entry in the enumeration, so:

$$\Lambda_{e,0} = K_e \cap \Lambda_0 = K_e.$$

Let $m = \Lambda(e)$. So m is least such that $K_e \cap \overline{\Lambda}_m \neq \emptyset$. Suppose that there exists no $i < \lceil\Lambda\rceil$ such that $K_e \cap \Lambda_i \subset K_e$. Then $\overline{\Lambda}_{e,0} = K_e$ and $m = 0$, so

$$\overline{\Lambda}_{e,0} = K_e = K_e \cap K = K_e \cap \overline{\Lambda}_0 = K_e \cap \overline{\Lambda}_m.$$

Now suppose that there exists $i \leq \lceil\Lambda\rceil$ such that $K_e \cap \Lambda_i \subset K_e$. Let $n =$ the least such i. Hence, $n > 0$ and $K_e \cap \Lambda_{n-1} = K_e$ and $\overline{\Lambda}_{e,0} = K_e \cap \Lambda_{n-1} \setminus \Lambda_n$. Furthermore, it follows that $m = n - 1$, since

$$K_e \cap (\Lambda_{n-1} \setminus \Lambda_n) = K_e \setminus \Lambda_n \neq \emptyset,$$

but for each $i < n$,

$$K_e \setminus \Lambda_i = K_e \setminus K_e = \emptyset.$$

Thus,

$$\overline{\Lambda}_{e,0} = K_e \cap \Lambda_m \setminus \Lambda_{m+1} = K_e \cap \overline{\Lambda}_{e,m}.$$

■

2.1 Data-driven Sphere Systems

Let (K, Θ) be a problem and let Λ be a sphere system for K. Say that Λ has **verifiable lower bounds** if and only if for each $i < \lceil\Lambda\rceil$:

$$\Lambda_i \text{ is open given } \Lambda_0.$$

This condition ensures that increasing information from world $w \in K$ eventually verifies that the rank of w is at least $\Lambda(w)$. Say that Θ is **decidable at each level** of ranking Λ if and only if for each $i < \lceil\Lambda\rceil$ and for each $H \in \Theta$:

$$(H \cap \overline{\Lambda}_i) \text{ is open given } \overline{\Lambda}_i.$$

The condition that $H \cap \overline{\Lambda}_i$ is open given $\overline{\Lambda}_i$ is equivalent to:

$$(\exists \text{ open } S \text{ given } \Lambda_0)\ S \cap \overline{\Lambda}_i = H \cap \overline{\Lambda}_i.$$

This condition ensures that, after it is verified that the complexity of w is at least $\Lambda(w)$, further information from w verifies the correct answer H_w given $\overline{\Lambda}_{\Lambda(w)}$. Say that Λ is **data-driven** for (K, Θ) if and only if:

1. Λ has verifiable lower bounds;

2. Θ is decidable at each level of Λ.

The motivation for for the terminology "data-driven" is recorded in the following proposition.

Proposition 5 (data-driven convergence) *Let (K, Θ) be a problem with data-driven sphere system Λ. Then:*

$$(\forall w \in K)(\exists i)(\forall j \geq i)\ w \in \overline{\Lambda}_{(w|j),0} \subseteq H_w.$$

Proof of proposition 5. Let $w \in K_e$. Since Λ_e is a sphere system for (K, Θ), there exists k such that $w \in \overline{\Lambda}_{e,k}$. Since Λ is data-driven for (K, Θ), it follows that Λ_k is open given K, so there exists n_0 such that for each $i \geq n_0$:

$$w \in K_{w|i} \subseteq \Lambda_k.$$

Let $i \geq n_0$. Then:

$$\Lambda_{(w|i),0} = K_{w|i} \cap \Lambda_k.$$

Hence, by nesting,

$$w \notin K_{w|i} \setminus \Lambda_k.$$

Furthermore:

$$\bigcup_{k'>0} \Lambda_{(w|i),k'} = K_{w|i} \cap \bigcup_{k'>k} \Lambda_{k'}.$$

Thus:

$$\begin{aligned}
\overline{\Lambda}_{(w|i),0} &= \Lambda_{(w|i),0} \setminus \bigcup_{k'>0} \Lambda_{(w|i),k'} \\
&= \left(K_{w|i} \cap \Lambda_k\right) \setminus \left(K_{w|i} \cap \bigcup_{k'>k} \Lambda_{k'}\right) \\
&= K_{w|i} \cap \left(\Lambda_k \setminus \bigcup_{k'>k} \Lambda_{k'}\right) \\
&= K_{w|i} \cap \overline{\Lambda}_k.
\end{aligned}$$

Furthermore, data-driven-ness implies that H_w is open given $\overline{\Lambda}_k$, so there exists n_1 such that for each $j \geq n_1$:

$$K_{w|j} \cap \overline{\Lambda}_k \subseteq H_w \cap \overline{\Lambda}_k.$$

Let $j \geq \max(n_0, n_1)$. Then:

$$w \in \overline{\Lambda}_{(w|j),0} \subseteq H_w.$$

■

Another nice feature of data-driven sphere systems is that the property of being data-driven is preserved under the arrival of new information.

Proposition 6 *Suppose that ranking Λ is data-driven for (K, Θ) and $e \in \widehat{K}$. Then Λ_e is data-driven for (K_e, Θ).*

Proof of proposition 6. Suppose that Λ is data-driven for (K, Θ). Let $e \in \widehat{K}$. (1) It is immediate from the definition that $K_e \in \Lambda_e$.

(2) Since each $S \in \Lambda$ is open given K, it follows that $S \cap K_e$ is open given K_e, so Λ_e has verifiable lower bounds.

(3) Let $H \in \Theta$, $i < \lceil \Lambda_e \rceil$. Then for some j,

$$\Lambda_{e,i} = \Lambda_j \cap K_e.$$

Since Λ is a simplicity concept for (K, Θ), there exists open S given K such that

$$H \cap \overline{\Lambda}_j \cap S = H \cap \overline{\Lambda}_j.$$

Then $(S \cap K_e)$ is open given K_e and:

$$\begin{aligned} H \cap \overline{\Lambda}_{e,i} \cap (S \cap K_e) &= H \cap (\overline{\Lambda}_j \cap K_e) \cap (S \cap K_e) \\ &= H \cap \overline{\Lambda}_j \cap S \cap K_e \\ &= H \cap \overline{\Lambda}_j \cap K_e \\ &= H \cap \overline{\Lambda}_{e,i}. \end{aligned}$$

■

2.2 Retractions and Data-driven Sphere System Existence

A third nice feature of data-driven sphere systems is that every solvable problem has one. Say that M **retracts** at e if and only if e is non-empty and

$$? \neq M(e_-) \neq M(e).$$

Then define:

$$\begin{aligned} R_M(w) &= \{i \in \omega : M \text{ retracts at } i \text{ on } w\}; \\ r_M(w) &= |R_M(w)|; \\ S_{M,i} &= \{w \in K : r_M(w) \geq i\}; \\ \Lambda_M &= \{S_{M,i} : i < \omega\} \smallsetminus \{\emptyset\}. \end{aligned}$$

There is no guarantee that M retracts sensibly, so it is not necessarily the case that $S_{M,i} \neq S_{M,i+1}$, even if both are nonempty.

Proposition 7 *If M solves (K, Θ), then Λ_M is data-driven.*

Proof of proposition 7. Let M solve (K, Θ). Let $w \in K$. (1) Then since M converges to H_w in e, $r_M(w)$ is finite, so $K = S_{M,0} \in \Lambda_M$.

(2) Next, observe that:

$$\Lambda_i = \bigcup\{K_e : r_M(e) \geq i\},$$

so Λ_i is open.

(3) Finally, let $H \in \Theta$ and let $i < \lceil \Lambda \rceil$. Then there exists j such that

$$\overline{\Lambda}_i = S_{M,j}.$$

Define open:

$$W_j = \bigcup\{K_e : r_M(e) = j \text{ and } M(e) = H\}.$$

Since M is a solution,

$$W_j \cap S_{M,j} \cap H = S_{M,j} \cap H.$$

■

Proposition 8 (forcible path update) *Let $\Lambda|S$ be a forcible path through Λ in Θ and $e \in \widehat{K}$ and $\overline{\Lambda}_{e,0} \cap S \neq \emptyset$. Then for each i such that $\Lambda(e) \leq i < \lceil \Lambda \rceil$:*

$$(\overline{\Lambda}_{e,(i-\Lambda(e))} \cap S) = (\overline{\Lambda}_i \cap K_e \cap S) \neq \emptyset.$$

Proof of proposition 8, by induction on $i - \Lambda(e)$. Suppose that $\Lambda|S$ is a forcible path through Λ in Θ and $e \in \widehat{K}$ and $\overline{\Lambda}_{e,0} \cap S \neq \emptyset$. The base case is by hypothesis and proposition 4. Now, suppose that $\Lambda(e) \leq i + 1 < \lceil \Lambda \rceil$. By induction hypothesis, let:

$$(*)\ w \in (\overline{\Lambda}_{e,(i-\Lambda(e))} \cap S) = (\overline{\Lambda}_i \cap K_e \cap S).$$

Since $\Lambda \cap S$ traverses Λ and $\Lambda(e) \leq i + 1 < \lceil \Lambda \rceil$, it follows that $\overline{\Lambda}_{i+1} \cap S \neq \emptyset$, so since $\Lambda|S$ is forcible in Θ, w is a limit point of $\overline{\Lambda}_{i+1} \cap S \setminus H_w$. Thus, since $w \in K_e$ and K_e is open given K:

$$\overline{\Lambda}_{i+1} \cap S \cap K_e \neq \emptyset.$$

So by (*) and the definition of $\overline{\Lambda}_{e,((i+1)-\Lambda(e))}$,

$$(\overline{\Lambda}_{e,((i+1)-\Lambda(e))} \cap S) = (\overline{\Lambda}_{i+1} \cap S \cap K_e) \neq \emptyset.$$

■

Proposition 9 (forcible path preservation) *Let $\Lambda|S$ be a forcible path through Λ in Θ and let $e \in \widehat{K}$. Then:*

$$\overline{\Lambda}_{e,0} \cap S \neq \emptyset \textit{ implies that } \Lambda_e|S \textit{ is a forcible path through } \Lambda_e.$$

Proof of proposition 9. Suppose that $\Lambda|S$ is a forcible path through Λ in Θ and $e \in \widehat{K}$ and $\overline{\Lambda}_{e,0} \cap S \neq \emptyset$. By proposition 8, one obtains that for each i such that $\Lambda(e) \leq i < \lceil\Lambda\rceil$:

$$(\overline{\Lambda}_{e,(i-\Lambda(e))} \cap S) = (\overline{\Lambda}_i \cap K_e \cap S) \neq \emptyset.$$

Each level of Λ_e is the restriction of some level in Λ to K_e, so:

$$\lceil\Lambda_e\rceil = \lceil\Lambda\rceil - \Lambda(e).$$

Hence, for each i such that $0 \leq i < \lceil\Lambda\rceil$:

$$(\overline{\Lambda}_{e,i} \cap S) = (\overline{\Lambda}_{(i+\Lambda(e))} \cap K_e \cap S) \neq \emptyset.$$

Therefore, $\Lambda_e|S$ traverses Λ_e.

Suppose that $0 \leq i \leq \lceil\Lambda_e\rceil$ and that $w \in \overline{\Lambda}_{e,i} \cap S$. Suppose, further, that $\Lambda_{e,(i+1)} \cap S \neq \emptyset$. Since $\overline{\Lambda}_{e,0} \cap S \neq \emptyset$, applications of proposition 8 to the preceding two claims yield, respectively:

$$w \in (\overline{\Lambda}_{(\Lambda(e)+i)} \cap K_e \cap S),$$

and

$$(\overline{\Lambda}_{(\Lambda(e)+i+1)} \cap K_e \cap S) \neq \emptyset.$$

So, since $\Lambda|S$ is forcible in Θ, it follows that w is a limit point of $(\overline{\Lambda}_{(\Lambda(e)+i+1)} \cap K_e \cap S) \setminus H_w$. So, by proposition 8, again, w is a limit point of $\overline{\Lambda}_{e,(i+1)} \cap S$. Hence, $\Lambda_e|S$ is forcible in Θ. ■

2.3 Answer-Preservation

Say that Λ **preserves answers** if and only if:

$$(\forall H \in \Theta)(\exists i < \lceil\Lambda\rceil)\ H \subseteq \overline{\Lambda}_i.$$

Otherwise, it **splits** answers across levels. A data-driven sphere system might split answers across levels. For example, suppose that input streams are Boolean, let K be the set of all streams that converge to zero and let the question be whether the total number of unit bits is even or odd. Every data-driven ranking for this problem has to split answers across levels. Skipping worlds would violate condition 1. Putting each answer on its own level would violate condition 2. Putting both answers at the same level would violate condition 3. But in such cases, one can always force the sphere system to preserve answers by asking the **refined** question:

$$\Theta \wedge \Lambda = \{H \cap \overline{\Lambda}_i : H \in \Theta \text{ and } i < \lceil\Lambda\rceil\}.$$

Proposition 10 *If Λ is a data-driven sphere system for (K, Θ), then Λ is a data-driven, answer-preserving sphere system for $(K, (\Theta \wedge \Lambda))$.*

Proof of proposition 10. Conditions (1) and (2) for a data-driven sphere system make no reference to Θ. Also, if H is open given $\overline{\Lambda}_i$, then $H \cap \overline{\Lambda}_i$ is open given $\overline{\Lambda}_i$. ■

One could view the refinement of the original question either ontologically or epistemically. Ontologically, simplicity is a relevant fact that it would be hard to argue is of no interest at all: surely, "even" because of a billion units is interestingly different than "even" because of no units. Epistemically, insofar as Ockham's razor "justifies" inductive conclusions, believing in "even" *because* one believes in "zero" even though the truth is "one billion" is a kind of Gettier case that is not penalized by merely counting retractions. Forcing the scientist to estimate the complexity of the actual world allows retractions to penalize these Gettier episodes.

The point of answer-preservation and forcible paths is summarized as follows.

Proposition 11 *Let (K, Θ) be a problem, for which Λ is an answer-preserving, data-driven sphere system for (K, Θ). Let $e \in \widehat{K}$ and let $\lceil \Lambda_e \rceil \geq n+1$. Let M solve (K, Θ) from e onward. Let $\Lambda_e|S$ be a forcible path through Λ_e. Let j be arbitrary. Then there exists $w_n \in (\overline{\Lambda}_{e,n} \cap S)$ and list of distinct answers $(H_0, \ldots, H_n)$ such that for each $m \leq n$, $H_m \subseteq \overline{\Lambda}_{e,m}$ and such that, in w_n, solution M produces answer H_0 for j successive times starting no sooner than $|e|+1$ and $\ldots$ and produces answer H_n for j successive times starting no sooner than $|e|+1+nj$.*

Proof of proposition 11, by induction on n. Let $n = 0$, so $\lceil \Lambda_e \rceil = 1$. Then $\overline{\Lambda}_{e,0} \cap S \neq \emptyset$, so let $w_0 \in \overline{\Lambda}_{e,0} \cap S \neq \emptyset$. Since M is a solution, there exists $i_0 \geq |e|+1$ such that for all $i \geq i$, $M(w_0|i) = H_{w_0}$. Hence, M produces H_{w_0} for j successive times starting no sooner than $|e|+1$ in w_0. So let the list be w_0 and the unit list (H_{w_0}) witness the claim for $n = 0$.

Now, suppose that $\lceil \Lambda_e \rceil \geq n+2$. Let $w_n \in \overline{\Lambda}_{e,n}$ and list $L = (H_0, \ldots, H_n)$ of answers be as specified by the induction hypothesis. By proposition 9, $\Lambda_e|S$ is a forcible path through Λ_e. Since $\Lambda_e|S$ is forcible in Θ and $w_n \in \overline{\Lambda}_{e,n}$ and $\lceil \Lambda_e \rceil \geq n+2$, it follows that w_n is a limit point of $\overline{\Lambda}_{e,n+1}$. Let e_n be the least initial segment along w_n by which M has produced each entry in list L for the j times promised by the induction hypothesis. Then, since K_{e_n} is open, it follows that there exists $w_{n+1} \in \overline{\Lambda}_{e_n,n+1} \cap S$. Since M is a solution, there exists $i_{n+1} \geq |e_n|+1$ such that for each $i \geq i_{n+1}$, $M(w|i) = H_{w_{n+1}}$. Since Ω preserves answers, $H_{w_{n+1}} \subseteq \overline{\Lambda}_{e,(n+1)}$, so $H_{w_{n+1}}$ is distinct from all the preceding answers in the list. Then w_{n+1} and list $L * H_{w_{n+1}}$ witness the claim for $n+1$. ■

2.4 Empirical Simplicity

Data-driven sphere systems factor solvable problems into layers. In many cases, these layers look like what is intuitively regarded as degrees of empirical complexity. But a further property is required to rule out gratuitous distinctions in complexity. Say that sphere system Λ is **forcible** if and only if for each $i < \lceil \Lambda \rceil - 1$:

$$\overline{\Lambda}_i \subseteq \mathrm{bdry}(\overline{\Lambda}_{i+1}).$$

This property says that for each world $w \in \Lambda_0$ that is not at the last level of Λ, an arbitrary amount of information from w is compatible with the possibility that the true world is at the next level of Λ and makes H_w false. So one might say that each non-terminal level poses the problem of induction with respect to the succeeding level. Let $S \subseteq \Lambda_0$. Say that $\Lambda|S$ **traverses** Λ if and only if for each $i < \lceil \Lambda \rceil$:

$$\overline{\Lambda|S}_i \cap \Lambda_i \neq \emptyset.$$

In other words, S includes at least one element of each level of Λ. Say that $\Lambda|S$ is a **forcible path through** Λ if and only if:

1. $\Lambda|S$ is forcible;

2. $\Lambda|S$ traverses Λ.

Say that Λ is a **simplicity concept** for (K, Θ) if and only if:

1. Λ is a data-driven sphere system for (K, Θ);

2. Λ preserves answers in Θ;

3. $(\forall e \in \widehat{K})(\exists S \subseteq K_e)$ $\Lambda|S$ is a forcible path through Λ_e in Θ.

If, in addition, Λ is a forcible path through Λ in Θ, then say that Λ is a **strong** simplicity concept for (K, Θ). If (K, Θ) is a problem with [strong] simplicity concept Λ, say that (K, Θ, Λ) is a [strongly] **simplified problem**.

Proposition 12 (simplicity preservation) *Let (K, Θ, Λ) be a simplified problem. Then (K_e, Θ, Λ_e) is a simplified problem.*

Proof of proposition 12. Suppose that ranking Λ is a simplicity ranking for (K, Θ). Then Λ is data-driven for (K, Θ), so Λ_e is data-driven for (K_e, Θ) by proposition 6. Λ_e preserves answers if Λ does, since restriction never separates levels. Finally, suppose that $e' \in \widehat{K}_e$. Then there exists $K' \subseteq K_e$ such that $\Lambda|K'$ is a forcible path through $\Lambda|K_e$. ■

An important open question is:

Question 1 *Let (K, Θ) be a solvable problem. Under what conditions does there exist a simplicity concept for (K, Θ)?*

According to proposition 7, the question reduces to the forcible path condition. But obtaining forcible paths may require some surgery on Λ_M. Alternatively, one might try to impose some extra conditions on the solution M that ensure the the forcible path condition but that can be shown not to preclude convergence to the truth.

2.5 Univocal Problems

Say that simplified problem (K, Θ, Λ) is **univocal** if and only if for each $e \in \widehat{K}$, there exists $H \in \Theta$ such that:

$$\overline{\Lambda}_{e,0} \subseteq H.$$

One way for a problem to be univocal is for it to be **linear**, in the sense that for each $i < \lceil \Lambda \rceil$,

$$\overline{\Lambda}_i \in \Theta.$$

Otherwise, the problem is **branching**. Less restrictively, accounting problems are univocal if effects arrive unambiguously (i.e., no disjunctive announcements of effects) and no possible sets of effects are ruled out.

2.6 Ockham's Razor

Let (K, Θ, Λ) be a simplified problem. Let $e \in \widehat{K}$.

Define:

$$\mathrm{Simp}_\Lambda(e) = \{H \in \Theta : \overline{\Lambda}_{e,0} \cap H \neq \emptyset\}.$$

Say that $H \in \Theta$ is **simplest** in (K, Θ, Λ) if and only if

$$H \in \mathrm{Simp}_\Lambda(e).$$

Say that $H \in \Theta$ is **Ockham** given e in (K, Θ, Λ) if and only if:

$$\mathrm{Simp}_\Lambda(e) = \{H\}.$$

Proposition 13 (Ockham answer characterization) *Let (K, Θ, Λ) be a simplified problem, let $H \in \Theta$, let $e \in \widehat{K}$ and let $K_e \cap H \neq \emptyset$. Then:*

$$Simp_\Lambda(e) = \{H\} \text{ iff } \overline{\Lambda}_{e,0} \subseteq H.$$

Proof of proposition 13. Suppose that $\mathrm{Simp}_\Lambda(e) = \{H\}$. Suppose for contradiction that $w \notin H$. Then since $K \subseteq \bigcup \Theta$, there exists $H' \in \Theta$ such that $H' \neq H$ and $H' \in \mathrm{Simp}_\Lambda(e)$, so $\mathrm{Simp}_\Lambda(e) \neq \{H\}$. Hence, $w \notin \overline{\Lambda}_{e,0}$.

Conversely, suppose that $\overline{\Lambda}_{e,0} \subseteq H$. Then $H \in \mathrm{Simp}_\Lambda(e)$. Let $H' \in \Theta$ such that $H' \neq H$. Then $H \cap H' = \emptyset$, since (K, Θ) is a problem. Hence, $\overline{\Lambda}_{e,0} \cap H' = \emptyset$, so $H' \notin \mathrm{Simp}_\Lambda(e)$. So $\mathrm{Simp}_\Lambda(e) = \{H\}$. ■

Say that M is **Ockham** at e if and only if:

$$M(e) = ? \text{ or } M(e) \text{ is Ockham at } e.$$

Say that M is **weakly Ockham** at e if and only if:

$$M(e) = ? \text{ or } M(e) \text{ is simplest at } e.$$

If M is not Ockham at e, say that M **violates** Ockham's razor at e and that M violates Ockham's razor **strongly** at e if M is not simplest at e.

2.7 Normal Ockham Strategies

Strategy M is **responsive** at e in (K, Θ) if and only if

$$(\forall w \in K_e)(\forall i > |e|)(\exists j > i)\ M(e|j) \neq ?.$$

Strategy M is **stalwart** at e if and only if:

$$\text{if } e \neq () \text{ and } ? \neq M(e_-) \neq M(e), \text{ then } M(e_-) \text{ is not Ockham at } e.$$

Finally, say that M is **normally Ockham** at e in (K, Θ, Λ) if and only if

M is Ockham, stalwart, and responsive at e in (K, Θ, Λ).

2.8 Looking Forward

For each relation $\Phi(K, \Theta, \Lambda, e)$, one can say that Φ holds from e **onward** if and only if:

$$(\forall e' \in \widehat{K}_e)\ \Phi(K, \Theta, \Lambda, e').$$

Furthermore, Φ holds **everywhere** if and only if:

$$(\forall e \in \widehat{K})\ \Phi(K, \Theta, \Lambda, e).$$

Thus, M can be a solution to (K, Θ) from e onward, Ockham from e onward or everywhere, stalwart from e onward or everywhere, etc.

3 Normal Ockham Convergence

One reason to follow a normal Ockham strategy is that it converges to the truth. The trouble is that many possible strategies converge to the truth. It remains to argue that normal Ockham strategies do so more efficiently than all other strategies.

Proposition 14 (normal Ockham convergence) *Let (K, Θ, Λ) be a simplified problem, let $e \in \widehat{K}$ and let M be normally Ockham from e onward in (K, Θ, Λ). Then M solves (K_e, Θ, Λ_e).*

Proof of proposition 14. Let $e \in \widehat{K}$ and let $w \in K_e$. So by proposition 5, there exists n_0 such that for each $i \geq n_0$:

$$w \in \overline{\Lambda}_{(w|j),0} \subseteq H_w.$$

Let $i \geq n_0$. So by proposition 13, H_w is Ockham given $w|i$ in (K, Θ, Λ). Since M is responsive from e onward, there exists $i > n_1$ such that $M(w|i) \neq ?$. Since M is Ockham from e onward, $M(w|i) = H_w$. Since M is stalwart from e onward, for each $j \geq i$, $M(w|j) = H_w$. ■

3.1 Normal Ockham Strategy Existence

Here is an obvious way to construct a normal Ockham solution for an arbitrary, simplified problem (K, Θ, Λ). Notice that this is the **most aggressive** such solution, in the sense that it leaps to the simplest answer immediately. Stalling with ? for an arbitrary amount of time is also permitted.

$$M_\Lambda(e) \quad = \quad \begin{cases} H & \text{if } \mathrm{Simp}_\Lambda(e) = \{H\}; \\ ? & \text{otherwise.} \end{cases}$$

Proposition 15 (normal Ockham existence) *Let (K, Θ, Λ) be a simplified problem. Then M_Λ is normally Ockham for (K, Θ, Λ).*

Proof of proposition 15. The Ockham and Stalwart properties are immediate. Responsiveness follows from propositions 5 and 13. ■

Say that $M \simeq_e M'$ if and only if for each $n < |e|$,

$$M(e|n) = M'(e|n).$$

This relationship is **agreement along** e. It is not too hard to see that each strategy agrees along e with a strategy that is normally Ockham thereafter. If M_0 is an arbitrary strategy for (K, Θ), define:

$$\tilde{M}_\Lambda(e') \quad = \quad \begin{cases} M(e') & \text{if } e' \leq e_-; \\ M_\Lambda(e') & \text{otherwise.} \end{cases}$$

Proposition 16 (normal Ockham extension) *Let (K, Θ, Λ) be a simplified problem and let $e \in \widehat{K}$. Let M be a strategy. Then $\tilde{M}_\Lambda \simeq_{e_-} M$ and $\tilde{M}_\Lambda$ is normally Ockham from e onward in (K, Θ, Λ).*

Proof of proposition 16. By proposition 15, $\tilde{M}_{0\Lambda}$ is normally Ockham from e onward. ■

4 Efficiency

Let the total number of errors committed by M in w be denoted by $\epsilon(M, w)$ and let $\rho(M, w)$ denote the sequence τ of times at which M retracts in w (which is finite for convergent strategies) is the total number of retractions performed. The cost ρ is called **timed retractions**. Let the **composite loss** of strategy M on $w \in K$ be given by the pair:

$$\lambda(M, w) = (\epsilon(M, w), \rho(M, w)).$$

If σ, τ are finite, ascending sequences of natural numbers, define:

$$\sigma \preceq \tau \quad \text{iff} \quad \text{there exists a subsequence } \gamma \text{ of } \tau \text{ such that for each } i \leq \mathrm{length}(e), e(i) \leq \gamma(i).$$

Hence, $(1, 3, 7) \leq (2, 3, 4, 8)$ in virtue of sub-sequence $(2, 3, 8)$. Then if (b, σ) and (c, τ) are both cumulative costs, define:

$$
\begin{array}{lll}
(b, \sigma) \preceq (c, \tau) & \text{iff} & b \preceq c \text{ and } \sigma \preceq \tau. \\
(b, \sigma) \cong (c, \tau) & \text{iff} & (b, \sigma) \preceq (c, \tau) \text{ and } (c, \tau) \preceq (b, \sigma); \\
(b, \sigma) \prec (c, \tau) & \text{iff} & (b, \sigma) \preceq (c, \tau) \text{ and } (c, \tau) \npreceq (b, \sigma).
\end{array}
$$

5 Worst-case Cost Bounds

A **potential cost bound** is a pair (b, σ), where $b \leq \omega$ and σ is a finite or infinite, non-descending sequence of entries $\leq \omega$ in which no finite entry occurs more than once. If (b, σ) is a cost vector and (c, τ) is a cost bound, then $(b, \sigma) \leq (c, \tau)$ can be defined just as for cost vectors, themselves. Cost bounds $(c, \tau), (d, \gamma)$ may now be compared as follows:

$$
\begin{array}{lll}
(c, \tau) \leq (d, \gamma) & \text{iff} & \text{for each cost vector} (b, \sigma), \text{ if } (b, \sigma) \leq (c, \tau) \text{ then } (b, \sigma) \leq (d, \gamma); \\
(c, \tau) \equiv (d, \gamma) & \text{iff} & (c, \tau) \leq (d, \gamma) \text{ and } (d, \gamma) \leq (c, \tau); \\
(c, \tau) < (d, \gamma) & \text{iff} & (c, \tau) \leq (d, \gamma) \text{ and } (d, \gamma) \not\leq (c, \tau).
\end{array}
$$

It follows, for example, that $(4, (2)) < (\omega, (2, \omega)) < (\omega, (0, 1, 2, \ldots)) \equiv (\omega, (\omega, \omega, \omega, \ldots))$. Now, each set C of cost vectors has a unique (up to equivalence) least upper bound $\sup(C)$ among the potential upper bounds (Kelly 2006). Then one may define the **worst-case cost** of strategy M given e over $\overline{\Lambda}_n$ as follows:

$$\lambda_n(M, e) = \sup_{w \in \overline{\Lambda}_n} \lambda(M, w).$$

The same approach can be taken toward worst-case bounds over retraction sequences alone.

5.1 Upper Cost Bound for Normal Ockham Solutions

Proposition 17 (normal Ockham retractions) *Let (K, Θ, Λ) be a simplified problem, let $e \in \widehat{K}$ and let M be normally Ockham for (K, Θ, Λ) from e onward. Then if M retracts at $e' \in K_e$,*

$$\Lambda(e'_-) < \Lambda(e').$$

Proof of proposition 17. Under the proposition's hypothesis, suppose that M retracts at $e' \in K_e$. Then, since M is Ockham, propositions 4 and 13 imply that for some $H \in \Theta$:

$$K_{e'_-} \cap \overline{\Lambda}_{\Lambda(e'_-)} \subseteq H,$$

and since M is stalwart, proposition 13 implies that:

$$K_{e'} \cap \overline{\Lambda}_{\Lambda(e')} \not\subseteq H.$$

Hence,

$$\Lambda(e'_-) \neq \Lambda(e').$$

So by proposition 2:

$$\Lambda(e'_-) < \Lambda(e').$$

■

Let $a(M, e, w)$ denote the number of errors committed along e by M in world w. Then define:

$$a_0(M, e) = \max_{w \in \overline{\Lambda}_{e,0}} a(M, e_-, w).$$

Also, let: $\sigma(M, e)$ denote the sequence of times at which M retracts along e. Define:

$$\omega^{(n)} = (\underbrace{\omega \dots \omega}_{n}).$$

Let $*$ denote sequence concatenation.

Proposition 18 (normal Ockham upper bound) *Let (K, Θ, Λ) be a simplified problem, let $e \in \widehat{K}$, let M be normally Ockham for (K, Θ, Λ) from e onward. Let $a = a_0(M, e)$ and let $\sigma = \sigma_{e_-}(M, e_-)$. Then:*

$$\lambda_0(M, e) \leq \begin{cases} \textit{1.} \quad (a, \sigma * |e|) & \textit{always;} \\ \textit{2.} \quad (a, \sigma) & \textit{if } M \textit{ does not retract at } e. \end{cases}$$

$$\lambda_{n+1}(M, e) \leq \begin{cases} \textit{3.} \quad (\omega, (\sigma * |e| * \omega^{(n)})) & \textit{always;} \\ \textit{4.} \quad (\omega, (\sigma * \omega^{(n)})) & \textit{if } M \textit{ does not retract at } e. \end{cases}$$

Proof of proposition 18.

Proof of 1. Let $w \in \overline{\Lambda}_{e,0}$ and let $i \geq |e|$. Suppose that $M(w|i) = H \neq ?$. Then, since M is Ockham from e onward, proposition 13 yields that $\overline{\Lambda}_{(w|i),0} \subseteq H$. Hence, $H = H_w$, so M commits no error at $w|i$. Furthermore, since $w \in \overline{\Lambda}_{e,0}$, it follows that $\overline{\Lambda}_{(w|(i+1)),0} = \overline{\Lambda}_{(w|i),0} \cap K_{w|(i+1)} \subseteq H_w$, so H_w is also Ockham at $w|(i+1)$, by proposition 13. Since M is stalwart, $M(w|(i+1)) = H_w$, so M does not retract at $i+1$ in w. Then: $\lambda(M, w) \leq (a_0(M, e), \sigma(M, e_-) * |e|)$, so since w is an arbitrary element of $\overline{\Lambda}_{e,0}$:

$$\lambda_0(M, e) \leq (a_0(M, e), (\sigma(M, e_-) * |e|)).$$

Proof of 3. Let $w \in \overline{\Lambda}_{e,(n+1)}$ and let $i > |e|$. Suppose that M retracts at i. Then by propositions 17 and 2,

$$\Lambda_e(w|(i-1)) < \Lambda_e(w|i).$$

So by proposition 3, M retracts at most $n+1$ times after e in w. Hence:

$$\lambda_{n+1}(M, e) \leq (\omega, (\sigma(M, e_-) * |e| * (\underbrace{\omega \dots \omega}_{n+1}))).$$

Proofs of 2,4. If M does not retract at e, the arguments are the same as for 1, 3, except that the "extra" retraction at $|e|$ is dropped. ■

5.2 Lower Cost Bound for Arbitrary Solutions

Proposition 19 (invariance of Ockham errors) *Let (K, Θ, Λ) be a simplified problem and let $e' < e \in \widehat{K}$ and let $H \in \Theta$ be Ockham in (K, Θ, Λ) given e'. Then:*

$$\overline{\Lambda}_{e,0} \subseteq H \text{ or } \overline{\Lambda}_{e,0} \cap H = \emptyset.$$

Proof. Let $e' < e \in \widehat{K}$ and suppose that H is Ockham given e'. Then, by proposition 13, $\overline{\Lambda}_{e',0} \subseteq H$.

Case: $\Lambda_{e'} = \Lambda_e$. Then, by proposition 8, $\overline{\Lambda}_{e,0} \subseteq \overline{\Lambda}_{e',0}$, so $\overline{\Lambda}_{e,0} \subseteq H$. Hence, for each $w' \in \overline{\Lambda}_{e,0}$, H is not an error in w'.

Case: $\Lambda_{e'} \neq \Lambda_e$. Then, by proposition 8, $\overline{\Lambda}_{e,0} \cap \overline{\Lambda}_{e',0} = \emptyset$, so since Λ_e preserves answers, H is an error in each $w' \in \overline{\Lambda}_{e,0}$. ■

Proposition 20 (lower bound for arbitrary solutions) *Let (K, Θ, Λ) be a simplified problem, let $e \in \widehat{K}$, let M be a solution to (K, Θ) from e onward. Let $a = a_0(M, e)$ and $\sigma = \sigma_{e_-}(M, e_-)$. Then:*

$$\lambda_0(M,e) \quad \geq \quad \begin{cases} \textit{1.} \quad ((a+1), (\sigma * (|e|+1))) & \textit{if } M(e) \textit{ is not simplest at } e; \\ \textit{2.} \quad ((a+1), (\sigma * (|e|+1))) & \textit{if } M \textit{ does not retract at } e \textit{ and} \\ & \quad e \textit{ is the least } e' \leq e \textit{ such that} \\ & \quad M(e') \textit{ is not Ockham at } e'; \\ \textit{3.} \quad ((a+1), (\sigma * |e| * (|e|+1))) & \textit{if } M \textit{ retracts at } e \textit{ and} \\ & \quad e \textit{ is the least } e' \leq e \textit{ such that} \\ & \quad M(e') \textit{ is not Ockham at } e'; \\ \textit{4.} \quad (a, (\sigma * |e|)) & \textit{if } M \textit{ retracts at } e; \\ \textit{5.} \quad (a, \sigma) & \textit{always.} \end{cases}$$

If $n + 1 < \lceil \Lambda_e \rceil$, then:

$$\lambda_{n+1}(M,e) \quad \geq \quad \begin{cases} \textit{6.} \quad (\omega, (\sigma * (|e|+1) * \omega^{(n+1)})) & \textit{if } e \textit{ is not simplest at } e; \\ \textit{7.} \quad (\omega, (\sigma * (|e|+1) * \omega^{(n+1)})) & \textit{if } M \textit{ does not retract at } e \textit{ and} \\ & \quad e \textit{ is the least } e' \leq e \textit{ such that} \\ & \quad M(e') \textit{ is not Ockham at } e'; \\ \textit{8.} \quad (\omega, (\sigma * |e| * \omega^{(n+1)})) & \textit{if } M \textit{ retracts at } e; \\ \textit{9.} \quad (\omega, (\sigma * \omega^{(n+1)})) & \textit{always.} \end{cases}$$

Proof of proposition 20. Let M be a solution from e onward and suppose that $e \in \widehat{K}$.

Proof of 1. Suppose that $H = M(e)$ is not simplest at e. So $\overline{\Lambda}_{e,0} \cap H = \emptyset$. Choose $w \in \overline{\Lambda}_{e,0}$ so that M commits $a_0(M, e)$ errors along e_- in w. Then, since $\overline{\Lambda}_{e,0} \cap H = \emptyset$, answer H is another error in w, for a total of $a + 1$ errors. Furthermore, since M is a solution, there exists $i > |e|$ such that $M(w|i) \neq H$, so $\lambda_0(M, e) \geq ((a+1), (\sigma*(|e|+1)))$.

Proof of 2. Suppose that $H = M(e)$ is not Ockham at e and that $H = M(e) = M(e_-)$ and that for each $e' < e$, either $M(e') =?$ or $M(e')$ is Ockham at e. Hence, $M(e_-) = H$ and H is Ockham at e_-. So $\overline{\Lambda}_{e_-,0} \subseteq H$ and $\overline{\Lambda}_{e,0} \not\subseteq H$. Hence, $\overline{\Lambda}_{e,0} \neq \overline{\Lambda}_{e_-,0}$. So, since Λ preserves answers, $\overline{\Lambda}_{e,0} \cap H = \emptyset$, so H is not simplest at e. Apply (1).

Proof of 3. Suppose that $H = M(e)$ is not Ockham at e and that M retracts at e and that for each $e' < e$, either $M(e') =?$ or $M(e')$ is Ockham at e. By proposition 13, there exists $w \in H \setminus \overline{\Lambda}_{e,0}$. By proposition 19, $a(M, e_-, e) = a_0(M, e)$. So $a(M, e, e) = a_0(M, e) + 1$. Since $w \notin H$ and M is a solution, M converges to $H_w \neq H$ at some stage $i > |e|$. Hence, $\lambda(M, w) \geq ((a + 1), (\sigma * |e| * (|e| + 1)))$.

Proof of 4. Choose $w \in \overline{\Lambda}_{e,0}$ such that M commits $a_0(M, e)$ errors along e_- in w. Then, since M retracts at e, $\lambda_0(M, e) \geq (a, (\sigma * |e|))$.

Proof of 5. Like the proof of (4), ignoring retraction at $|e|$.

Proof of 6. Suppose that $\lceil \Lambda_e \rceil > n + 1$, so that $\overline{\Lambda}_{e,(n+1)} \neq \emptyset$. Since Λ is a simplicity concept for (K, Θ), there exists S such that $\Lambda_e|S$ is a forcible path through Λ_e. Since S traverses Λ_e and $\overline{\Lambda}_{e,(n+1)} \neq \emptyset$, it follows that $\overline{\Lambda}_{e,0} \cap S \neq \emptyset$, so let $w_0 \in \overline{\Lambda}_{e,0} \cap S \neq \emptyset$. Suppose that $H = M(e)$ is not simplest at e. Let $j \in \omega$ be arbitrary. Since H is not simplest at e, $\overline{\Lambda}_{e,0} \cap H = \emptyset$, so $H_{w_0} \neq H$. Since M is a solution, there exists $i_0 > |e|$ such that $M(e|i_0) = H_w \neq H$. Let $e_0 = w|i_0$. Since $w_0 \in \overline{\Lambda}_{e,0} \cap S$, $\Lambda_e(e_0) = 0$, so by proposition 8, it is the case that $\lceil \Lambda_{e_0} \rceil \geq \lceil \Lambda_e \rceil$. By proposition 9, $\Lambda_{e_0}|S$ is a forcible path through Λ_{e_0}. So, by proposition 11, there exists $w_{n+1} \in (\overline{\Lambda}_{e_0,n+1} \cap S)$ and a list of distinct answers $(H_0, \ldots, H_{n+1})$ such that for each $m \leq n+1$, $H_m \subseteq \overline{\Lambda}_{e_0,m}$ and such that M produces each successive answer in the list j times in immediate succession along w_{n+1} after the end of e_0. Since $w_{n+1} \in \Lambda_{e_0,n+1}$, for each $m < n + 1$, answer $H_m \neq H_{w_{n+1}}$, so since $n + 1 > 0$, M produces at least j errors in w_{n+1} and, after the retractions σ already produced by M along e_-, retracts at least $n + 1$ more times no sooner than:

$$((|e| + 1), (|e| + 1 + j), (|e| + 1 + 2j), \ldots, (|e| + 1 + (n + 1)j))).$$

Thus:

$$\lambda(M, w_{n+1}) \geq (\omega, \sigma * ((|e| + 1), (|e| + 1 + j), (|e| + 1 + 2j), \ldots, (|e| + 1 + (n + 1)j))).$$

Since $w_0 \in (\overline{\Lambda}_{e,0} \cap S \cap K_{e_0})$, it follows that $\Lambda(e_0) = 0$. Hence, by proposition 8:

$$\overline{\Lambda}_{e_0,n+1} \cap S = \overline{\Lambda}_{e,n+1} \cap S \cap K_{e_0},$$

so $w_{n+1} \in \overline{\Lambda}_{e,n+1}$. So, since j is arbitrary:

$$\lambda_{n+1}(M, e) \geq (\omega, (\sigma * |e| * \omega^{(n+1)})).$$

Proof of 7. The argument for (2) establishes that $M(e)$ is not simplest at e. Then apply (6).

Proofs of 8, 9. Apply proposition 11 at e and proceed as in the proof of (6). ■

5.3 Relative Efficiency

Now it is possible to compare strategies in terms of their worst-case costs over problem instances of various sizes. Several such comparisons come immediately to mind:

$$\begin{array}{lll} M \leq_e M' & \text{iff} & (\forall n)\ \lambda_n(M,e) \leq \lambda_n(M',e); \\ M <_e M' & \text{iff} & M \leq_e M' \text{ and } M' \not\leq_e M; \\ M \prec_e M' & \text{iff} & (\forall n) \text{ if } C_e(n) \neq \emptyset \text{ then } \lambda_e(M,n) < \lambda_e(M',n). \end{array}$$

When $M \leq_e M'$, say that M is **as efficient as** M' given e. If $M <_e M'$ say that M is (weakly) **more efficient than** M' given e. Finally, in the event that $M \prec_e M'$, say that M **is strongly more efficient than** M' given e.

5.4 Efficiency and Optimality

Define optimality relative to simplified problem (K, Θ, Λ) as follows:

> M is **optimal** at e iff M is a solution at e and for each strategy $M' \simeq_{e\_} M$ that is a solution at e, $M \leq_e M'$.
>
> M is **optimal** from e onward iff for each e' extending e, M is optimal at e'.

When e is empty, say simply that M is **optimal**.

5.5 Beating

Define beatings relative to simplified problem (K, Θ, Λ) as follows:

> M is (weakly) **beaten** at e iff there exists $M' \simeq_{e\_} M$ such that M' is a solution at e and $M' <_e M$;
>
> M is **strongly beaten** at e iff there exists $M' \simeq_{e\_} M$ such that M' is a solution at e and $M' \prec_e M$;
>
> M is [**strongly**] **beaten** iff there exists $e \in \widehat{K}$ such that M is [strongly] beaten at e;
>
> M is **never** [**strongly**] **beaten** iff there exists no $e \in \widehat{K}$ such that M is [strongly] beaten at e'.

5.6 Normal Ockham Optimality Theorem

Proposition 21 (Ockham optimality) *Let (K, Θ, Λ) be a simplified problem and let M be a normally Ockham strategy always. Then M is an optimal solution to (K, Θ, Λ) always.*

Proof of proposition 21. Let M be a normally Ockham strategy always. Then M is normally Ockham from e onward so, by proposition 14, M solves (K, Θ). Let $M' \simeq_{e_-} M$ solve (K, Θ). Then $a_0(M, e) = a_0(M', e)$ and $\sigma_{e_-}(M, e_-) = \sigma_{e_-}(M', e_-)$.

Case 1: M does not retract at e. Then, by propositions 18.2 and 20.5:

$$\lambda_0(M, e) \leq (a, \sigma) \leq \lambda_0(M', e),$$

and if $n + 1 < \lceil \Lambda \rceil$, then propositions 18.4 and 20.9 yield:

$$\lambda_{n+1}(M, e) \leq (\omega, (\sigma * \omega^{(n)})) \leq \lambda_{n+1}(M', e).$$

Case 2: M' retracts at e. Then, by propositions 18.1 and 20.4:

$$\lambda_0(M, e) \leq (a, (\sigma * |e|)) \leq \lambda_0(M', e),$$

and if $n + 1 < \lceil \Lambda \rceil$, then propositions 18.3 and 20.8 yield:

$$\lambda_{n+1}(M, e) \leq (\omega, (\sigma * |e| * \omega^{(n)})) \leq \lambda_{n+1}(M', e).$$

Case 3: M retracts at e and M' does not retract at e. Since M is normally Ockham always and M retracts at e, $M(e_-)$ is not Ockham at e. Since $M' \simeq_{e_-} M$, it follows that $M'(e) = M(e_-)$ and that is the first Ockham violation by M' along e. Then, by propositions 18.1 and 20.3:

$$\lambda_0(M, e) \leq (a, (\sigma * |e|)) < (a + 1, (\sigma * (|e| + 1))) \leq \lambda_0(M', e),$$

and if $n + 1 < \lceil \Lambda \rceil$, then propositions 18.3 and 20.7 yield:

$$\lambda_{n+1}(M, e) \leq (\omega, (\sigma * |e| * \omega^{(n)})) < (\omega, (\sigma * (|e| + 1) * \omega^{(n)})) \leq \lambda_{n+1}(M', e).$$

■

6 Necessity of the Normal Ockham Property

It is not enough that principles jointly suffice for optimality. The clincher is that they be necessary for optimality.

Proposition 22 (necessity of responsiveness) *Let (K, Θ) be a problem, let $e \in \widehat{K}$ and let M' be a solution that either violates responsiveness at e. Then M is not a solution from e onward.*

Proof of proposition 22. Immediate. ■

Proposition 23 (stable necessity of stalwartness) *Let (K, Θ, Λ) be a simplified problem, let $e \in \widehat{K}$ and let M' be a solution that either violates stalwartness or that strongly violates Ockham's razor at e. Then every strategy $M \simeq_{e_-} M'$ that is normally Ockham from e onward is a more efficient solution than M' at e.*

Proof of proposition 23. Let M' violate stalwartness at e and let $M \simeq_{e_-} M'$ be normally Ockham from e onward. Then M' retracts at e and M does not. So by propositions 18.1 and 20.4,

$$\lambda_0(M, e) \leq (a, \sigma) < (a + 1, (\sigma * |e|)) \leq \lambda_0(M', e),$$

and if $n + 1 < \lceil \Lambda \rceil$, then propositions 18.4 and 20.8 yield:

$$\lambda_{n+1}(M, e) \leq (\omega, \omega^{(n)})) < (\omega, (\sigma * |e| * \omega^{(n)})) \leq \lambda_{n+1}(M', e).$$

■

Proposition 24 (stable necessity of weak Ockham) *Let (K, Θ, Λ) be a simplified problem, let $e \in \widehat{K}$ and let M' be a solution that strongly violates Ockham's razor at e. Then every strategy $M \simeq_{e_-} M'$ that is normally Ockham from e onward is a more efficient solution than M' at e.*

Proof of proposition 24. Let M' violate weak Ockham's razor at e and let $M \simeq_{e_-} M'$ be normally Ockham from e onward. Then $M(e)$ is not simplest given e. So by propositions 18.2 and 20.1,

$$\lambda_0(M, e) \leq (a, (\sigma * |e|)) < (a + 1, (\sigma * (|e| + 1))) \leq \lambda_0(M', e),$$

and if $n + 1 < \lceil \Lambda \rceil$, then propositions 18.4 and 20.6 yield:

$$\lambda_{n+1}(M, e) \leq (\omega, \sigma * |e| * \omega^{(n)})) < (\omega, (\sigma * (|e| + 1) * \omega^{(n)})) \leq \lambda_{n+1}(M', e).$$

■

Proposition 25 (necessity of Ockham) *Let (K, Θ, Λ) be a simplified problem, let $e \in \widehat{K}$ and let M' be a solution that violates Ockham's razor for the first time at e. Then every strategy $M \simeq_{e_-} M'$ that is normally Ockham from e onward is a (weakly) more efficient solution than M' at e.*

Proof of proposition 25. Let M' violate Ockham's razor for the first time at e and let $M \simeq_{e_-} M'$ be normally Ockham from e onward.

Case 1: M' does not retract at e. Then, since $M \simeq_{e_-} M'$ and M is Ockham along e_-, M' is Ockham along e_- as well, so e is the first Ockham violation by M', so by propositions 18.1 and 20.2:

$$\lambda_0(M, e) \leq (a, (\sigma * |e|)) < (a + 1, (\sigma * (|e| + 1))) \leq \lambda_0(M', e),$$

and if $n+1 < \lceil\Lambda\rceil$, then propositions 18.3 and 20.2 yield:

$$\lambda_{n+1}(M,e) \leq (\omega, (\sigma * |e| * \omega^{(n)})) < (\omega, (\sigma * (|e|+1) * \omega^{(n)})) \leq \lambda_{n+1}(M,e).$$

Case 2: M' retracts at e. Then, by propositions 18.1 and 20.3:

$$\lambda_0(M,e) \leq (a, (\sigma * |e|)) < (a+1, (\sigma * |e| * (|e|+1))) \leq \lambda_0(M',e),$$

and if $n+1 < \lceil\Lambda\rceil$, then propositions 18.3 and 20.8 yield:

$$\lambda_{n+1}(M,e) \leq (\omega, (\sigma * |e| * \omega^{(n)})) \leq \lambda_{n+1}(M',e).$$

■

Theorem 1 (efficiency characterization) *Let (K, Θ, Λ) be a simplified problem. The following are equivalent:*

1. *M is always an optimal solution;*
2. *M is always a normal Ockham srategy;*
3. *M is never (weakly) beaten.*

Proof of proposition 1. By definition, (1) implies (3). By propositions 16, 22, 23, 25, one obtains that (3) implies (2). By proposition 21, it follows that (2) implies (1). ■

6.1 Univocal Case

Proposition 26 (stable Ockham optimality in univocal problems) *Let (K, Θ, Λ) be a nested, simplified problem, let $e \in \widehat{K}$ and let M be a normally Ockham strategy from e onward. Then M is an optimal solution to (K, Θ, Λ) from e onward.*

Proof of theorem 26. Let M be a normally Ockham strategy from e onward. It suffices to revisit case 3 in the proof of proposition 21, since the assumption that M is normally Ockham from e onward suffices for cases 1 and 2. In case 3, M retracts at e and M' does not retract at e. Since, in case 3, M is normally Ockham always and M retracts at e, it follows that $M(e_-) = M'(e)$ is not Ockham at e. So, since (K, Θ, Λ) is univocal, $M'(e)$ is not simplest at e. Then, by propositions 18.1 and 20.1:

$$\lambda_0(M,e) < (a+1, (\sigma * (|e|+1))) \leq \lambda_0(M',e),$$

and if $n+1 < \lceil\Lambda\rceil$, then propositions 18.3 and 20.6 yield:

$$\lambda_{n+1}(M,e) < (\omega, (\sigma * (|e|+1) * \omega^{(n)})) \leq \lambda_{n+1}(M',e).$$

■

Proposition 27 (stable necessity of Ockham in univocal problems) *Let (K, Θ, Λ) be a univocal, simplified problem, let $e \in \widehat{K}$ and let M' be a solution that violates Ockham's razor for the first time at e. Then every strategy $M \simeq_{e_-} M'$ that is normally Ockham from e onward is a more strongly more efficient solution than M' at e.*

Proof of proposition 27. Let M' violate Ockham's razor at e and let $M \simeq_{e_-} M'$ be normally Ockham from e onward. By proposition 13, $\overline{\Lambda}_{e,0} \not\subseteq M(e)$, so since (K, Θ, Λ) is univocal, $\overline{\Lambda}_{e,0} \cap M(e) = \emptyset$, so $M(e)$ is not simplest given e. Proceed as in the proof of proposition 24. ∎

Theorem 2 (univocal efficiency characterization) *Let (K, Θ, Λ) be a univocal, simplified problem and let $e \in \widehat{K}$. The following are equivalent:*

1. *M is an optimal solution from e onward;*
2. *M is a normal Ockham srategy from e onward;*
3. *M is never strongly beaten from e onward.*

Proof of proposition 1. Proof of proposition 1. By definition, (1) implies (3). By propositions 16, 22, 23, 27, one obtains that (3) implies (2). By proposition 26, it follows that (2) implies (1). ∎

6.2 Strong Simplicity Case

Univocality is a strong condition that is not met in interesting applications like learning from effects (e.g., causal graph inference), the problem is non-uinivocal but there exists a strong simplicity concept. One cannot strengthen the general results for errors and retractions jointly. For example, suppose that the problem is to determine the total counts of blue effects and red effects. Strategy M has seen no red effects but guesses "one red effect" twenty times along e_- and changes its tune to "one blue effect" right at e, when it is announced that an effect of some color is coming. Now one can only force an extra retraction from M by presenting a red effect (the option of presenting no effects is ruled out). But presenting just one blue effect would maximize errors.

Proposition 28 (lower bound for strong simplicity) *Let (K, Θ, Λ) be a strongly simplified problem, let $e \in \widehat{K}$, let M be a solution to (K, Θ) from e onward such that $M(e)$ is not Ockham at e. Let $\sigma = \sigma_{e_-}(M, e_-)$. Then:*

1. *$\rho_0(M, e) \geq (\sigma * (|e| + 1))$;*
2. *$\rho_{n+1}(M, e) \geq (\sigma * (|e| + 1) * \omega^{(n+1)})$, if $n + 1 < \lceil \Lambda_e \rceil$.*

Proof of proposition 28. Since $M(e)$ is not Ockham at e, it follows from proposition 13 that there exists $w_0 \in \overline{\Lambda}_{e,0} \not\subseteq H$. Since M is a solution, there exists $i > |e|$ such that $M(w_0|i) = H_{w_0} \neq M(e)$. Hence:

$$\rho_0(M, e) \geq (\sigma * (|e| + 1)).$$

Let $n+1 < \lceil \Lambda_e \rceil$. Since Λ is a strong simplicity concept, Λ_e is a forcible path through Λ_e. Let $e_0 = w_0|i$. Since $w_0 \in \overline{\Lambda}_{e,0} \cap S$, it follows that $\Lambda_e(e_0) = 0$, so $\lceil \Lambda_{e_0} \rceil \geq \lceil \Lambda_e \rceil$. Continue as in the proof of proposition 20.6. ■

Proposition 29 (stable Ockham optimality with strong simplicity) *Let (K, Θ, Λ) be a strongly, simplified problem, let $e \in \widehat{K}$ and let M be a normally Ockham strategy from e onward. Let the cost be timed retractions (i.e., ρ). Then M is an optimal solution to (K, Θ, Λ) from e onward.*

Proof of theorem 26. Let M be a normally Ockham strategy from e onward. It suffices to revisit case 3 in the proof of proposition 21, since the assumption that M is normally Ockham from e onward suffices for cases 1 and 2. In case 3, M retracts at e and M' does not retract at e. Since M is stalwart from e onward, it follows that $M(e_-) = M'(e_-) = M'(e)$ is not Ockham at e So, by propositions 18.1 and 28.1:

$$\lambda_0(M, e) \leq (\sigma * |e|)) < (\sigma * (|e| + 1)) \leq \lambda_0(M', e),$$

and if $\lceil \Lambda_e \rceil > n + 1$ then by propositions 18.3 and 28.2:

$$\lambda_{n+1}(M, e) \leq (\sigma * |e| * \omega^{(n+1)}) < (\sigma * (|e| + 1) * \omega^{\omega}) \leq \lambda_0(M', e).$$

■

Proposition 30 (stable necessity of Ockham with strong simplicity) *Let (K, Θ, Λ) be a nested, simplified problem, let $e \in \widehat{K}$ and let M' be a solution that violates Ockham's razor for the first time at e. Then every strategy $M \simeq_{e_-} M'$ that is normally Ockham from e onward is a more strongly more efficient solution than M' at e.*

Proof of proposition 27. Let Λ be a strong simplicity concept for (K, Θ). Let M' violate Ockham's razor at e and let $M \simeq_{e_-} M'$ be normally Ockham from e onward. So, by propositions 18.1 and 28.1:

$$\lambda_0(M, e) \leq (\sigma * |e|)) < (\sigma * (|e| + 1)) \leq \lambda_0(M', e),$$

and if $\lceil \Lambda_e \rceil > n + 1$ then by propositions 18.3 and 28.2:

$$\lambda_{n+1}(M, e) \leq (\sigma * |e| * \omega^{(n+1)}) < (\sigma * (|e| + 1) * \omega^{\omega}) \leq \lambda_0(M', e).$$

■

Theorem 3 (efficiency characterization for strong simplicity) *Let (K, Θ, Λ) be a strongly simplified problem and let $e \in \widehat{K}$. Let the cost be timed retractions (i.e., ρ). The following are equivalent:*

1. *M is an optimal solution from e onward;*
2. *M is a normal Ockham srategy from e onward;*
3. *M is never strongly beaten from e onward.*

Proof of proposition 1. Proof of proposition 1. By definition, (1) implies (3). It follows from propositions 16, 22, 23, 30 that (3) implies (2). By proposition 29 one obtains that (2) implies (1). ■

7 Simplicity in Topological Spaces

It is often the case that one thinks of an empirical problem not in terms of infinite input streams themselves but, rather, in terms of points in some space, as in the case of curve-fitting with real-valued parameters. In that problem, answers are *models* corresponding to finite sets of parameters. One could first translate such a problem into the input streams one might encounter, but one could also define simplicity directly on the problem as presented. The latter approach will now be developed.

8 Topological Problems

A **topological empirical problem** is a triple $\mathcal{P} = (K, \Omega, \Theta)$ such that $\mathcal{T} = (K, \Omega)$ is a topological space and Θ is a partition on K, where Ω is the collection of **open sets**. Let H_x denote the unique cell Ω that contains $x \in K$. The possibilities in X are now just points, rather than infinite input streams. As before, think of open sets as verifiable propositions. The definitions that follow are all relative to $\mathcal{P}$. An **information stream** for $x \in K$ is a downward-nested ω-sequence of elements of Ω. Let S denote the set of all information streams drawn from elements of Ω. Define, for all $x \in X, w \in S$:

$$I(x, w) \text{ iff } (\forall S \in \Omega_x)(\exists i)\ w(i) \subseteq S,$$

in which case w is **informative** for x. Then define:

$$\dot{K}_x = \{w \in S_\Omega : I(x, w)\}.$$

If $H \in \Theta$, define:

$$\dot{H} = \{w \in \dot{K}_x : x \in H\},$$

and let the **induced question** be:

$$\dot{\Theta} = \{\dot{H} : H \in \Theta\}.$$

Then the **induced problem**:

$$\dot{\mathcal{P}} = (\dot{K}, \dot{\Theta})$$

is an empirical problem of the sort defined earlier, so all of the earlier definitions and results apply. In particular,it makes sense to write $\widehat{\dot{K}}_{\mathcal{P}}$, $\widehat{\dot{K}}_{\mathcal{P},e}$, etc. Let $\dot{\mathcal{T}} = (\dot{K}, \dot{\Omega}^*)$ denote the branching future space on $\dot{K}$.

8.1 Erotemorphism

Let R be a relation in $X \times X'$: Then define the **direct** and **inverse** projections of R as follows, where $S \subseteq X, S' \subseteq X'$:

$$R(S) = \{x' \in X' : (\exists x \in S)\ R(x, x')\};$$
$$R^{-1}(S') = \{x \in X : (\exists x' \in S')\ R(x, x')\}.$$

Suppose that $\mathcal{T} = (X, \Omega), \mathcal{T}' = (X', \Omega')$ are topological spaces and $R \subseteq X \times X'$. Then say that R is **directly continuous** if and only if R^{-1} takes open sets in $\mathcal{T}'$ to open sets in $\mathcal{T}$. Say that R is **directly total** if and only if $R(X) = X'$. If Θ partitions X and Θ' partitions X', then say that R **preserves cells directly** if and only if:

$$(\forall H \in \Theta)(\exists H' \in \Theta')\ R(H) \subseteq H'.$$

Each of these properties holds of R **inversely** if and only if it holds directly for R^{-1}. Each property is held **symmetrically** if and only if it holds directly and inversely (so symmetrical continuity is bi-continuity). Say that $R \subseteq X \times X'$ is an **erotemorphism** if and only if:

1. R is symmetrically total;
2. R is symmetrically continuous;
3. R is symmetrically cell-preserving.

Then write $\mathcal{P} \equiv \mathcal{P}'$ and say that $\mathcal{P}$ is **erotemorphic** to $\mathcal{P}'$. Say that $\mathcal{P} = (X, \Omega, \Theta)$ is **presentable** if and only if I is directly total.

Proposition 31 *If $\mathcal{P}$ is presentable, then relation I witnesses $\mathcal{P} \equiv \dot{\mathcal{P}}$.*

Proof.

8.2 Strategies

A strategy gets to draw conclusions from finite initial segments of open sets. However, the fact that the open sets may arrive by any schedule means that there is no information encoded in the order in which information arrives, so one may as well look only at the last open set presented. Let the **content** of $e \in \widehat{K'}$ be given as:

$$c(e) = \begin{cases} K & \text{if} |e| = 0; \\ e(|e| - 1) & \text{otherwise.} \end{cases}$$

Say that strategy M is **content-driven** if and only if for each $e, e' \in \widehat{K'}$:

$$c(e) = c(e') \text{ implies } M(e) = M(e').$$

Thus, a content-driven strategy may be expressed as a mapping directly from open sets to answers:

$$M(e) = M'(c(e)).$$