## 2 THE FREEWAY TO THE TRUTH

It is no fault of simplicity that it fails to point out or indicate the true theory, since nothing possibly could. General theories or models can always be overturned in the future by the discovery of subtle effects missed earlier even by the most diligent probing. So science is not an uneventful voyage along a compass course to the truth. It is more like an impromptu road trip through the mountains, with numerous hairpin twists and detours along the way. Taking this more appropriate metaphor seriously is the key to the simplicity puzzle.

Suppose that, on your way to a distant city, you exit the freeway for a rest stop and become lost in the neighboring town. If you ask for directions, you will be told the shortest route back to the freeway entrance ramp even before you say which city you are headed to, because the freeway is the best route to anywhere a stranger might wish to go (Figure 4.1). That remains true even if the shortest route to the entrance ramp takes you west for a few miles when your ultimate destination is east.

Suppose that you disregard the local resident's advice. You find yourself on small dirt tracks headed nowhere and, after enough of this, you make a U-turn and head back toward the entrance ramp. Your hubris is rewarded by the addition of one gratuitous course reversal to your route before you even begin the real journey on the freeway, with all of its unavoidable curves through the mountains. So even if directions to the freeway take you directly away from your ultimate goal at first, you ought to follow them.

The journey to the truth likewise occasions reversals and detours: revolutions or revisions in which one theory is retracted and replaced by another and the textbooks are rewritten accordingly [13]. Some retractions are unavoidable in principle given that one finds the truth at all, since accepting a general theory always occasions a risk of being surprised by an unanticipated anomaly later. In that case, retracting the theory is not merely excusable but virtuous—the alternative would be dogmatic commitment to error for

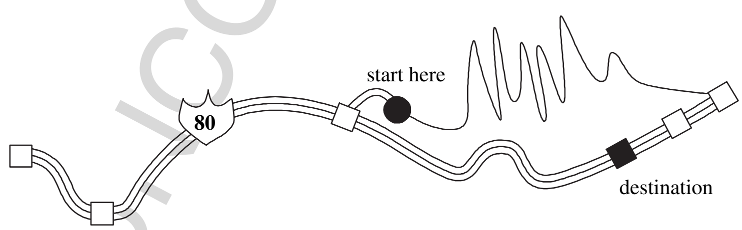


*Figure 4.1.* Entrance ramp

eternity, as Popper [16] emphasized. But gratuitous reversals in the course of inquiry are another matter entirely: it would be better to avoid them.[2]

Suppose that you violate Ockham's razor by selecting a theory more complex than experience requires. Then the simple experience up to now can be extended for eternity with equally uniform, simple experience, devoid of "effects" whose detection would indicate the need to postulate more causes or free parameters. If you refuse ever to retract to a simple hypothesis, you never arrive at the truth at all, so you have to take the bait, eventually, and fall back to the simplest theory. Now you are essentially where you would have been had you never violated Ockham's razor, except that you have already retracted once; and you are still subject to the future appearance of any number of subtle empirical effects that could not be detected at current sample sizes or using current instrumentation. Each such effect may occur sufficiently late to result in an unavoidable retraction. So you are stuck with an extra retraction at the outset added to all of these. Therefore, always presuming that the world is simple keeps you on the straightest path to the truth even though the truth may be arbitrarily complex! So both the realist and the anti-realist are right, since simplicity keeps one on the straightest path to the truth, but the straightest path may point in the wrong direction for the time being and for any finite number of times in the future as well, assuming that you converge to the truth at all.

## 3 ILLUSTRATION: COUNTING MARBLES

Suppose that you are studying a marble-emitting device that occasionally emits a marble (a new empirical effect). Your job is to determine how many marbles it will ever emit (how many free parameters the true theory has). You know nothing about when the marbles will be emitted (empirical effects may be arbitrarily small and hard to notice) but you do know on general grounds that at most finitely many marbles will be emitted (every theory under consideration has at most finitely many free parameters). Call the situation just described the *counting problem*.

In this simplistic setting, it seems that when exactly $k$ marbles have been seen so far, $k$ is the simplest answer compatible with experience. First,



[2] Retractions have been studied extensively in computational learning theory. For a review cf. [9]. The first version of the U-turn argument, albeit restricted to problems in which at most $k$ marbles may be seen, is presented in [17]. An infinite ordinal version of the argument, based loosely on ideas in [3] is presented in [10], but that idea still can't handle the marble counting problem described below.

Ockham strategy retract once, but you retract later than the Ockham strategy. That is worse, for one's state after the retraction is more enlightened than one's state prior to it (think of the Newtonians before and after they lost their faith that an ether drift would be detected) and needlessly delaying a retraction allows more subsidiary conclusions to accumulate that must be flushed when it finally occurs.

So instead of simply counting retractions, let the cost of inquiry in a given world $w$ be represented by a possibly empty, finite sequence of ascending natural numbers $(r_1, \ldots, r_k)$ such that the strategy retracts exactly $k$ times in $w$ and for each $i$ from 1 to $k$, the strategy retracts at moment $r_i$. It is necessary to rank such cost sequences. It would be unfortunate if Ockham's razor were to depend upon some fussy weighting of time against overall retractions so that, say, $(9) > (1, 2)$. Happily, it suffices in the following argument to restrict attention to weak Pareto dominance with respect to overall retractions and the times of occurrence thereof, which yields only a partial order over cost sequences. Accordingly, if $c, c'$ are both cost vectors, let $c \leq c'$ if and only if there exists a sub-sequence $d$ of $c'$ whose length matches that of $c$ such that the successive entries in $d$ are at least as great as the corresponding entries in $c$. Then define $c < c'$ if and only if $c \leq c'$ but $c' \square c$. For example:

$$(1, 3, 8) < (1, 5, 9) < (1, 2, 5, 9).$$

Refer to the cost concept just defined as *timed retractions*.

Next, consider bounds on sets of timed retraction cost sequences. Recall that $\omega$ is the least (infinite) ordinal upper bound on the natural numbers. A potential *timed retraction bound* is the result of substituting $\omega$ from some point onward in a cost sequence: e.g., $(1, 2, \omega, \omega)$. If $S$ is a set of cost sequences and $b$ is a potential bound, then $b$ bounds $S$ (written $S \leq g$) if and only if for each $c$ in $S$, $c \leq b$. Thus, $(1, \omega)$ bounds the set of all sequences $(1, k)$ such that $k$ is an arbitrary natural number.

Finally, say that a strategy is *Ockham* just in case it never chooses an answer other than the current count (or possibly '?'). Then one obtains the following, strengthened result.

**Proposition 52** *Let a solution to the counting problem be given. Then:*

1. *if the solution violates either the Ockham property or stalwartness at e, then the solution is strongly beaten in terms of timed retractions at e;*
2. *if the solution satisfies stalwartness and the Ockham property at e, then the solution is efficient in terms of timed retractions at e.*

**Proof.** Suppose that you over or under count at $e$, which presents exactly $k$ marbles. As before, let hybrid strategy $\sigma$ be just like you along $e_-$ and then always return the current count from $e$ onward. Consider answer $k + k'$, where $k'$ is an arbitrary natural number. Suppose that you retract at $e$ if $\sigma$ does. Then the cost sequence for $\sigma$ along $e$ is no worse than yours, which is, say, $(c_1, \ldots, c_r)$. Then since $\sigma$ retracts at most once, for each of the additional marbles that appear after $e$ in answer $k + k'$, the worst-case cost bound for $\sigma$ over answer $k + k'$ is at most $(c_1, \ldots, c_r, \omega, \ldots, \omega)$, with $k'$ repetitions of $\omega$. Nature can withhold marbles after $e$ until you eventually retract your answer (say, at stage $i$) in preparation for convergence to $k$. Furthermore, after you converge to $k$, nature can continue to withhold marbles until you say $k$ an arbitrary number of times before presenting another marble. Eventually, you drop $k$ in preparation for convergence to $k + 1$, etc. So your bound in answer $k + k'$ is at least $(c_1, \ldots, c_r, i, \omega, \ldots, \omega)$, with $k'$ repetitions of $\omega$. That is worse than the bound for $\sigma$ because the bound for $\sigma$ is a proper sub-sequence of your bound.

Now suppose that you don't retract at $e$ but $\sigma$ does. Then let your cost through $e$ be $(c_1, \ldots, c_r)$, in which case the cost of $\sigma$ through $e$ is $(c_1, \ldots, c_r, i)$, where $i$ is the length of $e$. Then since $\sigma$ retracts at least once, arbitrarily late, for each of the additional marbles that appear after $e$ in answer $k + k'$, the worst-case cost bound for $\sigma$ over answer $k + k'$ is at most $(c_1, \ldots, c_r, i, \omega, \ldots, \omega)$, with $k'$ repetitions of $\omega$. But since you do not produce $k$ at the end of $e$, nature can withhold marbles until (say, at stage $i' > i$) you retract your answer at $e$ in preparation for convergence to $k$. Then nature can exact one retraction out of you, arbitrarily late, for each of the $k'$ marbles that appears after $e$ in answer $k + k'$. Hence, your worst case bound is at least $(c_1, \ldots, c_r, i', \omega, \ldots, \omega)$, where $i' > i$. So your bound is worse than that of $\sigma$. The beating argument for stalwartness and the efficiency argument for stalwart, Ockham solutions are similar.[7] ■

So when retraction delays are taken into account, every solution is either efficient, stalwart, and Ockham or strongly beaten. Again, there is no middle ground.

**Corollary 53** *Let $\sigma$ be a solution to the counting problem and let the cost be timed retractions. Then the following are equivalent:*[8]

1. *$\sigma$ is efficient at each $e$;*
2. *$\sigma$ is weakly beaten at no $e$;*

[7] In any event, more general arguments are provided in the appendix for propositions 56 and 58 below.

[8] Corollary 53 is an instance of corollary 59 below.

*3. $\sigma$ is strongly beaten at no e;*
*4. $\sigma$ is stalwart and Ockham at each e.*

## 6 GENERALIZING THE ARGUMENT

In order to argue, in general, that Ockham's razor is necessary for minimizing timed retractions, one must say, in general, what Ockham's razor amounts to. That may seem like a tall order compared to counting marbles. First, simplicity has such manifold characteristics—e.g., uniformity, unity, testability, and reduction of free parameters, causes, or ontological commitments—that one wonders if there is a single notion that underlies them all. Second, it seems that some aspects of simplicity are a mere matter of description. For example, if one describes inputs as marbles or non-marbles, then marble-free worlds are most uniform. But if an "*n*-ble" is a marble at each time other than *n*, when it is a non-marble, then uniformly marble-free experience is not uniformly *n*-ble free experience. Nor can one complain that the definition of *n*-ble is strange, since marbles are *n*-bles at each stage but *n*, when they are non-*n*-bles Goodman 1983 [6]. So with respect to the syntactic complexity of definitions, the situation is entirely symmetrical. These sorts of observations have led to widespread skepticism about the prospects for a general, unified, objective account of simplicity . But the skepticism is premature, for in the marble counting problem, the question at hand concerns marbles rather than *n*-bles and simplicity may depend upon the structure of the problem one is trying to solve. Indeed, if simplicity is to have anything to do with efficiency, it must somehow reflect the structure of the problem one is trying to solve.

In the marble counting problem, answers positing more marbles are more complex. Presumably, then, worlds that present more marbles are more complex, assuming that simpler answers are answers satisfied by simpler worlds. One might plausibly say that each marble is an *anomaly* relative to the counting problem, since the previously simplest (best) explanation is no longer simplest after the marble appears. Some insight is gained into the nature of anomalies by characterizing the occurrence of a marble entirely in terms of the structure of the marble counting problem, itself.[9]

One structural feature of the marble counting problem is that, prior to seeing a third marble, nature can *force* an arbitrary solution to the problem to produce successive answers 2, 3, 4, ... by presenting no marbles until the

[9] For a critique of this idea and a response, see [1] and [19].

solution converges to 2, one marble followed by no more until the method converges to 3, and so forth. But after seeing the third marble, nature can only force the solution to produce successive answers 3, 4, 5, .... So as a working hypothesis, it seems that an anomaly occurs when the sequence of answers nature can force is truncated (from the front). This might be expressed by saying that an anomaly occurs when nature uses up an opportunity to force the scientist to change her mind or, more colorfully, when nature leads the scientist one exit further down the freeway to the truth.

Nature may be capable of taking more than one step down the freeway at a time (e.g., modify the marble counting problem so that several marbles can be emitted at one time), in which case nature takes two steps down the forcible path $(0, 1, 2, \ldots)$ when two marbles are presented at one time, for after these marbles are seen, only $(2, 3, 4, \ldots)$ is forcible.

Also, there may be more than one freeway to the truth, in which case there may be several simplest answers to select among. For example (Figure 4.5), modify the counting problem so that marbles come in two colors, white and black, and you have to determine $(i, j)$, where $i$ is the total number of white marbles and $j$ is the total number of black. If no marbles have been seen so far, then patterns of form $((0, 0), (1, 0), \ldots)$ and $((0, 0), (0, 1), \ldots)$ are forcible. Suppose you now hear a rumble in the machine, which guarantees that another marble is coming, but you don't see the color. Now $(0, 0)$ is no longer forcible (the rumble can't be "taken back") so only patterns of form $((1, 0), \ldots)$ and $((0, 1), \ldots)$ are forcible. That is a step by nature down both possible paths, so the rumble constitutes an anomaly. Suppose that the announced marble is black. Now only patterns of form $((0, 1), \ldots)$ are forcible. No step is taken down path $((0, 1))$, however, so seeing the black marble after hearing the noise is not an anomaly— intuitively, the anticipated marble has to have some color or other. The same is true if a black marble is seen. So no world in which just one more marble is seen presents any anomalies after the sound, so all such worlds are maximally simple in light of

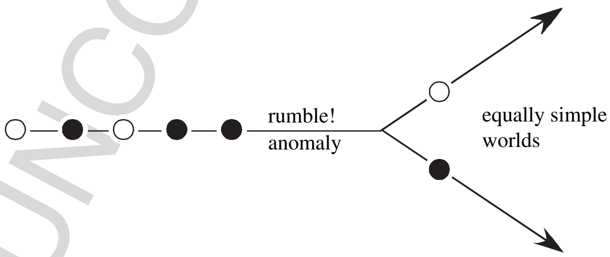


*Figure 4.5.* nature chooses a path without stepping down it

the sound. Hence, answers $(1, 0)$ and $(0, 1)$ are both simplest after the sound, whereas answers that entail more than one marble are more complex than necessary. That is intuitive, since Ockham's razor seems to govern number rather than color in this example.[10]

As it is usually formulated, Ockham's razor requires that one never presume a more complex hypothesis than necessary, which allows for selection among simplest answers when the noise is heard: e.g., $(3, 3)$ over $(2, 4)$. Answers positing extra marbles—e.g., $(3, 4000)$ are plausibly ruled out. But there is still something odd about guessing one color rather than another before seeing what the color is: after the noise, it seems that one should simply wait to see what color the announced marble happens to be. Indeed, the problem's future structure is entirely symmetrical with respect to color, so there could be no efficiency advantage in favoring one color over another until one sees which color it is. Say that a method has the *symmetry* property at a given stage if it produces no answer other than the uniquely simplest answer at that stage.

In the counting problem, worst-case cost bounds were assessed over possible answers to the question. In the general theory presented below, worst-case bounds are assessed over complexity classes of worlds. One reason for this is to "break up" coarse answers sufficiently to recover the U-turn argument. For example, recall the problem in which you must count the marbles if the total count is odd and must return "even" if the total count is even. In this problem, Ockham violators are not necessarily strongly beaten because the retractions of an arbitrary solution are unbounded in answer "even", both in terms of retractions and in terms of timed retractions. In the general theory, the answer "even" is partitioned into anomaly complexity classes corresponding to each possible even count and retractions are bounded over these complexity classes so that the strong beating arguments rehearsed earlier for the counting problem can be lifted to this coarser problem. This agrees with standard practice in the theory of computational complexity, in which one examines an algorithm's worst-case resource consumption over sets of inputs of equal size [4].

[10] That is because color does not lead to unavoidable retractions in the example under discussion. If each white marble could spontaneously change color, just once, from white to black at an arbitrary time after being emitted, then white would be simpler than black. The same is true if a continuum of gray-tones between white and black is possible and marbles never get brighter. Then Ockham should say "presume no more darkness than necessary".

**Restriction 2 (well-foundedness of forcibility)** *If pattern b is forcible at e, then there exists pattern $b'$ of which b is a sub-pattern such that $b'$ is in $\triangle_e$.*

One would expect that if $(A, B, C)$ is in $\triangle_e$, then there should be further experience $e'$ such that $(B, C)$ is in $\triangle_e$; but that is not necessarily the case.[13] It simplifies the following theory to ignore those cases as well. Let $*$ denote concatenation.

**Restriction 3 (graceful decrementation)** *If $A * B * c$ is in $\triangle_e$, then there exists proper extension $e'$ of e compatible with K such that $B * c$ is in $\triangle_{e'}$ and exactly one anomaly occurs along $e'$ properly after the end of e.*

If $g$ is an answer pattern, let $g * \triangle_e$ denote the set of all $g * g'$ such that $g'$ is an element of $\triangle_e$. Say that an *anomaly* occurs at finite, non-empty input sequence $e$ compatible with $K$ if and only if there exists a non-empty, finite answer pattern $A * g$ such that:

1. $A * g * \triangle_e \subseteq \triangle_{e\_}$;
2. no $g'$ in $\triangle_e$ begins with answer $A$.

Suppose that two marbles are seen simultaneously at stage $e$ in the counting problem. This anomaly is represented in figure (Figure 4.6).The fact that answer pattern $A * g$ is non-empty ensures that nature moves down some path in $\triangle_e$. Thus, seeing a black marble is not an anomaly after the noise that announces it.

If $w$ is a world in $K$, then let $c(w, e)$ denote the number of anomalies that occur along $w$ properly after $e$. If $A$ is an answer, let $c(A, e)$ denote the least $c(w, e)$ such that $w$ is in $K_e \cap A$. Call $c(w, e)$ the *conditional anomaly complexity* of $w$ (or of $A$) given $e$, and similarly for $c(A, e)$.

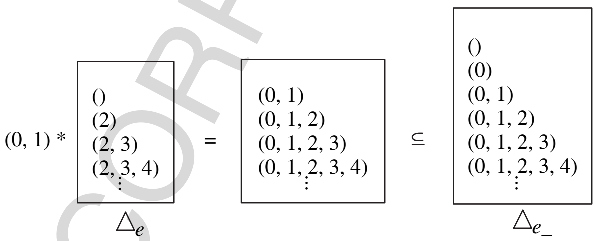


*Figure 4.6.* simultaneous observation of two marbles

[13] Suppose that you have to determine the total number of marbles and the time of the first marble if the total count is 2. If no marbles appear in $e$ yet, then we have that $(0, 1, 3)$ is in $\triangle_e$. But upon seeing the first marble at stage $k$ in $e'$, $(1, (2, k), 3)$ is in $\triangle_{e'}$, so $(1, 3)$ is not in $\triangle_{e'}$.

Then let *unconditional* anomaly complexity be given by $c(w) = c(w, ())$ and $c(A) = c(A, ())$, where () is the empty input sequence.

Marbles are still anomalies in the marble-counting problem, but the preceding definitions don't see the marbles; they see only the structural "shadow" each marble occurrence casts against the branching topology of the marble counting problem. The noise announcing a marble is anomalous, but seeing a marble after the noise is not. Seeing two marbles after the noise is anomalous, however. If several marbles are visible and some of them might disappear permanently at any time, then disappearances of marbles count as anomalies and simple worlds have more marbles than complex ones. Refutations of lower polynomial degrees and the discovery that a linear function depends upon an independent variable also count as anomalies in the corresponding problems (assuming that the data consist of ever-tighter open intervals around the dependent variable).

## 8 OCKHAM'S RAZOR, SYMMETRY, AND STALWARTNESS

Answer $A$ is *simplest* at $e$ if and only if

$$c(A, e) = \min_{B \in \Pi_{\mathrm{P}i}} c(B, e).^{14}$$

A method satisfies *Ockham's razor* at $e$ just in case the answer output by the method at $e$ is '?' or is simplest at $e$. *Symmetry* at $e$ requires that the method output at $e$ either "?" or the unique answer that minimizes $c(A, e)$. *Stalwartness* at $e$ requires that if the scientist's output $A$ at $e_-$ is uniquely simplest at $e$, then the scientist produces $A$ also at $e$.

Ockham's razor may be defined in terms of simplicity rather than complexity, using a standard rescaling trick familiar from information theory. Define *conditional simplicity* as:

$$s(A, e) = \exp(-c(A, e)).$$

This definition reveals an interesting connection between Ockham's razor and Bayesian updating, for it follows immediately from the definition of $c(A, e)$ that:

$$c(A, e) = c(A \cap K_e) - c(K_e).$$

[14] In light of lemma 67 in the appendix, this condition is equivalent to $c(A, e) = 0$.

Applying the definition of $s(A, e)$ to both sides of the preceding equation yields:

$$s(A, e) = \frac{s(A \cap K_e)}{s(K_e)},$$

which is the usual definition of Bayesian updating. Then Ockham's razor requires that one choose the uniquely simplest hypothesis, where simplicity degree is updated by conditionalization. Nothing about coherence or probability has been presupposed, however, so Bayesians who seek Ockham's razor in prior probabilities updated by conditioning put the arbitrary cart before the essential horse.

## 9 SYMMETRICAL SOLVABILITY

Not every problem has a symmetrical solution. For example, suppose that the problem is to say not only how many marbles appear, but when each of them appears. In this problem, every answer compatible with $e$ is simplest at $e$, since only patterns of unit length are forcible. That may seem counter-intuitive, since particle counts are analogous to free parameters and times of appearance are analogous to settings of those parameters, so it would seem that answers involving more free parameters are more complex. But it must be kept in mind that the same possibilities could be parameterized in different ways, and simplicity depends upon which parametrization the question asks about. If the problem is to count marbles, then worlds with more marbles are more complex, whenever the marbles arrive. If it is to count $n$-bles, then worlds with more $n$-bles are more complex, regardless of when the marbles arrive. If the problem is to identify particular worlds, the parametric structure of the problem disappears and complexity is flattened. Examples of the latter sort are excluded from consideration by the following restriction.

**Restriction 4 (symmetrical solvability)** *Only problems with symmetrical solutions are considered in the results that follow.*

In typical applications, restriction 4 can be sidestepped by coarsening or refining the question in a manner that disambiguates the intended parametrization. It is also worth mentioning that restrictions 2 and 4 are logically independent given restriction 1.[15]

[15] The problem of identifying individual worlds in which at most finitely many marbles occur satisfies the determinacy assumption (restriction 1) and the well-foundedness assumption (restriction 2) but not the symmetrical-solvability assumption (restriction 4), whereas

bound for $\sigma$ over $C_e(k)$ is at least $r * j * \omega[k]$. But since $\sigma'$ retracts after $e$ only at anomalies (by lemma 68), the worst-case timed retraction bound for $\sigma'$ over $C_e(k)$ is at most $r * i * \omega[k] < r * j * \omega[k]$. ■

**Proof of proposition 58.1**. Let $\sigma'$ be a solution to a nested problem that is Ockham and stalwart from $e$ onward. Since the problem is nested, $\sigma'$ is also symmetrical from $e$ onward. Let $\sigma$ agree with $\sigma'$ along $e_-$. Suppose that $C_e(k)$ is non-empty. There exists a pattern $b$ of length at least $k + 1$ in $\triangle_e$ (by lemma 63).

*Case* A: $\sigma$ retracts at $e$ if $\sigma$ does. Then let $r$ denote the identical costs of $\sigma$ and $\sigma'$ along $e$. Since $\sigma'$ is symmetrical and stalwart from $e$ onward, the worst-case timed retraction bound for $\sigma'$ over $C_e(0)$ is less than or equal to $r * \omega[k]$ (by lemma 68). There exists $w'$ in $C_e(k)$ along which $\sigma$ can be made to repeat each successive entry in $b$ an arbitrary number of times (by lemma 69), so the worst-case timed retraction bound for $\sigma$ over $C_e(0)$ is at least $r * \omega[k]$.

*Case* B: $\sigma'$ retracts at $e$ and $\sigma$ does not. Since $(K, \Pi)$ is nested, there exists a uniquely simplest answer $B$ at $e$. So every pattern in $\triangle_e$ begins with $B$ (by lemma 66), so $b$ begins with $B$. Let $i$ be the length of $e$. Then since $\sigma'$ retracts only at anomalies after $e$ (by lemma 68), the worst-case timed retraction bound for $\sigma'$ over $C_e(k)$ is less than or equal to $r * i * \omega[k]$. Since $\sigma'$ is stalwart at $e$ and retracts at $e$, answer $A = \sigma'(e_-) = \sigma(e_-)$ is not uniquely simplest at $e$, so $A \neq B$. There exists $w$ in $B \cap K_e$ along which $b$ remains forcible after $e$ (by lemma 64). Since $\sigma$ is a solution, $\sigma$ must retract $A$ in $w$ after $e$, say by $e'$. Now $b$ is still forcible given $e'$, so there exists $w'$ in $C_{e'}(k)$ along which $\sigma$ can be made to repeat each successive entry in $b$ an arbitrary number of times (by lemma 69). Since $b$ is forcible at $e'$, no anomaly occurs along $e'$ after $e$ (by lemma 61). Hence, $w'$ is in $C_e(k)$. So letting $j > i$ be the length of $e'$, the worst-case timed retraction bound for $\sigma$ over $C_e(k)$ is at least $r * j * \omega[k] > r * i * \omega[k]$. ■

**Proof of proposition 58.2**. Let $\sigma'$ be a stalwart, symmetrical solution at every $e$. Let $\sigma$ agree with $\sigma$ along $e_-$. Now consider non-empty complexity class $C_e(k)$, for arbitrary $k > 0$ and let $w$ be in $C_e(k)$. Then there exists pattern $b$ in $\triangle_e$ of length at least $k + 1$ (by lemma 63).

*Case* A: $\sigma$ retracts at $e$ if $\sigma'$ does. Follow the argument for case A in the proof of proposition 57, observing that a stalwart, symmetrical solution retracts only at anomalies (by lemma 68).

*Case* B: $\sigma'$ retracts at $e$ and $\sigma$ does not. Since $\sigma'$ is always symmetrical and stalwart and $\sigma'$ retracts at $e$, answer $A = \sigma'(e_-) = \sigma(e_-)$ is uniquely simplest at $e_-$ but not at $e$. Pick up from here in case B of the proof of proposition 57, again observing that a stalwart, symmetrical solution retracts only at anomalies (by lemma 68). ■

**Proof of corollary 59**. (1) implies (2) implies (3) by definition. (3) implies (4) by propositions 55 and 56. (4) implies symmetry and stalwartness since the problem is nested. Symmetry and stalwartness imply (1) by proposition 58.1. ■

**Proof of corollary 60.** (1) implies (2) by definition. (2) implies (3) by propositions 57 and 56. (3) implies (1) by proposition 58.2. ■

**Lemma 61 (anomaly freedom)** *Let $b$ be in $\triangle_e$ and let $b$ be forcible at $e'$ properly extending $e$. Then for all $e''$ properly extending $e$ and extended by $e'$:*

1. *$b$ is in $\triangle_{e''}$ and*
2. *$e''$ is not an anomaly.*

**Proof.** Suppose that $b$ is in $\triangle_e$ and $b$ is forcible at $e'$ properly extending $e$. Let $e''$ properly extend $e$ and be extended by $e'$. Then $b$ is forcible at $e''$ since $b$ is still forcible at $e'$. Suppose for contradiction that $b$ is not in $\triangle_{e''}$. Then since $b$ is forcible at $e''$, there exists $b'$ forcible at $e''$ such that $b$ is a subsequence of $b'$ but $b$ is not an initial segment of $b'$. But then $b'$ is forcible at $e$, so $b$ is not in $\triangle_e$. Contradiction. So $b$ is in $\triangle_{e'}$. Again, let $e''$ be an arbitrary input sequence properly extending $e$ and extended by $e'$. Then it has just been shown that $b$ is in both $\triangle_{e''}$ and $\triangle_{e''_-}$. Suppose that $e''$ is an anomaly. Then there exists $A * g$ such that $A * g * \triangle_{e''} \subseteq \triangle_{e''_-}$ and no element of $\triangle_{e''}$ begins with $A$. So $A * g * b$ is in $\triangle_e$. But $b$ does not begin with $A$, so $b$ is not an initial segment of $A * g * b$. Hence, $b$ is not in $\triangle_e$. Contradiction. ■

**Lemma 62 (forcibility is asymptotic)** *Let $A * a$ be forcible given $e$. Then there exists a world $w$ in $K_e \cap A$ extending $e$ such that for each finite initial segment $e'$ of $w$, $A * a$ is forcible given $e'$.*

**Proof.** Suppose $A * b$ is forcible given $e$. Suppose for contradiction that the consequent of the lemma is false. Then for each $w$ in $A \cap K_e$ there exists $e'$ extending $e$ and extended by $w$ such that $A * b$ is not forcible given $e'$. For each $w$ in $A \cap K_e$, let $e_w$ be the shortest such $e'$. For each $e_w$, $A * b$ is not forcible at $e_w$, so since the forcing games in $(K, \Pi)$ are all determined (by restriction 1), there exists a solution $\sigma_w$ for $(K_{e_w}, \Pi_{e_w})$ that never produces $A * b$ after $e_w$. Let $\sigma$ solve $(K, \Pi)$ and let $\sigma^*$ be just like $\sigma$ except that control is shifted permanently to $\sigma_w$ when $e_w$ is encountered. So $\sigma^*$ is a solution that never produces $A * b$ after seeing some $e_w$. Let $\sigma^\dagger$ be like $\sigma^*$ except that

$\sigma^\dagger$ produces "?" along each $e_w$ and at each $e$ not extended by some $e_w$ such that $\sigma$ returns $A$ at $e$. Then $\sigma^\dagger$ is still a solution, since $\sigma^*$ converges to the truth over $K_e \cap A$ (the question marks eventually end in each $w$ in $K_e \cap A$) and over $K_e - A$ ($\sigma$ does not converge to $A$ in any such world, so again, the question marks end eventually in each $w$ in $K_e - A$). But $\sigma^\dagger$ doesn't produce $A * b$ after $e$ along any $e'$ extending $e$. So $A * b$ is not forcible given $e$. Contradiction. ■

**Lemma 63 (forcible pattern existence)** *Suppose that $C_e(n)$ is non-empty. Then there exists a finite pattern in $\triangle_e$ of length at least $n + 1$.*

**Proof.** Let $w$ be in $C_e(0)$. In the base case, nature can force the answer $A$ true in $w$ from an arbitrary solution. For induction, suppose that $w$ is in $C_e(n + 1)$. Let $e'$ be the first anomaly along $w$ after $e$. So there are $n$ anomalies occurring in $w$ after $e'$. By the induction hypothesis, there exists pattern $a$ in $\triangle_{e'}$ of length at least $n + 1$. Since $e'$ is an anomaly, there exists pattern $A * b$ such that $A * b * a$ is a pattern in $\triangle_{e'\_}$. Hence, $A * b * a$ has length at least $n + 2$. Since $A * b * a$ is forcible at $e'\_$, $A * b * a$ is forcible at $e$ as well. So there exists some pattern $d$ in $\triangle_e$ of which $A * b * a$ is a sub-pattern (by restriction 2), so $d$ has length at least $n + 2$. ■

**Lemma 64 (nature's starting point)** *Let $A * a$ be in $\triangle_e$. Then there exists a world $w$ in $C_e(0) \cap A$ such that for each finite initial segment $e'$ of $w$ that extends e, $A * a$ is in $\triangle_{e'}$.*

**Proof.** Let $A * a$ be in $\triangle_e$. So $A * a$ is forcible given $e$. By lemma 62, there exists $w$ in $K_e \cap A$ such that $A * a$ is forcible along each initial segment of $w$ extending $e$. Let $e'$ properly extend $e$ and be extended by $w$. Then $A * a$ is in $\triangle_{e'}$ and $e'$ is not an anomaly (by lemma 61). Hence, $w$ is in $C_e(0)$. ■

**Lemma 65 (simplest answer forcible first)** *Let answer A be the first entry in some pattern in $\triangle_e$. Then A is a simplest answer.*

**Proof.** Suppose that $A * b \in \triangle_e$. Then there exists $w$ in $A \cap C_e(0)$ (by lemma 64). So $c(A, e) = 0$. ■

**Lemma 66 (uniquely simplest answer and forcibility)** *Let answer A be uniquely simplest at e. Then each pattern in $\triangle_e$ begins with A.*

**Proof.** Suppose that for some answer $B \neq A$, pattern $B * a$ is in $\triangle_e$. Then by lemma 65, $B$ is simplest at $e$. So $A$ is not uniquely simplest. ■

**Lemma 67 (simple world existence)** *Let $K_e$ be non-empty. Then there exists a world $w$ in $C_e(0)$.*

**Proof.** Suppose there exists $w$ in $K_e$. If $c(w, e) = 0$, we are done. So suppose $c(w, e) = k > 0$. Then (by lemma 63) there exists $A * a$ in $\triangle_e$ of length $k + 1$. So there exists $w'$ in $A \cap C_e(0)$ (by lemma 64). ■

**Lemma 68 (simplest answer defeated only by anomalies)** *Let $K_e$ be non-empty, let e be non-empty, and let A be an answer in* $\Pi$ *such that A is uniquely simplest at $e_-$ and A is not uniquely simplest at e. Then e is an anomaly.*

**Proof.** Let $K_e, e$ be non-empty. Then $K_{e_-}$ is non-empty, so by lemma 67, $C_{e_-}(0)$, $C_e(0)$ are non-empty. So since $A$ is uniquely simplest at $e_-$ but not at $e$, we have $C_{e_-}(0) \subseteq A$ but $C_e(0) \nsubseteq A$. So there exists $w$ in $C_e(0) - C_{e_-}(0)$. Hence, $c(w, e_-) > 0$ and $c(w, e) = 0$, so $e$ is an anomaly. ■

**Lemma 69 (forcing lemma)** *Let $\sigma$ be a solution and let pattern a of length at least $k + 1$ be in $\triangle_e$ and let m be a natural number. Then there exists w in $C_e(k)$ such that after e, $\sigma$ produces $a_0$ successively for m times and then $a_1$ successively for m, times, ... and finally $a_k$ successively for m times.*

**Proof.** Let natural number $m$ be given. In the base case, let pattern $(A)$ be in $\triangle_e$. Then there exists world $w \in A \cap C_e(0)$ such that $(A)$ remains in $w$ from $e$ onward (by lemma 64). Since $A$ is true in $w$ and $\sigma$ is a solution, $\sigma$ converges to $A$ in $w$, so $\sigma$ produces $A$ at least $m$ times in succession after $e$ in $w$.

For induction, let $A * a$, be forcible at $e$, where $a$ is a finite answer pattern of length $k + 1$. There exists a world $w$ in $A \cap K_e$ such that $A * a$ is in $\triangle_{e'}$, for each finite, initial segment $e'$ of $w$ extending $e$ (by lemma 64). Since $\sigma$ is a solution, $\sigma$ converges to $A$ in $w$. Nature can wait $m$ steps after the onset of convergence until $\sigma$ produces $A$ at least $m$ times after $e$ in $w$. Let $e'$ extend $e$ such that $a$ is in $\triangle_{e'}$ and exactly one anomaly occurs along $e'$ after $e$ (by restriction 3). So by the induction hypothesis, there exists $w'$ in $C_e(k)$ such that, after $e$, $\sigma$ produces $a_0$ successively for $m$ times and then $a_1$ successively for $m$, times, ... and finally $a_k$ successively for $m$ times. Hence, $\sigma$ produces $A$ successively for $m$ times followed by $a_0$ for $m$ times, etc. Since exactly one anomaly occurs along $e'$ after the end of $e$ and $k$ anomalies occur along $w$ after $e'$, $w'$ is in $C_e(k + 1)$. ■

## REFERENCES

[1] Chart, D. (2000). "Discussion: Schulte and Goodman's Riddle", *The British Journal for the Philosophy of Science* 51, 147–149.

[2] Forster, M. and Sober, E. (1994). "How to Tell When Simpler, More Unified, or Less

ad Hoc Theories Will Provide More Accurate Predictions", *The British Journal for the Philosophy of Science* 45, 1–35.

[3] Freivalds, R. and Smith, C. (1993). "On the Role of Procrastination in Machine Learning", *Information and Computation* 107, 237–271.

[4] Garey, M. and Johnson, D. (1979). *Computers and Intractability*, New York: Freeman.

[5] Glymour, C. (1980). *Theory and Evidence*, Princeton: Princeton University Press.

[6] Goodman, N. (1983). *Fact, Fiction, and Forecast*, 4th ed., Cambridge (Mass.): Harvard University Press.

[7] Harman, G. (1965). "The Inference to the Best Explanation", *Philosophical Review* 74, 88–95.

[8] Hitchcock, C. (ed.) (2004). *Contemporary Debates in the Philosophy of Science*, Oxford: Blackwell.

[9] Jain, S., Osherson, D., Royer, J.S. and Sharma A. (1999). *Systems That Learn: An Introduction to Learning Theory*, 2nd ed., Cambridge (Mass.): MIT Press.

[10] Kelly, K. (2002). "Efficient Convergence Implies Ockham's Razor", in *Proceedings of the 2002 International Workshop on Computational Models of Scientific Reasoning and Applications* (CMSRA 2002), Las Vegas, USA, June 24–27.

[11] Kelly, K. (2004). "Uncomputability: The Problem of Induction Internalized", *Theoretical Computer Science* 317, 227–249.

[12] Kelly, K. and Glymour, C. (2004). "Why Probability Does Not Capture the Logic of Scientific Justification", in Hitchcock, C. [8], 94–114.

[13] Kuhn, T.S. (1970). *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

[14] Laudan, L. (1981). "A Confutation of Convergent Realism", *Philosophy of Science* 48, 19–48.

[15] Mitchell, T. (1997). *Machine Learning*, New York: McGraw-Hill.

[16] Popper, K. (1968). *The Logic of Scientific Discovery*, New York: Harper and Row.

[17] Schulte, O. (1999). "Means-Ends Epistemology", *The British Journal for the Philosophy of Science* 50, 1–31.

[18] Schulte, O. (2000a). "Inferring Conservation Principles in Particle Physics: A Case Study in the Problem of Induction", *The British Journal for the Philosophy of Science* 51, 771–806.

[19] Schulte, O. (2000b). "What to Believe and What to Take Seriously: A Reply to David Chart concerning the Riddle of Induction", *The British Journal for the Philosophy of Science* 51, 151–153.

[20] Sklar, L. (1977). *Space, Time, and Spacetime*, Berkeley: University of California Press.

[21] Van Fraassen, B. (1981). *The Scientific Image*, Oxford: Clarendon Press.