

Kludged*

Jeffrey C. Ely[†]

March 5, 2007

Abstract

I develop a formal model which illustrates a fundamental limitation of adaptive processes: improvements tend to come in the form of *kludges*. A kludge is a marginal adaptation that compensates for, but does not eliminate fundamental design inefficiencies. When kludges accumulate the result can be perpetually sub-optimal behavior. This is true even in a model of evolution in which mutations of any size occur infinitely often with probability 1. This has implications for traditional defenses of both positive and normative methodology.

Keywords: kludge.

*Preliminary (cite anyway.) This paper was not previously circulated under the titles “Bad Adaptation” or “Ex-Post Regrettable Organism Design.”

[†]Department of Economics, Northwestern University. jeffely@northwestern.edu. I thank Larry Samuelson and Jeroen Swinkels and Marcin Peski for early conversations on this subject and advice which steered me in a good direction. Conversations with Sandeep Baliga, George Mailath, Bill Zame, Ilya Segal, Joel Sobel, Alvaro Sandroni, Daron Acemoglu and Tomasz Strzalecki were also influential. Ben Handel was a valuable research assistant at an early stage. For some of the technical results, I took inspiration from a paper by Sandholm and Pauzner (1998). There are no mistakes in this paper which are not mine.

1 Introduction

In July of 2004, Microsoft announced that the release of Vista, the next generation of the Windows operating system, would be delayed until late 2006. Jim Allchin famously walked into the office of Bill Gates and proclaimed, "It's not going to work." Development of Windows had become unmanageable and Allchin decided that Vista would have to be re-written essentially from scratch.

Mr Allchin's reforms address a problem dating to Microsoft's beginnings. . . . PC users wanted cool and useful features quickly. they tolerated – or didn't notice – the bugs riddling the software. Problems could always be patched over. With each patch and enhancement, it became harder to strap new features onto the software since new code could affect everything else in unpredictable ways.¹

The Alternative Minimum Tax was introduced by the Tax Reform Act of 1969. It was intended to prevent taxpayers with very high incomes from exploiting numerous tax exemptions and paying little or no tax at all. Over time, the shortcomings of the AMT as a solution to the proliferation of exemptions have begun to appear. However, over this same time, the federal tax and budgeting system has come to depend on the AMT to the point that many observers think that changing the AMT, without complicated accompanying adjustments elsewhere, would be worse than leaving it as is.

Flat fish inhabit the sea floor. When their ancestors moved to the sea floor, they adapted by changing their orientation from swimming "up-right" to on their sides. This rendered one eye useless so by a further adaptation, many of today's species of flatfish migrate one eye to the opposite side of their body during development.

As beautifully documented the film *The March of the Penguins*, emperor penguins spend a nearly 9 month breeding and nurturing cycle which involves walking up to 100 KM away from any food source in order to avoid predators. The problem for penguins is that they are birds, and hence lay eggs; but they are flightless birds, so they find it inconvenient to move to areas where the eggs can be easily protected. They adapted not by recti-

¹"Code Red: Battling Google, Microsoft Changes How it Builds Software." *The Wall Street Journal*, Robert Guth, September 2005.

fyng either of these two basic problems,² but instead by compensating for them by an extremely costly and risky behavior.

Each of these examples represents a *kludge*: an improvement upon a highly complex system that solves an inefficiency but in a piecemeal fashion and without addressing the deep-rooted underlying problem. There are three ingredients to a kludge. First the system must be increasing in complexity so that new problems arise that present challenges to the internal workings of the system. Second, a kludge addresses the problem by patching up any mis-coordination between the inherited infrastructure and the new demands. Third, the kludge itself— because it makes sense only in the presence of the disease it is there to treat— intensifies the internal inefficiency, necessitating either further kludges in the future or else eventually a complete revolution.³

Microsoft Windows is a complex system whose evolution is guided by a forward-looking dynamic optimizer. It is not surprising therefore that, after two decades worth of kludges that accompanied the expansion from DOS to Windows to 32 bit and eventually 64 bit architecture, revolution was the final solution. In the case of the US Tax Code, or for that matter any sufficiently complex body of contracts that govern interactions among diverse interests, while the evolution may be influenced by forward-looking considerations, full dynamic optimization is more tenuous as a model of the long-run trade-offs.

But the story is very different for flat fish and penguins, and, to come to the point, for human brains, whether we are considering the evolution of the brain across generations or the development of the decision-making apparatus within the life of a single individual. Here, progress is *adaptive*. An adaptive process is not forward-looking and certainly not governed by dynamic optimization. An adaptive process inherits its raw material from the past, occasionally modifies it by chance (mutation or experimentation), and selects among variants according to success *today*.

Nevertheless there is the possibility, not completely fanciful, that an adaptive process can produce complex systems that perform as well today as those that were designed by an optimizer given the same set of raw materials. Indeed, there is a tradition in economics that accepts the distinction between adaptation and optimization, but rationalizes a positive

²Incidentally, it has happened in evolutionary history that oviparous (egg-laying) species have adapted to vivipary (giving birth to live offspring.) Some species of sharks are important examples. Vivipary enables a long internal gestation so that the developing offspring is protected and nourished within the body of the mother.

³See [wikipedia](#) for the history and pronunciation of the word kludge.

methodology based on unfettered optimization by an appeal to this unwritten proposition.⁴

In this paper I present a model intended to suggest that this hope was a longshot at best. I analyze a simple single-person decision problem. An *organism* is a procedure for solving this problem. I parameterize a family of such algorithms which includes the optimal algorithm in addition to algorithms that perform less well. An adaptive process alters the organism over time, favoring improvements. I show conditions under which no matter how long the adaptive process proceeds, an engineer, at any point in time, working only with the raw materials that presently make up the organism, could eliminate a persistent structural inefficiency and produce a significant improvement. In the model, kludges arise naturally and are the typical adaptations that improve the organism. A kludge always improves the organism at the margin, but also increases both its complexity and its internal complementarity and as a by-product makes it harder and harder for adaptation to undo these inefficiencies in the future.

In the model, a resource is available at a randomly determined location. The organism evolves a procedure for collecting and processing information about the location. Two trade-offs govern the design of the optimal organism. First, a fixed number of computational steps must be allocated between estimation of the location and exploitation of the resource. More precise estimates come at the expense of reduced intensity of exploitation. Second, the organism must evolve the optimal protocol for processing the information. The pitfall is that the organism may adapt an inefficient protocol which requires too many processing steps to achieve a given precision. The cost is reduced intensity. However, once this inefficient protocol is in place, future evolution (modeled as expansion of computational power) continues to "invest" in it making it increasingly difficult to re-optimize.

The problem in the model is not due to "local optima." The model admits arbitrarily large mutations with positive probability, so they occur infinitely often. Given enough time, the process would escape any non-global static optimum. Indeed I present a benchmark model (see 1) in which there is an artificial upper bound on the complexity of the organism. In this model the optimally adapted organism eventually appears with probability 1. Also, the effect is not due to altered evolutionary incentives that come from strategic interactions with other agents. The model analyzes the performance of a single agent solving an isolated decision problem.

⁴The classic defense is Friedman (1966).

Structurally inefficient decision-makers present a problem not just for positive methodology, but normative as well. Much of welfare economics is founded on revealed preference and agent sovereignty. The principle is that the choices we observe reveal what benefits the agent. But when the adaptive process creates a wedge between the underlying objective it is designing the agent to satisfy and the agent's actual observed behavior, there is a corresponding wedge between revealed preference and true preference. Put differently, if we grant that there is some underlying objective that guides the adaptive process, then at best we can view the organism as an agent whose efforts at achieving that objective are the result of a second-best solution designed by nature, the principal. We can no better infer that underlying objective from the choice behavior of the organism than we can identify the distorted choices made by an incentivized agent with the principal's first-best solution.⁵

2 Overview of the Model and Results

An organism is designed to solve a fixed decision problem, instances of which are presented to the organism repeatedly over time. The decision problem has the following interpretation. A resource is available at a certain location. The location is realized independently in each period. Signals which reveal the location of the resource are available to the organism. The problem for the organism is to input these signals, interpret them, and then choose a location in attempt to exploit the resource. The fitness of the organism is determined by the distance between the actual location of the resource and the location chosen.

The organism is described by an algorithm for inputting and processing signals. The components of this algorithm adapt over time according to a general evolutionary process which selects for improvements in overall fitness. We describe the long run behavior of this evolutionary process.

2.1 The Decision Problem

One aspect of the environment is fixed throughout. An infinite sequence $\lambda = \lambda_1, \lambda_2, \dots \in \{-1, +1\}^\infty$ is determined at the beginning of time according to an i.i.d. process with $\text{Prob}(\lambda_j = 1) = l > 1/2$. We will refer to λ as the *environment*.

⁵Indeed, this metaphor is behind the methodology of Samuelson and Swinkels (2006)

In each period of the process, a location $\theta \in [-1, +1]$ is selected. Condition on θ , an infinite sequence $\tau = \tau_1, \tau_2, \dots$ is selected in i.i.d. fashion from $\{-1, +1\}^\infty$ with probability

$$\text{Prob}(\tau_j = 1) = \frac{\theta + 1}{2}.$$

Next, a sequence $\sigma = \sigma_1, \sigma_2, \dots$ of signals is produced by setting

$$\sigma_j = \tau_j \lambda_j \quad \text{for all } j.$$

Thus, σ is an infinite sequence from $\{-1, +1\}^\infty$ which is an encoding of the raw data τ using the environmental “key” λ . If $\lambda_j = -1$, then we say that the j th signal is *inverted*. The organism will have available a sample of the physical signal σ . The problem it faces as it evolves is to learn about the environment λ so that in each period it can infer the raw data τ . The raw data can then be used to estimate the location of the resource in that period.

We view the organism as an algorithm for locating and exploiting the resource. The organism will be parameterized by the total number of steps it is able to perform. This number x will be called the *complexity* of the organism. Each use of the following operations requires a single step: observing a signal σ_j , multiplying the signal by -1 , applying a decision rule which selects a location based on the processed data, and taking an action to exploit that location. Therefore, an organism of complexity x which uses l steps to select a location, can use the remaining steps to take actions. The total payoff to the organism is the sum of the payoffs from each action taken.

Locating the Resource In each period, the organism processes a sample consisting of the first k signals from σ . The parameter k is called the *precision* of the organism. The probability distribution which governs θ is such that the true conditional expected value of θ based on a sample of $\sigma_1, \dots, \sigma_k$ has the following simple formula

$$\bar{\theta}_k := \mathbf{E}(\theta | \sigma_1, \sigma_2, \dots, \sigma_k) = \frac{1}{k+2} \sum_{n=1}^k \lambda_n \sigma_n.$$

(See section 3 for the details of how the distribution of θ is specified.)

The above formula is the one that would be used by an optimally adapted organism of precision k . We now consider a broader class of organisms, each associated with its own estimation formula. An organism of

precision k is defined by a sequence π_1, \dots, π_k which encodes the algorithm used by the organism to process signals. Specifically, upon observing sample $\sigma_1, \dots, \sigma_k$, it is assumed that the organism produces the following estimate of θ .

$$\beta(k, \pi) = \frac{1}{k+2} \sum_{j=1}^k \pi_j \sigma_j. \quad (1)$$

If $\pi_j = -1$ then the organism spends one computation step to invert the input σ_j . Let $|\pi| := |\{j : \pi_j = -1\}|$ denote the total number of these *pre-processor* steps.

We can think of the sequence π as part of the genetic code of the organism. Obviously when $\pi = \lambda$, the organism is using the optimal formula.

Exploiting the Resource Once the organism has observed and processed the sample $\sigma_1, \sigma_2, \dots, \sigma_k$ to form its estimate, it earns fitness by choosing a location $a \in [-1, 1]$ to exploit. The organism's *decision rule* a translates the estimate β into a location $a(\beta)$. The payoff to exploiting location a when the resource is located at θ is defined to be

$$u(a, \theta) = 2a\theta - a^2.$$

Notice that an optimally adapted organism who observes the sample $\sigma_1, \sigma_2, \dots, \sigma_k$ would maximize fitness by choosing

$$a = \mathbf{E}(\theta \mid \sigma_1, \sigma_2, \dots, \sigma_k) = \bar{\theta}_k$$

In fact there are two types of organisms which implement this optimal strategy. A *positively-aligned* organism is one with $\pi_j = \lambda_j$ for $j = 1, \dots, k$ and decision rule a^+ where

$$a^+(\beta(k, \pi)) = \beta(k, \pi).$$

A *negatively-aligned* organism is one with $\pi_j = -\lambda_j$ for $j = 1, \dots, k$ and decision rule a^- where

$$a^-(\beta(k, \pi)) = -\beta(k, \pi).$$

Both types of organism select the conditional expected fitness maximizing location given a sample size of k . Any other organism of equal precision chooses an inferior location.

I will say that an organism is *positive* or *negative* depending on whether it uses a^+ or a^- . Because these are the only potentially optimal decision

rules, I will keep the model simple by assuming that a^+ and a^- are the only decision rules. Nothing would change if the model were extended to include a richer set of possible decision rules. Say that input j is *aligned* or *misaligned* according to whether the product $\text{sign}(a)\pi_j\lambda_j$ is equal to $+1$ or -1 . With this terminology, e.g. a positively aligned organism is a positive organism for which all inputs are aligned.

Fitness An organism of complexity x which has precision k , and uses $|\pi|$ steps to process inputs, has $x - k - |\pi| - 1$ steps remaining to take actions. This number, denoted i , is called the *intensity* of the organism. The total payoff of the organism in a period when the resource is located at θ and the signal is σ is equal to the sum of the payoffs of each action:

$$i [2a(\beta(k, \pi))\theta - \beta(k, \pi)^2]$$

The fitness of the organism is defined as the expected value of this payoff with respect to the distributions of θ and σ .

So, while positive and negatively aligned organisms of the same precision select the same location $a(\beta(k, \pi))$, they typically require a different number of steps to do it and therefore they will differ in the intensity i with which they are able to exploit the resource. This means that, for a given total complexity x , only one of these two types of organism will achieve the maximum fitness.

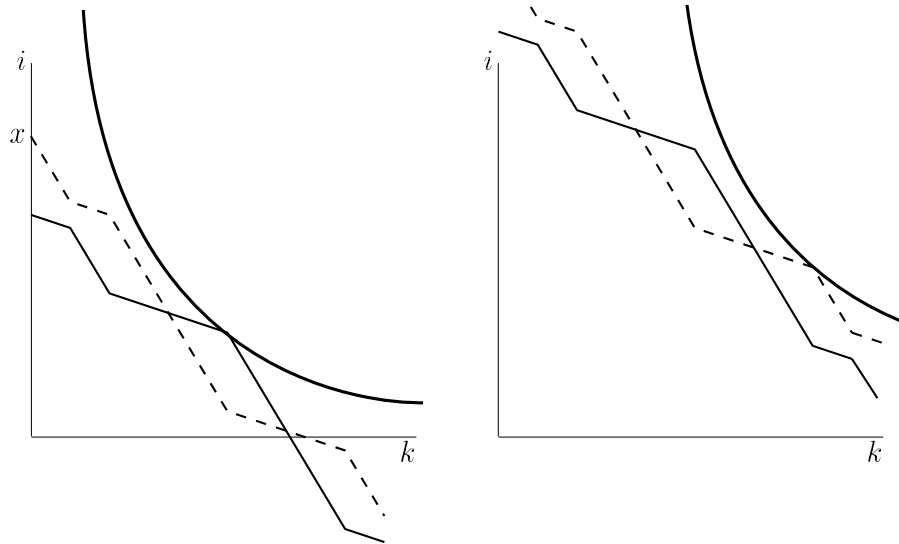
The diagrams in [Figure 1](#) illustrate the optimal organism for a fixed complexity x . The “budget” lines capture the tradeoff between intensity and precision for positively- (dashed) and negatively- (solid) aligned organisms respectively. Adding the j th unit of precision requires a sacrifice of one or two units of intensity, depending on the alignment and the value of λ_j . This yields the following budget equations

$$x = i + k \left(\frac{3}{2} - \frac{1}{k} \sum_{j=1}^k \frac{\lambda_j}{2} \right)$$

for positive alignment and

$$x = i + k \left(\frac{3}{2} + \frac{1}{k} \sum_{j=1}^k \frac{\lambda_j}{2} \right)$$

for negative alignment. The “indifference curve” is the set of pairs (i, k) which achieve the same fitness.



(a) Low x . Negative alignment (solid line) is optimal. (b) Higher x . Budget lines shift upward and now positive alignment is optimal.

Figure 1: Optimal organism for a fixed level of complexity x .

Figure 1(a) shows a case in which the optimal organism is negatively aligned. As the organism increases in complexity, the budget lines shift up, potentially switching the alignment of the optimal organism. This is illustrated in Figure 1(b). Indeed, the optimal alignment depends on the sign of the moving average

$$L(k) := \frac{1}{k} \sum_{j=1}^k \lambda_j > 0.$$

If it is positive, then the fraction of inverted signals up to k is greater than $1/2$, and the optimal organism will be positively aligned. The negatively aligned organism is optimal in the alternative case.

Recall that we have assumed that $l > 1/2$. This implies that for sufficiently complex organisms, positive alignment is optimal. A convenient way to visualize this is to consider k sufficiently large so that $L(k) \approx 2l - 1$ and the two budget lines are approximately

$$x \approx i + k(2 - l)$$

and

$$x \approx i + k(1 + l).$$

This is illustrated in [Figure 2](#).

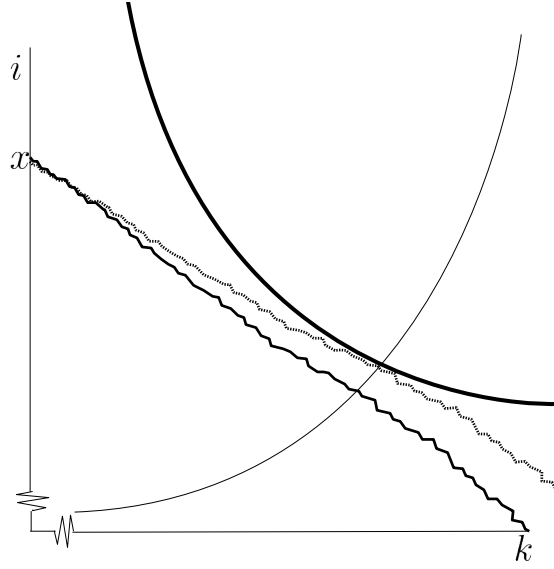


Figure 2: Optimal alignment for large k .

Kludge Note that for sufficiently complex organisms, positive alignment yields a greater budget. Once this is the case, any negatively aligned organism is attempting to implement the optimal decision rule via an inefficient protocol. For this reason and reasons developed further below, we refer to such an organism as a *kludge*.

Definition 1. Suppose that the fraction of inverted signals up to k exceeds $1/2$, i.e.

$$\frac{1}{k} \sum_{j=1}^k \lambda_j > 0.$$

Then we say that a negatively aligned organism with precision k is a kludge.

We can quantify the inefficiency of a kludge of complexity x . A switch to positive alignment would produce an organism of the same precision

but strictly higher intensity. Indeed the intensity and therefore the fitness can be increased by a number which (on average) increases linearly in k .

However, this measure may be hard to interpret as it depends on a cardinal interpretation of payoffs. As an alternative, let us define the following ordinal concept of inefficiency of an organism. Say that the organism is *asymptotically structurally inefficient* if there is a given component of the organism (here, a subset of tokens) such that at point in time, and forever thereafter, this component should be altered as a part of some improvement to the organism, but nevertheless the component remains fixed forever.⁶

2.2 The Adaptive Process

The organism is completely specified by the tuple $\mathcal{O} = \langle k, \pi, i, a \rangle$. Over time, the organism will adapt. I adopt a simple model of mutation and natural selection designed to capture the effects of a general class of adaptive processes. The specific assumptions are chosen mostly for expositional and analytical convenience.

Each period t , the organism \mathcal{O}_t is evaluated according to its overall fitness $V(\mathcal{O}_t)$. With positive probability, a mutation occurs which results in a variant \mathcal{O}' of the organism. If the variant \mathcal{O}' produced by a mutation is more fit, i.e. $V(\mathcal{O}') > V(\mathcal{O}_t)$, then the variant replaces the existing version and survives to date $t + 1$, that is $\mathcal{O}_{t+1} = \mathcal{O}'$. If not, then the existing version survives, i.e. $\mathcal{O}_{t+1} = \mathcal{O}_t$.

Mutations come in two varieties. With probability q , the organism increases in complexity. It keeps the analysis simple to assume that when complexity increases it increases by two, and the two additional computational steps are allocated optimally taking as given the existing allocation. On the other hand, with probability $(1 - q)$ the organism does not increase complexity, but some (possibly empty) subset of existing computational steps are re-allocated.

One simple and natural model of this latter component of the mutation

⁶A virtue of this definition is that it excludes “marginal inefficiencies” where at any point in time some inefficiencies are present, but every inefficiency, once it appears, is eventually eliminated. For example, we may imagine that the most recently developed features of the organism might begin in an inefficient state, but eventually as the organism matures, these features are improved to their optimal state and align optimally with the rest of the organism. By contrast, asymptotic structural inefficiency identifies persistent mis-alignments. It would be desirable to sharpen the definition even further by considering dynamic efficiency issues. Without going into the details of such a definition, I note that the kludges in this paper represent static as well as dynamic inefficiencies. Positively aligned organisms grow in intensity and precision faster than kludges.

process would be as follows. Think of each of the x steps as a gene. There is a fixed mutation probability $\mu > 0$ and each gene is subject to mutation with independent probability μ . When a gene other than a mutates, it can take on any value (input step, action step, preprocessor for input j) with some fixed positive probability. When the gene for a mutates, it changes sign. This model is useful for building intuition but far less structure is required. In the process of proving the main result we will establish a general class of mutation probabilities that deliver identical results (see [Definition 2.](#))

2.3 Analysis

The main result of the paper is

Theorem 1. *Suppose $\mu < 1/6$. When $q > 0$ there is a positive probability that the organism will be forever kludged and thus asymptotically structurally inefficient.*

In the remainder of this section, I will give an informal sketch of the proof. Recall that we have assumed that $l > 1/2$. The parameter l determines the probability that each $\lambda_j = +1$. As discussed above, what matters for the optimal design of the organism is the sign of the moving average

$$L(k) = \frac{1}{k} \sum_{j=1}^k \lambda_j > 0.$$

Because $l > 1/2$, with probability 1 there will exist some \bar{k} such that $L(k)$ will be positive for all $k > \bar{k}$. Let us consider a path in which for values of k immediately preceding \bar{k} , the value of $L(k)$ is negative.

Imagine that the organism has precision $k < \bar{k}$ and that the organism is optimally adapted. The scenario described thus far arises with positive probability.⁷ An optimally adapted organism will be negatively aligned when $L(k) < 0$. When $q > 0$, an optimally adapted organism can improve by increasing in complexity. In particular, its precision will continue to increase beyond the threshold \bar{k} . Once beyond that point, the organism is no longer optimally adapted. The additional information obtained by increased precision will be processed according to a protocol that is inefficient. Nevertheless these incremental improvements are optimal given the organism's pre-existing structure. The organism improves by applying a kludge.

⁷Even if the organism does not begin the process optimally adapted, there is always a positive probability at any date that a sufficiently large mutation occurs to make it so.

Because, $L(k)$ will be positive forever after \bar{k} , this negatively aligned protocol will remain inefficient forever. The question is whether the organism will ever become positively aligned, and thus optimally adapted, or remain a kludge forever, i.e. asymptotically structurally inefficient. Two forces are at work in opposite directions. First, the necessary mutation always occurs with positive probability. On the other hand, the organism improves by increasing complexity and we will show that a consequence of this is that the size of the change necessary for re-alignment increases, correspondingly decreasing its probability.

To analyze this model, consider the following simplified stochastic process. Let the states of the process correspond to the levels of overall complexity x of the organism. At each state, three transitions are possible. With probability q , the value of x increases by two. With probability $(1 - q)\eta_x$, the value of x is unchanged. Finally, with the remaining probability, the process terminates. Figure 3 illustrates. We set the initial value to be \bar{x} , the complexity of the organism at the stage in which its precision crosses \bar{k} .

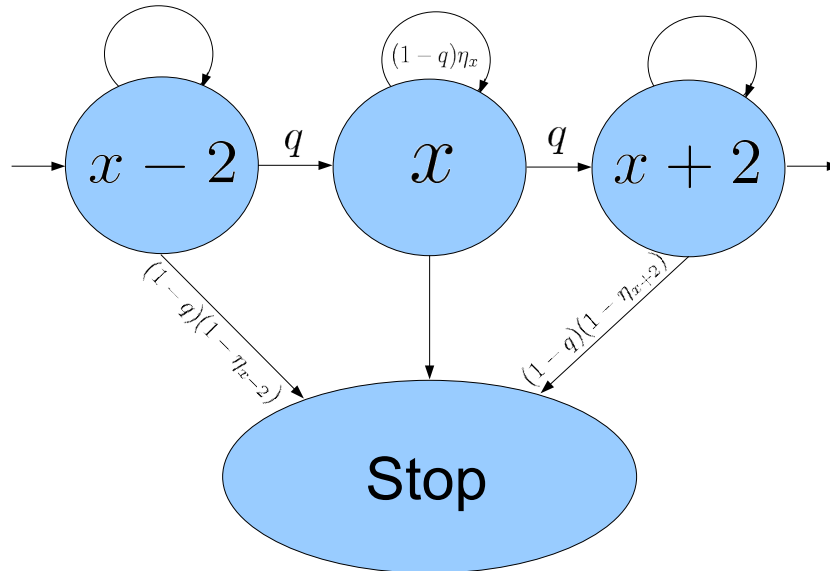


Figure 3: Stochastic Termination Process.

A *drastic mutation* is any mutation of sufficient size to profitably change the organism's alignment. We shall set $(1 - \eta_x)$ so that it bounds from above the probability that a kludge of complexity x will undergo a drastic mutation. Then, we can use this simplified process to place a lower bound on the probability that the organism remains kludged indefinitely. That probability will be no smaller than the probability that the simplified process never terminates.

I show that this probability is positive if and only if

$$\sum_x (1 - \eta_x) < \infty.$$

Next I show that the probability of a drastic mutation is determined by the precision k of the organism. Let $M(k)$ denote this probability. Say that a kludged organism \mathcal{O} is an *optimal kludge* if \mathcal{O} maximizes $V(\mathcal{O}')$ among all negatively aligned organisms \mathcal{O}' . Let us denote by $k(x)$ the precision of an optimal kludge which has complexity x .⁸ Then we can change variables

$$\begin{aligned} \sum_{x=\bar{x}}^{\infty} (1 - \eta_x) &= \sum_{x=\bar{x}}^{\infty} M(k(x)) \\ &= \sum_{k=\bar{k}}^{\infty} C(k) M(k) \end{aligned}$$

where $C(k) := |\{x : k(x) = k\}|$.

There are two steps to showing that this series converges. First, the probabilities of drastic mutations shrink very quickly. I show that the probability of a drastic mutation is bounded above by a function $\bar{M}(k)$ and

$$\limsup_k \frac{\bar{M}(k+1)}{\bar{M}(k)} < 1. \quad (2)$$

Second, along the optimal growth path of a kludge, the number of steps $C(k)$ the organism spends at a fixed level of precision k does not grow too fast. In particular, $C(k)$ is bounded by a function $\bar{C}(k)$ and

$$\lim_k \frac{\bar{C}(k+1)}{\bar{C}(k)} = 1. \quad (3)$$

Because it contains the key intuition for the result, I focus the discussion here on [Equation 2](#). Let \mathcal{O}^* be an optimal kludge with intensity i^* and precision k^* . As long as no drastic mutation has yet occurred, the organism will

⁸The value $k(x)$ is found by a discrete "first-order" condition as illustrated in [Figure 1](#).

remain an optimal kludge: additional computational steps from increased complexity will be allocated optimally.

Let us define the *effective precision* of an organism $\mathcal{O} = \langle k, \pi, i, a \rangle$ as follows.

$$\tilde{k}(\mathcal{O}) = \sum_{j=1}^k \text{sign}(a) \pi_j \lambda_j$$

Note that a positively aligned organism is equivalent to a positive organism whose precision is equal to its effective precision. Furthermore, the fitness of a positive organism with effective precision \tilde{k} is no greater than that of a positively aligned organism with precision k equal to \tilde{k} . Indeed, when \tilde{k} is strictly less than k , it is because some inputs are misaligned ($\pi_j \neq \lambda_j$), and each of these misaligned inputs "cancels out" the benefit of one properly aligned input. The number of "un-canceled" inputs is the effective precision \tilde{k} .

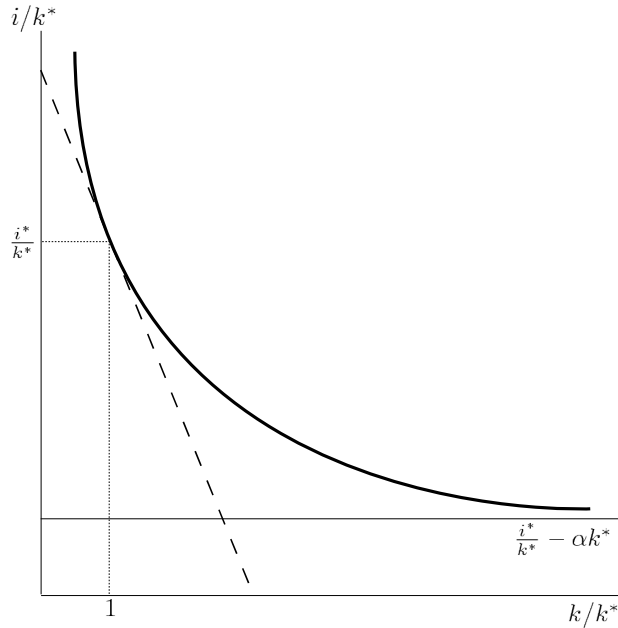


Figure 4: The optimal kludge \mathcal{O}^* . The axes have been rescaled. The asymptote represents the minimum intensity of any organism which achieves at least the fitness of \mathcal{O}^* .

Figure 4 illustrates the situation. The optimal kludge \mathcal{O}^* achieves the maximum fitness among all points on the budget-line for negative alignment. The horizontal axis is now the effective fitness of an organism. A necessary condition for an organism to achieve a higher fitness than \mathcal{O}^* is for the (effective fitness, intensity) pair to lie above the budget line. It is convenient to normalize the axes by dividing by k^* , yielding Figure 4.

As illustrated, the indifference curve has a horizontal asymptote. It represents the minimum intensity of any organism which achieves a higher fitness than \mathcal{O}^* . A key result is that the (normalized) difference between this minimum intensity and the intensity of an optimal kludge is bounded by a constant, α . That is, increasing the effective precision to infinity is not worth sacrificing more than αk^* units of intensity. Underlying this calculation is the observation that when the organism's estimate is already very precise, additional units of precision reduce the variance of the estimate (and thereby increase the per-action payoff) by at most a second-order magnitude.

Clearly, a drastic mutation requires that the organism switch to the positive decision rule. However, if a mutation changes only the decision rule from negative to positive, the intensity is unchanged and the organism's effective precision drops to $-k^*$, which in the normalized coordinate system is a horizontal movement to the point -1 . See Figure 5. Therefore a drastic mutation requires accompanying changes that increase the effective precision. There are three types of mutations that can increase effective precision.

1. Change an action step into an input step.
2. Change a pre-processor step to an input step.
3. Change a pre-processor step to an action step.

To find a drastic mutation, we must find some combination of these which results in an overall improvement over \mathcal{O}^* . Refer to Figure 5. Each mutation of type 1 moves one unit down and one unit to the right. Type 2 mutations move three steps to the right, and type 3 mutations move two steps to the right and one step upward. A drastic mutation is a path which combines these movements and moves (at least) beyond the dashed budget line.

I show that paths consisting entirely of type 1 mutations cannot constitute a drastic mutation. The reason is illustrated in Figure 5. Moving along

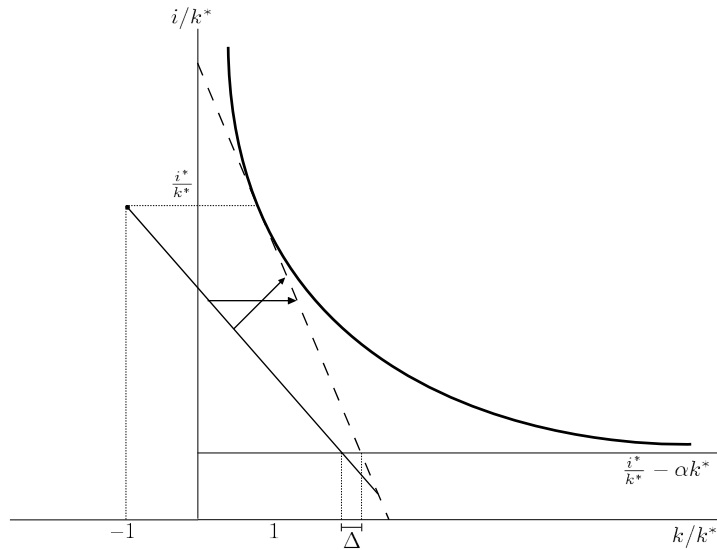


Figure 5: Paths to improvement. After the change in alignment, the (normalized) effective precision is -1. The downward sloping, horizontal, and upward sloping paths represent the three types of mutations.

the downward sloping line reaches the budget line only after crossing below the asymptote. Thus, the necessary increases in effective precision cost too much in terms of sacrificed intensity when using only mutations of type 1. Any drastic mutation must therefore involve some mutations of types 2 and 3. In fact I show that there is a constant Δ such that a drastic mutation to a kludge with precision k^* must involve at least Δk^* mutations of either type 2 or type 3. The constant Δ is found by identifying the path to the budget line which minimizes the number of these mutations. That path, illustrated in the figure, uses mutations of type 1 to reach the asymptote and mutations of 2 and then Δk^* mutations of type 2.

The proof is now concluded by applying a large-deviation result. We have shown that the probability of a drastic mutation is no greater than the probability that a proportion Δ of the “genes” from the set of pre-processor genes, of which there are at most k^* . When each gene has an independent probability $\mu < \Delta$ of mutating, a standard result from large-deviation theory is that this probability shrinks to zero exponentially fast as k^* increases. This immediately implies Equation 2, and establishes the theorem once we calculate that $\Delta > 1/6$. In fact, we can see from this sketch why the result

does not depend on the independent mutation model. Any distribution of mutations that satisfies such a large-deviation condition will deliver the same conclusion.

3 The Details

The Dirichlet Process The Dirichlet family of distributions is a convenient framework for modeling the organism's inference problem. We begin by reviewing some details about Dirichlet priors and multinomial sampling. A Dirichlet measure on the set $[0, 1]$, is parameterized by a pair z of non-negative integers (z^-, z^+) . We denote it by D_z . Consider a sampling process in which first, a probability p is secretly drawn from the prior D_z and then data from $\{-1, +1\}$ are sampled in i.i.d. fashion with p being the probability of observing the value 1. Suppose that the value -1 was observed y^- times and the value 1 was observed y^+ times. This yields a posterior distribution over p . The Dirichlet process D_z has the following properties, where $\bar{z} = z^+ + z^-$.

1. The unconditional expectation is

$$\mathbf{E}D_z = \frac{z^+}{\bar{z}}$$

2. The variance is

$$\text{var } D_z = \frac{z^+z^-}{(\bar{z})^2(\bar{z} + 1)}$$

3. Conditional on a sample $y = (y^-, y^+)$, the Bayesian posterior is

$$D_{(z+y)}.$$

Thus, a Dirichlet prior updated on the basis of observations from any finite sample remains in the Dirichlet family.

We fix $D = D_{(1,1)}$ and specify that the location θ is drawn from the distribution

$$F = 2D - 1$$

so that

$$\mathbf{E}(\theta | \sigma_1, \sigma_2, \dots, \sigma_k) = \frac{1}{k+2} \sum_{n=1}^k \lambda_n \sigma_n. \quad (4)$$

Thus, a fully adapted organism who knew the environment λ would use the formula on the right-hand-side to estimate the location from a sample of size k of signals. Recall that $\bar{\theta}_k$ denotes this Bayesian estimate of the location based on k observations. (Dependence on the specific observations is left implicit in the notation.)

Also, given that the unconditional expectation of θ is equal to zero, variances have a simple form

$$\text{var}(\bar{\theta}_k) = \mathbf{E}_\sigma (\bar{\theta}_k^2) . \quad (5)$$

and the variance of the estimate increases monotonically in the precision

$$\text{var}(\bar{\theta}_{k-1}) < \text{var}(\bar{\theta}_k) < \text{var}(\theta) = 1/3$$

for all k . The following lemma records some additional useful facts about the process that will be used later.

Lemma 1. *For any level of precision k ,*

$$\text{var}(\bar{\theta}_k) - \text{var}(\bar{\theta}_{k-1}) = \frac{1 - \text{var}(\bar{\theta}_{k-1})}{(k+2)^2}$$

and

$$\text{var}(\theta) - \text{var}(\bar{\theta}_k) < \frac{1}{k+3}$$

Proof. From Equation 4,

$$\bar{\theta}_k = \left(\frac{k+1}{k+2} \right) \bar{\theta}_{k-1} + \frac{\tau_j}{k+2},$$

so

$$\begin{aligned} \text{var}(\bar{\theta}_k) &= \mathbf{E} \left[\left(\frac{k+1}{k+2} \right)^2 \bar{\theta}_{k-1}^2 + 2 \left(\frac{\tau_j}{k+2} \right) \left(\frac{k+1}{k+2} \right) \bar{\theta}_{k-1} + \frac{1}{(k+2)^2} \right] \\ &= \mathbf{E}_{\tau_1, \dots, \tau_{j-1}} \mathbf{E} \left[\left(\frac{k+1}{k+2} \right)^2 \bar{\theta}_{k-1}^2 + 2 \left(\frac{\tau_j(k+1)}{(k+2)^2} \right) \bar{\theta}_{k-1} + \frac{1}{(k+2)^2} \mid \tau_1, \dots, \tau_{j-1} \right] \end{aligned}$$

By the properties of the Dirichlet process, $\mathbf{E}(\tau_j \mid \tau_1, \dots, \tau_{j-1}) = \bar{\theta}_{k-1}$, so

$$= \mathbf{E}_{\tau_1, \dots, \tau_{j-1}} \left[\frac{1}{(k+2)^2} + \frac{2(k+1) + (k+1)^2}{(k+2)^2} \bar{\theta}_{k-1}^2 \right]$$

and we have

$$\begin{aligned}
\text{var}(\bar{\theta}_k) - \text{var}(\bar{\theta}_{k-1}) &= \mathbf{E} \left[\frac{1}{(k+2)^2} - \left(1 - \frac{2(k+1) + (k+1)^2}{(k+2)^2} \right) \bar{\theta}_{k-1}^2 \right] \\
&= \mathbf{E} \left[\frac{1}{(k+2)^2} - \frac{1}{(k+2)^2} \bar{\theta}_{k-1}^2 \right] \\
&= \frac{1 - \text{var}(\bar{\theta}_{k-1})}{(k+2)^2}
\end{aligned}$$

To show the second part, note that

$$\begin{aligned}
\text{var}(\theta) - \text{var}(\bar{\theta}_k) &= \sum_{j=k}^{\infty} \text{var}(\bar{\theta}_{j+1}) - \text{var}(\bar{\theta}_j) \\
&= \sum_{j=k}^{\infty} \frac{1 - \text{var}(\bar{\theta}_j)}{(j+3)^2} \\
&< \sum_{j=k+3}^{\infty} \left(\frac{1}{j} \right)^2 \\
&< \frac{1}{k+3}
\end{aligned}$$

□

Actions and Payoffs The organism attempts to exploit the resource by choosing actions $a \in [-1, 1]$. Each action taken earns payoffs

$$u(a, \theta) = 2a\theta - a^2.$$

A fully adapted organism who observes $\sigma_1, \sigma_2, \dots, \sigma_k$ would optimally choose

$$a = \mathbf{E}(\theta \mid \sigma_1, \sigma_2, \dots, \sigma_k) = \bar{\theta}_k$$

and obtain conditionally expected payoff

$$\mathbf{E}(2\bar{\theta}_k\theta - \bar{\theta}_k^2 \mid \sigma_1, \sigma_2, \dots, \sigma_k) = \bar{\theta}_k^2$$

for each action taken. Recall that the number of actions taken is defined as the *intensity* of the organism. Thus, the total unconditional expected payoff (fitness) of a fully adapted organism with precision k and intensity i is

$$i \cdot \text{var}(\bar{\theta}_k).$$

3.1 The Organism

If an organism with precision k could be designed in a single step in order to maximize expected fitness, for each action it takes it would earn payoff $\text{var}(\hat{\theta}_k)$. In this section we describe a general class of algorithms (organisms) for solving the decision problem. Within this class there are algorithms which implement the optimal decision rule and algorithms that do less well. The organism is described by a list of instructions (its “genetic code”) for collecting, processing, and acting on information.

Tokens There is a collection of *tokens* which are labeled. The number of tokens is x , the complexity of the organism. There is one distinguished token, the *alignment* token, which is labeled with either $+$ or $-$. The sign of the alignment token indicates which decision-rule a the organism uses. Each of the remaining tokens can be labeled in one of three ways, $*$, \circ , or $-j$ where j is a positive integer. Tokens labeled $*$ are *action* tokens and the number of these tokens is the intensity of the organism. That is, the organism will take a number of actions equal to the number of action tokens. The remaining tokens dictate how the organism inputs and processes information.

Each token labeled with \circ is an *input* token and the number of these tokens is the precision of the organism. That is, the organism will observe a number of signals equal to the number of input tokens. Finally a token labeled $-j$ is a *pre-processor* token and it indicates that the value of the j th input should be multiplied by -1 .

The organism locates the resource by collecting a sample of size k , changing the sign of each input j for which a pre-processor token is present, adding these together and normalizing. Specifically, if we represent the pre-processor steps by

$$\pi_j = \begin{cases} -1 & \text{if } j \in \pi \\ 1 & \text{otherwise,} \end{cases}$$

then the estimate is

$$\beta(k, \pi) = \frac{1}{k+2} \sum_{j=1}^k \pi_j \sigma_j.$$

and the expected fitness of the organism is

$$V(\mathcal{O}) = i\mathbf{E}_\theta \mathbf{E}_{\sigma_1, \sigma_2, \dots, \sigma_k} [a\beta(k, \pi)\theta - \beta(k, \pi)^2 \mid \theta].$$

Mutation and Natural Selection Each period, with probability q , the organism increases in complexity by adding two new tokens. The labels on these tokens will be chosen to maximize the fitness of the resulting organism, taking as given the existing configuration of tokens. In the complementary event (probability $(1 - q)$) a mutation occurs and some (possibly empty) subset of existing tokens are re-labeled. A selected token is re-labeled with each label having equal probability, with the proviso that the alignment token must be labeled with $+$ or $-$, and no other token has these labels. We assume that tokens are selected according to a probability measure which satisfies monotonicity (larger sets have weakly smaller probability) and the following large-deviation condition.

Definition 2. Let M_x be a probability distribution over subsets of x tokens. We say that the family of distributions $\{M_x\}_{x \in \mathbb{N}}$ satisfies a large-deviation condition if there exists $\mu \in (0, 1)$ and a function $\delta : (\mu, 1] \times \mathbb{N} \rightarrow (0, 1)$ such that if T is any subset of $\{1, \dots, x\}$, and $m \geq \mu$, then the probability under M_x of selecting a mutation set which includes more than a fraction m of elements from T is no greater than $\delta_m(|T|)$ and

$$\limsup_N \frac{\delta_m(N+1)}{\delta_m(N)} \leq \beta(m) < 1.$$

This is a large-deviation property which limits the probability of selecting a large fraction m (greater than μ) of tokens from any given large subset. This probability must shrink to zero at a rate which is asymptotically faster than some $\beta(m) < 1$. Note that by a standard result from large-deviation theory, the independent mutation model discussed previously is a special case.

3.2 Analysis

First, we consider an instructive benchmark case in which $q = 0$. In this case, the complexity of the organism is fixed and cannot increase. Then, because the mutation probabilities are strictly positive, with probability 1 the organism will be optimally adapted after some finite timespan.

Benchmark with $q = 0$ Consider an arbitrary organism \mathcal{O} of complexity x . Let \mathcal{O}^* be an optimal organism of the same complexity. There is a lower bound on the probability of a mutation large enough to transform any such \mathcal{O} into \mathcal{O}^* . In the worst case, a change to the entire genetic structure will be required and the probability of such a large mutation is strictly positive by

assumption. When $q = 0$, the organism will never increase in complexity and so this remains forever a lower bound on the probability of reaching an optimally adapted organism in a single step. It follows that with probability 1 the optimal organism will appear eventually. Moreover an optimal organism can never be replaced if the complexity of the organism cannot increase.

Proposition 1. *When $q = 0$, with probability 1 the organism is eventually optimally adapted, regardless of the initial complexity.*

The proposition shows that any asymptotic inefficiency that arises when $q > 0$ is not due to a simple problem of local optima. The model allows for arbitrarily large mutations at any point in time. Thus, any improvement, of any fixed size, if available for sufficiently long, will eventually be realized. On the other hand, this argument does not apply to improvements which require larger and larger mutations. Potentially, the organism can gradually improve at the margin by increasing in complexity, all the while intensifying the complementarity among its components. This would mean that substantial improvements decline in probability. Whether such improvements will be realized will depend on the rate at which this probability declines.

Proof of Theorem 1 The remainder of the paper fills in the details of the proof sketched previously.

With probability 1, there exists a date after which positive alignment is optimal. With positive probability the organism is an optimal kludge at that date. The conditional probability that this remains true forever thereafter is bounded below by the probability that the process depicted in [Figure 3](#) never terminates. Let us calculate that probability. A standard tool from the theory of countable-state Markov chains⁹ indicates an analysis of the following system of equations in unknowns Z_0, Z_2, \dots

$$\begin{aligned} Z_0 &= qZ_2 + (1 - q)(\eta_{\bar{x}})Z_0 \\ Z_2 &= qZ_4 + (1 - q)(\eta_{\bar{x}+2})Z_2 \\ &\vdots \\ Z_x &= qZ_{x+2} + (1 - q)(\eta_{\bar{x}+x})Z_x \\ &\vdots \end{aligned}$$

⁹See (Billingsley, 1995, Theorems 8.4 and 8.5)

If the system can be solved by a bounded sequence Z_x , then the probability is strictly positive that the system will never terminate.

We can set $Z_0 = 1$ and then solve the system recursively, first writing

$$Z_{x+2} = \left(\frac{1 - (1-q)\eta_{\bar{x}+x}}{q} \right) Z_x$$

for each x , or

$$Z_{x+2} = \left(1 - \frac{(1-q)(1-\eta_{\bar{x}+x})}{q} \right) Z_x$$

then recursively substituting to obtain

$$Z_{x+2} = \prod_{n=2}^x \left[1 - \frac{(1-q)(1-\eta_{\bar{x}+n})}{q} \right].$$

We wish to show that $\lim Z_x < \infty$ which is equivalent to the convergence of the following series.¹⁰

$$\sum_{n=\bar{x}+2}^{\infty} \frac{(1-q)(1-\eta_n)}{q}.$$

which is convergent iff $\sum(1-\eta_n) < \infty$. So our focus is on the probability of a drastic mutation as a function of the complexity of the organism.

We can write

$$\sum_{x=\bar{x}}^{\infty} (1-\eta_x) = \sum_{k=k^*}^{\infty} C(k)M(k) \quad (6)$$

where $C(k)$ is the number of steps the process remains at a fixed level of precision k .

Suppose that an organism has intensity i and precision k . Then if the organism is an optimal kludge the following inequality must be satisfied.

$$i [\text{var}(\bar{\theta}_k) - \text{var}(\bar{\theta}_{k-1})] > \text{var}(\bar{\theta}_{k-1}).$$

The left-hand side is the marginal increment to fitness from an increase in precision, while the right-hand side is the marginal increment to fitness from instead increasing intensity. In fact, when $\lambda_j = -1$, two tokens are required to increase precision, so in that case the left-hand side must exceed

¹⁰Note that for any sequence of positive numbers R_n , $1 + \sum_1^x R_n \leq \prod_1^x (1 + R_n) \leq \exp(\sum_1^x R_n)$.

twice the right-hand side. An optimal kludge with precision k will increase intensity until it reaches the smallest level i which satisfies the corresponding inequality. It follows that the organism will increase its the level of precision from $k - 1$ to k as soon as

$$i = \left(1 + \mathbb{1}_{\lambda_j = -1}\right) \frac{\text{var}(\bar{\theta}_{k-1})}{\text{var}(\bar{\theta}_k) - \text{var}(\bar{\theta}_{k-1})}$$

(up to an integer,) and thus will spend at most

$$C(k) \leq \frac{2 \text{var}(\bar{\theta}_k)}{\text{var}(\bar{\theta}_{k+1}) - \text{var}(\bar{\theta}_k)} - \frac{\text{var}(\bar{\theta}_{k-1})}{\text{var}(\bar{\theta}_k) - \text{var}(\bar{\theta}_{k-1})}$$

steps of the process with precision k .

Applying [Lemma 1](#), we have

$$\begin{aligned} C(k) &\leq \bar{C}(k) := \frac{2 \text{var}(\bar{\theta}_k)}{1 - \text{var}(\bar{\theta}_k)} (k+3)^2 \\ &< \frac{2 \text{var}(\theta)}{1 - \text{var}(\theta)} (k+3)^2 \end{aligned}$$

so that

$$\lim \frac{\bar{C}(k+1)}{\bar{C}(k)} = \lim \frac{(k+4)^2}{(k+3)^2} = 1.$$

With this result, we can prove that the series in [Equation 6](#) converges by showing

$$\limsup_k \frac{\bar{M}(k+1)}{\bar{M}(k)} < 1. \quad (7)$$

I show [Equation 7](#) as a consequence of the following lemma which is the central result about the model.

Lemma 2. *An optimal kludge with precision $k^* > 14$ can be improved only by a mutation that includes at least the following fraction of the pre-processor tokens.*

$$\frac{2}{3} - \frac{1}{2} \left(\frac{k^* + 3}{k^*} \right).$$

Proof. (Preliminary. The proof covers the case of $l = 1$.) The effective precision of a positive organism is

$$\tilde{k}(\mathcal{O}) = \sum_{j=1}^k \pi_j \lambda_j$$

Note that a positively aligned organism is equivalent to a positive organism whose precision is equal to its effective precision. Moreover, the fitness of a positive organism with effective precision \tilde{k} is no greater than that of a positively aligned organism with precision equal to \tilde{k} . To see this, note first that for any even number z the estimator

$$\frac{1}{\tilde{k} + z + 2} \sum_{j=1}^{\tilde{k}} \lambda_j \sigma_j$$

is strictly worse than the optimal estimator from a sample of size \tilde{k} (see Equation 4). And when $z = 2$, we can show that the above estimator is strictly better than the estimate produced by an organism of precision $\tilde{k} + 2$ and effective precision \tilde{k} . The difference between the two estimators is that the latter incorporates two additional inputs, one of which is misaligned. When the signals from these two inputs have the same sign, the two estimators produce identical estimates. When the signals from these two inputs have opposite signs, the displayed estimator produces the optimal estimate while the latter does not. Now by induction, we can show that the displayed estimator is strictly better for any even number z .

Because \mathcal{O}^* is an optimal kludge, the intensity and precision satisfy the “first-order condition”

$$i^* \cdot [\text{var } \bar{\theta}_{k^*+1} - \text{var } \bar{\theta}_{k^*}] < 2 \text{var } \bar{\theta}_{k^*}.$$

Applying Lemma 1 and rearranging,

$$i^* < 2 \left[\frac{\text{var } \bar{\theta}_{k^*}}{1 - \text{var } \bar{\theta}_{k^*}} \right] (k^* + 3)^2.$$

Let us define α by the following equation. It gives the maximum amount by which intensity can be reduced and still produce a drastic mutation.

$$(i^* - \alpha k^*) \text{var } \theta = i^* \text{var } \bar{\theta}_{k^*}$$

$$\alpha k^* \text{var } \theta < 2 \left[\frac{\text{var } \bar{\theta}_{k^*}}{1 - \text{var } \bar{\theta}_{k^*}} \right] (k^* + 3)^2 (\text{var } \theta - \text{var } \bar{\theta}_{k^*})$$

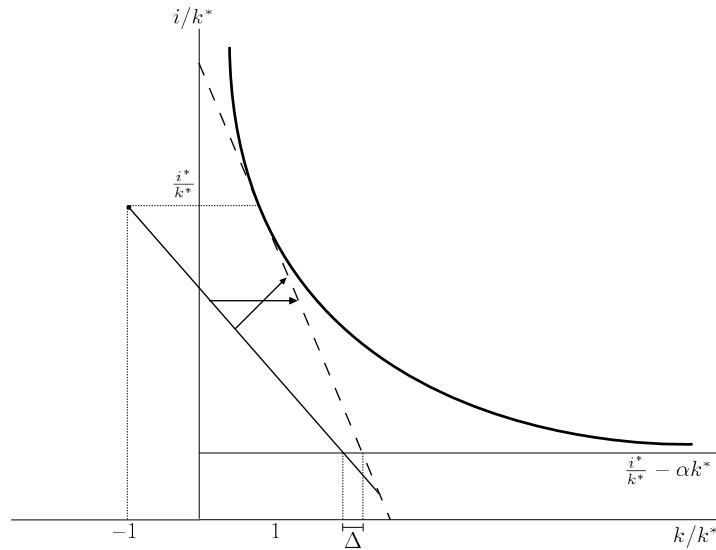
Applying lemma 1,

$$\alpha < \frac{2}{1 - \text{var } \theta} \frac{k^* + 3}{k^*}. \quad (8)$$

By the definition of an optimal kludge, a mutation results in an improvement only if the alignment is changed. Once the alignment is changed, the effective precision of the organism is $-k^*$. Therefore, in order for a mutation to result in an improvement, it must include accompanying changes that raise the effective precision. Recall the three types of mutations that can increase effective precision.

1. Change an action token into an input token.
2. Change a pre-processor token to an input token.
3. Change a pre-processor token to an action token.

To prove the lemma, we search for the path to improvement which involves the fewest mutations of types 2 and 3. We first show that mutations involving only changes of type 1 will not improve upon \mathcal{O}^* . Each change of type 1 reduces intensity by one unit and increases effective precision by one unit. Refer to Figure 5, reproduced below. Mutations of type 1 move along the solid line with slope -1. It improves upon \mathcal{O}^* only if it moves past the intersection point with the dashed budget line.



Noting that the slope of the solid line is -1 and the slope of the budget line is -2, the vertical coordinate of their intersection is $\frac{i^*}{k^*} - 4$. When $k^* >$

14, this is below the horizontal asymptote and hence no organism with such a low intensity could achieve a fitness higher than that of \mathcal{O}^* . It follows that when $k^* > 14$, no mutation consisting only of changes of type 1 can improve.

Next we can rule out paths that involve mutations of type 3. Any improvement must move from the solid line to the right of the budget line. Each type 3 step moves two units to the right and one unit upward. By comparison, each type 2 step moves three steps to the right. Because of the relative slopes of the two lines, type 2 steps close the gap more quickly.

Thus, in trying to find an improvement we can confine to paths that involve only mutations of type 1 and 2. Among these paths, we calculate the minimum number of type 2 steps required for an improvement. Any improvement must reach the budget line before falling below the horizontal asymptote of the indifference curve. Thus, the horizontal distance is bounded below by the minimum among points above the asymptote of the distance between the solid line (which is where the initial type 1 steps land) and the budget line. Because of the relative slopes of these lines, this minimum is obtained at the asymptote where the horizontal distance, denoted Δ in the figure, is equal to $\frac{4-\alpha}{2}$. (Moving a distance $4 - \alpha$ upward from the intersection point puts this much distance between the two lines because of their relative slopes.) Applying Equation 8, Δ is at least

$$\frac{1}{2} \left[4 - \left(\frac{2}{1 - \text{var } \theta} \right) \left(\frac{k^* + 3}{k^*} \right) \right] = 2 - \left(\frac{1}{1 - \text{var } \theta} \right) \left(\frac{k^* + 3}{k^*} \right)$$

which, multiplied by k^* gives the total increase in effective precision resulting from the horizontal type 2 steps. Recalling that $\text{var } \theta = 1/3$ and each type 2 step increases effective precision by 3, the total number of type 2 steps required is

$$\frac{k^*}{3} \left[2 - \frac{3}{2} \left(\frac{k^* + 3}{k^*} \right) \right]$$

and since the total number of pre-processor tokens is at most k^* , this yields the statement of the lemma. \square

Based on the lemma, we can bound $M(k)$ by the probability that a large fraction of the pre-processor tokens are selected for mutation. We apply the large-deviation property (recall Definition 2.) Pick m to satisfy

$$\mu < m < 1/6.$$

Suppose that for an organism of complexity x , the set of tokens subject to mutation will be selected from a distribution M_x which satisfies a

large deviation condition and monotonicity: larger subsets have smaller probability. Then since there are no more than k pre-processor tokens, the probability that a fraction m of these are selected for mutation is no greater than $\delta_m(k)$. We therefore set $\bar{M}(k) = \delta_m(k)$. By Lemma 2, $M(k) < \bar{M}(k)$ and since $\mu < m$, the large deviation property implies

$$\limsup_N \frac{\bar{M}(k+1)}{\bar{M}(k)} \leq \beta(m) < 1,$$

concluding the proof of Equation 7 and the Theorem.

References

- BILLINGSLEY, P. (1995): *Probability and Measure*. Wiley and Sons.
- FRIEDMAN, M. (1966): “The Methodology of Positive Economics,” in *Essays in Positive Economics*. University of Chicago Press.
- SAMUELSON, L., AND J. SWINKELS (2006): “Information, Evolution and Utility,” *Theoretical Economics*, 1(1), 119–142.
- SANDHOLM, W., AND A. PAUZNER (1998): “Evolution, Population Growth, and History Dependence,” *Games and Economic Behavior*, 22, 84–120.