Adaptation Is Not Just Improvement over Time

Simon McGregor*
University of Sussex
s.mcgregor@sussex.ac.uk

Pedro A. M. Mediano
Imperial College London
pmediano@imperial.ac.uk

Abstract The idea that an agent's actions can impact its actual long-term survival is a very appealing one, underlying influential treatments such as Di Paolo's (2005). However, this presents a tension with understanding the agent and environment as possessing specific objective physical microstates. More specifically, we show that such an approach leads to undesirable outcomes, for example, all organisms being maladaptive on average. We suggest that this problematic intuition of improvement over time may stem from Bayesian inference. We illustrate our arguments using a recent model of autopoietic agency in a model protocell, showing the limitations of previous approaches in this model and specific instantiations of Bayesian inference by ignorant observers in certain scenarios.

Keywords
Adaptivity, viability, Bayesian inference

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

1 Introduction

Most scientists believe that organisms, like other physical systems, can be construed as possessing a specific *objective* microstate, and interacting with their environment according to purely mechanical laws. Philosophers have long recognized that this view creates tensions with our intuitive views of agency, as reflected in debates about free will and determinism.

From this perspective, phenotypic plasticity corresponds to transitions in the microstate of an organism, and this plasticity (e.g., changes in synaptic weights in the brain) will often be beneficial to the organism. Hence, we might imagine that the benefit of phenotypic plasticity could be understood in terms of the organism being mechanically constituted so that its own future survival prospects improve *diachronically* with beneficial plasticity, that is, that if we plotted a graph of its prospects over time, the graph would show an increase at the time of the transition.

This intuition finds theoretical expression in formulations such as Di Paolo's notion of adaptivity [6], recently explored using a concrete model of protocell behavior [1]. However, we analyze this intuition mathematically, and conclude that it is logically problematic, because the future behavior of a known mechanical system should already be factored into our assessment of its current prospects. We explain, using Bayesian methods, how an organism can *appear* to improve its future survival prospects through its actions, but only in the eyes of an observer, who implicitly uses the actions to infer previously unknown information about the organism.

* Corresponding author.

© 2018 Massachusetts Institute of Technology

1.1 Companion Articles

This article is the first of two companion articles considering the notion of adaptation from an abstract, formal perspective:

1. "Adaptation Is Not Just Improvement over Time" critiques an influential view of adaptation, under which an organism can be mechanically constituted so as to improve its future prospects over time.
2. "Measuring Fitness Effects of Agent-Environment Interactions" proposes a quantitative measure for the degree to which an agent's internal dynamics capture beneficial information from its sensory stimuli.

The two articles can be read independently, but deal with closely related themes.[1]

1.2 Structure of This Article

Section 2 provides an informal introduction to the intuition that we consider problematic: that an organism can actively change the future prospects implied by its current microstate. We explore in Section 3 how this theme has influenced the theoretical literature on adaptation, and especially Di Paolo's influential treatment of adaptivity. In particular, Section 3.1 introduces a formal definition of adaptivity, which has previously been applied in a model of protocell behavior [1]. In Section 4 we expose the flaws of this approach to defining and quantifying adaptivity, and in Section 5 describe how these flaws may be understood as a manifestation of Bayesian reasoning and deceptive human intuitions. Section 6 concludes the main body of the article.

2 On Actual Prospects Improving over Time

Intuitively, it seems that an organism can be constituted so as to tend to improve its own future prospects. For instance, consider a prey animal at time t_1, being stalked by a predator. Between t_1 and t_2, it spots the predator, which causes it to behave more cautiously than usual at time t_2. We might naively suppose that the animal's prospects have "objectively" improved between time t_1 and time t_2, since behaving cautiously is a wise move in the presence of predators (i.e., other things being equal, life expectancy is higher for alert organisms than non-alert ones when predators are around).

But according to physicalism, the organism and its environment are two coupled physical systems with respective microstates x_1 (a non-alert state) and y_1 at time t_1; the dynamics of this system cause a transition to states x_2 (an alert state) and y_2 at time t_2. Logically, one of the two following conditions must hold:

1. State x_1 (in environment state y_1) might have evolved into some non-alert state at time t_2, had stochastic events unfolded differently (e.g., thermal noise inside the animal's brain).
2. State x_1 (in environment state y_1) was predestined to evolve into an alert state at time t_2, in which case its prospects at time t_1 were far better than we might first have realized.

In case 1, the organism's life expectancy may indeed increase between time t_1 and time t_2, but only because of a lucky chance event; in case 2, the organism was already predestined at time t_1 to become alert, which needs to be taken into account in a calculation of its life expectancy. Either way, there is no way that the organism could have been constituted so as to systematically increase its own objective life expectancy over time.

Briefly, physicalist assumptions imply that an organism's future prospects already depend on all relevant aspects of its internal state and environment. The organism-environment system's future

1 Both articles are found in this issue of *Artificial Life*.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

will unfold mechanically over time, with its immediate future becoming its present, and stochastic effects will cause fluctuations in the organism's future prospects.

We can certainly compare our animal with a different organism, constituted so as to show no caution in the face of danger, and they are very likely to have different prospects; but despite what our intuitions tell us,

- the difference in the prospects of these organisms does not arise at the time of the actions, but already exists before then, because the organism's prior constitution determines its *disposition* to display caution (or not) under an appropriate trajectory of environmental influences; and
- an organism's behavior cannot change what its *own* future will hold, since the behavior simply forms part of its state trajectory over time.

These misleading intuitions may stem from a fact about observers of an organism: If an observer does not know what action the organism will take, it is indeed possible (and in some cases even rational) for their *subjective judgment* of the organism's prospects to improve in a nontrivial manner, as a consequence of observing the organism's actions. This is because those actions can provide new evidence about whether it is the sort of organism that has good prospects or bad prospects. We discuss this further in Section 5.

However, if we want to treat the organism as a physical system, then our theoretical treatment of it will typically assume that it has a particular internal state and that the agent-environment system has some particular dynamical laws (even if we do not specify what that state or those laws are).

This has a counterintuitive character: Surely when an animal's state changes in such a way as to become increasingly robust to repeated perturbations, the animal's later states must have better prospects than its previous states did? In short, the surprising answer is no: An animal cannot possess behavioral propensities that cause its average long-term prospects to improve over time; the animal's later states will typically have *worse* prospects than the previous ones; and any deviation from this rule can only be due to random perturbations, rather than systematic properties of the animal.

3 Related Work and Common Descriptions of Adaptivity

The notion of *adaptivity* plays a major role in artificial life research, and has been the subject of much thought for several decades.[2] A significant part of the modern ideas in the philosophy and science of adaptivity can be traced back to W. Ross Ashby's seminal work in cybernetics [3]. Some of the concepts introduced by Ashby were formalized and extended by Di Paolo's 2005 essay [6], which has had a significant impact in the literature (with over 400 citations, according to Google Scholar, at the time of our writing).

For both Di Paolo and Ashby, the notion of adaptivity relies on the existence of a *viability boundary*—a boundary in the organism's state space beyond which it is dead or irreversibly bound to die. Then, an adaptive system can be informally defined as one that tends to move away from such a boundary and to remain within its viability zone despite external perturbations.

More precisely, Di Paolo [6, p. 438] defines adaptivity as:

> A system's capacity, in some circumstances, to regulate its states and its relation to the environment with the result that, if the states are sufficiently close to the boundary of viability,
> 1. Tendencies are distinguished and acted upon depending on whether the states will approach or recede from the boundary and, as a consequence,

2 We note that the notion of adaptivity we treat in this article is different from the technical notion of adaptation in evolutionary biology.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

2. Tendencies of the first kind are moved closer to or transformed into tendencies of the second and so future states are prevented from reaching the boundary with an outward velocity.

According to this definition, an adaptive organism must be able to adjust its own future prospects to maintain itself within the limits of viability. In turn, this also relies on the existence of two different kinds of variables within the organism: *essential variables*, which the system wants to preserve within a physiologically safe limit (such as blood pressure or nutrient concentration), and *control variables*, which the organism can manipulate to achieve its goals (which may include, e.g., reproduction, in addition to self-preservation).

This definition suffers from several problems, including the distinction between essential and control variables, and the definition of the viability boundary itself. To this respect, Egbert and Barandiaran call upon metabolism—understood as the process that maintains a set of processes relevant to the organism's life in equilibrium—to define the viability boundary [7]. States within the viability boundary are states in which the metabolism can operate. Along similar lines, Barandiaran and Moreno [5] argue that a basic metabolic organization provides a conceptual framework for adaptive behavior, but it is incomplete without some regulatory subsystem that, again, tends to maintain the whole system in equilibrium amidst an ongoing flux of matter and energy.

We can see recurrent themes in previous definitions of adaptivity. The terms regulation, normativity, adjustment, and viability are often used. The organism adapts if it can steer itself away from certain fatal regions of its state space defined by its internal metabolic dynamics. We note this definition has a strong implicit *survival-normative* assumption—namely, the only event relevant to the organism is whether it crosses the viability boundary (i.e., whether it lives or dies).

On a separate line of inquiry, and much closer to our view, early work by Klyubin defines adaptivity from an *organization-centric perspective* as "increased robustness to repeated perturbation" [10]. For Klyubin, a system is adaptive insofar as it is able to settle into different stable organizations in different external environments. That is, a process of adaptation is a process of structural change, specific to the organism's present environment, that contributes to the organism's stability if it is exposed to that environment. Note that Klyubin and Di Paolo both follow in the autopoietic tradition of focusing on some notion of homeostasis or self-maintenance; however, while Di Paolo's notion considers *diachronic* differences in viability (i.e., differences in viability at separate points in time along the same physical trajectory), Klyubin's (as we interpret it, at least) considers *counterfactual* differences in viability (i.e., differences in viability under counterfactually different physical conditions). We explore this in more detail in our companion article.

There is of course great nuance in the implications of any definition of adaptivity. It is indeed a pervasive idea and has been linked, among others, to concepts like autopoiesis [12], autonomy, agency [4], and intelligence [11]. These are outside the scope of this article, and we refer the reader to the original articles for more information.

As we discuss extensively later in this article (and as has been suggested before; see, e.g., [13, 14]), there are fundamental problems with the notions of adaptivity stemming from Ashby's legacy. Historically, the concept of adaptation has not always been associated with survival normativity; sometimes the more general case of an arbitrary fitness measure has been considered (see, e.g., [15]). The approach found in our companion article treats the more general case, removing the focus on survival normativity found in the autopoietic literature.

3.1 Formalizing and Illustrating Di Paolo's Adaptivity

3.1.1 Viability and Adaptivity

Agmon, Gates, and Beer (henceforth, AGB) have used a particular concrete model to explore Di Paolo's abstract notion of adaptivity [1, 6]. An organism is taken to have a state that can be either alive or dead; in the AGB model, there are many live states and one dead state.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

AGB, following di Paolo, define the following terms, which we rephrase in a more mathematically explicit notation:

Viable. All states x are defined as *viable*, with the exception of a single designated absorbing *dead* state d.

Viability. The *viability* $V(x)$ of a state x is defined as the expected number of time steps for a system starting in state x to enter the dead state d, under a time-independent distribution over external perturbations.

Adaptivity. The *adaptivity* $A_{x \to x'}$ of a transition from a state x to another state x', after a single time step, is defined as

$$A_{x \to x'} = V(x') - V(x). \tag{1}$$

Adaptive transitions are defined as those that have positive adaptivity, and *maladaptive transitions* as those that have negative adaptivity. N.B. AGB restrict the calculation of adaptivity to cases where $x \neq x' \wedge x' \neq d$, but do not provide any principled reason why the transitions $x \to x$ or $x \to d$ should be treated in a mathematically special manner; hence, we have dropped AGB's requirements that $x \neq x'$ and $x' \neq d$.

This abstract definition does not depend on the details of the simulation, but applies to any system capable of undergoing an irreversible transition that can be interpreted as death. It has the virtue that it makes the notion of the system's effects on its own prospects both quantifiable and readily calculable. We believe that this is the correct measure of change in an organism's survival prospects. However, as we show in Section 4, it suffers from multiple drawbacks that make it an inappropriate measure to quantify adaptivity.

3.1.2 Details of the Simulation Model

AGB's model of adaptivity is a two-level hierarchical model simulated at two different time scales; viability is defined in terms of time steps at the higher level of this model. The lower level is a deterministic reaction-diffusion-repulsion system where each state takes the form of a grid of cells with real-valued concentrations of several chemical species in each cell. The system has a variety of attractor states; these attractor states resemble a spatially heterogeneous internal chemistry, surrounded by a chemical membrane, in a homogeneous external environment. There is also a distinguished homogeneous *dead state* attractor. In each higher-level time step, an attractor state is subjected to one of several different instantaneous perturbations, and then permitted to go to equilibrium;[3] transitions at the higher level of dynamics are therefore from one lower-level stable state x to another lower-level stable state x' (Figure 1).

In the original AGB article, the authors simulated point perturbations to two different chemical species, at three different magnitudes, in nine distinct locations, for a total of 72 distinct perturbations [1]. Beginning with a single initial stable microstate, they computed a network of other microstates reachable from this initial state by some series of perturbation-relaxation cycles. They arbitrarily terminated their search at 16 cycles, discovering 267 distinct attractors in total within this distance (including the dead state), including 113 attractors whose successors were not computed. For the purposes of calculating viability, they assumed that the uncomputed successors of all these 113 attractors were exactly the dead state.

This produced a graph with 267 nodes, whose edges were labeled with one of the 72 perturbations. Placing a probability distribution over the perturbations makes it possible to formulate the model as an irreducible discrete Markov process. For more technical details we refer the reader

3 AGB do not report finding any non-point attractors.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

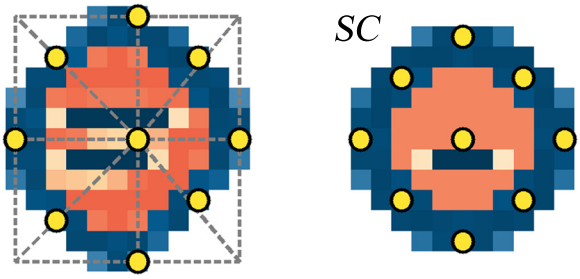


Figure 1. Illustration of system states in the AGB protocell model. The left-hand image shows the nine locations for perturbations, and the right-hand image shows the initial state for the Markov process, labeled as SC (stable configuration) by the authors. Reproduced with permission from the authors [1].

to the original articles [1, 2]. AGB provided us with the data for this graph; we used this data without running the underlying simulation.

Throughout the rest of this article we will refer back to the AGB model and use it as a practical numerical example to illustrate notions of adaptivity. Technically, it constitutes an absorbing Markov chain [8, p. 416]. However, our results are fully general, apply in the non-Markov case, and do not depend on the details of the AGB protocell model.

4 Intrinsic Limitations of Prospective Survival-Based Approaches

If one can measure the survival prospects (viability) of an organism, it may seem reasonable to ask questions about how these prospects change over time: in other words, to compare the organism's viability at one time step with its viability at the next time step. This is in fact the core of the proposals by AGB [1, 6], among others (see Equation 1).

However, in this section we show that such an approach has several intrinsic limitations, which arise directly from the mathematical properties of survival measures. We prove in Section 4.1 that if one identifies viability directly with expected life span (as in [1]), then the average change in viability at each time step will be -1, regardless of the form of the stochastic process underlying the organism-environment system's dynamics; and we explain why we consider this unsatisfactory.

In Section 4.2, we identify a general problem that must affect any (reasonable) scalar measure of future survival prospects: The average change in this measure over an organism's entire life span must be negative. Finally, in Section 4.3, we discuss technical issues relating to the potentially infinite quantities that life span measures can induce.

Of course, there are important considerations for biological agents beyond averages of measures, since the average does not tell us whether random perturbations causing abnormally high (or low) measure values are possible at all. In an evolutionary context, the range of possible variation may be as important as the average. However, when discussing the case of individual (lifetime) adaptation, we contend that expected (mean) adaptivity is a fair measure to consider.

4.1 Adaptivity Always Averages to −1

Consider an arbitrary discrete stochastic (not necessarily Markovian) process over random variables X_i with $i \in \mathbb{N}$ (for a one-way infinite process) or $i \in \mathbb{Z}$ (for a two-way infinite process), each X_i taking values in $\mathcal{X}$, and a designated *dead* state $d \in \mathcal{X}$ such that

$$\forall i,j \in \mathbb{N} : j > i, \forall \mathbf{h} \in \mathcal{X}^n, \quad P\big(X_j = d \mid X_{i+1} = d, X_{\leq i} = \mathbf{h}\big) = 1, \tag{2}$$

where $X_{\leq i}$ represents the complete history of the organism up to and including time index i, and n is the *size* of the history, with $n = i$ in the one-way infinite case, and $n = \aleph_0$ in the two-way infinite case.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

The previous equation merely states that the dead state invariably transitions to itself with probability one, regardless of the system's prior history. The AGB model as described above fits naturally into this description.

Define the *expected life span* $L(\mathbf{h})$ of a history $\mathbf{h} \in \mathcal{X}^n$ as the expected number of time steps before the system enters the dead state d, after following history $\mathbf{h}$. Consider the system's future at time index i: In particular, consider how many times it will ever again occupy a non-d state (including h_i itself, if $h_i \neq d$). This is identical to the number of transitions that will occur before it dies, that is, its life span. Hence, its expected life span is the same as the expected number of times that any of its subsequent states will be alive, which is easy to write, and is well known from the study of absorbing Markov chains:

$$L(\mathbf{h}) = \sum_{j \geq i}^{\infty} \mathbb{P}(X_j \neq d \mid X_{\leq i} = \mathbf{h}) = \sum_{t=0}^{\infty} \mathbb{P}(X_{i+t} \neq d \mid X_{\leq i} = \mathbf{h}). \tag{3}$$

Call a history $\mathbf{h}$ *properly mortal* iff this sum converges. Call $\mathbf{h}$ *terminated* iff $h_i = d$. An obvious, but nonetheless crucial result is presented in the following theorem.

Theorem 1. *A properly mortal system's life span falls (on average) at the unit rate until death, regardless of state.*

Proof. Assume that a system's first n states are observed to be $\mathbf{h} \in \mathcal{X}^n$ and that $\mathbf{h}$ is properly mortal. If $\mathbf{h}$ is terminated, then $L(\mathbf{h}) = 0$; otherwise, we have

$$L(\mathbf{h}) = \mathbb{P}(X_i \neq d \mid X_{\leq i} = \mathbf{h}) + \sum_{t=1}^{\infty} \mathbb{P}(X_{i+t} \neq d \mid X_{\leq i} = \mathbf{h}) \tag{4}$$

$$= 1 + \sum_{t=1}^{\infty} \sum_{x' \in \mathcal{X}} \mathbb{P}(X_{i+t} \neq d \mid X_{i+1} = x', X_{\leq i} = \mathbf{h}) \mathbb{P}(X_{i+1} = x' \mid X_{\leq i} = \mathbf{h}) \tag{5}$$

$$= 1 + \sum_{x' \in \mathcal{X}} L(\mathbf{h}x') \mathbb{P}(X_{i+1} = x' \mid X_{\leq i} = \mathbf{h}) \tag{6}$$

$$= 1 + \langle L(\mathbf{h}x') \rangle_{|\mathbf{h}}, \tag{7}$$

where $\langle \cdot \rangle_{|\mathbf{h}}$ is the expectation operator conditioned on $\mathbf{h}$, and $\mathbf{h}x'$ is the sequence $\mathbf{h}$ with the symbol x' concatenated. This gives

$$\langle L(\mathbf{h}x') \rangle_{|\mathbf{h}} - L(\mathbf{h}) = \begin{cases} 0 & \text{if } \mathbf{h} \text{ is terminated,} \\ -1 & \text{otherwise.} \end{cases} \tag{8}$$

In other words, the expected change in the system's expected life span, following one time step, is -1 unless the system is already in the dead state, in which case it is 0. □

The implications of this theorem for AGB's abstract definition of adaptivity are immediate. For every non-dead attractor state x with $L(x) < \infty$, the mean adaptivity of its transitions must be exactly -1. The adaptivity of every transition from the dead state must be exactly 0. This would imply the absurd conclusion that (ignoring immortal systems) every biological organism is systematically (i.e., averaged over stochastic noise) maladaptive at every point in its life span, and that dead organisms are (again, averaged over stochastic noise) more adaptive than living ones. For instance, suppose a bacterium enters a state in which the expression of certain genes changes in response to the detection of an external toxin (thereby preserving the bacterium's capacity for replication): Presumably, whatever adaptivity means, we would want to say that this environment-sensitive state is associated

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

on average with positive adaptivity. By contrast, we have proved above that the "adaptivity" measure A that will be associated with this state is sampled randomly from some distribution with mean -1.

We note that, naturally, in particular instantiations of the stochastic process an observer might measure adaptivity values different from -1. Also, as reported in [1, section "Quantifying viability and adaptivity"], particular (x, x') pairs may have $A_{x \to x'} \neq -1$. The problem only becomes apparent when talking about average adaptivity. If we consider an ensemble of organisms in a given state, they will always on average be 1 step closer to the dead state.

Furthermore, this reasoning extends trivially to the whole state space of the organism. Since the average adaptivity of state x is identically -1, the average adaptivity of all non-dead states is also -1. This means that any organism, with any internal dynamics, in any possible environment, has the same average adaptivity as any other organism.

This property of the definition makes it difficult to rank organisms or states in terms of how adaptive they are. One could compare one particular organism or transition with an ensemble, and as we show above, the ensemble average is consistently -1; parenthetically, this means that a transition can still be "maladaptive," in the sense of having $A < 0$, even if it is better than the ensemble average.

For completeness, we derive exact values for expected viability and lifetime-mean viability (with an arbitrary viability function) for Markov processes in the Appendix.

4.2 Lifetime Mean Viability Changes

Our previous results regarding the definitions of adaptivity in AGB's [1] and Di Paolo's [6] work apply where the viability measure is identical with the expected life span. These results do not necessarily hold if viability is redefined to mean something other than that. In this section, however, we demonstrate that adaptivity has another undesirable property under *any* reasonable viability measure. Under very weak constraints on a measure of viability, any increases in viability made at time t must logically be canceled out by reductions in viability later on in life, with a residual negative term.

Suppose the function $V_{\mathbb{P}} : \mathcal{X}^* \to \mathbb{R}$ is an arbitrary measure of the viability $V_{\mathbb{P}}(\mathbf{h})$ of a variable-length history $\mathbf{h}$ in a (not necessarily Markov) stochastic process defined by $\mathbb{P}$. We will require that

$$V_{\mathbb{P}} \text{ is a monotonically decreasing function in } \mathbb{P}(X_{i+t} = d \mid X_{\leq i} = \mathbf{h})$$

for all $t \in \mathbb{N}$, $\mathbf{h} \in \mathcal{X}^*$. In other words, we require that if any two histories $\mathbf{h_1}$, $\mathbf{h_2}$ entail the same probability of being dead at every future time step except one, the history that carries the higher probability of being dead at that time step must have a lower viability. This would seem intuitively to be a minimal necessary condition for measures of viability.

Consider an organism transitioning from state h_i (after following history $\mathbf{h}$) to state $x \in \mathcal{X}$. Following AGB, the adaptivity A of this transition is defined as

$$A(\mathbf{h} \to x) = V_{\mathbb{P}}(\mathbf{h}x) - V_{\mathbb{P}}(\mathbf{h}), \tag{9}$$

that is, as the difference in viability between $\mathbf{h}$ alone, and $\mathbf{h}$ with x appended to it. Note that this generalizes the Markovian case, as for a Markov process $A(\mathbf{h}x \to x') = A(x \to x')\ \forall \mathbf{h}$. We will now consider the average value $\hat{A}_{\mathbf{h}}(\mathbf{j})$ of A over every transition in a sequence of states $\mathbf{j} = (j_1 \ldots j_m) : \mathbf{j} \in \mathcal{X}^m$, given a preceding history $\mathbf{h} \in \mathcal{X}^n$:

$$\hat{A}_{\mathbf{h}}(\mathbf{j}) = \frac{1}{m} \sum_{i=1}^{m} A(\mathbf{h}\mathbf{j}_{1 \ldots i-1} \to j_i). \tag{10}$$

The definitions above allow us to formulate the following theorem.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

Theorem 2. *Lifetime-mean adaptivity is always negative.*

Proof. Trivially, since $\hat{A}$ is a sum of consecutive differences (i.e., a telescopic sum), it reduces to

$$\hat{A}_{\mathbf{h}}(\mathbf{j}) = \frac{1}{m}(V_{\mathbb{P}}(\mathbf{hj}) - V_{\mathbb{P}}(\mathbf{h})). \tag{11}$$

Now suppose that $j_m \neq d$ and $j_{m+1} = d$, such that $\mathbf{j}$ consists of the entire life trajectory of the system after following $\mathbf{h}$, until one step before the system enters the dead state. It seems reasonable to require that

$$\forall \mathbf{h} \in \mathcal{X}^* \quad V_{\mathbb{P}}(\mathbf{h}d) = c, \tag{12}$$

that is, the viability of the system once it has reached the dead state does not depend on its previous history (this should be particularly persuasive in the Markov case). Consequently, the mean lifetime adaptivity $\hat{A}_{\mathbf{h}}(\mathbf{j}d)$ must satisfy

$$\hat{A}_{\mathbf{h}}(\mathbf{j}d) = \frac{1}{m}(c - V_{\mathbb{P}}(\mathbf{h})). \tag{13}$$

From the monotonicity of $V_{\mathbb{P}}$ and the fact that d is an absorbing state, we have that

$$\forall \mathbf{k} \in \mathcal{X}^* \quad V_{\mathbb{P}}(\mathbf{k}) \geq c \tag{14}$$

(with equality when $\mathbf{k}$ is terminated), giving

$$\hat{A}_{\mathbf{h}}(\mathbf{j}d) < 0. \tag{15}$$

□

In other words, the lifetime-mean adaptivity of every transition from any point until the organism's death must be negative, regardless of the current organism's state, its history, and even the precise form of the viability function (provided that the function is monotonic and assigns the same viability to all terminated histories). This, unlike the problem described in the previous section, does not depend on ensembles or averages of various organisms, but affects every possible life trajectory of each instantiation of the process. It occurs essentially because the sum of a series of consecutive differences telescopes into the difference between the final term and the initial term. Since the inevitable end of the lifetime trajectory (the dead state) is the worst possible state from a survival point of view, lifetime-mean adaptivity will always be negative.

This property is at best counterintuitive, and at worst absurd. Phenotypic plasticity of the sort described as "adaptive" is considered evolutionarily valuable, so presumably adaptivity is supposed to be beneficial, in some persistent sense, during an organism's lifetime. But consider the very first state transition in the organism's life history $\mathbf{h}$, from h_1 to h_2, and let us suppose that its adaptivity $A(h_1 \to h_2) > 0$. Let us write $A(h_1 \to h_2)$ as the lifetime mean adaptivity, as expressed in Equation 13, plus some residual term $R(h_1 \to h_2)$:

$$R(h_1 \to h_2) = A(h_1 \to h_2) - \frac{1}{m}(c - V_{\mathbb{P}}(\mathbf{h})). \tag{16}$$

This residual term is, so to speak, merely a loan that must be repaid in full later on: Any positive benefit it represents is guaranteed to be exactly canceled by corresponding maladaptive transitions in the subsequent life trajectory of the agent; moreover, it does not represent any tangible resource that the agent can exploit better in the short term than in the long term. In some sense, then, the only persistently meaningful contribution that the supposedly adaptive transition $h_1 \to h_2$ can make is via

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

the factor $\frac{1}{m}$ and the term $V_{\mathbb{P}}(\mathbf{h})$ in Equation 13; even if one believes these dependences capture something important, the specific measure $A(h_1 \to h_2)$ can in general be dominated by the ephemeral residual $R(h_1 \to h_2)$.

This difficulty, or one like it, will apply to all purely diachronic definitions of adaptivity: ones that rely on comparing successive points in the organism's trajectory—that is, the organism's past with its future. Instead, as we describe in our companion article, to solve this problem we propose a measure that relies on comparing the organism's trajectory with other possible, counterfactual trajectories.

4.3 Immortality

Expected life span is a potentially problematic measure because, in principle, an organism may live forever. Infinities "percolate" through additive expressions (such as the expectation operator), meaning that even agents that eventually die can have infinite expected life spans while they are still alive. We show in the Appendix that viability, and hence adaptivity, is undefined for any state in a Markov process that can lead to an immortal state with probability different from 0.

This fact is awkward when dealing with analytical or simulation models such as the protocell model [1],[4] but it should also be remarked that certain biological organisms such as *Hydra* appear to be potentially immortal [9].

5 Human Intuitions and Bayesian Reasoning

As has been discussed above, a large body of commonly used definitions of adaptivity are based on the intuition that an organism's actions can improve its future expected life span. However, as has been repeatedly argued in Section 4, these definitions have unintuitive mathematical consequences. Where, then, might the underlying intuitions come from?

There is a scenario in which our intuitions about an organism's actions improving its expected life span do hold true: the scenario in which the agent's internal details are imperfectly known by the theorist. In such a case, the expectation is an average over a subjective Bayesian probability distribution that, as the theorist observes the behavior of the system, will evolve in line with the theorist's accumulated knowledge. That is, if the theorist has only incomplete knowledge, it is perfectly rational to believe that an organism can improve its expected life span. However, it should be noted that this phenomenon is an artefact of the observer having incomplete knowledge, and not a feature of the objective agent-environment dynamics.

To explore this scenario we turn to Bayesian statistics. Bayesian statistics can be seen as the mathematical idealization of reasoning about observable evidence under uncertainty. The underlying principle is that a probability distribution may be interpreted as capturing a rational observer's subjective state of uncertainty, and that when relevant evidence is observed, that state of uncertainty should change in the way prescribed by Bayes' rule.

We present now a minimal example that illustrates this point. Consider a system whose evolution is determined by a Markov process on a finite set of states $\mathcal{X}$. We can calculate the expected life span of an agent in an unknown state at time t, under an observer's subjective probability distribution over the agent's state. We formulate this as the average life span of an ensemble of agents, distributed according to a probability distribution $\mathbb{P}$, that is,

$$\langle L_t \rangle = \sum_{n \in \mathbb{N}} n \mathbb{P}(L_t = n), \tag{17}$$

4 The Markov process derived from AGB's model that we analyze in this article contains no immortal states, but this could be due to the arbitrary search cutoff (after 16 perturbations) that was described in Section 3.1.2.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

where L_t is the agent's life span at time t. This may be decomposed into a weighted sum involving the specific possible states of each agent in the ensemble, as

$$\langle L_t \rangle = \sum_{n \in \mathbb{N},\ x \in \mathcal{X}} n \mathbb{P}(L_t = n \mid X_t = x) \mathbb{P}(X_t = x). \tag{18}$$

Importantly, now $\mathbb{P}(X_t)$ represents the observer's subjective uncertainty about the organism's state. Next, suppose the observer sees the agent undergo some particular transition to state x'. This transition may provide some life-span-relevant information about the agent's current state, such that the observer's updated expectation of the agent's life span is

$$\langle L \mid X_{t+1} = x' \rangle = \sum_{n \in \mathbb{N},\ x \in \mathcal{X}} n \mathbb{P}(L = n \mid X_t = x, X_{t+1} = x') \mathbb{P}(X_t = x \mid X_{t+1} = x'). \tag{19}$$

Even if the new state x' provides the observer with no new information about the environment—and hence $\mathbb{P}(L_t = n \mid X_t = x, X_{t+1} = x') = \mathbb{P}(L_t = n \mid X_t = x)$—something has changed: the observer's belief state. That is, the observer-assigned probability $\mathbb{P}(X_t = x)$ that the agent's state is x, before observing x', is not the same as $\mathbb{P}(X_t = x \mid X_{t+1} = x')$, the observer's belief that the agent's state is x, after observing x'. Hence, a perfectly rational observer will change her estimate in light of new evidence of the organism's state. We show empirically in this section that this induces a tendency to overestimate the organism's life span, that is, in general $\langle L_t \mid X_{t+1} = x' \rangle > L_t(x)$.

5.1 The Cautious Prey

For instance, consider the prey animal described in the introduction. When the animal transitions to the cautiously behaving state x_2, this transition does not (except by luck) improve the animal's own prospects, as we have extensively argued in Section 4. However, suppose an observer did not know how (if at all) the animal responded to the presence of predators, prior to observing the cautious behavior. Then, the cautious behavior will provide evidence that the organism is the sort of animal that tends to detect and avoid predators, rather than the sort of animal that does not; the former sort of animal typically has a longer life span than the latter, so the observer's subjective expectation of the animal's life span will increase.

It is important to appreciate the following detail in this scenario. Upon observing the cautious behavior, the Bayesian observer should retrospectively update her estimate of the animal's life span before the behavior was observed, as well as updating her estimate of the animal's present life span. In other words, evidence of this sort that the organism's life span will be longer than previously estimated is (typically) evidence that the organism's life span in the past was also correspondingly longer. However, in the following subsection we show numerically, based on the AGB, that even if the observer updates her estimate in hindsight, the transition is still more likely to be considered adaptive.

5.2 The AGB Model

We will now show numerically how observations of the protocell in the AGB model can affect a rational observer's estimates of the protocell's survival prospects.

Assume the observer has perfect knowledge of the model's dynamics in the form of a transition probability $\mathbb{P}(X_{t+1} \mid X_t)$, but is uncertain about which state the protocell is in. The observer can express her belief about the state of the protocell in a distribution $q(x)$. The life span measured by the observer is then not a function of the protocell's real state, but of the observer's belief state, and can be written as

$$\hat{L}(X) = \mathbb{E}_{q(x)}[L(x)] = \sum_{x \in \mathcal{X}} L(x) q(X = x), \tag{20}$$

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

where $L(x)$ is the true expected life span of the agent in state s, calculated as shown in the Appendix. Assume the observer is completely ignorant of the state of the agent, that is, $q(x) = |\mathcal{X}|^{-1}$. The observer's estimation of the agent's life span under this prior is $\hat{L}(X) = 3.35$, that is, the mean life span of all states. Suppose she now observes the system transition to a particular state x'. A rational observer would then update her beliefs about the previous state of the system using standard Bayesian inference:

$$q(x \mid x') = \frac{p(X' = x' \mid X = x)q(x)}{\sum_r p(X' = x' \mid X = r)q(r)}. \tag{21}$$

This belief, updated after the observation of x', will lead to a corrected estimation of the viability of the previous (unknown) state X, which we can write as

$$\hat{L}(X \mid X' = x') = \mathbb{E}_{q(x' \mid x)}[L(x)] = \sum_{x \in \mathcal{X}} L(x)q(x \mid x'). \tag{22}$$

This is the posterior estimate of the life span of the creature in the previous time step, given an observation of x' in the current time step. With this we can now define the *perceived adaptivity* of the transition $X \to x'$ as

$$\hat{A}(X \to x') = L(x') - \hat{L}(X \mid X' = x'), \tag{23}$$

which we can compare against the real adaptivity of the transition, given by

$$A(x \to x') = L(x') - L(x). \tag{24}$$

Figure 2 shows the relation between the expected life span $L(x')$ of the current state and the estimated life span $\hat{L}(X \mid X' = x')$ of the previous state, for every possible state in the model. We observe that these two quantities are strongly correlated and lie close to the identity line $\hat{L}(X \mid X' = x') \approx L(x')$. In less mathematical terms, we say that if the observer sees the agent transitioning into a "good" state, she will tend to think the previous state was also "good."

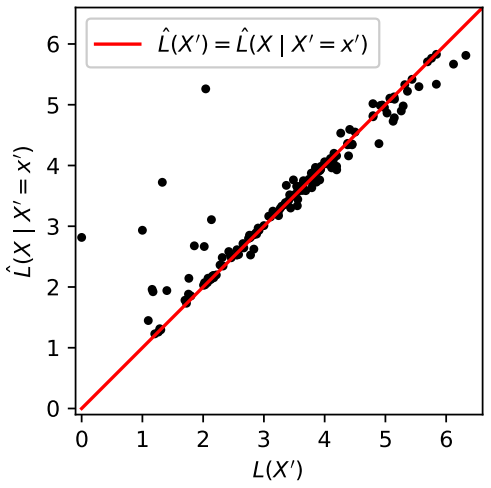


Figure 2. Expected life span of a current state $L(x')$, plotted against the conditional expectation of the previous state's life span $\hat{L}(X \mid X' = x')$, as calculated by a Bayesian observer with a uniform prior on X. Identity line shown in red.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

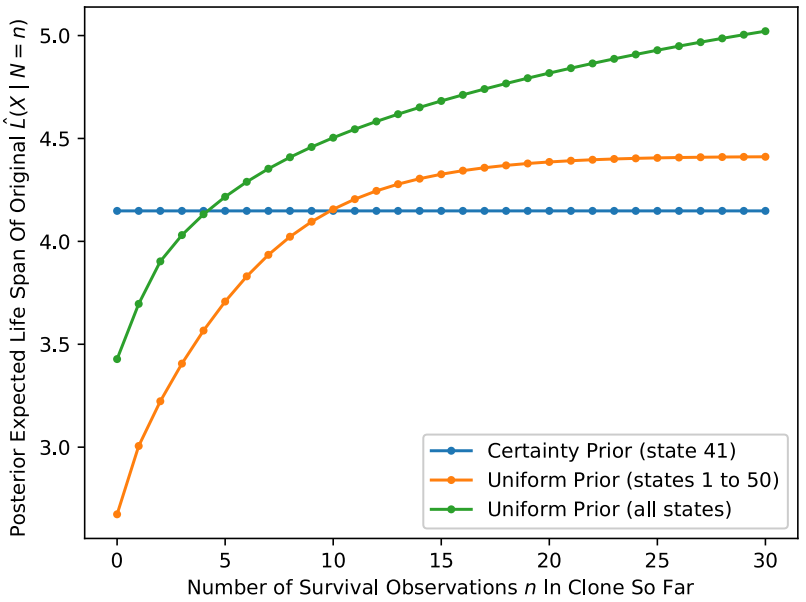


Figure 3. Observer-expected (posterior) life span for states in the AGB model, after observing n non-lethal transitions in a clone of the original state. Results are shown for three different priors over states.

If the observer has perfect knowledge of the state of the system, she is able to calculate $A(x \to x')$. With this metric, 33% of all possible transitions in the protocell model are considered adaptive, in the sense of having a positive adaptivity $A(x \to x') > 0$. Conversely, an ignorant (but rational) observer estimating the perceived adaptivity $\hat{A}(X \to x')$ will consider 39% of the transitions to be adaptive. In other words, an observer ignorant of the agent's past is more likely to believe the organism improved its survival prospects, and therefore more likely to judge a transition as adaptive.

For the second part of this argument we consider the case where the observer also has incomplete knowledge of the new state x'. Imagine that we *clone* the original state of the organism (making a perfect copy), "freeze" time for the original, and subject the clone to a number of unknown perturbations. After each perturbation, the observer is not told in which state the clone is, but only whether it is alive or dead. What effect will a series of non-lethal transitions in the clone have on the observer's expectation of the original organism's life span?

Figures 3 and 4 illustrate how the answer to this question will depend on how much the observer knows about the organism's original state. In particular, Figure 3 shows the mean life span ascribed

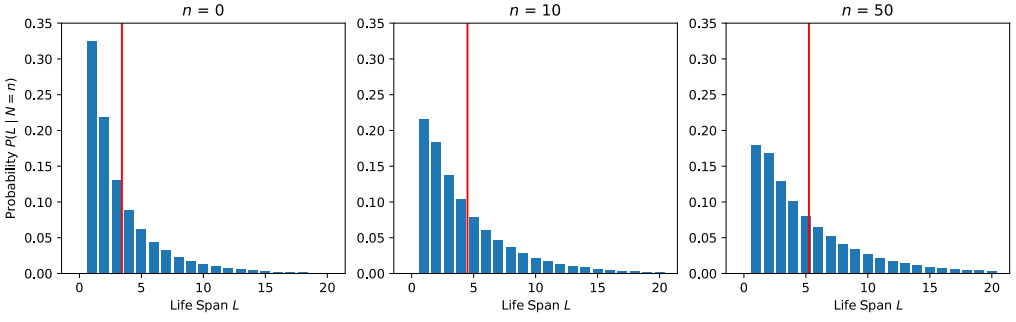


Figure 4. Posterior probabilities for the original state's life span, after observing n non-lethal transitions in a clone of the original state, assuming a uniform prior over states.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

to the original organism as the clone is seen to survive n consecutive perturbations. We see that if the original state is already known with certainty, the observer can directly calculate the original state's expected life span from the transition matrix, and information about the clone's survival is irrelevant. Otherwise, her estimate of the original's life span will increase as the clone lives longer. Exactly how the estimates change will depend on the prior: If the original state of the organism is partially known (e.g., known to be in states 1 to 50, in the AGB terminology), the estimates grow more slowly than if the observer is completely ignorant of the true state. This is because observations of the clone's survival provide less new information, the more the observer knows already about the organism.

More specifically, Figure 4 shows in detail how the observer's full subjective probability distribution over possible life spans changes with n observations of non-lethal transitions in the clone, beginning from a uniform prior over states (i.e., the dots on the green line in Figure 3 are the means of the histograms in Figure 4). Here we can clearly see that the more perturbations the clone endures, the more peaked and more right-shifted the distribution of possible life spans becomes.

In summary, we have shown that in general an observer's estimate of an organism's life span can change as she observes life-preserving behavior on the part of the organism; however, this occurs only if the behavior provides the observer with previously unknown information about the agent's state. Moreover, the change of estimate should also be applied retrospectively to the observer's estimate of the agent's original state, before the behavior was emitted.

This principle was demonstrated by applying it to the AGB protocell model. Observer-relative improvements in expected life span occur after an observation of a non-lethal transition, even if that observation occurs in a clone of the original system, and hence the observed event does not necessarily occur in the original system's state trajectory. These effects may explain the intuitions behind some of the previously proposed measures of adaptivity.

6 Conclusion

The intuitive notion that an organism can improve its own objective long-term survival prospects through its actions is ubiquitous in the artificial life literature. However, these intuitions do not appear to withstand thorough mathematical examination. Most importantly, we have proved here that definitions of adaptivity based on this idea of improvement over time lead to counterintuitive (and in our opinion problematic) results—for example, that the lifetime mean adaptivity of any organism is always negative, or that the average expected adaptivity of any organism in any state is always -1.

We have shown that these intuitions could have a well-founded basis in Bayesian reasoning, which in cases of imperfect knowledge does predict that the estimate of an organism's viability should increase as the organism is seen to display life-preserving behavior. This is, however, an artefact of an ignorant observer, and not a feature of the objective agent-environment dynamics.

To solve the problems with other adaptivity measures exposed in this article, we advocate a view of adaptivity as a causal property of the agent-environment system. In a companion article we present a family of such measures and show how they can be used to study the adaptivity of a system from a theoretically grounded standpoint. In particular, we adopt a *counterfactual* rather than purely *diachronic* approach, with the conclusion that positive changes in (arbitrarily construed) fitness over time should not necessarily be regarded as adaptive change.

Acknowledgments

The authors would like to thank Eran Agmon for providing data for the simulation of the protocell model used in Section 5, as well as discussion and images. We would also like to thank anonymous reviewers who provided helpful feedback, and the editor of *Artificial Life* for permitting an unconventional submission in the form of dual companion articles.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

References

1. Agmon, E., Gates, A. J., & Beer, R. D. (2015). Ontogeny and adaptivity in a model protocell. In P. Andrews, L. Caves, R. Doursat, S. Hickinbotham, F. Polak, S. Stepney, T. Taylor, & J. Timmis (Eds.), *Proceedings of the European Conference on Artificial Life 2015* (pp. 216–223). Cambridge, MA: MIT Press.
2. Agmon, E., Gates, A. J., Churavy, V., & Beer, R. D. (2016). Exploring the space of viable configurations in a model of metabolism–boundary co-construction. *Artificial Life*, *22*(2), 153–171.
3. Ashby, W. R. (1960). *Design for a brain: The origin of adaptive behavior* (2nd ed.). London, UK: Chapman and Hall.
4. Barandiaran, X. E., Di Paolo, E., & Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, *17*(5), 367–386.
5. Barandiaran, X. E., & Moreno, Á. (2008). Adaptivity: From metabolism to behavior. *Adaptive Behavior*, *16*(5), 325–344.
6. Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, *4*(4), 429–452.
7. Egbert, M. D., & Barandiaran, X. E. (2011). Quantifying normative behaviour and precariousness in adaptive agency. In T. Lenaerts, M. Giacobini, H. Bersini, P. Bourgine, M. Dorigo, & R. Doursat (Eds.), *Proceedings of the 11th European Conference on Artificial Life* (pp. 210–217). Cambridge, MA: MIT Press.
8. Grinstead, C. M., & Snell, J. L. (2003). *Introduction to probability*. Providence, RI: American Mathematical Society.
9. Khokhlov, A. N. (2014). On the immortal hydra. Again. *Moscow University Biological Sciences Bulletin*, *69*(4), 153–157.
10. Klyubin, A. (2002). *An organization-centric approach to viewing adaptation in complex adaptive systems*. Master's thesis, Tallinn Technical University, Tallinn, Estonia.
11. Martín H. J. A., de Lope, J., & Maravall, D. (2009). Adaptation, anticipation and rationality in natural and artificial systems: Computational paradigms mimicking nature. *Natural Computing*, *8*(4), 757–775.
12. Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht, Holland: Reidel.
13. McGregor, S. (2006). Is adaptiveness a coherent idea? In L. M. Rocha, L. S. Yeager, M. Bedau, D. Floreano, R. L. Goldstone, & A. Vespignani (Eds.), *Artificial life X: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems* (pp. 534–540). Cambridge, MA: MIT Press.
14. McGregor, S. (2016). A more basic version of agency? As if! In L. Tuci, A. Giagkos, M. Wilson, & J. Hallam (Eds.), *SAB 2016: Proceedings of the 14th International Conference on the Simulation of Adaptive Behavior* (pp. 183–194). London: Springer.
15. Zadeh, L. A. (1962). On the definition of adaptivity. *Proceedings of the IEEE*, *51*(3), 469–470.

Appendix. Exact Calculations for Discrete Markov Chains

This appendix considers the case where the stochastic process mentioned in Section 4.1 is a discrete Markov process over a finite set of states $\mathcal{X} = \{x_1, \ldots, x_K\}$, with transition probabilities $\mathbb{P}$. In this instance, the definitions of *expected life span* and *properly mortal* in Section 4.1 can be applied to individual states $x = \mathcal{X}$ instead of histories $\mathbf{h} \in \mathcal{X}^i$, with conditionings on $X_{\leq i} = \mathbf{h}$ replaced by conditionings on $X_i = x$.

We extend the standard expression for the expected life span of states in an absorbing Markov chain (see, e.g., [8]) to the general case of possibly non-absorbing Markov processes, and derive an expression for the lifetime-mean change in viability for an arbitrary viability function V.

Lemma 1. *If any immortal state is accessible from x, then x is not properly mortal.*

Proof. Define a state x as *immortal* if d is not accessible from x. If any immortal state u is accessible from a mortal state x, then there must exist some $m \in \mathbb{N}$ and some $p \in (0, 1)$ such that $\mathbb{P}(X_m = u \mid X_0 = x) = p$.

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

Let

$$a = \sum_{n=0}^{m} \mathbb{P}(X_n \neq d \mid X_0 = x), \quad b = \sum_{n=m}^{\infty} \mathbb{P}(X_n \neq d \mid X_m \neq u, X_0 = x); \tag{25}$$

then

$$L(x) = a + (1-p)b + p\sum_{n=m}^{\infty} \mathbb{P}(X_n \neq d \mid X_m = u, X_0 = x) \tag{26}$$

$$= a + (1-p)b + p\sum_{n=m}^{\infty} 1. \tag{27}$$

This clearly does not converge, as each term in the infinite sum over n is equal to one, and all terms on the right-hand side are non-negative. □

Derivation 1. *Expected life span $L(x)$ for Markov processes.*

Define $\mathcal{M} = \{m_1, \ldots, m_T, d\}$ as the set of all states in $\mathcal{X}$ from which no immortal state is accessible. $\mathcal{M}$ is exactly the set of states $x \in \mathcal{X}$ that have a well-defined expected life span $L(x)$, for the following reasons:

1. The state d is an absorbing state that is accessible from every element of $\mathcal{M}$; hence, $\mathbb{P}$ defines an absorbing Markov chain over $\mathcal{M}$, and the expected time $f(x)$ for each state $x \in \mathcal{M}$ to enter a closed communicating class[5] is well defined.
2. The set $\{d\}$ can be the only closed communicating class in $\mathcal{M}$, so every state $x \in \mathcal{M}$ has a well-defined expected life span $L(x) = f(x)$.
3. $L(x)$ is not well defined for any state $x \in \mathcal{X} \setminus \mathcal{M}$, by Lemma 1.

The expected time to enter an absorbing state from a state $m_i \in \mathcal{M} \setminus \{d\}$ is given by the elements of the vector $\mathbf{v}$, defined by the standard equation

$$\mathbf{v} = (I - Q)^{-1}\mathbf{1}, \tag{28}$$

where Q is the submatrix of $\mathbb{P}$ containing only the rows and columns corresponding to the $m_i \in \mathcal{M} \setminus \{d\}$ states. Hence, the expected life span of any state $s \in S$ is given by

Result 1.1. *Expected life span $L(x)$ for Markov processes:*

$$L(x) = \begin{cases} 0 & \text{if } x = d, \\ v_i & \text{if } x \in \mathcal{M}\ \{d\}, \\ \perp & \text{(undefined) otherwise.} \end{cases} \tag{29}$$

5 A closed communicating class C is a maximal set of states such that every state in C is accessible from any state in C, and no state outside C is accessible from any state in C [8].

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024

Derivation 2. *Expected lifetime-mean adaptivity for Markov processes with an arbitrary viability function.*

We will assume that the dead state d has finite viability $V(d)$, and hence serves as a constant offset; without loss of generality, we may set it to zero. Hence, Equation 13, describing the mean adaptivity during a future life span that starts with the history x_i, followed by some series $\mathbf{j}$ of non-dead states, and ending in death, can be written as

$$\hat{A}_{x_i}(\mathbf{j}d) = -\frac{1}{m} V(x_i), \tag{30}$$

where the c term in Equation 13 is set to zero and $m = |\mathbf{j}d|$. We can now define the expected lifetime-mean adaptivity $\hat{A}$ of a particular state x_i as

$$\hat{A}(x_i) = -V(x_i) \sum_{n=1}^{\infty} \frac{1}{n} P(L = n), \tag{31}$$

where L is the observed life span of the organism. This is simply the expectation of Equation 30 over all future trajectories. If we write the transition matrix P as

$$P = \begin{pmatrix} \underset{\sim}{Q} & \mathbf{r} \\ \mathbf{0} & 1 \end{pmatrix}, \tag{32}$$

then we can write the lifetime-mean adaptivity $\hat{a}_i$ of a particular state as the ith element of the vector $\hat{\mathbf{a}}$,

$$-\hat{\mathbf{a}} = V \sum_{n=1}^{\infty} \frac{1}{n} \underset{\sim}{Q}^n \mathbf{r}, \tag{33}$$

where V is a diagonal matrix such that $V_{ii} = V(x_i)$ and $\mathbf{r}$ is a vector such that $r_i = \mathbb{P}(X_{t+1} = d \mid X_t = x_i)$. Writing $\underset{\sim}{Q}$ in diagonal form as UWU^{-1}, we have

$$\hat{\mathbf{a}} = -VUMU^{-1}\mathbf{r}, \tag{34}$$

where

$$M = \sum_{n=1}^{\infty} \frac{1}{n} W^n, \tag{35}$$

and it is easy to verify that M is a diagonal matrix whose elements m_{ij} are

$$m_{ij} = \begin{cases} -\log(1 - w_{ij}) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{36}$$

Downloaded from http://direct.mit.edu/artl/article-pdf/24/3/182/1666990/artl_a_00268.pdf by UNIVERSITY COLLEGE LONDON user on 07 November 2024