Emergence of Darwinian evolution in Bayesian collectives

Márton Csillag[1], Hamza Giaffar[2], Eörs Szathmáry[1,3,4], Dániel Czégel[1,5]

[1]Institute of Evolution, Centre for Ecological Research, 1121 Budapest, Hungary
[2]Halıcıoğlu Data Science Institute, University of California, San Diego, La Jolla, CA 92093, USA
[3]Parmenides Foundation, Center for the Conceptual Foundations of Science, 82343 Pöcking, Germany
[4]Dept. of Plant Systematics, Ecology and Theoretical Biology, Eötvös Loránd University, 1117 Budapest, Hungary
[5]Beyond Center for Fundamental Concepts in Science, Arizona State University, Tempe, AZ 85287, USA

The emergence of self-replication in chemical space has led to an explosive diversification of form and function. It is hypothesized that a similar process underlies human action selection in complex combinatorial spaces, such as the space of simulated action sequences. Furthermore, the spontaneous appearance of a non-predesigned evolutionary search in artificial agents might lead to a higher degree of open-endedness, arguably a key missing component of current machine intelligence. In this paper we design a computational model to show that Darwinian evolutionary dynamics over informational units can emerge if collectives of such units need to infer statistics of changing environments. We build our argument on a series of equivalences between Bayesian computations and replicator dynamics to demonstrate that the selective advantage of higher information transmission ability between units and of larger population size appear very early on, already at no consistently shared information between the population size of two units. Further selection for statistical inference at the collective level leads to a continuous increase of transmission fidelity and population size until the population reaches the ability to maintain and iteratively improve combinatorial information, a transition to the regime of Darwinian evolution. Candidate systems include prebiotic collectives of non-replicating molecules, collectives of neural ensembles representing competing action plans, and reinforcement learning agents with parallel policy search.

Introduction

Darwinian evolution has been interpreted and formalized as learning in various ways[1]. Besides the apparent similarity that they both perform adaptation to high-dimensional and complex environments[2], evolution exhibits hallmarks of memory across scales, formalized as autoassociative networks[3]. Such memory storage and adaptive retrieval aid development when implemented by gene regulatory networks[4–6]; at a much larger temporal and spatial scale it also provides adaptive explanations at the level of ecosystems, when implemented by evolutionary ecological dynamics among interacting species[7].

Another class of adaptive behavior that is ubiquitous across the animal kingdom and learning machines alike is statistical inference[8–10]. Environments are noisy and changing at various timescales - extracting relevant information from such environments that can drive behavior by predicting future environments is essential for survival and reproduction, as well as adaptive behavior by artificial agents[11–13]. It has been shown that basic statistical inference problems, such as Bayesian inference[14–16], filtering in hidden Markov models[17], and inference in hierarchical models [18], can be implemented by simple systems of Darwinian replicators. Viewed these dynamical-structural analogies from evolutionary theory perspective, they suggest that replicator systems as wholes perform Bayesian computations despite not being units of selection themselves[19].

This partial unification of replicator dynamics and Bayesian computations makes it possible to bridge the two domains[19,20]. In practice, it opens the possibility to hybrid architectures and mechanisms, exhibiting both Darwinian and Bayesian properties[21,22]. In this paper, we propose such a unified mechanism for the emergence of self-replication. We suggest that when collectives are under selection of computing accurate statistical inference over changing environments, they might undergo a Darwinian transition akin to a critical step during the origin of life, whereby high-accuracy self-replication of information emerges, providing unlimited hereditary potential[23,24]. Such a transition entails an adaptive trajectory that starts at low information transmission among a few units and leads to high-accuracy heredity between many units, as illustrated by Figure 1A.

In animal brains, such representational units might be neural ensembles, scaling up from a low number of loosely functionally connected ensembles to a large number with a high-accuracy information copying ability between them[25–29]. In artificial agents, a small number of appropriate architectural and hyperparameter choices might lead to informational replication and Darwinian evolution over representations[30–33]. We suggest that these choices would include a modular representation of the environment with modules that are interconnected and architecturally similar, required for high-capacity information transmission and for shared representational capacity.

The final ingredient that is needed for our proposed mechanism to work is a need for statistical inference by such systems as wholes, and a coupling between hyperparameters and performance in statistical inference tasks. When such coupling is provided by an evolutionary competition *between* animals or agents (this is a different, outer loop evolutionary competition), accurate and efficient statistical computations correlate with survival and reproduction, imposing a selection pressure. However, we do not exclude the possibility of foresighted, sequential, gradient-like learning of (architectural) hyperparameters in artificial agents as an outer loop optimizer. In both cases, outer loop optimization over collective(s) might lead to an emergence of Darwinian evolution despite the hyperparameter cost only evaluating the agent's ability to perform statistical computations. In short, we propose that Darwinian dynamics might follow from the system's need to do statistical inference effectively and accurately, as a form of *exaptation*[34,35].

While statistical inference is a ubiquitous ability of agents that move as they need to predict a changing environment to plan their immediate actions, and consequently, animal and artificial minds are a natural place where our proposed mechanism might unfold, there is another possibility. When collectives are Darwinian, but not the individuals within them (for example, because they transmit hereditary information among them at a low accuracy or not at all), this mechanism leads

to the emergence of Darwinian evolution at the level of individuals. In the language of multilevel selection, this is a scenario of *downwards* transition in Darwinian individuality, i.e., the new evolutionary level that emerges is not at the level of collectives but at the level of parts. What is needed for this downwards transition is a selection pressure acting on the Darwinian collectives to perform statistical inference over past environments, in order to predict future environmental changes, and, as a consequence, adapt to new environments faster.

The core of our proposed mechanism for the emergence of Darwinian evolution via Bayesian computations is the algorithmic equivalence between replicator-mutator dynamics and filtering in hidden Markov models, see Figure 1B and C, as well as ref.[19] for details. This equivalence extends to both Marr's computational and algorithmic level[36]. On the computational level, dynamics of relative frequencies in the discrete-time replicator-mutator equation is equivalent to the dynamics of the filtering distribution in hidden Markov models (HMM)[17,19]. On the algorithmic level, the Darwinian mechanism of multiplication, inheritance, and selection, operating on a population, maps to a sampling-based approximation of the filtering distribution of HMMs, called particle filters[19,20,37,38]. A fundamental difference between the two domains is the parameter regimes that give rise to "good" computations. In HMMs, the task is to infer the posterior (filtering) distribution accurately. Evolutionary dynamics, on the other hand, is frequently considered, as well as applied as, an optimizer in large combinatorial spaces with many local optima. We combine these two ideas in this paper to show that a continuous adaptive trajectory from no information replication between two units to Darwinian dynamics over many units is achievable if the selection pressure initially comes from statistical inference of stochastic environmental dynamics.

There are at least three distinctive contexts in which the mechanism proposed in this paper potentially applies: (i) Darwinian neurodynamics, the hypothesis that Darwinian dynamics unfolds in human or animal brains[25–29], (ii) emergence of Darwinian dynamics *within* artificial agents[30–33], and (iii) Downwards (or filial) transitions of individuality, from Darwinian groups of non-Darwinian individuals to Darwinian groups of Darwinian individuals[26,39,40]. Figure 1A illustrates the mechanism and the three proposed contexts.

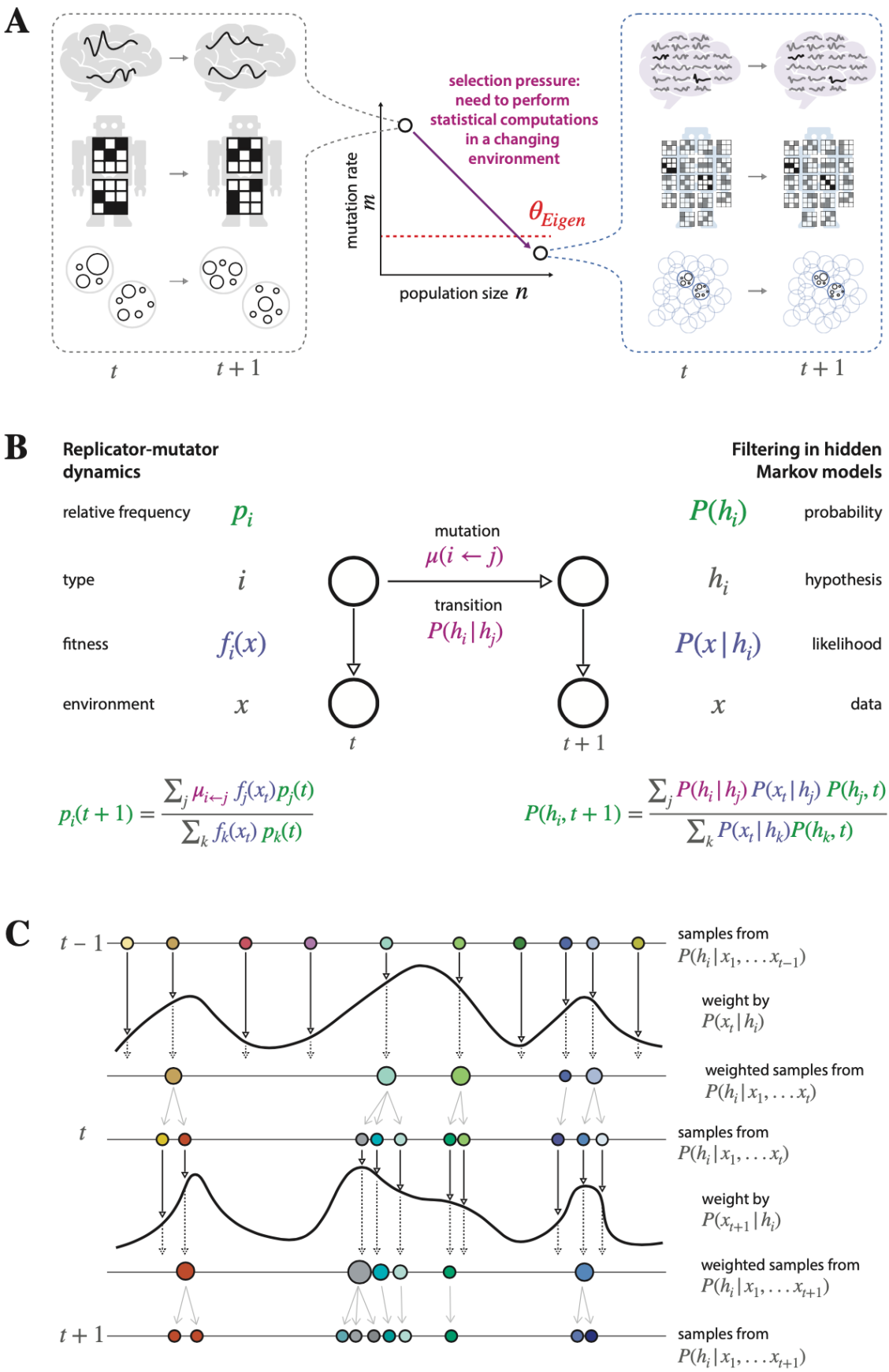
A
selection pressure: need to perform statistical computations in a changing environment
θ_{Eigen}
mutation rate m
population size n
t
$t+1$
B
Replicator-mutator dynamics
Filtering in hidden Markov models
relative frequency
p_i
$P(h_i)$
probability
mutation
$\mu(i \leftarrow j)$
type
i
h_i
hypothesis
transition
$P(h_i \mid h_j)$
fitness
$f_i(x)$
$P(x \mid h_i)$
likelihood
environment
x
x
data
t
$t+1$
$p_i(t+1) = \frac{\sum_j \mu_{i \leftarrow j} f_j(x_t) p_j(t)}{\sum_k f_k(x_t) p_k(t)}$
$P(h_i, t+1) = \frac{\sum_j P(h_i \mid h_j) P(x_t \mid h_j) P(h_j, t)}{\sum_k P(x_t \mid h_k) P(h_k, t)}$
C
$t-1$
samples from $P(h_i \mid x_1, \ldots x_{t-1})$
weight by $P(x_t \mid h_i)$
weighted samples from $P(h_i \mid x_1, \ldots x_t)$
t
samples from $P(h_i \mid x_1, \ldots x_t)$
weight by $P(x_{t+1} \mid h_i)$
weighted samples from $P(h_i \mid x_1, \ldots x_{t+1})$
$t+1$
samples from $P(h_i \mid x_1, \ldots x_{t+1})$

Figure 1. A Bayesian objective provides a possible route towards the emergence of self-replication in brains, artificial agents, and prebiotic molecular collectives. A) When collectives of representational units (neural ensembles, policy representations, or prebiotic molecular systems) are selected for estimating the Bayesian posterior of changing environments, inheritance of information between particles within collectives can emerge. Continued selection for Bayesian inference at the level of collectives leads to Darwinian evolutionary dynamics of a large population of high-fidelity replicators within collectives. Even though never selected for it, such replicator populations can perform efficient combinatorial search when they unfold over longer timescales. B) Computational level equivalence between replicator-mutator dynamics and the dynamics of the filtering distribution in Hidden Markov Models. C) Algorithmic equivalence between the Wright-Fisher model of population genetics and particle filters, an efficient sampling-based approximation of the filtering distribution in Hidden Markov Models. Together, B and C entail a partial unification of Bayesian and Darwinian dynamics, which forms the basis of the scenario depicted in A.

Results

We formulate a model that contrasts statistical inference with optimization within a single framework. Importantly, our model spans the space from no information transmission between units to almost perfect copying, and it sets the number of participating units in the collective computation as a free and potentially adaptive hyperparameter. The model intends to be as agnostic as possible regarding both specific representational choices (e.g., how and where information is stored and transmitted) and statistical inference or optimization tasks.

NK landscapes. We choose NK landscapes[41–43] to represent both statistical inference and optimization tasks with tunable complexity (ruggedness). NK landscapes form a class of combinatorial optimization problems over the space of binary strings of length *N*. Inspired by the genotype to phenotype to fitness maps in biological evolution, where multiple genes collectively determine phenotypic traits and thus fitness, called epistasis, NK landscapes assign a scalar fitness value Φ to each binary string (genotype) $\boldsymbol{g} = (g_1, g_2, \ldots, g_N)$ as $\Phi = \sum_{i=1}^{M} f_i$, such that each of the *M* fitness components f_i depends on exactly *K* arguments (genes) of $\boldsymbol{g}$. $K = 1$ encodes a linear landscape, $\Phi(g_1, g_2, \ldots, g_N) = \sum_{i=1}^{M} f_i\,(g_j)$, with a single global maximum and no other local maxima, whereas $K > 1$ describes rugged, nonlinear surfaces with multiple local maxima. In our simulations, we set $N = 20$, $M = N$, and $K = 1$ or $K = 3$, the latter corresponding to $\Phi(g_1, g_2, \ldots, g_{20}) = \sum_{i=1}^{20} f_i\,(g_j, g_k, g_l)$, i.e., each fitness component depends on exactly $K = 3$ genes. Fitness components f_i are random variables drawn from a uniform distribution over [0,1], which makes different realizations of NK landscapes with the same parameters *N* and *K* numerically different.

In order to be able to compare NK landscapes across different *K* ruggedness parameters, we transform numerical fitness values to ordinal ones by ranking them in decreasing order, $\Phi(\boldsymbol{g}) \mapsto r(\boldsymbol{g})$, and then we transform rank to a logarithmic scale, which we call *information gain* $I(\boldsymbol{g}) = -\log_2(r(\boldsymbol{g})\,/r_{max}) \in\ [0, N]$. Information gain $I(\boldsymbol{g})$ quantifies the amount of information (binary

decisions or halving the remainder of the space) needed to reach to a state that is at least as good as $\boldsymbol{g}$.

Figure 2A and B illustrates the effect of the ruggedness parameter K on the local neighborhood of states: the number of mutational neighbors of $\boldsymbol{g}$, that differ in exactly one bit from $\boldsymbol{g}$, is plotted against information gain $I(\boldsymbol{g})$. As the right panel of Figure 2B shows, rugged landscapes ($K = 3$) are still far from being random: fitness of neighbors are highly correlated, reflected by the fact that states with high fitness ($I(\boldsymbol{g}) \gtrsim 2$) have many neighbors with even higher fitness, compared to landscapes where fitness assignments are randomly reshuffled (dashed grey line). However, ruggedness has a strong effect on local neighborhoods: increasing K decreases the mean number of better neighbors but increases the variability of it across states. This is particularly apparent at the region of high fitness. In other words, adaptive mutational paths *from* high-fitness states become less available as ruggedness increases, and they also become strongly dependent on the state itself and not only its fitness. In the rest of the paper, we use NK landscapes with $K = 1$ and $K = 3$, representing linear and rugged surfaces, respectively.

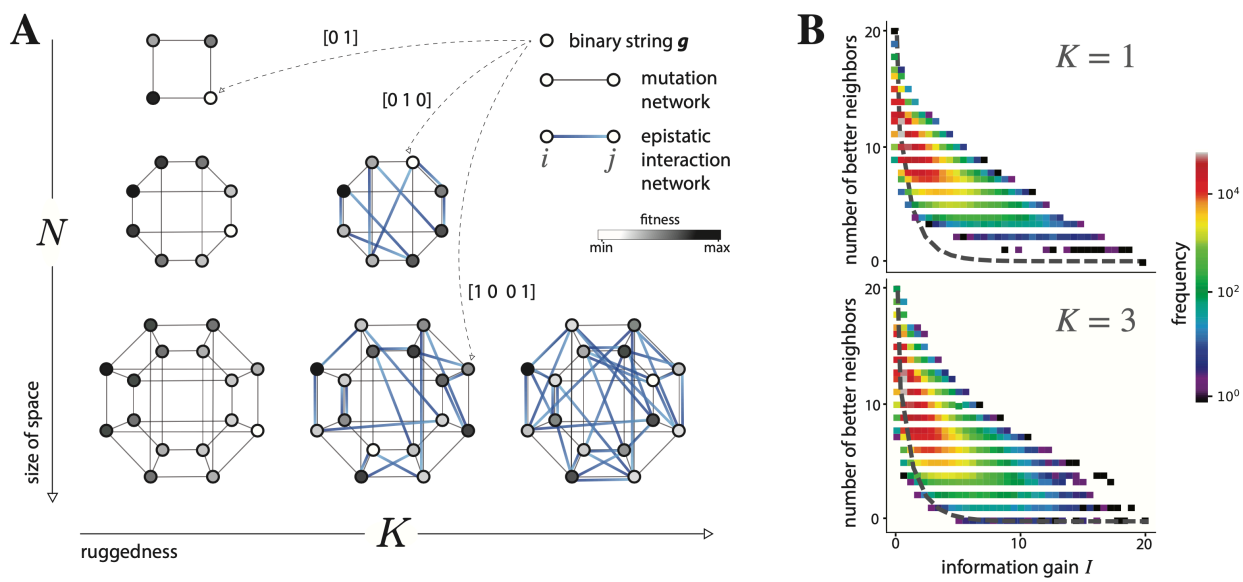

Figure 2. Tuneably rugged combinatorial surfaces model complex statistical inference and combinatorial optimization tasks. A) NK landscapes define statistical classes of scalar functions over binary strings, parametrized by the length of the string N and the ruggedness of the landscape K. B) The local statistics of the landscape as a function of ruggedness K. The number of better mutational neighbors (local adaptive paths) is strongly determined by fitness when the landscape is linear ($K = 1$), but is highly variable when the landscape is rugged ($K = 3$). Specifically, when $K = 1$, only the global maximum ($I = 20$) is a local maximum (no better neighbors), while at $K = 3$, local maxima are already present at fitness $I \approx 5$. Yet local statistics of both linear ($K = 1$) and rugged ($K = 3$) landscapes are very different from landscapes with randomly reshuffled fitness assignments, with mean number of better neighbors shown by the dashed grey line.

Statistical inference and optimization tasks. We define the classes of tasks that we will refer to as *statistical inference* and *optimization* below. In both cases, we exponentialize fitness values to make them positive, amenable to the interpretation of exponentialized fitness as expected number of offspring, while preserving the complexity of the problem by using a monotonous transformation. Note that information gain $I(\boldsymbol{g})$ and the local neighborhood statistics shown by Figure 2B are invariant under this (and any monotonous) transformation.

The *statistical inference* task is defined here as follows. Starting from a prior distribution $P_{\text{prior}}(\boldsymbol{g}) \propto \exp(\alpha\Phi_{\text{prior}}(\boldsymbol{g}))$, the dynamics unfolds on the likelihood (of the environment e) $P_{\text{likelihood}}(e|\boldsymbol{g}) \propto \exp(\tau\alpha\Phi_{\text{likelihood}}(\boldsymbol{g}))$ to compute the posterior distribution,

$$P_{\text{posterior}}(\boldsymbol{g}|e) \propto P_{\text{likelihood}}(e|\boldsymbol{g}) \times P_{\text{prior}}(\boldsymbol{g}) = \exp(\alpha\Phi_{\text{prior}}(\boldsymbol{g}) + \tau\alpha\Phi_{\text{likelihood}}(\boldsymbol{g})) \quad (1)$$

with Φ_{prior} and $\Phi_{\text{likelihood}}$ being different realizations of NK landscapes with the same parameters *N* and *K*. *Optimization*, on the other hand, is simply climbing a single landscape, which we choose to be $\exp(\tau\alpha\Phi_{\text{likelihood}}(\boldsymbol{g}))$ for maximal comparability. Intuitively, statistical inference entails an optimal integration of the past (prior) and the present (likelihood), which is well approximated by the population when evolution on the likelihood landscape is neither too fast or too slow, and neither too directed or too explorative. Furthermore, it is always computed at the transient regime due to the importance of remembering the initial condition encoded by the prior.

Another difference between the two tasks is how performance is evaluated. In case of statistical inference, the task is to compute a probability distribution $P_{\text{posterior}}$ over all 2^N genotypes. Accordingly, we quantify performance on the statistical inference task as the accuracy *A* of the population-based estimation of the true (exact) Bayesian posterior $P_{\text{posterior}}$, measured as the Kolmogorov-Smirnov (KS) distance between the population-based estimate and the exact posterior. Optimization, on the other hand, consists of a search for maximum information gain in the population, denoted by $I_{\max}$. Note that both accuracy of Bayesian inference *A* and performance in optimization $I_{\max}$ are invariant under any monotonous transformation of fitness Φ, justifying the choice of using them as simple performance measures on the two tasks.

Dynamics/algorithmic solution to the two tasks. The dynamics, known as the Wright-Fisher model in population genetics and as particle filters or sequential Monte Carlo algorithms in Bayesian statistics, is the following. At every timestep, called generation, the population of n individuals, each represented by their genotype $\boldsymbol{g}$, is resampled according to the individuals' fitness/likelihood $P_{\text{likelihood}} = \exp(\tau\alpha\Phi_{\text{likelihood}}(\boldsymbol{g}))$, and then mutated. Thus, a generation consists of *(i)* selecting an individual with genotype $\boldsymbol{g}$ for reproduction with probability proportional to its $P_{\text{likelihood}}$, *(ii)* mutating it, *(iii)* adding it to the pool of the next generation, *(iv)* repeating *(i)-(iii)* n times. In our model, mutation is simply implemented as a point mutation (flipping each bit) with probability m, called the (per bit) mutation rate. Since individuals are represented by binary strings of length *N*, the per individual mutation rate is Nm; the expected total number of mutations in the population in a single generation is nNm.

Since the dynamics is assumed to be fully internal to the animal, agent, or collective, its parameters, mutation rate m and population size n are evolvable: over a timescale that is longer than that of

the internal dynamics (for example, genetic evolutionary timescales versus millisecond-timescale statistical computations in animal brains), evolutionary parameters m and n follow an adaptive path on a cost surface given by their (average) performance in statistical inference or optimization. We do not model parameter evolution directly here, instead, we compare systems with different parameters (i.e., we evaluate the cost surface over the parameters) and assume that populations, at least approximately, follow an adaptive path on this cost surface.

Since population size n serves as sample size in estimating the Bayesian posterior, varying n introduces a bias in Bayesian accuracy A. We factor out this bias by introducing an *action selection* step when constructing population-based estimates of the posterior distribution. Action selection corresponds to selecting a *single* particle uniformly from the population of size n at the end of each run; then constructing the sample-based estimation of the posterior from n_{actions} independent runs, each contributing one sample.

Another parameter that the animal, agent, or collective, is able to optimize is the way it breaks down the task to subtasks, each corresponding to one generation of the dynamics above. This is a strong deviation from conventional evolutionary models since it assumes that there exists a collective that can make an adaptive decision of this partitioning. On one end of the scale, it imposes a fitness landscape $P_{\text{likelihood}}(e|\boldsymbol{g}) \propto \exp(\tau\Phi_{\text{likelihood}}(\boldsymbol{g}))$ and performs the computation (either statistical inference or optimization) in a single timestep (generation). The other end of the scale is $T >> 1$ generations, which necessitates a proportionally less selective fitness landscape *at each generation*, $\exp(\tau\Phi_{\text{likelihood}}(\boldsymbol{g}))/T)$, so that the full computation, corresponding to an iterative update T times, provides an unbiased estimate of the posterior,

$$\begin{aligned} P_{\text{posterior}}(\boldsymbol{g}|e) &\propto \exp(\alpha\Phi_{\text{prior}}(\boldsymbol{g})) \times [\exp(\tau\alpha\Phi_{\text{likelihood}}(\boldsymbol{g}))/T)]^T \\ &= \exp(\alpha\Phi_{\text{prior}}(\boldsymbol{g}) + \tau\alpha\Phi_{\text{likelihood}}(\boldsymbol{g})) \end{aligned}$$

(2)

In our analyses below, we use the more intuitive parameter $\tau\alpha/T = s$, called selection strength, instead of the number of generations T directly. Correspondingly, the likelihood fitness landscape, parameterized by s, is given by $\exp(s\Phi_{\text{likelihood}}(\boldsymbol{g}))$. τ, on the other hand, is defined by the task itself, interpreted as the ratio of importance of the likelihood (present) and the prior (past). Figure 3A illustrates the relationship between parameters s, α, τ, and T.

For statistical inference tasks, we set $\tau = 1$, which makes the importance of the prior, $\exp(\alpha\Phi_{\text{prior}}(\boldsymbol{g}))$, equal to that of the likelihood, $\exp(\tau\alpha\Phi_{\text{likelihood}}(\boldsymbol{g}))$, since Φ_{prior} and $\Phi_{\text{likelihood}}$ are both NK landscapes with the same parameters N and K. For optimization tasks, on the other hand, we set $\tau = 10$, approximately evaluating long-term dynamics on a single landscape $\exp(\tau\alpha\Phi_{\text{likelihood}}(\boldsymbol{g}))$. As before, we model both tasks and the corresponding dynamics in a unified framework wherever possible, which includes setting the initial population to be an unbiased sample of the prior, $\exp(\alpha\Phi_{\text{prior}}(\boldsymbol{g}))$. This is important for statistical inference as it evaluates the integration of the prior and the likelihood against an optimal one given by the true posterior, and it also guarantees that all statistics between the initial population and the likelihood landscape is fixed across tasks - i.e., there is no preadaptation for optimization.

Overall, we propose this as a minimal, proof of principle model that unifies Bayesian inference and Darwinian dynamics in a single framework to investigate adaptive hyperparameter trajectories over both Bayesian and optimization objectives. Most important, we *(i)* choose to be task agnostic with no explicit model of what the genotypes represent, using only a statistical model of the fitness or likelihood function and *(ii)* we evaluate performance in a way that is invariant over any monotonic transformation of fitness or likelihood to ensure maximal comparability across statistical classes. Figure 3B illustrates the iterative evaluation of Bayesian inference (A) and optimization ($I_{\max}$) as the population dynamics unfolds in rescaled time τ (measured in the units of selection). In particular, accuracy A might significantly decrease over time as the population composed of particles with only local information tracks an ideal posterior distribution with optimal global reallocation of probabilities. We investigate this phenomenon in detail below.

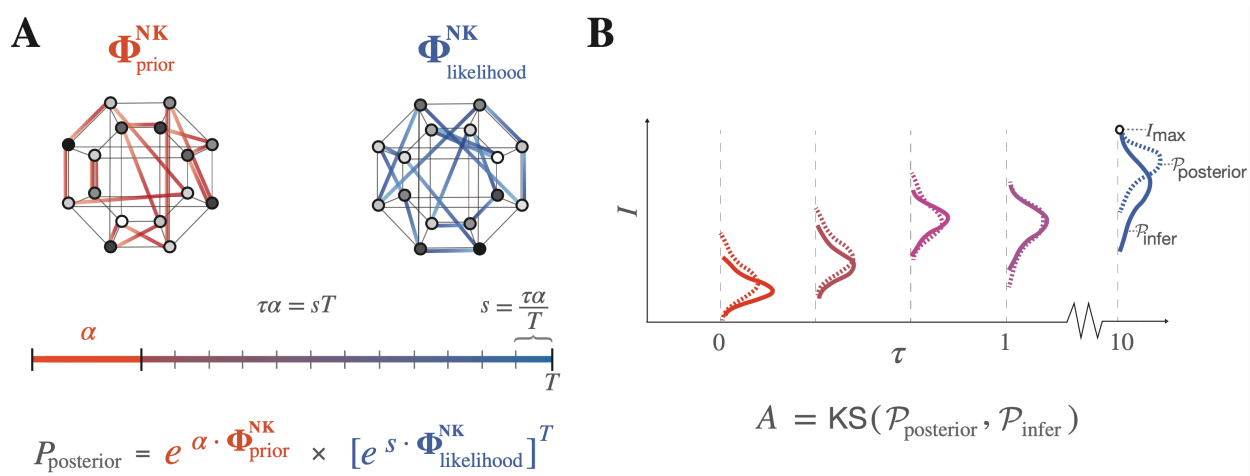

Figure 3. Bayesian inference and optimization in a single framework. A) Bayesian inference entails the integration of the prior and the likelihood, both represented by a particular realization of an NK landscape, $\Phi^{\mathrm{NK}}_{\mathrm{prior}}$ and $\Phi^{\mathrm{NK}}_{\mathrm{likelihood}}$. The selection strength is set by α and sT, respectively, with $\tau = sT/\alpha$ parametrizing the relative selection strength imposed by the likelihood versus the prior landscape. The integration of the likelihood is adaptively subdivided into T computational steps, or generations, each experiencing a selection pressure s. Thus, τ is interpreted as rescaled time, measured in the units of selection imposed by the likelihood landscape. B) We refer to the transient ($\tau = 1$) regime as Bayesian inference, and the approximately stationary ($\tau = 10$) regime as combinatorial optimization. Crucially, they also differ in how performance is evaluated: optimization performance is measured as the maximum information gain $I_{\max}$ in the population, while the accuracy A of Bayesian inference is evaluated by comparing the *distribution* of information gain values in the population P_{infer} with the exact posterior distribution $P_{\mathrm{posterior}}$, measured by the Kolmogorov-Smirnov distance between them, $A = \mathrm{KS}(P_{\mathrm{posterior}}, P_{\mathrm{infer}})$.

Approximating the Bayesian posterior. To understand when and why populations estimate the true Bayesian posterior well, we first map evolutionary parameters, mutation rate per bit m and selection strength per generation s, to population summary statistics, mean information gain μ_I and standard deviation of information gain σ_I (see Figure 4A). Then, on Figure 4B, we show the

relationship between summary statistics μ_I and σ_I, and accuracy A of the population-based estimate of the true Bayesian posterior. In all panels of Figure 4, population size is set to $n = 100$. Although the parameter space of μ_I and σ_I is narrowly explored even when evolutionary parameters m and s span multiple decades, high accuracy ($A \approx 0.2$) Bayesian inference is relatively easy to achieve at short timescales ($\tau = 1$). Long timescale computations ($\tau = 10$), however, do not yield high accuracy Bayesian estimates by the population due to its inability to explore *most* high fitness (high information gain) genotypes as they are scattered across the landscape. Timescale of environmental change is, therefore, a strong determinant of the computational capacities of Bayesian-Darwinian populations even when parameters such as mutation rate m and selection strength s span a large range from very low correlation between generations to high-fidelity replication, and from weak to strong selection.

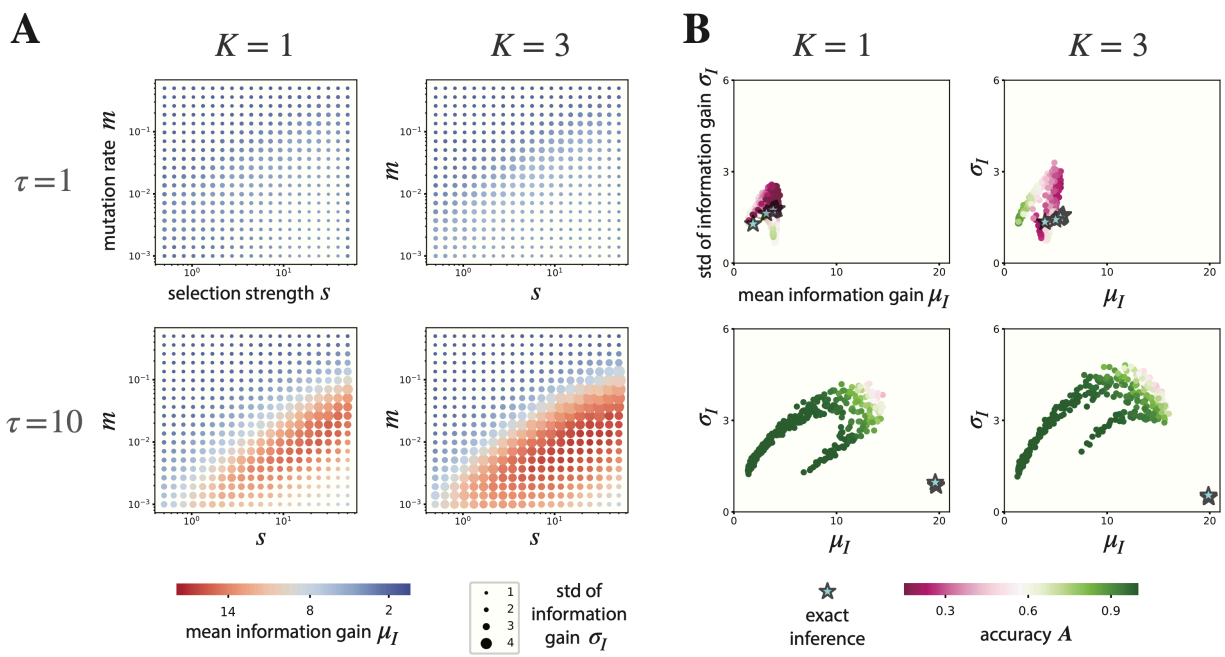

Figure 4. Darwinian dynamics on an appropriate timescale can estimate the Bayesian posterior. A) The population mean information gain μ_I (color) and its standard deviation σ_I (dot size) on the likelihood landscape, $P_{\text{likelihood}} \propto \exp(\tau\alpha\Phi_{\text{likelihood}})$, as a function of mutation rate m and selection strength s. B) μ_I and σ_I is further mapped to the accuracy A of the estimation of Bayesian posterior. High accuracy (low A) Bayesian inference can be realized by various μ_I and σ_I population measures, reflecting a near-optimal arbitration between remembering the past and fitting the present, encoded by the prior and the likelihood, respectively. Under wide parameter regimes of dynamical parameters m and s, high accuracy Bayesian inference can be achieved over short timescales ($\tau = 1$) and high mean fitness can be reached over long timescales ($\tau = 10$), but not vice versa. On both panels, population size is $n = 100$, and samples are constructed from $n_{\text{actions}} = 100$ runs.

Adaptive hyperparameter trajectories. We further investigate the direct relationship between hyperparameters mutation rate m, selection strength s, and population size n, on the accuracy A of Bayesian inference by the population. Since $m, s,$ and n are all assumed to be evolvable at the level of collectives, adaptive paths in their 3-dimensional space might correspond to possible evolutionary scenarios. We focus on two main questions. *(i)* What is the smallest population size and the highest mutation rate (lowest information transmission) where a gradient in hyperparameter space is already present, driving the system towards lower mutation rate and larger population size? (*ii)* Can a population, with hyperparameters m, s, and n, evolved for Bayesian inference, be used as an efficient optimizer in combinatorial spaces? In other words, can performance in Bayesian inference, as a computational trait of a collective, serve as a preadaptation to combinatorial optimization?

As Figure 5 shows, a hyperparameter gradient is already present at $n = 2, m \approx 0.1$ if selection is imposed by Bayesian inference at the level of collectives. This is in contrast with the observed neutral landscape at $n = 2$ under the selection of combinatorial optimization. Following a gradient on the Bayesian objective bridges serial ($n = 2$, one comparison per generation) to population-based ($n \approx 10$) computation, while it also lowers mutation rate and increases the number of generations T, to approximately $n, m, T \approx 10, 0.03, 10$. After this Bayesian regime, both the Bayesian and the optimization objective leads to a further increase in population size and decrease in mutation rate, perfecting the system in both, achieving a functional flexibility, where function is only set by the effective timescale of likelihood integration τ. This implies that efficient population-based combinatorial optimizers could evolve following an adaptive trajectory *only* on a Bayesian objective, starting from serial, high-noise, low iteration number computation, $n, m, T \approx 2, 0.1, 2$, to a population-based dynamics with low mutation rate and many generations, $n, m, T \approx 100, 0.01, 20$, which might be called a Darwinian regime. Taken together, this adaptive path leads to a "free" Darwinian combinatorial optimizer when the same dynamics is employed on longer timescales ($\tau = 10$ vs 1), a functional exaptation of Bayesian inference.

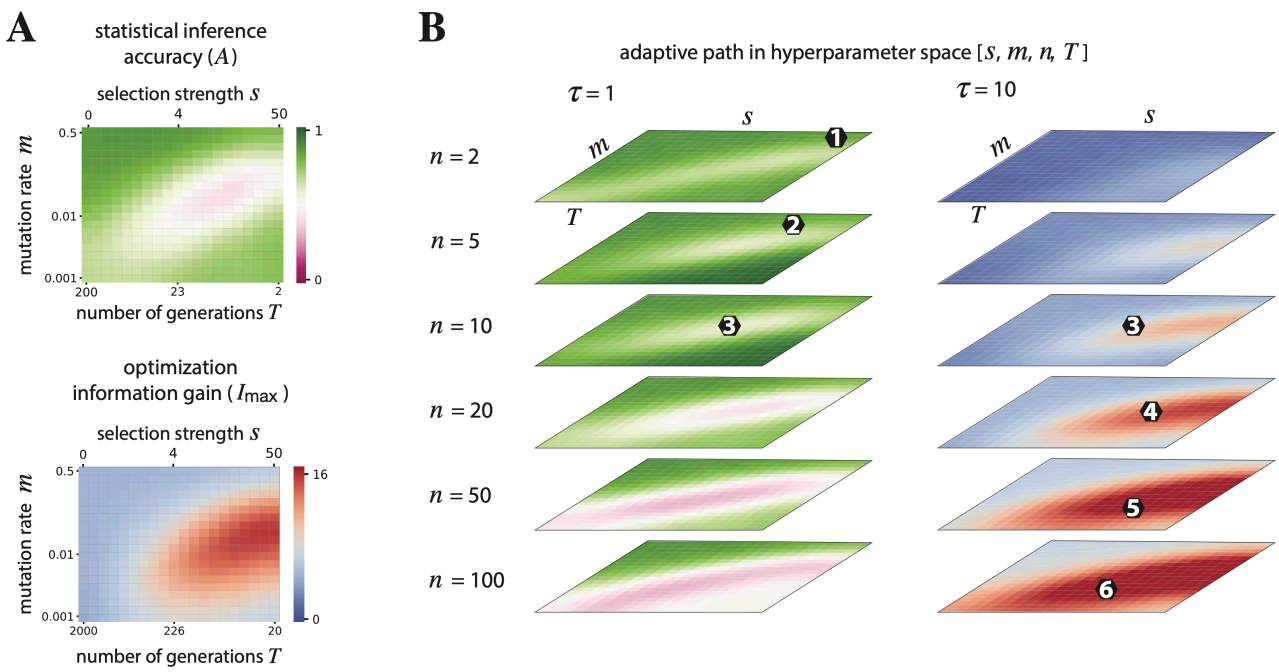

Figure 5. Darwinian exaptation of Bayesian inference. Adaptive trajectory in hyperparameter space (s, m, n, T) connects a serial ($n = 2$, one comparison) sampling-based approximation of the Bayesian posterior, with low inherited information between subsequent samples (high m), to Darwinian dynamics of a large population (large n, low m) that performs well on a combinatorial optimization task (steps *1-6*). The Bayesian objective (left column) selects for more inherited information (lower m) before the combinatorial optimization objective (right column) does, bridging the critical gap from serial ($n = 2$) to population-based ($n \approx 10$) computations. Further hyperparameter evolution on either objective leads to further decrease in mutation rate and increase in population size, a regime where the same dynamics can be flexibly employed to compute Bayesian inference or combinatorial optimization, set only by the timescale of the process τ. The results shown here correspond to a rugged landscape ($K = 3$); linear ($K = 1$) problems display similar behavior (see SI).

Discussion

Bayesian statistical inference is a ubiquitous functional adaptation in natural or artificial minds that compete to predict changing environments. Better extraction of relevant statistical features of the past and optimal integration with present noisy sensory input allows for faster and better action selection. A much less ubiquitous feature across the tree of life is combinatorial action planning. This is partly explained by the fact that its effect on survival and reproduction, especially in its impoverished, not fully evolved form, is believed to be low compared to other computations under similar energy constraints. Adaptive evolutionary trajectories, leading to minds that perform efficient search in high-dimensional combinatorial spaces must overcome this barrier.

In this paper, we propose that combinatorial action planning in human and animal minds might have been a co-adaptation of the simpler and more widespread computational function of statistical

inference. What makes such a scenario plausible is the fact that Bayesian statistical computations and various forms of Darwinian replicator dynamics are analogous; there is no need for an invention of a new algorithmic solution to transition from solving statistical inference tasks to solving combinatorial optimization (search) tasks. This can be seen as an adaptive evolutionary route to the algorithmic emergence of what is called *Darwinian neurodynamics*, the hypothesis that Darwinian dynamics over neural representations unfold in human or animal brains. This proposed mechanism complements previous work on *(i)* models of neural information copying[27,29,44–46], *(ii)* experimental findings suggesting that *in vitro* information transfer between neural ensembles can emerge without explicit training[47,48], *(iii)* proposed cognitive markers of a Darwinian search in hypothesis space[28].

Beyond neural and in silico implementations, understanding how replicator systems adapt to stochastic environments in Bayesian terms is potentially crucial to both the origin of life itself and to the origin of adaptive immune systems. We discuss these possibilities below.

Accurate replication emerged during the origin of life process on Earth. Living systems operate away from thermodynamic equilibrium and sustain and reproduce themselves by autocatalytic chemical processes. The latter operate in cycles. Such cycles have remarkable self-regulating capacities; they are reactive but not proactive. They do not infer the future based on past experience. Simple autocatalysis does not provide informational replication, for the latter at least limited heredity is necessary, a small number of states that can reliably be passed on to the offspring and be selected on. Limited heredity can be realized by a set of alternative autocatalytic cycles. Such a population might be capable of modest transgenerational Bayesian adaptation, but the full capacity must have evolved concomitant with unlimited heredity served by nucleic acid-like template polymers[23,24,49,50]. Further genetic evolution of cells enabled them to perform sequential Bayesian inference[51–56]. Cells live in a stochastic environment; they must decide, for example, whether to turn on or off the biosynthesis of an enzyme, transforming a substrate into a product. Typically, cells import the molecule in question, measure its intracellular concentration to infer extracellular concentration, and decide accordingly. These processes require autocatalytic components, sigmoid response curves[51,53,56], and memory, implemented by e.g., specific receptors for particular signals in the cell membrane.

The latter capacity has been co-opted by innate immunity, in which natural killer cells and macrophages possess the type of cellular reactivity and memory discussed above, and use these to react to invading challenges. A Bayesian approach of this capacity[57] seems to explain several interesting dynamical facts, including a trade-off between precision and adaptation. Remarkably, this system seems to learn what is self, and attacks everything else that does not look like it. Note that adaptation is happening at the cellular level, and the population level is just a direct outcome of the former. The adaptive immune system of vertebrates can do more than that: it uses a generative grammar and clonal selection to build up a vast antibody repertoire against invading antigens. This system learns about antigens directly by variation and selection, and can be characterized also as a Bayesian inference device[58]. The population, as a whole, is an adapting Darwinian entity, the Bayesian behavior of which rests on the analogy between the Bayesian update rule and replicator dynamics. An explicitly Bayesian model of the functioning of memory within the adaptive immune system shows how an optimal balance between remembering old and collecting new pathogen evidence can minimize the risks of future infections[59].

Methods

NK-landscapes. Epistasis, defined by the number of arguments of fitness components f_i entails a many-to-one mapping between genes and phenotypic traits. Pleiotropy, on the other hand, is the one-to-many nature of this map. For simplicity, we choose *NK* landscapes to have uniform pleiotropy in a statistical sense, each gene influencing exactly K fitness components by imposing f_i to depend only on g_i when $K = 1$ (no epistasis, no pleiotropy), and on $g_{i-1}, g_i, g_{i+1}$when $K = 3$. Fitness component values f_i are sampled from the uniform distribution over the unit interval $[0,1]$ independently for each of the 2^K combination of arguments. consequently, a specific realization of an NK-landscape is defined by $2^K M$ parameters. Figure 2B was constructed by sweeping through all 2^N states. The dashed lines, corresponding to randomly reshuffled fitness assignments, are analytical approximations based on the definition of information gain, calculated as the number of better neighbors $= N\, 2^{-I}$.

Sampling the prior. We generate the population representing the prior, $\exp(\alpha\Phi_{\text{prior}}(\boldsymbol{g}))$, by independently sampling each particle via a Markov chain Monte Carlo (MCMC) algorithm. A single particle corresponds to the 1000th step (stationary regime) of a Metropolis-Hastings algorithm with proposal distributions set by point mutations.

KS-distance. We compare two distributions over information gain, the exact Bayesian posterior $P_{\text{posterior}}(I)$, constructed from all 2^N values, and the population-based estimate of it, $P_{\text{infer}}(I)$, based on n_{actions} values. Accuracy A is calculated as the Kolmogorov-Smirnov (KS) distance between these two empirical distributions, given by the maximal difference of their cumulative density function across I. On Figure 4B, the stars correspond to the mean and standard deviation of information gain, μ_I and σ_I, of $P_{\text{posterior}}$, computed over all 2^N states of the landscape. By definition, they correspond to $A = 0$.

Parameters. Table 1 lists the simulation parameters that are used to generate the results of Figure 4 and 5.

	Fig. 4.	Fig. 5.
N	20	20
K	$1, 3$	3
α	100	100
τ	$1, 10$	$1, 10$
m	$0.001, \ldots, 0.5$	$0.001, \ldots, 0.5$
s	$0.5, \ldots, 50$	$0.5, \ldots, 50$
n	100	$2, 5, 10, 20, 50, 100$
n_{actions}	100	100

Table 1. Parameters used to generate the results of Figures 4 and 5.

Code availability

Python code for all simulations is available at https://github.com/csillagm/BayesToDarwin.

Author contributions

Conceptualization: HG, ES, DC. Model: MC, HG, DC. Simulations: MC. Figures: MC, HG. First draft: DC. Final manuscript: MC, HG, ES, DC.

Acknowledgements
We acknowledge financial support from the Hungarian National Research, Development, and Innovation Office (NKFIH, grant KKP129848), the Templeton World Charity Foundation (grant TWCF0268), and the John Templeton Foundation (grant 62231).

References

1. Watson, R. A. & Szathmáry, E. How Can Evolution Learn? *Trends in Ecology & Evolution* **31**, 147–157 (2016).
2. Valiant, L. *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World*. (Basic Books, 2013).
3. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A* **79**, 2554–2558 (1982).
4. Kouvaris, K., Clune, J., Kounios, L., Brede, M. & Watson, R. A. How evolution learns to generalise: Using the principles of learning theory to understand the evolution of developmental organisation. *PLoS computational biology* **13**, e1005358 (2017).
5. Szilágyi, A., Szabó, P., Santos, M. & Szathmáry, E. Phenotypes to remember: Evolutionary developmental memory capacity and robustness. *PLOS Computational Biology* **16**, e1008425 (2020).

6. Biswas, S., Manicka, S., Hoel, E. & Levin, M. Gene regulatory networks exhibit several kinds of memory: Quantification of memory in biological and random transcriptional networks. *iScience* **24**, 102131 (2021).

7. Power, D. A. *et al.* What can ecosystems learn? Expanding evolutionary ecology with learning theory. *Biol Direct* **10**, 69 (2015).

8. Knill, D. C. & Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences* **27**, 712–719 (2004).

9. Körding, K. P. & Wolpert, D. M. Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences* **10**, 319–326 (2006).

10. Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. How to Grow a Mind: Statistics, Structure, and Abstraction. *Science* **331**, 1279–1285 (2011).

11. Friston, K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci* **11**, 127–138 (2010).

12. Huang, Y. & Rao, R. P. N. Predictive coding. *WIREs Cognitive Science* **2**, 580–593 (2011).

13. Oord, A. van den, Li, Y. & Vinyals, O. Representation Learning with Contrastive Predictive Coding. Preprint at https://doi.org/10.48550/arXiv.1807.03748 (2019).

14. Harper, M. The Replicator Equation as an Inference Dynamic. Preprint at https://doi.org/10.48550/arXiv.0911.1763 (2010).

15. Shalizi, C. R. Dynamics of Bayesian updating with dependent data and misspecified models. *Electronic Journal of Statistics* **3**, 1039–1074 (2009).

16. Campbell, J. O. Universal Darwinism As a Process of Bayesian Inference. *Frontiers in Systems Neuroscience* **10**, (2016).

17. Akyıldız, Ö. D. A probabilistic interpretation of replicator-mutator dynamics. Preprint at https://doi.org/10.48550/arXiv.1712.07879 (2017).
18. Czégel, D., Zachar, I. & Szathmáry, E. Multilevel selection as Bayesian inference, major transitions in individuality as structure learning. *Royal Society open science* **6**, 190202 (2019).
19. Czégel, D., Giaffar, H., Tenenbaum, J. B. & Szathmáry, E. Bayes and Darwin: How replicator populations implement Bayesian computations. *Bioessays* **44**, 2100255 (2022).
20. Suchow, J. W., Bourgin, D. D. & Griffiths, T. L. Evolution in Mind: Evolutionary Dynamics, Cognitive Processes, and Bayesian Inference. *Trends in Cognitive Sciences* **21**, 522–530 (2017).
21. Geisler, W. S. & Diehl, R. L. Bayesian natural selection and the evolution of perceptual systems. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **357**, 419–448 (2002).
22. Lo, A. W. & Zhang, R. The evolutionary origin of Bayesian heuristics and finite memory. *iScience* **24**, 102853 (2021).
23. Maynard Smith, J. & Szathmáry, E. *The major transitions in evolution*. (Freeman Oxford, 1995).
24. Vasas, V., Fernando, C., Santos, M., Kauffman, S. & Szathmáry, E. Evolution before genes. *Biology Direct* **7**, 1 (2012).
25. Aunger, R. *The Electric Meme: A New Theory of How We Think*. (Simon and Schuster, 2002).
26. Fernando, C., Szathmáry, E. & Husbands, P. Selectionist and Evolutionary Approaches to Brain Function: A Critical Appraisal. *Frontiers in Computational Neuroscience* **6**, (2012).

27. Szilágyi, A., Zachar, I., Fedor, A., de Vladar, H. P. & Szathmáry, E. Breeding novel solutions in the brain: a model of Darwinian neurodynamics. *F1000Res* **5**, 2416 (2017).
28. Fedor, A. *et al.* Cognitive Architecture with Evolutionary Dynamics Solves Insight Problem. *Frontiers in Psychology* **8**, (2017).
29. Czégel, D., Giaffar, H., Csillag, M., Futó, B. & Szathmáry, E. Novelty and imitation within the brain: a Darwinian neurodynamic approach to combinatorial problems. *Sci Rep* **11**, 12513 (2021).
30. Moriarty, D. E., Schultz, A. C. & Grefenstette, J. J. Evolutionary Algorithms for Reinforcement Learning. *Journal of Artificial Intelligence Research* **11**, 241–276 (1999).
31. Busoniu, L., Babuska, R. & De Schutter, B. A Comprehensive Survey of Multiagent Reinforcement Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **38**, 156–172 (2008).
32. Stanley, K. O., Clune, J., Lehman, J. & Miikkulainen, R. Designing neural networks through neuroevolution. *Nat Mach Intell* **1**, 24–35 (2019).
33. Abrantes, J. P., Abrantes, A. J. & Oliehoek, F. A. Mimicking Evolution with Reinforcement Learning. Preprint at https://doi.org/10.48550/arXiv.2004.00048 (2020).
34. Gould, S. J. & Lewontin, R. C. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc R Soc Lond B Biol Sci* **205**, 581–598 (1979).
35. Gould, S. J. Exaptation: A Crucial Tool for an Evolutionary Psychology. *Journal of Social Issues* **47**, 43–65 (1991).
36. Marr, D. & Poggio, T. From Understanding Computation to Understanding Neural Circuitry. (1976).

37. Doucet, A., Godsill, S. & Andrieu, C. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing* **10**, 197–208 (2000).

38. Kwok, N. M., Fang, G. & Zhou, W. Evolutionary particle filter: re-sampling from the genetic algorithm perspective. in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* 2935–2940 (2005). doi:10.1109/IROS.2005.1545119.

39. Szathmáry, E. Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences* **112**, 10104–10111 (2015).

40. Müller, V., de Boer, R. J., Bonhoeffer, S. & Szathmáry, E. An evolutionary perspective on the systems of adaptive immunity. *Biol Rev Camb Philos Soc* **93**, 505–528 (2018).

41. Kauffman, S. A. *The Origins of Order: Self-organization and Selection in Evolution*. (Oxford University Press, 1993).

42. Kauffman, S. & Levin, S. Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology* **128**, 11–45 (1987).

43. Kaznatcheev, A. Computational Complexity as an Ultimate Constraint on Evolution. *Genetics* **212**, 245–265 (2019).

44. Fernando, C., Karishma, K. K. & Szathmáry, E. Copying and Evolution of Neuronal Topology. *PLOS ONE* **3**, e3775 (2008).

45. Fernando, C., Goldstein, R. & Szathmáry, E. The Neuronal Replicator Hypothesis. *Neural Computation* **22**, 2809–2857 (2010).

46. Fernando, C., Vasas, V., Szathmáry, E. & Husbands, P. Evolvable Neuronal Paths: A Novel Basis for Information and Search in the Brain. *PLOS ONE* **6**, e23534 (2011).

47. Roberts, T. P. *et al.* Encoding Temporal Regularities and Information Copying in Hippocampal Circuits. *Sci Rep* **9**, 19036 (2019).

48. Barral, J., Wang, X.-J. & Reyes, A. D. Propagation of temporal and rate signals in cultured multilayer networks. *Nat Commun* **10**, 3969 (2019).

49. Wächtershäuser, G. Groundworks for an evolutionary biochemistry: The iron-sulphur world. *Progress in Biophysics and Molecular Biology* **58**, 85–201 (1992).

50. Fernando, C. & Rowe, J. The origin of autonomous agents by natural selection. *Biosystems* **91**, 355–373 (2008).

51. Libby, E., Perkins, T. J. & Swain, P. S. Noisy information processing through transcriptional regulation. *Proceedings of the National Academy of Sciences* **104**, 7151–7156 (2007).

52. Naoki, H., Sakumura, Y. & Ishii, S. Stochastic control of spontaneous signal generation for gradient sensing in chemotaxis. *Journal of Theoretical Biology* **255**, 259–266 (2008).

53. Perkins, T. J. & Swain, P. S. Strategies for cellular decision-making. *Molecular Systems Biology* **5**, 326 (2009).

54. Ginsburg, S. & Jablonka, E. Epigenetic learning in non-neural organisms. *Journal of biosciences* **34**, 633–646 (2009).

55. Sivak, D. A. & Thomson, M. Environmental Statistics and Optimal Regulation. *PLOS Computational Biology* **10**, e1003826 (2014).

56. Kobayashi, T. J. Implementation of Dynamic Bayesian Decision Making by Intracellular Kinetics. *Phys. Rev. Lett.* **104**, 228104 (2010).

57. Qin, Y., Mace, E. M. & Barton, J. P. A quantitative model for the regulation of innate immune activation. 2020.04.26.061465 Preprint at https://doi.org/10.1101/2020.04.26.061465 (2020).

58. Mayer, A., Balasubramanian, V., Mora, T. & Walczak, A. M. How a well-adapted immune system is organized. *Proceedings of the National Academy of Sciences* **112**, 5950–5955 (2015).

59. Mayer, A., Balasubramanian, V., Walczak, A. M. & Mora, T. How a well-adapting immune system remembers. *Proceedings of the National Academy of Sciences* **116**, 8815–8823 (2019).