Selective Evolutionary Generation: A Model for Optimally Efficient Search in Biology

Amor A. Menezes and Pierre T. Kabamba

Abstract— This paper describes the biological principles underlying a recently proposed optimization technique, Selective Evolutionary Generation Systems (SEGS), and concludes a novel, fundamental result about the process of evolution in Nature. A systems-theoretic framework from the emerging field of self-reproducing systems is utilized in this work to illustrate the parallels between biological processes and SEGS. The SEGS technique is useful for tackling a generalization of the standard global optimization problem; the generalization depends on a parameter referred to as the level of selectivity, which restores traditional optimization when the parameter equals infinity. The SEGS technique has been shown to produce responsiveness efficiently, and to also be a generalization of both the canonical genetic algorithm with fitness proportional selection and the (1+1) evolutionary strategy. This paper explains how the SEGS technique models biological responsiveness and search, and the result is a Markov chain Monte Carlo method that has connections with statistical mechanics. The implication of the analysis is that natural evolution is an optimally efficient search process under certain technical conditions, which are often satisfied in Nature.

I. INTRODUCTION

A. Background and Motivation

THE technique of Selective Evolutionary Generation Systems (SEGS) [1] was proposed in response to the problem of efficiently designing an agent's behavior from a search space of possible actions such that the designed behavior is "good" and also responsive to changes in what constitutes good behavior. The concepts of behavior design, efficiency, goodness and responsiveness are defined precisely in the following section. Reference [1] emphasized an optimization version of this problem.

An alternative biological formulation of this problem is the efficient determination of a viable species from a search space of possible organisms such that the species is fit and adapts to variations in its fitness landscape. This problem interpretation suggests evolution; however, 'Darwinian evolution may appear inefficient' [2]. For optimization, Nature-based approaches may involve reinforcement learning [3] and simulated annealing [4], both of which are off-line and non-responsive. Techniques inspired by evolution (e.g., genetic algorithms [5], evolutionary strategies [6], and variations of these two) have the following features: 1) [7] notes that the use of 'the Darwinian principle does not guarantee successful optimization,' 2) [8] states that evolution optimization theories (on which these techniques are based) are an attempt to understand the diversity of life rather

A. A. Menezes and P. T. Kabamba are with the Department of Aerospace Engineering at the University of Michigan, Ann Arbor, MI 48109, USA amenezes@umich.edu; kabamba@umich.edu

than demonstrate that organisms optimize, 3) [9] opines that 'searching for peaks depicts evolution as a slowly advancing, tedious, uncertain process,' and 4) evolutionary computation for dynamic fitness landscapes is a relatively new area of study (for an overview, see [10]).

This paper shows that one possible rationale for evolution in Nature is to solve exactly the biological problem above. The purpose of this paper is to emphasize the biological parallels of SEGS and describe the process using a systems-theoretic abstraction, rather than fully detail the SEGS optimization technique; full proofs and comparative performance results are available in [1]. Since the canonical genetic algorithm with fitness proportional selection and the (1+1) evolutionary strategy are particular cases [1] of the SEGS technique, these approaches share some biological parallels. SEGS is linked [1] to a Markov chain Monte Carlo (MCMC) [11] method with known connections to statistical mechanics; here, it is indicated that this MCMC method is optimal with respect to an efficient search criterion. The implication and chief novel claim of this paper is that evolution is an optimally efficient search process under certain technical conditions.

B. Problem Definition

Let X be a search space, the set of genotypes [12]. The problem of *behavior design* seeks 1) a probability density function (referred to as the *behavior*) $\phi_X : X \to \mathbb{R}^+$ that accomplishes specified objectives, and 2) dynamic transition laws that cause the variable x to be distributed according to ϕ_X, i.e., to exhibit the behavior specified by ϕ_X.

Let $z : X \to Z$ be an unknown, computable, and possibly changing function that we are interested in. Here, Z is a metric space, the set of phenotypes [12]. Suppose that we are given an element z_{des} in the image of z, and we wish to find $x \in X$ such that $z(x) = z_{des}$, or such that $||z(x) - z_{des}||$ is small. Formally, we want to design a behavior ϕ_X that achieves a known expected value Y, i.e., $\mathrm{E}_{\phi_X}[||z(x) - z_{des}||] = Y$, and we refer to this expectation as *goodness*. Let $y(x) = ||z(x) - z_{des}||$.

We also desire the behavior ϕ_X to be responsive to perturbations in z, i.e., $\frac{\partial \phi_X}{\partial z} \neq 0$.

Let $f : Z \to \mathbb{R}^+$. We allow the behavior design method to employ a real-valued, positive fitness function $F : X \to \mathbb{R}^+ : x \mapsto F(x) = (f \circ z)(x) = f(z(x))$. The scheme to find ϕ_X should be *efficient* in that it trades off prior information about X for search effort savings as quickly as possible.

C. Related Literature

The SEGS technique arose out of studies of biologically-inspired self-reproducing systems, a field resulting from the work of John von Neumann [13]. A comprehensive overview of self-replication is documented in [14]–[16].

Biological responsiveness was first examined as resilience in the seminal work [17], and a recent survey of the many definitions of resilience in the literature is available in [18]. There are numerous instances of autonomous robustness as well as resilience to small and large environment fluctuations in complex natural systems. Examples include physiological regulation in multi-cellular organisms [19], [20]; group regulation in colonies of social insects [21]–[23]; the evolution of species through adaptation and natural selection [24]–[26]; and the rebounding of complex systems from earthquakes, tsunamis, hurricanes, asteroid strikes, etc. [27], [28].

Reference [29] on rational behavior, an important SEGS principle, sought to explain a remarkable property of the collectives that appeared in nature. These collectives, which had different fractions of professions (as in beehives for example), maintained an appropriate fractional distribution among the various social functions even if one of the castes was removed. Using fractional interactions, the theory examines the behavior of a collective and identifies the properties of systems of many elements. This is still an important topic; recent articles on the subject now incorporate evolution and natural selection [30], [31].

D. Paper Outline

The remainder of this paper is as follows. Section II explains the biological relationships that are embodied in the theoretical framework of SEGS. Section III describes the physical and biological processes contained in the theory that SEGS is based on. Section IV highlights the biology connections of the SEGS scheme itself. Section V highlights the physics connections of the MCMC method that is equivalent to the SEGS scheme. Section VI describes the final result on evolution that is implied using a SEGS model.

II. THE BIOLOGICAL PARALLELS IN SEGS DEFINITIONS

In behavior design, a *cell* is any element of the domain of a reward function, and a *resource* is any input that facilitates a transition between cells. Cells may also be referred to as states or candidate optimizers. A cell utilizes a resource to *reproduce* and generate an offspring, i.e., transition to another cell. Furthermore, it is possible that resources are chosen probabilistically.

Definition 1: An *evolutionary generation system* is a quadruple $\mathscr{E} = (X, R, P, G)$, where

- X is a *set of n cells*, $X = \{x_1, x_2, \ldots, x_n\}$;
- R is a *set of m resources*, $R = \{r_1, r_2, \ldots, r_m\}$, that can be utilized for cell reproduction;
- $P : R \to (0, 1]$ is a *probability mass function on R*, given by $P(r_i) = \Pr[\mathscr{R} = r_i] = p_i$, $\sum_{k=1}^{m} p_k = 1$; and
- $G : X \times R \to X$ is a *generation function* that maps a parent cell and a resource into a descendant cell.

Use of the adjective *evolutionary* here is consistent with biology [12], where evolution is defined as the genetic changes in a biological population that occur every generation due to genetic changes from parent to descendant.

Let $(r_\mu) = (r_1, r_2, \ldots, r_\mu)$ be a sequence of μ resources from R. SEGS theory defines the notation

$$G(x, (r_\mu)) := G(\ldots G(G(x, r_1), r_2) \ldots, r_\mu) \qquad (1)$$

to denote the cell produced by x using sequence (r_μ).

Definition 2: The set of cells, X, of the evolutionary generation system $\mathscr{E} = (X, R, P, G)$ is *reachable* through G and R if, for all pairs $(x_1, x_2) \in X^2$, there exists $k \in \mathbb{N}$ and a sequence $(r_k) \in R$ such that $x_2 = G(x_1, (r_k))$.

In Definition 1, the restriction that the offspring of a cell be itself a cell implies that the set of cells is *closed* [32], since there is no feasible transition to any element outside X. If the set of cells is also reachable, then X is said to be *irreducible* [32].

SEGS theory associates each cell with a non-zero, positive performance index that is a measure of the fitness of the cell, $F : X \to \mathbb{R}^+$. The notion of fitness facilitates the following novel mathematical definition of selection.

Definition 3: Given a cell set, X, and a fitness function $F : X \to \mathbb{R}^+$, let $Select : X \times X \times \mathbb{N} \to X$ be a random function such that if $x_1 \in X$ and $x_2 \in X$ are any two cells, and $N \in \mathbb{N}$ is the *level of selectivity*, then

$$Select(x_1, x_2, N) = \begin{cases} x_1 & \text{with probability } \frac{F(x_1)^N}{F(x_1)^N + F(x_2)^N}, \\ x_2 & \text{with probability } \frac{F(x_2)^N}{F(x_1)^N + F(x_2)^N}. \end{cases}$$

$$(2)$$

Definition 4: A *selective evolutionary generation system* is a quintuple
$\Gamma = (X, R, P, G, F)$, where

- (X, R, P, G) is an evolutionary generation system;
- $F : X \to \mathbb{R}^+$ is a function that evaluates cell fitness;
- the set of cells, X, is reachable through G and R; and
- the dynamics of the system are given by

$$\mathscr{X}(t+1) = Select(\mathscr{X}(t), G(\mathscr{X}(t), \mathscr{R}(t)), N). \qquad (3)$$

In (3), $\mathscr{X}(t)$ denotes the realization of a random cell variable at time t, $\mathscr{R}(t)$ denotes the realization of a random resource variable at time t, $G(\mathscr{X}(t), \mathscr{R}(t))$ denotes the offspring of the realized random cell utilizing the realized random resource at time t, and $\mathscr{X}(0)$ has a known probability mass function.

Also in (3), the probability of a cell realization at some future time given the present cell realization is conditionally independent of the past time history of cell realizations. Thus, the dynamics of a SEGS form a discrete-time homogeneous Markov chain [11]. This property is useful for the SEGS analysis conducted in [1].

The two central tenets of Darwin's theory of evolution [12] are embodied in Definition 4.

1) *Undirected variation* via the generation function.
2) *Natural selection* via the *Select* function.

The *Select* function has a number of interesting properties [1], including:

- For all N,

$$\frac{\Pr[Select(x_1, x_2, N) = x_1]}{\Pr[Select(x_1, x_2, N) = x_2]} = \left(\frac{F(x_1)}{F(x_2)}\right)^N. \quad (4)$$

That is, the ratio of the probabilities of selecting any two cells is equal to the ratio of their respective fitnesses raised to the power N. This property is called *local rationality*.

The level of selectivity, N, has a biological interpretation as well. Suppose that the fitness of a cell is measured by the total number of descendants produced over k generations, $k \geq 1$. This prolificity is typically called *future reproductive value* or *fecundity* [12]. When a colony is initiated by two self-reproducing progenitors x_1 and x_2, the ratio of the descendant population fractions after k generations equals the ratio of the respective future reproductive values,

$$\left(\frac{F(x_1)}{F(x_2)}\right). \quad (5)$$

After k generations, the ratio of the probability of choosing, by random sampling, a descendant of x_1 to the probability of choosing a descendant of x_2 is equal to the ratio of the descendant population fractions (5). Correspondingly, the ratio of the probability of selecting x_1 at the initial time to the probability of selecting x_2 at the initial time, (4), is identical to the ratio of the respective prolificities, (5), with $N = 1$.

Now consider the following sequence of operations.

1) Initiate a colony with two self-reproducing progenitors x_1 and x_2, and let descendants be produced for k generations.
2) Extract a sample from the resulting population. Use the sample to initiate a second colony, and let descendants be produced for k generations.
3) Iterate the sample and colony initiation procedure until an Nth colony is produced.

Then, the ratio of the probability of selecting a descendant of x_1 to the probability of selecting a descendant of x_2 using this multi-step process becomes

$$\left(\frac{F(x_1)}{F(x_2)}\right)\left(\frac{F(x_1)}{F(x_2)}\right)\cdots\left(\frac{F(x_1)}{F(x_2)}\right) = \left(\frac{F(x_1)}{F(x_2)}\right)^N, \quad (6)$$

and it is now clear that N represents the number of selections that are made, assuming a k-generation fecundity interpretation of fitness. A recent, well-publicized, biological experiment that fits this multi-selection model is [33].

III. The Physical and Biological Parallels Underlying SEGS Efficiency and Responsiveness

A. Efficiency

Let $(X, \mathbf{P})$ be a time-homogeneous, irreducible, ergodic Markov chain, where $X = \{x_1, x_2, \ldots, x_n\}$ is the set of states of a Markov process, $\mathbf{P} \in \mathbb{R}^{n \times n}$ is the matrix of transition probabilities for these states, and $n < \infty$ is the number of states. Assume that the initial probability distribution over the states is known, i.e., we are given an n-vector $\mathbf{p}(0)$ having elements $p_i(0) = \Pr[\mathscr{X}(0) = x_i]$ for all $x_i \in X$, where

$\mathscr{X}(0)$ denotes the state realization at time 0, and we have $\sum_{i=1}^{n} p_i(0) = 1$. Since we have assumed that the states in X are ergodic and irreducible, they admit a unique stationary probability distribution [11], [32]. Let $\pi = \begin{bmatrix} \pi_1 & \pi_2 & \ldots & \pi_n \end{bmatrix}$ be the row vector of these stationary probabilities, satisfying the constraints $\pi_i > 0 \ \forall i$, and $\sum_{i=1}^{n} \pi_i = 1$. Let $F : X \to \mathbb{R}^+$ be a positive fitness function. Let $N \in \mathbb{N}$ be a natural number.

Definition 5: The time-homogeneous, irreducible, ergodic Markov chain $(X, \mathbf{P})$ is said to *behave rationally* with respect to fitness F with level N if

$$\frac{\pi_i}{\pi_j} = \left(\frac{F(x_i)}{F(x_j)}\right)^N, \ 1 \leq i \leq n, \ 1 \leq j \leq n. \quad (7)$$

This is a definition of *global rationality*.

Each stationary probability can also be explicitly characterized to ensure Markov chain rational behavior.

Theorem 1: The time-homogeneous, irreducible, ergodic Markov chain $(X, \mathbf{P})$ behaves rationally with respect to fitness F with level N if and only if

$$\pi_i = \frac{F(x_i)^N}{\sum_{k=1}^{n} F(x_k)^N}, \ 1 \leq i \leq n. \quad (8)$$

Proof: See [1]. ∎

Here, we have a more general, probabilistic version of the optimization of an objective function. A Markov chain that behaves rationally selects the state of maximum fitness with the highest stationary probability, and, in the limit as N approaches ∞, this probability is 1. The problem and solution then revert to one of standard optimization. Remarkably, rational behavior in Markov chains is the result of a subsidiary optimization.

Theorem 2: The stationary distribution π of the time-homogeneous, irreducible, ergodic Markov chain $(X, \mathbf{P})$ that behaves rationally with respect to fitness F with level N solves the optimization problem

$$\min_{\pi_1, \ldots, \pi_n} U(\pi) = -\sum_{i=1}^{n} \varphi_i \ln(\pi_i), \quad (9)$$

subject to the constraints $\sum_{i=1}^{n} \pi_i = 1$, and $\pi_i > 0$, $\forall i$, utilizing the fitness distribution

$$\varphi_i = \frac{F(x_i)^N}{\sum_{k=1}^{n} F(x_k)^N}, \ 1 \leq i \leq n. \quad (10)$$

Proof: See [1]. ∎

Furthermore, Theorem 2 states that at the optimum, the stationary distribution agrees with the fitness distribution, i.e., $\pi = \varphi$.

Using the notion of entropy, we can interpret (9) as follows. First, we recognize the term $-\ln(\pi_i)$ as the information content of state x_i [34]. Hence, the right hand side of (9) represents the "fitness-expectation of information." Moreover, we have the following.

Corollary 1: The time-homogeneous, irreducible, ergodic Markov chain $(X, \mathbf{P})$ behaves rationally with respect to fitness F with level N if and only if its stationary probability

distribution minimizes the fitness-expectation of information. At the optimum, this fitness-expectation of information is the entropy of the fitness distribution, i.e.,

$$U^* = H(\varphi) = -\sum_{i=1}^{n} \varphi_i \ln(\varphi_i). \qquad (11)$$

Entropy maximization is important for search: 'in making inferences on the basis of partial information, the maximum entropy probability distribution subject to whatever is known is the only unbiased assignment we can make; to use any other would amount to arbitrary assumption of information which by hypothesis we do not have' [35]. The relationship between entropy maximization and optimal search is clarified in [36]. The optimal search policy for cells with exponential "sizes" 'appears very much like an irreversible process in thermodynamics, in which an initially non-equilibrium state relaxes in the the equilibrium state of maximum entropy. But now it is only our state of knowledge that relaxes to the "equilibrium" condition of maximum uncertainty' [36].

Applying the results from [36] and [35], an exponential normalized fitness function relates rational behavior, entropy and optimal search through the following theorem.

Theorem 3: Let $y : X \to \mathbb{R}$ be an unknown function for which an expected value, $\mathrm{E}[y(x)]$, is a known number Y. The normalized fitness

$$\varphi_i = \alpha e^{-\beta y(x_i)}, \ 1 \le i \le n, \qquad (12)$$

and the stationary distribution π of the time-homogeneous, irreducible, ergodic Markov chain $(X, \mathbf{P})$ that behaves rationally with respect to fitness F with level N solves the optimization problem

$$\max_{\varphi \in \mathbb{D}_n} \min_{\pi \in \mathbb{D}_n} U(\varphi, \pi) = -\sum_{i=1}^{n} \varphi_i \ln(\pi_i), \qquad (13)$$

subject to the constraint

$$\mathrm{E}[y(x)] = Y. \qquad (14)$$

Proof: See [1]. ∎

Hence, a scheme with underlying Markov chain dynamics that behave rationally also maximizes the entropy of the fitness distribution when the fitness function is exponential. The implication is that a fitness function like

$$F(x_i) = e^{-\left((z(x_i) - z_{des})^2\right)} \qquad (15)$$

together with a scheme that makes use of rational behavior (e.g., SEGS, see Section IV) guarantees "good" behaviors efficiently.

Exponential fitness functions arise in nature if one considers fecundity as the measure of fitness (recall that population growth is an exponential function). Other examples include the beak depth of the Galapagos finches [37] and instances when *directional selection* [12] is prevalent.

B. Responsiveness

Responsiveness in Markov chain rational behavior theory is defined as the sensitivity of the stationary distribution to changes in fitness.

Definition 6: For any time-homogeneous, irreducible, ergodic Markov chain $(X, \mathbf{P})$ with a positive fitness function for all the states in X, the *extrinsic resilience* of state x_i to changes in the fitness of state x_j, $j \ne i$, is defined as

$$\rho_{ij} = \frac{\partial \pi_i}{\partial F(x_j)}, \qquad (16)$$

and the *intrinsic resilience* of state x_i to changes in its own fitness is taken to be

$$\rho_{ii} = \frac{\partial \pi_i}{\partial F(x_i)}. \qquad (17)$$

We say that the Markov chain $(X, \mathbf{P})$ is *resilient* if $\rho_{ij} \ne 0$ for all i and j.

The level of selectivity has the following asymptotic effect on resilience.

Theorem 4: For the time-homogeneous, irreducible, ergodic Markov chain $(X, \mathbf{P})$ that behaves rationally with respect to fitness F with level N,

$$\rho_{ij}\Big|_{\substack{N=0 \\ j \ne i}} = \rho_{ii}\Big|_{N=0} = 0, \qquad (18)$$

and

$$\lim_{\substack{N \to \infty \\ j \ne i}} \rho_{ij} = \lim_{N \to \infty} \rho_{ii} = 0. \qquad (19)$$

Proof: See [1]. ∎

As a result of Theorem 4, we have quantification that standard optimization ($N \to \infty$) is non-resilient. Moreover, recall that if we assume a k-generation fecundity interpretation of fitness as in Section II, then $N \to \infty$ also represents an infinite number of selections made over k generations. There is much biological evidence to confirm that prolonged selective breeding yields non-resilient strains [38]–[42].

Resilience is a direct outcome of Markov chain rational behavior, as stated below.

Theorem 5: The time-homogeneous, irreducible, ergodic Markov chain $(X, \mathbf{P})$ is resilient if the chain behaves rationally.

Proof: See [1]. ∎

IV. THE BIOLOGICAL PARALLELS IN THE SEGS SCHEME

Definition 7: Let $\Gamma = (X, R, P, G, F)$ be a selective evolutionary generation system. Let $x_i \in X$ and $x_j \in X$ be any two cells, and $r_k \in R$ be a resource. The *descendancy tensor*, δ, has elements

$$\delta_{ijk} = \begin{cases} 1 & \text{if } x_j = G(x_i, r_k), \\ & \quad 1 \le i \le n, \ 1 \le j \le n, \ 1 \le k \le m, \qquad (20) \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the descendancy tensor indicates whether it is possible to produce cell x_j in one step from cell x_i, using resource r_k. We can use this tensor to create a matrix that represents the conditional probability of generating x_j given that the progenitor is x_i, by utilizing the probability of selecting each available resource and summing over all m resources.

Definition 8: For the SEGS $\Gamma = (X, R, P, G, F)$, the *matrix of generation probabilities*, γ, also called the unselective matrix of transition probabilities, has elements

$$\gamma_{ij} = \Pr[\text{offspring is } x_j \mid \text{progenitor is } x_i], \quad (21)$$

$$= \sum_{k=1}^{m} \delta_{ijk} p_k, \ 1 \leq i \leq n, \ 1 \leq j \leq n. \quad (22)$$

This matrix is a stochastic matrix.

Recall that a SEGS follows the stochastic Markov process described by (3). Therefore, we can find a matrix of transition probabilities to describe the cell-to-cell transitions that occur as a result of the selection dynamics. For the SEGS $\Gamma = (X, R, P, G, F)$, the *matrix of transition probabilities*, **P**, has elements

$$P_{ij} = \Pr[\mathscr{X}(t+1) = x_j \mid \mathscr{X}(t) = x_i], \quad (23)$$

$$= \Pr[Select(x_i, x_j, N) = x_j \mid \mathscr{X}(t) = x_i] \times$$

$$\Pr[\text{offspring is } x_j \mid \text{progenitor is } x_i] \quad (24)$$

$$= \begin{cases} \dfrac{1}{1 + \left(\dfrac{F(x_i)}{F(x_j)}\right)^N} \gamma_{ij}, & \forall j \neq i, \\[4ex] \gamma_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \dfrac{1}{1 + \left(\dfrac{F(x_j)}{F(x_i)}\right)^N} \gamma_{ij}, & \text{if } j = i. \end{cases} \quad (25)$$

Note that the matrix of transition probabilities in (25) is also a stochastic matrix.

In addition to irreducibility, if we assume that the selection dynamics of the SEGS is ergodic, then a unique stationary probability distribution over the set of cells exists.

Theorem 6: For the ergodic SEGS $\Gamma = (X, R, P, G, F)$, assume that the matrix of generation probabilities, γ, is symmetric. Then the Markov chain representing the stochastic dynamics of the ergodic SEGS behaves rationally with fitness F and level N. That is, the row vector $\pi = \begin{bmatrix} \pi_1 & \pi_2 & \dots & \pi_n \end{bmatrix}$, where π_i satisfies (8), is a left eigenvector of **P**, the matrix of transition probabilities for Γ, with corresponding eigenvalue 1 (i.e., $\pi\mathbf{P} = \pi$). Hence, π is the vector of stationary probabilities for the SEGS.

Proof: See [1]. ∎

As a result of Theorem 5, the stochastic dynamics of the ergodic SEGS with symmetric matrix of generation probabilities, γ, are resilient. Hence, a SEGS is a computationally inexpensive on-line technique to achieve these characteristics because only local decisions between two candidate optimizers are made at any time. The need to evaluate the fitness of all elements in the domain of the objective function, or even in a sub-population of candidate optimizers (as in genetic algorithms or evolutionary strategies), is avoided.

The symmetry condition on the matrix of generation probabilities, γ, implies that there exists equiprobable forward and reverse transitions between any pair of cells prior to the selection process. More specifically, symmetry of γ is a requirement that mutations be reversible. This reversibility requirement is satisfied in biology, and such mutations are called *true back mutations* [43], [44].

Theorem 7: For the ergodic SEGS $\Gamma = (X, R, P, G, F)$, assume that the matrix of generation probabilities, γ, is symmetric. Then the Markov chain representing the stochastic

dynamics of the ergodic SEGS is time-reversible, i.e.,

$$\pi_i P_{ij} = \pi_j P_{ji}, \ \forall i, j. \quad (26)$$

Proof: See [1]. ∎

As a consequence, the Markov chain representing the stochastic dynamics of the SEGS and its time reversed form are statistically the same.

The SEGS algorithm has many biological parallels. During asexual reproduction, an additional cell is generated through a division of genetic material in a process called mitosis [12]. Each cell's genetic identity is parameterized by a sequence of symbols from a four letter alphabet, its genotype, while the physical realization of a genotype is known as its phenotype. Although reproduction occurs through operations on the genotype, it is the functionality of the phenotype that determines cell fitness [12]. Hence, it is necessary to account for the genotype-phenotype mapping when evaluating cell fitness in a SEGS.

V. THE PHYSICS PARALLELS IN THE SEGS MCMC METHOD

The SEGS algorithm is an example of a Markov chain Monte Carlo (MCMC) algorithm. MCMC algorithms are useful for simulating large random fields through sampling, and are frequently employed in statistical mechanics applications [11]. MCMC algorithms utilize an irreducible, aperiodic, time-homogeneous Markov chain such that the stationary distribution, π, is the target distribution. Since convergence to the target distribution is easier to check for reversible Markov chains, these Markov chains are the most frequent case of MCMC algorithms [11].

A generic formulation is specified by the Hastings algorithm; special cases of the Hastings algorithm include the Metropolis algorithm, which is used in simulated annealing, and Barker's algorithm. The latter algorithm is derived from Barker's sampler [45], which computes radial distribution functions for plasmas over a wide range of temperatures and densities to calculate macroscopic thermodynamic variables. In [1], it is shown that the SEGS algorithm and Barker's algorithm are the same.

However, we arrived at Barker's algorithm in a non-traditional manner, i.e., we did not assume time-reversibility and begin at Hasting's algorithm. Instead, modeling Nature, we started with a self-reproducing process and selected according to local rationality. The aim was to achieve global rational behavior, thereby resulting in resilience. A required assumption was equiprobable forward and reverse transitions prior to selection, a fact borne out in Nature. This assumption resulted in the SEGS algorithm being time-reversible. Furthermore, efficient searching suggested exponential fitness functions, which are also documented in Nature. The combination of an exponential fitness function and a SEGS algorithm is the Barker sampler.

In light of the connection between rational behavior and statistical mechanics through entropy, it is perhaps unsurprising that selective evolutionary generation results in Barker's version of an MCMC method.

VI. THE IMPLICATION

Whereas the Metropolis algorithm is optimal with respect to asymptotic variance in the class of Hastings algorithms with fixed candidate-generating matrix $\mathbf{Q}$ [11], Barker's algorithm is optimal with respect to search efficiency under the technical conditions specified in Theorem 3.

Since the SEGS technique results in Barker's algorithm, efficient search is guaranteed under these same technical conditions. SEGS is a model for evolutionary processes that achieve responsive behaviors, and the implication is that evolution is also optimally search efficient when the technical conditions hold. As documented in this paper, the technical conditions (exponential fitness functions, true back mutations, etc.) have been verified in Nature. Even intermediate results of this model (rational behavior in collectives, repeated selections corresponding to the level of selectivity, losses in resilience due to prolonged selection) are corroborated by examples in Nature.

Lastly, the fact that Nature utilizes sexual reproduction pairs does not invalidate the above SEGS model and claim. This is because half of the reproductive pairing can be viewed as a cell in the model, and the other half can be viewed as a resource; thus, the set of cells and resources in the SEGS model have a non-empty intersection. Further, the protestation that more than one resource is required for reproduction can also be included in the model without significant changes; each resource can itself be considered a set containing the requisite water, nutrients, etc. required by a cell for reproduction.

REFERENCES

[1] A. A. Menezes, "Selective evolutionary generation systems: Theory and applications," Ph.D. dissertation, University of Michigan, 2010.

[2] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2009.

[3] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[4] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[5] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley, 1989.

[6] I. Rechenberg, "Evolutionsstrategie: Optimierung technischer systeme nach prinzipien der biologischen evolution," Ph.D. dissertation, Technical University of Berlin, 1971.

[7] H.-G. Beyer, *The Theory of Evolution Strategies*. Springer, 2001.

[8] J. M. Smith, "Optimization theory in evolution," *Annual Review of Ecology and Systematics*, vol. 9, pp. 31–56, November 1978.

[9] D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, 3rd ed. Wiley, 2006.

[10] I. Dempsey, M. O'Neill, and A. Brabazon, *Foundations in Grammatical Evolution for Dynamic Environments*. Springer, 2009.

[11] P. Brémaud, *Markov Chains: Gibbs fields, Monte Carlo Simulation and Queues*. Springer, 1999.

[12] S. Freeman, *Biological Science*, 2nd ed. Pearson Prentice Hall, 2005.

[13] J. von Neumann, *Theory of Self-Reproducing Automata*, A. Burks, Ed. University of Illinois Press, 1966.

[14] M. Sipper, "Fifty years of research on self-replication: An overview," *Artificial Life*, vol. 4, no. 3, pp. 237–257, 1998.

[15] R. A. Freitas, Jr. and R. C. Merkle, *Kinematic Self-Replicating Machines*. Landes Bioscience, 2004.

[16] P. D. Owens and A. G. Ulsoy, "Self-replicating machines: Prevention of degeneracy," University of Michigan, Tech. Rep. CGR-06-02, 2006.

[17] C. S. Holling, "Resilience and stability of ecological systems," *Annual Review of Ecology and Systematics*, vol. 4, pp. 1–23, November 1973.

[18] F. S. Brand and K. Jax, "Focusing the meaning(s) of resilience: Resilience as a descriptive concept and a boundary object," *Ecology and Society*, vol. 12, no. 1, 2007.

[19] J. D. Fernstrom and R. J. Wurtman, "Brain serotonin content: Physiological regulation by plasma neutral amino acids," *Science*, vol. 178, no. 4059, pp. 414–416, 27 October 1972.

[20] G. L. Stiles, M. G. Caron, and R. J. Lefkowitz, "Beta-adrenergic receptors: Biochemical mechanisms of physiological regulation," *Physiological Reviews*, vol. 64, no. 2, pp. 661–743, April 1984.

[21] G. F. Oster and E. O. Wilson, *Caste and Ecology in the Social Insects*. Princeton University Press, 1978.

[22] G. E. Robinson, "Regulation of division of labor in insect societies," *Annual Review of Entomology*, vol. 37, pp. 637–665, 1992.

[23] D. M. Gordon, "The organization of work in social insect colonies," *Nature*, vol. 380, no. 6570, pp. 121–124, 14 March 1996.

[24] C. Darwin, *The Origin of Species By Means of Natural Selection, or the Preservation of Favored Races in the Struggle for Life*, 6th ed. John Murray, 1876.

[25] D. N. Reznick and C. K. Ghalambor, "The population ecology of contemporary adaptations: What empirical studies reveal about the conditions that promote adaptive evolution," *Genetica*, vol. 112–113, no. 1, pp. 183–198, November 2001.

[26] A. Wagner, *Robustness and Evolvability in Living Systems*. Princeton University Press, 2005.

[27] L. R. Walker and R. del Morel, *Primary Succession and Ecosystem Rehabilitation*. Cambridge University Press, 2003.

[28] T. P. Hughes, A. H. Baird, D. R. Bellwood, M. Card, S. R. Connolly, C. Folke, R. Grosberg, O. Hoegh-Guldberg, J. B. C. Jackson, J. Kleypas, J. M. Lough, P. Marshall, M. Nyström, S. R. Palumbi, J. M. Pandolfi, B. Rosen, and J. Roughgarden, "Climate change, human impacts and the resilience of coral reefs," *Science*, vol. 301, no. 5635, pp. 929–933, 15 August 2003.

[29] S. M. Meerkov, "Mathematical theory of behavior — individual and collective behavior of retardable elements," *Mathematical Biosciences*, vol. 43, no. 1–2, pp. 41–106, 1979.

[30] E. Bonabeau, G. Theraulaz, J.-L. Deneubourg, S. Aron, and S. Camazine, "Self-organization in social insects," *Trends in Ecology and Evolution*, vol. 12, no. 5, pp. 188–193, 5 May 1997.

[31] S. A. Levin, "Ecosystems and the biosphere as complex adaptive systems," *Ecosystems*, vol. 1, no. 5, pp. 431–436, September 1998.

[32] C. G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*, 2nd ed. Springer, 2008.

[33] D. Burd, "Plastic not fantastic," in *Project Reports of the Canada Wide Science Fair*, 2008.

[34] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October 1948.

[35] E. T. Jaynes, "Information theory and statistical mechanics," *The Physical Review*, vol. 106, no. 4, pp. 620–630, 15 May 1957.

[36] ——, "Entropy and search theory," in *Proceedings of the First Maximum Entropy Workshop*, June 1981.

[37] D. Schluter, "Estimating the form of natural selection on a quantitative trait," *Evolution*, vol. 42, no. 5, pp. 849–861, September 1988.

[38] J. Bergelson and C. B. Purrington, "Surveying patterns in the cost of resistance in plants," *The American Naturalist*, vol. 148, no. 3, pp. 536–558, September 1996.

[39] A. R. Kraaijveld and H. C. J. Godfray, "Trade-off between parasitoid resistance and larval competitive ability in *Drosophila Melanogaster*," *Nature*, vol. 389, no. 6648, pp. 278–280, September 1997.

[40] A. Bult and C. B. Lynch, "Breaking through artificial selection limits of an adaptive behavior in mice and the consequences for correlated responses," *Behavior Genetics*, vol. 30, no. 3, pp. 193–206, May 2000.

[41] L. Xie and P. L. Klerks, "Responses to selection for cadmium resistance in the least killifish, *Heterandria Formosa*," *Environmental Toxicology and Chemistry*, vol. 22, no. 2, pp. 313–320, February 2003.

[42] J. M. S. Burger, M. Kolss, J. Pont, and T. J. Kawecki, "Learning ability and longevity: A symmetrical evolutionary trade-off in *Drosophila*," *Evolution*, vol. 62, no. 6, pp. 1294–1304, June 2008.

[43] K. J. Ryan and C. G. Ray, Eds., *Sherris Medical Microbiology: An Introduction to Infectious Diseases*, 4th ed. McGraw-Hill, 2004.

[44] E. J. Gardner, M. J. Simmons, and D. P. Snustad, *Principles of Genetics*, 8th ed. Wiley, 1998.

[45] A. A. Barker, "Monte Carlo calculations of the radial distribution functions for a proton-electron plasma," *Australian Journal of Physics*, vol. 18, pp. 119–133, 1965.