brief version

Learning Resembles Evolution – the Markov Case

Tom Westerdale

September 3, 2023

Abstract

A learning system can be said to learn strings of actions. An evolving population can be said to evolve strings of genes. We show here that in a broad class of adaptive plans, the string frequencies change according to the same simple formula whether the strings are strings of actions in time or strings of genes in space, just as John Holland foresaw. The result is new. We have shown this by first analyzing the basic case where the Markov property holds in the learning system and then extending the analysis to cases where it doesn't. This paper analyzes the basic case, where an action is simply a state transition. In these adaptive plans, a transition's frequency increase is proportional to the product of its frequency and its total value. Its total value is the sum of its pre-value, which reflects payoff received in time steps before the transition, and its post-value, which reflects payoff received in time steps after the transition. Total value is time symmetric. Machine learning procedures are correctly based on post-values alone. In this paper we reframe that learning in terms of total values. Then in these plans, the formulae simplify and come to resemble those of population genetics and evolutionary computation. The procedures have not changed, only the analysis. It establishes a formal link between the process of learning and the process of evolution.

1 Learning and Evolution

A brain learns. A population evolves. In his 1975 book, *Adaptation in Natural and Artificial Systems* [2], John Holland placed these adaptive systems and others in a common formal framework. He believed that learning and evolution are similar adaptive processes.

But a formal approach highlighted problems. Reinforcement Learning [3] deals with conditional probabilities of actions, whereas evolution deals with unconditional probabilities of genes and gene strings. These different probabilities seeemed to change and adapt differently. Learning and evolution seemed different. Reinforcement Learning and Evolutionary Computation separated.

Our work shows that if we take a time symmetric aproach in the formal framework, then the probability changes are not really different. Learning and evolution are similar. Reinforcement Learning and Evolutionary Computation can now get back in touch. Holland was right.

Through interaction with the environment, a learning system learns what to do, which actions to take when and in what combinations. An evolving population, through interaction with the environment, learns which genes are appropriate when and in what combinations. A learning system can be said to learn strings of actions. An evolving population can be said to evolve strings of genes. In a broad class of adaptive plans, which we call Natural Plans, the string frequencies change according to the same simple formula whether the strings are strings of actions in time or strings of genes in space. We have shown this by first analyzing the basic case where the Markov property holds in the learning system and then extending the analysis to cases where it doesn't. This paper analyzes the basic case, where the Markov property holds and the system can distinguish the various states. So an action is simply a Markov chain state transition.

In an evolving population of a haploid microorganism, the value v_σ of a gene σ is the average of the values of the individuals in the population that have that gene. The frequency φ_σ of the gene changes in this way.

$$\varphi'_\sigma \;=\; \varphi_\sigma K v_\sigma \tag{1}$$

In this paper, K is the adaptation rate constant, and $'$ is derivative with respet to time.

The equation holds roughly not only when σ is a single gene, but also when it is a string of several genes, and even when it is an entire genome.[1] This is the basis of the hierarchical adaptive structure of evolving populations. It is the basis of Evolutionary Computation. The frequencies of the more valuable strings increase.

[1] Sexual recombination distorts this for the longer strings, and there are many complications. What I call value is *sometimes* called fitness, but fitness sometimes means reproductive rate.

This paper shows that equation (1) also holds in the basic reinforcement learning context, where σ is a string of successive state transitions. This result is new, and depends on the correct definition of string total value v_σ . The *total value* v_{ij} of state transition $i \to j$ is the sum of its pre-value, which reflects payoff received in time steps before the transition, and its post-value, which reflects payoff received in time steps after the transition. Total value is time symmetric.[2]

Machine Learning systems tend to ignore pre-values. They adjust the state transition conditional probabilities P_{ij} on the basis of the post-values c_j alone. The broad class of Natural Adaptive Plans do this. This makes good sense, since two alternative transitions coming from the same state have the same pre-values. Post-values are usually called simply values.

This paper shows that in any Natural Plan, the change in unconditional transition frequencies F_{ij} follows a very simple formula $F'_{ij} = F_{ij}Kv_{ij}$ based on total value v_{ij} . The system is not using pre-values. But using pre-values and total values in analysis allows us to describe the system's action simply. The changes in transition string frequencies follow the simple equation (1), just as in evolving populations.

The Actor-Critic architecture in Reinforcement Learning conceptually separates the system into the critic, which calculates post-value estimates, and the actor, which uses these estimates to adjust the probabilities.[3] Generally speaking, the estimates are noisy and biased.[4]

This paper deals with only the basic situation. We have an adaptive finite state Markov chain, with a fixed payoff number attached to each state. When the chain enters a state, the system receives payoff equal to the payoff number attached to that state. I write $\bar{m}$ for the average payoff received per time step.

Each legal transition has an attached positive *availability* number s_{ij} . The transition probabilities P_{ij} are the normalized availabilities. The actor adjusts the availabilities in an attempt to increase $\bar{m}$. The actor and critic can distinguish the various states.[5] In this simple situation, the post-value estimates are always unbiased. This paper deals only with the actor, and we assume that when we are in state i , the actor receives the correct post-value from the critic. This paper derives formulae for F'_{ij} and φ'_σ in this basic adaptive Markov chain situation.

We have extended these results to more general systems in which the Markov property does not hold. In such systems a Temporal Difference Critic usually gives biased post-value estimates. Our analysis of these systems has already helped to make sense of the biases. The adaptive Markov chain results form the basis of all these extensions. This paper reports only the adaptive Markov chain results.

Section 2 is preparatory material. In it we introduce our terminology and define the probabilities, the values, and $\bar{m}$ as functions of the availabilities. We note that in fact they are all differentiable (in fact rational) functions of the availabilities. We then assume that the availabilities are differentiable functions of time, and this makes all the quantities differentiable functions of time. We exhibit seven needed equations that relate the time derivatives of the various quantities. Proof outlines are given for proofs that are not fairly obvious. Almost everything in section 2 is traditional time-asymmetric material. It is given here mainly to illustrate our notation. Only at the end of section 2 do we introduce the non-traditional time-symmetric quantities, the pre-values, total values, and string total values.

Section 3 proves the key equivalence (lemma 6). It says that equation $F'_{ij} = F_{ij}Kv_{ij}$ holds if and only if $P'_{ij} = P_{ij}Kh_{ij}$ holds. The forward implication (lemma 3) is easy, but turning the implication around requires constructing what I call the derivative relation. The construction is in subsection 3.1.

From the key equivalence, the rest follows. Section 4 defines Natural Adaptive Plan and proves Theorem 1. Actually, you could think of the whole paper as a proof of Theorem 1. Theorem 1 gives three equivalent definitions of Natural Plan. It follows that string frequencies change as in the evolution formula (1).

2 Adaptive Markov Chains

All vectors in this paper are row vectors. Their transposes are column vectors. Each of our vectors will be a row of N complex numbers. A vector is non-negative if all its entries are non-negative real numbers. A vector is called *stochastic* if its entries are all non-negative real numbers and the entries sum to 1. Of course a vector is nonzero if it has a nonzero entry.

In most cases we will use the following notation. A vector will be a bold lower case letter, for example, $\mathbf{v}$. Each entry in the vector will be written using the corresponding italic letter thus, v_i . A matrix will

[2]The pre-value and post-value are both undiscounted. In Reinforcement Learning literature, the word "reward" means payoff, whereas in Evolutionary Computation literature, "reward" usually means reinforcement.

[3]The estimates can be implicit, but Temporal Difference critics use explicit estimates.

[4]Temporal Difference estimates have less noise, but they are inherently biased.

[5]They have separate parameters for each state and they know which state they are in.

be an upper case italic letter, for example, M . Each of its entries will be written using that same letter thus, M_{ij} .

The vector $\mathbf{e}$ is the vector of all ones. The vector $\mathbf{e}_i$ is the vector whose i'th entry is 1 and whose other entries are all 0.

Our matrices will be $N \times N$ matrices of real numbers. A matrix is *non-negative* if all of its entries are non-negative real numbers. A matrix is row stochastic if each of its rows is a stochastic vector. So we see that a matrix A is row stochastic of and only if A is non-negative and $A\mathbf{e}^\top = \mathbf{e}^\top$. Of course a row is nonzero if it has a nonzero entry. Similarly for nonzero column.

To *normalize* a non-negative vector means to change its entries, dividing each entry by the sum of all the entries. So the normalized vector is stochastic. To *row normalize* a non-negative matrix means to normalize each of its rows, to divide each entry by the corresponding row sum.

If $\mathbf{v}$ is a vector, then $\operatorname{diag}(\mathbf{v})$ is the diagonal matrix whose ii'th entry is v_i , for all i .

We will be discussing a Markov chain that has N states. When I say Markov chain, I will mean a Markov chain with N states. $(N \geq 2)$

Attached to each legal chain transition $i \to j$ is a positive real number s_{ij} that I call its *availability.* If there is no legal transition from i to j then $s_{ij} = 0$. The matrix S is the $N \times N$ matrix of availabilities s_{ij} . Adaptation will change the availabilities.[6]

The matrix P is the row normalized S . The number $\bar{s}_i$ is the i'th row sum of S . So $P_{ij} = s_{ij}\bar{s}_i^{-1}$.

The matrix P is the chain's transition probability matrix. If $i \to j$ is a legal chain transition, then P_{ij} is the transition probability. By that I mean that P_{ij} is the conditional probability that the next state is j , given that the current state is i . If $i \to j$ is illegal then of course $P_{ij} = 0$. The transition probability of every legal transition is positive. We see that P is row stochastic.

Our Markov chain will be *strongly connected*, which means that for any ordered pair of states $\langle i, j \rangle$, there is a sequence of legal transitions that takes the chain from state i to state j .

Attached to each state i is a fixed unchanging real number m_i called its *payoff.* As S changes, P changes, but the payoff vector $\mathbf{m}$ remains the same. We see that P is a rational function of S .

We now define the traditional time assymetric quantities in terms of the availabilities. I use the term *excess payoff* to mean payoff minus $\bar{m}$. (I write $\sum\!\!\!\!\!\text{C}$ for Cesaro summation.)

Symbol	Meaning	Defining Equation
$\bar{s}_i$	availability row sum	$\bar{s}_i = \sum_j s_{ij}$ or $\bar{\mathbf{s}} = \mathbf{e}S^\top$
$\bar{S}$		$\bar{S} = \operatorname{diag}(\bar{\mathbf{s}})$
P_{ij}	probability of $i \to j$	$P_{ij} = s_{ij}\bar{s}_i^{-1}$ or $P = \bar{S}^{-1}S$
$\tilde{p}_i$	probability of state i	
D		$D = \operatorname{diag}(\tilde{\mathbf{p}})$
F_{ij}	frequency of $i \to j$	$F_{ij} = \tilde{p}_i P_{ij}$ or $F = DP$
$\hat{P}$		$\hat{P} = I - P + \mathbf{e}^\top\tilde{\mathbf{p}}$
$\bar{m}$	average payoff	$\bar{m} = \tilde{\mathbf{p}}\,\mathbf{m}^\top$
a_i	excess payoff of state i	$a_i = m_i - \bar{m}$
c_i	post-value of state i	$\mathbf{c}^\top = \hat{P}^{-1}\mathbf{a}^\top = \sum\!\!\!\!\!\text{C}_{k=0}^{\infty}(P^k\mathbf{a}^\top)$
h_{ij}	choice value of $i \to j$	$h_{ij} = c_j - c_i + a_i$
q_i		$q_i = \tilde{p}_i\bar{s}_i^{-1}$

The state probability vector $\tilde{\mathbf{p}}$ is a left eigenvector of P with eigenvalue 1. And $\tilde{\mathbf{p}}\,\mathbf{e}^\top = 1$. The Frobenius-Perron theorem [1] tells us that:

$$\left.\begin{array}{c}\text{The set of left eigenvectors of } P \text{ that have eigenvalue 1 forms a one-dimensional subspace.}\\ \text{The set of right eigenvectors of } P \text{ that have eigenvalue 1 forms a one-dimensional subspace.}\end{array}\right\} \tag{2}$$

And it tells us that every $\tilde{p}_i$ is positive. We note that $\hat{P}$ is nonsingular.[7]
So the definition $\mathbf{c}^\top = \hat{P}^{-1}\mathbf{a}^\top$ makes sense and we have the following facts.

$$\tilde{\mathbf{p}}\,\mathbf{a}^\top = 0 \qquad \text{and} \qquad \tilde{\mathbf{p}}\,\mathbf{c}^\top = 0 \tag{3}$$

[6] The matrix S is the only matrix whose entries I write with a lower case letter, s_{ij} .

[7] Suppose $\mathbf{v}\hat{P} = 0$ for some non-zero vector $\mathbf{v}$. Since $\hat{P}\,\mathbf{e}^\top = \mathbf{e}^\top$, we have $\mathbf{v}\,\mathbf{e}^\top = \mathbf{v}\hat{P}\,\mathbf{e}^\top = 0$. Therefore, $\mathbf{v}(I - P) = \mathbf{v}\hat{P} = 0$, and $\mathbf{v}$ is a left eigenvector of P with eigenvalue 1. So (2) tells us that $\mathbf{v} = \lambda\tilde{\mathbf{p}}$ for some scalar λ . Therefore, $\mathbf{v} = \lambda\tilde{\mathbf{p}} = \lambda\tilde{\mathbf{p}}\hat{P} = \mathbf{v}\hat{P} = 0$. Contradiction.

$$\tilde{\mathbf{p}} \;=\; \tilde{\mathbf{p}}P \;=\; \tilde{\mathbf{p}}\hat{P} \;=\; \tilde{\mathbf{p}}\hat{P}^{-1} \qquad \text{and} \qquad \mathbf{e}^\top \;=\; P\,\mathbf{e}^\top \;=\; \hat{P}\,\mathbf{e}^\top \;=\; \hat{P}^{-1}\mathbf{e}^\top \tag{4}$$

$$(I-P)\,\mathbf{c}^\top \;=\; \mathbf{a}^\top \qquad \text{and} \qquad c_i - a_i \;=\; \sum_j P_{ij}c_j \tag{5}$$

$$\sum_j P_{ij}h_{ij} \;=\; 0 \tag{6}$$

Since $(P^k - \mathbf{e}^\top\tilde{\mathbf{p}})\hat{P} \;=\; P^k - P^{k+1}$ we have

$$\sum_{k=0}^{n-1}(P^k - \mathbf{e}^\top\tilde{\mathbf{p}}) \;=\; (I-P^n)\,\hat{P}^{-1} \tag{7}$$

From this we have $\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}(P^k - \mathbf{e}^\top\tilde{\mathbf{p}}) \;=\; 0$, which gives us

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}P^k = \;\mathbf{e}^\top\tilde{\mathbf{p}} \tag{8}$$

The definition of Cesaro sum gives us $\sum\!\!\!\!\!\!\mathrm{C}_{n=0}^{\infty}(P^n - \mathbf{e}^\top\tilde{\mathbf{p}}) \;=\; \lim_{\ell\to\infty}\frac{1}{\ell}\sum_{n=0}^{\ell-1}\sum_{k=0}^{n-1}(P^k - \mathbf{e}^\top\tilde{\mathbf{p}})$, and then equations (7) and (8) give us this.

$$\sum\!\!\!\!\!\!\mathrm{C}_{n=0}^{\infty}(P^n - \mathbf{e}^\top\tilde{\mathbf{p}}) \;=\; \hat{P}^{-1} - \mathbf{e}^\top\tilde{\mathbf{p}} \tag{9}$$

Multiplying by $\mathbf{a}^\top$ on the right gives us $\sum\!\!\!\!\!\!\mathrm{C}_{n=0}^{\infty}(P^n\mathbf{a}^\top) \;=\; \hat{P}^{-1}\mathbf{a}^\top$, so the Cesaro sum in the definition of c_i converges and is correct.

We said that P is a rational function of S .
In fact, the following quantities are all rational functions of S .
P, $\tilde{\mathbf{p}}$, $\bar{m}$, $\mathbf{a}$, $\hat{P}$, $\hat{P}^{-1}$, F, and $\mathbf{c}$.
The only tricky part of showing this is showing that $\tilde{\mathbf{p}}$ is a rational function of P .

One way of doing this is via the *adj* of a matrix.[8] We define $M = adj(I-P)$.
By (2), the transformation $(I-P)$ is singular and has a one dimensional kernel. Hence it has a nonzero $(N-1)\times(N-1)$ minor and a non-zero cofactor. So M is not the zero matrix.

Now for any square matrix A we have $(adj(A))\,A \;=\; A\,(adj(A)) \;=\; |A|\,I$. So we have $M\,(I-P) \;=\; (I-P)\,M \;=\; 0$, and consequently we have $MP = M$ and $PM = M$. So M has the property that any nonzero row is a left eigenvector of P with eigenvalue 1, and hence[9] is a scalar multiple of $\tilde{\mathbf{p}}$. And any nonzero column is a right eigenvector of P with eigenvalue 1, an hence is a scalar multiple of $\mathbf{e}^\top$. So every row is identical, and since M is not the zero matrix, every row is nonzero and is a multiple of $\tilde{\mathbf{p}}$. So $\tilde{\mathbf{p}}$ is a rational function of P .

Note that if A is a nonsingular square matrix then A^{-1} is a rational function of A since it is $|A|^{-1}(adj(A))$.

The rest is obvious. All the quantities are rational functions of S .

We now suppose that each entry s_{ij} in S is a continuous and differentiable function of some real parameter t , which we can think of as time. It follows that all our quantities are continuous and differentiable functions of t . We write $'$ to mean derivative with respect to t . So the entries in the matrix S' are the derivatives s'_{ij} , and the entries in the vector $\tilde{\mathbf{p}}'$ are the derivatives $\tilde{p}'_i$, and so forth. The following equations show how the derivatives $S', \bar{S}', P', \tilde{\mathbf{p}}', F',$ and $\bar{m}'$ are related.

$$\bar{S}' \;=\; \mathrm{diag}(\mathbf{e}\;(S')^\top) \quad . \tag{10}$$

$$\tilde{\mathbf{p}}' = \mathbf{e}F' \quad . \tag{11}$$

$$F' = D'P + DP' \tag{12}$$

$$P' = D^{-1}(F' - D'P) \tag{13}$$

[8]If A is a square matrix, then we get the ij'th co-factor of A by deleting the i'th row and j'th column, taking the determinant of what is left, and then changing its sign if $i+j$ is odd. We get $adj(A)$ as follows. We take A , and for each element we replace it with the corresponding co-factor. Thus the ij'th element is replaced by the ij'th co-factor. When we are all done, we transpose the matrix. Suppose we look at the $(N-1)\times(N-1)$ minors of A . We see that the ij'th minor either equals the ij'th co-factor or equals its negation. So there is a non-zero $(N-1)\times(N-1)$ minor if and only if there is a non-zero co-factor.

[9]by (2)

$$P' = \bar{S}^{-1}(S' - \bar{S}'P) \quad . \tag{14}$$
$$\tilde{\mathbf{p}}' = \tilde{\mathbf{p}}P'\hat{P}^{-1} \tag{15}$$
$$\bar{m}' \;=\; \tilde{\mathbf{p}}\,\bar{S}^{-1}(S' - \bar{S}'P)\,\mathbf{c}^\top \quad . \tag{16}$$

Equations (10) and (11) are obvious. Equations(12) and (13) are just the product rule applied to $F = DP$, and equation (14) is the product rule applied to $\bar{S}P = S$. To derive equation (15) we first note that since $\tilde{\mathbf{p}}\,\mathbf{e}^\top \;=\; 1$, we have $\tilde{\mathbf{p}}'\mathbf{e}^\top \;=\; 0$. Therefore if we use the product rule on $\tilde{\mathbf{p}} = \tilde{\mathbf{p}}P$ we can obtain $\tilde{\mathbf{p}}' \;=\; \tilde{\mathbf{p}}P' + \tilde{\mathbf{p}}'P \;=\; \tilde{\mathbf{p}}P' + \tilde{\mathbf{p}}'P - \tilde{\mathbf{p}}'\mathbf{e}^\top\tilde{\mathbf{p}}$. Then transposing a couple of terms yields $\tilde{\mathbf{p}}'\hat{P} \;=\; \tilde{\mathbf{p}}P'$ and then (15). Now $\bar{m} \;=\; \tilde{\mathbf{p}}\,\mathbf{m}^\top$, so $\bar{m}' \;=\; \tilde{\mathbf{p}}'\,\mathbf{m}^\top \;=\; \tilde{\mathbf{p}}'(\mathbf{a}^\top + \bar{m}\,\mathbf{e}^\top) \;=\; \tilde{\mathbf{p}}'\,\mathbf{a}^\top$. So multiplying (15) on the right by $\mathbf{a}^\top$ gives us $\bar{m}' \;=\; \tilde{\mathbf{p}}\,P'\mathbf{c}^\top$. This and equation (14) give us equation (16).

We said that $\bar{m}$ is a differentiable function of the availabilities s_{ij} . We would like a formula for the partial derivative $\frac{\partial \bar{m}}{\partial s_{ij}}$. We proceed as follows.

We choose any legal transition $x \to y$. In the availability matrix we keep all entries s_{ij} constant except for the xy'th entry. The xy'th entry will be equal to the time parameter t . So t determines the xy'th entry, which determines $\bar{m}$. The current time t is s_{xy} , so the current availability matrix is S . The current time derivative $\bar{m}'$ is $\frac{\partial \bar{m}}{\partial s_{xy}}$. The time derivative S' is $\mathbf{e}_x^\top\mathbf{e}_y$, and the time derivative $\bar{S}'$ is $\mathbf{e}_x^\top\mathbf{e}_x$.

We plug those values into equation (16) and obtain a formula for $\frac{\partial \bar{m}}{\partial s_{xy}}$.
$\frac{\partial \bar{m}}{\partial s_{xy}} \;=\; (\tilde{\mathbf{p}}\,\bar{S}^{-1}\mathbf{e}_x^\top)(\mathbf{e}_y\mathbf{c}^\top \,-\, \mathbf{e}_x P\,\mathbf{c}^\top)$
We see that $\tilde{\mathbf{p}}\,\bar{S}^{-1}\mathbf{e}_x^\top \;=\; q_x$. And equation (5) tells us that $P\,\mathbf{c}^\top \;=\; \mathbf{c}^\top \,-\, \mathbf{a}^\top$.
Therefore, our formula simplifies to $\frac{\partial \bar{m}}{\partial s_{xy}} \;=\; q_x h_{xy}$.

The chosen transition could have been any legal transition.
Therefore, for any legal transition $i \to j$ we have

$$\frac{\partial \bar{m}}{\partial s_{ij}} \;=\; q_i h_{ij} \quad . \tag{17}$$

We now define the non-traditional time symmetric quantities.

We call a square matrix *row-column compatible* if every row sum equals the corresponding column sum. The matrix F is row-column compatible. The i'th row sum and the i'th column sum of F are both $\tilde{p}_i$. We note that P is the row normalized F . Let the matrix B be the row normalized $F^\top$. The entry B_{ij} is the conditional probability that the previous state was j , given that the current state is i . The matrix B is the transition probability matrix of a different chain. I call it the *backward chain* and call the original chain the forward chain. The state probabilities and state payoffs in the backward chain are the same as in the forward chain. (The backward chain is sometimes called the time reversed chain.)

Post-value in the backward chain is *pre-value* in the forward chain.
Total value is pre-value plus post-value.
So we have the following additional definitions

Symbol	Meaning	Defining Equation
B_{ij}	backward probability	$B_{ij} \;=\; F_{ji}\,\tilde{p}_i^{-1}$ or $B \;=\; D^{-1}F^\top$
$\hat{B}$		$\hat{B} \;=\; I - B + \mathbf{e}^\top\tilde{\mathbf{p}}$
b_i	pre-value of state i	$\mathbf{b}^\top \;=\; \hat{B}^{-1}\mathbf{a}^\top \;=\; \sum_{k=0}^{\infty}(B^k\mathbf{a}^\top)$
v_{ij}	total value of $i \to j$	$v_{ij} \;=\; b_i + c_j$
$\bar{v}_i$	total value of state i	$\bar{v}_i \;=\; b_i - a_i + c_i$

These quantities too are obviously rational functions of S .
It is straightforward to show the following unsurprising facts.

$$(I - B)\,\mathbf{b}^\top \;=\; \mathbf{a}^\top \qquad \text{and} \qquad b_i - a_i \;=\; \sum_j B_{ij}b_j \tag{18}$$

$$\sum_j F_{ij}v_{ij} \;=\; \tilde{p}_i\bar{v}_i \qquad \text{and} \qquad \sum_i F_{ij}v_{ij} \;=\; \tilde{p}_j\bar{v}_j \tag{19}$$

$$\sum_i \tilde{p}_i\bar{v}_i \;=\; 0 \qquad \text{and} \qquad \sum_{ij} F_{ij}v_{ij} \;=\; 0 \tag{20}$$

We can extend the notion of the total value of a transition and define what we can call the total-value of a string of transitions. We write υ_σ to mean the total-value of string σ. To avoid introducing unnecessary notation, I shall avoid giving a formal definition and proving the facts about it. I shall merely illustrate with an example and leave the proofs of the generalizations to the reader.

We choose an illustrative string σ.

$$\sigma = (\alpha \to \beta)(\beta \to i)(i \to j)(j \to \delta)(\delta \to \varepsilon) \tag{21}$$

The frequency φ_σ of string σ is this.
$\varphi_\sigma = \tilde{p}_\alpha P_{\alpha\beta} P_{\beta i} P_{ij} P_{j\delta} P_{\delta\varepsilon}$.
The *total value* of σ is this.
$\upsilon_\sigma = b_\alpha + a_\beta + a_i + a_j + a_\delta + c_\varepsilon$
The pre-value b_α and the post-value c_ε embody the interaction of σ with the environment, with the preceding and following strings.

These definitions make sense. For example, suppose we select a random five transition string, selecting each string σ with probability φ_σ. Then F_{ij} is the probability that the third transition in the selected sequence is $i \to j$. To see that this is the case, note that in general we have $\tilde{p}_i P_{ij} = F_{ij} = B_{ji}\tilde{p}_j$, so if σ is the example sequence above, then we can write the string frequency φ_σ as
$B_{\beta\alpha} B_{i\beta} F_{ij} P_{j\delta} P_{\delta\varepsilon}$.
Now summing over α, β, ε, and δ gives us F_{ij}. The transition frequency is the sum of the frequencies of the strings it is in, just as in population genetics.[10]

3 Adaptation

3.1 The Derivative Relation

Our task now is to show that the matrix P' determines F', and that conversely, F' determines P'. We do this by defining a relation from the possible P' matrices to the possible F' matrices and then showing it is one to one.

Select any availability matrix S. This matrix will remain fixed throughout section 3. We will examine various different S' matrices.

Since S is fixed throughout section 3, the following quantities are also fixed.
$\bar{S}$, P, $\tilde{\mathbf{p}}$, $\hat{P}$, $\hat{P}^{-1}$, D, F, B, c_i, b_i, h_{ij}, v_{ij}.

Although those quantities are fixed, their derivatives can differ depending on what S' is. We are particularly interested in P' and F'.

Let $\mathfrak{S}$ be the set of real $N \times N$ matrices that have a zero in every entry that corresponds to an illegal transition. It is easy to see that any S' matrix must be a member of $\mathfrak{S}$. In fact, any member of $\mathfrak{S}$ could be an S' matrix. The set $\mathfrak{S}$ is the set of possible S' matrices.

Now suppose we select an S' from the set $\mathfrak{S}$. Our selected S' gives us a matrix P' and a matrix F' via equations (10), (14), (15), and (12). Let's investigate what that P' and F' might be.

Let $\mathfrak{P}$ be the set of all real $N \times N$ matrices in which every row sum is zero and in which every entry corresponding to an illegal transition is zero. Obviously any P' must be a member of $\mathfrak{P}$. We now show that any member of $\mathfrak{P}$ might be P'.

Lemma 1
If $A \in \mathfrak{P}$ then there is an S' in $\mathfrak{S}$ such that $P' = A$.

Proof:
Let $S' = \bar{S}A$. We see that $S' \in \mathfrak{S}$.
Since $A\mathbf{e}^\top = 0$, we have $\bar{S}'\mathbf{e}^\top = S'\mathbf{e}^\top = \bar{S}A\,\mathbf{e}^\top = 0$. But $\bar{S}'$ is diagonal, so $\bar{S}' = 0$.
Plugging this and $S' = \bar{S}A$ into equation (14) gives us $P' = A$.
■

So we see that $\mathfrak{P}$ is exactly the set of all possible matrices P'.

Let's be quite clear what we have said. Remember that S is fixed. We have said that if we take any member A of $\mathfrak{P}$, that A could be P'. By that I mean that there is an S' that gives us a P' equal to A *with S having the value of our fixed S*.

[10]Furthermore, using equation (5) on our example string σ, we have $\sum_\varepsilon P_{\delta\varepsilon}\upsilon_\sigma = b_\alpha + a_\beta + a_i + a_j + c_\delta$. In this way, using equations (5) and (18), we can show that the total value of a substring is the average of the total values of the strings it is in, just as in population genetics.

Now let $\mathfrak{F}$ be the set of all real $N \times N$ matrices that are row column compatible, whose entries all sum to zero, and in which every entry corresponding to an illegal transition is zero. It is easy to see that every possible F' matrix must be a member of $\mathfrak{F}$. We now show it could be any member.

Lemma 2
If $G \in \mathfrak{F}$ *then there is an* S' *in* $\mathfrak{S}$ *such that* $F' = G$ *.*

Proof:
Given $G \in \mathfrak{F}$, let $\bar{G}$ be the diagonal matrix whose diagonal entries are the row sums of G .
Since G is row column compatible, they are also the column sums.
So we have $\bar{G}\mathbf{e}^\top = G\mathbf{e}^\top$ and $\mathbf{e}\bar{G} = \mathbf{e}G$, and we also have $(G - \bar{G}P)\,\mathbf{e}^\top = 0$.

Define $S' = \bar{S}D^{-1}(G - \bar{G}P)$. We see that $S' \in \mathfrak{S}$.
From the last two equations we have $S'\mathbf{e}^\top = 0$. But we have $\bar{S}'\mathbf{e}^\top = S'\mathbf{e}^\top$, so $\bar{S}'\mathbf{e}^\top = 0$.
Since $\bar{S}'$ is diagonal, this means that $\bar{S}' = 0$.
We now put our values for S' and $\bar{S}'$ in equation (14). It becomes
$P' = D^{-1}(G - \bar{G}P)$.

So we have this.
$\tilde{\mathbf{p}}P' = \mathbf{e}\,(G - \bar{G}P) = \mathbf{e}G - \mathbf{e}\bar{G}P = \mathbf{e}G - \mathbf{e}GP = \mathbf{e}G(I - P)$
Since $\mathbf{e}G\mathbf{e}^\top = 0$, we can write
$\tilde{\mathbf{p}}P' = \mathbf{e}G(I - P) = \mathbf{e}G(I - P) + (\mathbf{e}G\mathbf{e}^\top)\tilde{\mathbf{p}} = \mathbf{e}G\hat{P}$.
$\tilde{\mathbf{p}}P'\hat{P}^{-1} = \mathbf{e}G$
By equation (15), this becomes $\tilde{\mathbf{p}}' = \mathbf{e}G$.
Now $\mathbf{e}D' = \tilde{\mathbf{p}}'$ and $\mathbf{e}\bar{G} = \mathbf{e}G$, so we have $\mathbf{e}D' = \mathbf{e}\bar{G}$.
Since D' and $\bar{G}$ are diagonal, we have
$D' = \bar{G}$.

We insert our values of P' and D' in equation (12).
$F' = \bar{G}P + D(D^{-1}(G - \bar{G}P)) = G$
■

So we see that $\mathfrak{F}$ is exactly the set of all possible matrices F' .

Remember that we have a fixed S and we are investigating what P' and F' are for different matrices S' . We define a relation from $\mathfrak{P}$ to $\mathfrak{F}$ that I call the *derivative relation*.

> *Derivative Relation:*
> Suppose $A \in \mathfrak{P}$ and $G \in \mathfrak{F}$.
> Matrix A is related to matrix G just if there is an S' in $\mathfrak{S}$ that both gives us a P' equal to A and also gives us an F' equal to G .

In the above definition, when I said "if there is an S' in $\mathfrak{S}$ that" I meant if there is at least one such S' . If there is one there are probably lots.

Lemmas 1 and 2 tell us that the domain of the derivative relation is the whole of $\mathfrak{P}$ and that the range of the derivative relation is the whole of $\mathfrak{F}$.

Equations (15), (12), (11), and (13) tell us that whatever S' is, we have
$F' = (\,\mathrm{diag}(\tilde{\mathbf{p}}P'\hat{P}^{-1})\,)\,P + DP'$ and
$P' = D^{-1}(F' - (\,\mathrm{diag}(\mathbf{e}\,F')\,)\,P)$.
So if A is related to G by the derivative relation, then we have
$G = (\,\mathrm{diag}(\tilde{\mathbf{p}}A\hat{P}^{-1})\,)\,P + DA$ and
$A = D^{-1}(G - (\,\mathrm{diag}(\mathbf{e}\,G)\,)\,P)$.
In other words, A determines G , and G determines A . The derivative relation is one to one.

> The derivative relation is a function from $\mathfrak{P}$ one to one and onto $\mathfrak{F}$.

If at the start of section 3 we had selected a different fixed S then we would probably have had a different derivative relation. The set $\mathfrak{P}$ would have been the same, and the set $\mathfrak{F}$ would have been the same, and the new different derivative relation would still have been a function from $\mathfrak{P}$ one to one and onto $\mathfrak{F}$.

So in a sense we can say that the fixed matrix S determines the derivative relation. We selected a particular S at the start of this section and we are sticking with it. Our derivative relation is fixed and unchanging.

3.2 The Key Equivalence

Now at last we are ready to prove the Key Equivalence, lemma 6. We begin with the easy part, the forward implication.

Whatever S' is, it gives us a P' and F' for which the following lemma holds.

Lemma 3
If for all ij we have $F'_{ij} = F_{ij} K v_{ij}$,
then for all ij we have $\tilde{p}'_i = \tilde{p}_i K \bar{v}_i$ *and* $P'_{ij} = P_{ij} K h_{ij}$.

Proof:
In this proof I will assume $K = 1$. The reader can supply a more general K .
So we assume $F'_{ij} = F_{ij} v_{ij}$ and prove both $\tilde{p}'_i = \tilde{p}_i \bar{v}_i$ and $P'_{ij} = P_{ij} h_{ij}$.

We know that $\sum_j F_{ij} = \tilde{p}_i$, so taking derivatives and using the assumption and equation (19) gives us $\tilde{p}'_i = \sum_j F'_{ij} = \sum_j F_{ij} v_{ij} = \tilde{p}_i \bar{v}_i$.

$$\tilde{p}'_i = \tilde{p}_i \bar{v}_i \tag{22}$$

Now we derive an expression for P'_{ij} . Applying the product rule to $F_{ij} = \tilde{p}_i P_{ij}$ gives us $F'_{ij} = \tilde{p}'_i P_{ij} + \tilde{p}_i P'_{ij}$. From this and the assumption and (22) we obtain $\tilde{p}_i P'_{ij} = F'_{ij} - \tilde{p}'_i P_{ij} = F_{ij} v_{ij} - \tilde{p}_i \bar{v}_i P_{ij} = \tilde{p}_i P_{ij} (v_{ij} - \bar{v}_i) = \tilde{p}_i P_{ij} h_{ij}$.

$$P'_{ij} = P_{ij} h_{ij} \quad . \tag{23}$$

■

Now what we are going to do is turn the implication in Lemma 3 round the other way and say that if $P'_{ij} = P_{ij} K h_{ij}$ for all ij then $F'_{ij} = F_{ij} K v_{ij}$ for all ij . In a sense, it's obvious that the reverse implication follows directly from the fact that the derivative relation is one to one, but because this reverse implication is central we are going to be pedantic about proving it.

Lemma 4
Suppose for some positive real number K we define the matrices P^* *and* F^* *as follows;*

$P^*_{ij} = P_{ij} K h_{ij}$ *for all* ij
$F^*_{ij} = F_{ij} K v_{ij}$ *for all* ij

Then the following three things are true:

(A) $P^* \in \mathfrak{P}$
(B) $F^* \in \mathfrak{F}$
(C) P^* *is related to* F^* *by the derivative relation.*

Proof:
Assume that P^* and F^* are defined as in the lemma's supposition.
We note that in both P^* and F^* , every entry corresponding to an illegal transition is zero.
By equation (6), every row sum in P^* is zero.
Therefore, $P^* \in \mathfrak{P}$. (conclusion (A))
Equation (19) can be written $\sum_j F_{ji} v_{ji} = \tilde{p}_i \bar{v}_i = \sum_j F_{ij} v_{ij}$,
so we see that F^* is row column compatible.
Equation (20) tells us that the entries in F^* all sum to zero.
Therefore, $F^* \in \mathfrak{F}$. (conclusion (B))

Now since $F^* \in \mathfrak{F}$, lemma 2 tells us that there is some derivative matrix S' in $\mathfrak{S}$ such that $F' = F^*$.
We will use that derivative S' matrix. It gives us P' and F' .
We have $F' = F^*$, so for all ij we have $F'_{ij} = F_{ij} K v_{ij}$.
Lemma 3 then tells us that for all ij we have $P'_{ij} = P_{ij} K h_{ij}$.
That means $P' = P^*$.
So S' gives us a P' equal to P^* and also gives us an F' equal to F^* .
This means by definition that P^* is related to F^* by the derivative relation. (conclusion (C))
■

Whatever S' is, it gives us a P' and F' for which the following lemma holds.

Lemma 5
If for all ij *we have* $P'_{ij} = P_{ij}Kh_{ij}$,
then for all ij *we have* $F'_{ij} = F_{ij}Kv_{ij}$.

Proof:
Suppose we have an S' such that $P'_{ij} = P_{ij}Kh_{ij}$ for all ij .
With that same K , we define the matrix F^* such that
$F^*_{ij} = F_{ij}Kv_{ij}$ for all ij .
Our S' gives us F' , and we ask what matrix that is.
Lemma 4 holds with $P^* = P'$, so P' is related to F^* by the derivative relation.
And of course P' is also related to F' by the derivative relation.
The derivative relation is one to one, so $F' = F^*$.
It follows that
$F'_{ij} = F^*_{ij} = F_{ij}Kv_{ij}$ for all ij .
■

This is lemma 3 reversed.

In a sense lemma 5 is obvious, but in another sense it's astonishing. Let's re-write it using the definitions of h_{ij} and v_{ij} .

If $P'_{ij} = P_{ij}K(c_j - c_i + a_i)$ for all ij ,
then $F'_{ij} = F_{ij}K(b_i + c_j)$ for all ij .

The astonishing part is the pre-values b_i , which appear in the implication's consequent but not in the antecedent.

In lemma 3 the implication is the other way round, and all the proof of lemma 3 has to do is find a way of getting rid of the pre-values by having them cancel one another. But in lemma 5 the pre-values pop into existence, created out of thin air.

We can now sum up the results of this subsection.
Whatever S' is, it gives us a P' and F' for which the following lemma holds.

Lemma 6 The Key Equivalence
The equation $F'_{ij} = F_{ij}Kv_{ij}$ *holds for all* ij
if and only if
the equation $P'_{ij} = P_{ij}Kh_{ij}$ *holds for all* ij .

Proof:
The forward implication is lemma 3.
The backward implication is lemma 5.
■

Let's look at lemma 6. It holds whatever S' is. At the start of section 3 we selected an arbitrary matrix S . That S determines each F_{ij}, v_{ij}, P_{ij}, and h_{ij}, so it determines the right hand sides of the equations in the lemma. It also determines the derivative relation.

If at the start of section 3 we had selected a different S , then we would have had different F_{ij}, v_{ij}, P_{ij}, and h_{ij}. Sets $\mathfrak{P}$ and $\mathfrak{F}$ would have been the same, but the derivative relation would have been different. But lemma 6 would still hold whatever S' is.

So whatever S we select, lemma 6 holds whatever S' is. In other words, whatever S and S' are, lemma 6 holds. We can now forget about the derivative relation. Lemma 6 always holds.

4 Natural Adaptive Plans

We now define a class of adaptive plans that I call *Natural Plans*.[11]

Natural Adaptive Plan
There are numbers $\beta_1, \beta_2, \beta_3, ..., \beta_N$ such that
for all ij we have $s'_{ij} = s_{ij}K(\beta_i + c_j)$.

[11] Natural in the sense of a natural population.

Plan $s'_{ij} = s_{ij}Kc_j$ and plan $s'_{ij} = s_{ij}Kh_{ij}$ are both Natural Plans.[12]
We first show that Natural Plans tend to increase payoff.

Lemma 7
In any Natural Plan,
$\bar{m}' = K\sum_{ij} F_{ij}h_{ij}^2$.

Proof:
If we write $\sum_{ij}$ to mean sum over all legal transitions $i \to j$, then we have
$\bar{m}' = \sum_{ij} \frac{\partial \bar{m}}{\partial s_{ij}} s'_{ij}$.
Using equation (17) and the definition of Natural Plan, this becomes
$\bar{m}' = \sum_{ij} q_i h_{ij} s_{ij} K(\beta_i + c_j)$.
If there is no legal transition from i to j , then $s_{ij} = 0$,
so we see that the last equation holds for all ij .
We now use $q_i s_{ij} = \tilde{p}_i P_{ij}$ and $c_j = h_{ij} + c_i - a_i$.
$\bar{m}' = K\sum_{ij} \tilde{p}_i P_{ij} h_{ij}(h_{ij} + c_i - a_i + \beta_i)$
$\bar{m}' = K\sum_{ij} \tilde{p}_i P_{ij} h_{ij}^2 + K\sum_i \tilde{p}_i(c_i - a_i + \beta_i)\sum_j P_{ij}h_{ij}$
Equation (6) tells us that the second term on the right is zero.
■

Lemma 7 is an analog of Fisher's Fundamental Theorem of Natural Selection.
We see that $\bar{m}'$ can be zero only if h_{ij} is zero for every legal transition $i \to j$.
Equation (17) says that this can happen only if the ground is level.
So in a Natural Plan, the climb rate $\bar{m}'$ is positive unless the ground is level. The quantity $\bar{m}$ is a function of S , and we are heading uphill on that function unless the ground is level.

We now show that it is the Natural Plans for which the equations in the Key Equivalence hold.

Lemma 8
If there are numbers $\beta_1, \beta_2, \beta_3, ..., \beta_N$ *such that for all* ij *we have* $s'_{ij} = s_{ij}K(\beta_i + c_j)$,
then for all ij *we have* $P'_{ij} = P_{ij}Kh_{ij}$.

Proof:
As usual, we simplify by assuming $K = 1$.
We assume $s'_{ij} = s_{ij}(\beta_i + c_j)$ and prove $P'_{ij} = P_{ij}h_{ij}$.
Now we have $\bar{s}_i = \sum_j s_{ij}$, so using equation (5) gives us this.
$\bar{s}'_i = \sum_j s'_{ij} = \sum_j s_{ij}(\beta_i + c_j) = \bar{s}_i \sum_j P_{ij}(\beta_i + c_j) = \bar{s}_i(\beta_i + \sum_j P_{ij}c_j) = \bar{s}_i(\beta_i + c_i - a_i)$
We have $s_{ij} = \bar{s}_i P_{ij}$, so
$s'_{ij} = \bar{s}'_i P_{ij} + \bar{s}_i P'_{ij}$. Substituting for s'_{ij} and $\bar{s}'_i$ gives us
$s_{ij}(\beta_i + c_j) = \bar{s}_i(\beta_i + c_i - a_i)P_{ij} + \bar{s}_i P'_{ij}$.
Dividing by $\bar{s}_i$ gives us $P_{ij}(\beta_i + c_j) = P_{ij}(\beta_i + c_i - a_i) + P'_{ij}$, which simplifies to $P_{ij}h_{ij} = P'_{ij}$.
■

Lemma 9
If for all ij *we have* $P'_{ij} = P_{ij}Kh_{ij}$,
then there are numbers $\beta_1, \beta_2, \beta_3, ..., \beta_N$ *such that for all* ij *we have* $s'_{ij} = s_{ij}K(\beta_i + c_j)$.

Proof:
Again we assume $K = 1$. We assume $P'_{ij} = P_{ij}h_{ij}$ for all ij .
For each i , we define $\beta_i = \bar{s}'_i \bar{s}_i^{-1} - c_i + a_i$. Then $\beta_i + c_j = \bar{s}'_i \bar{s}_i^{-1} + h_{ij}$.
Since $s_{ij} = \bar{s}_i P_{ij}$, we have
$s'_{ij} = \bar{s}'_i P_{ij} + \bar{s}_i P'_{ij} = \bar{s}'_i \bar{s}_i^{-1} s_{ij} + \bar{s}_i P_{ij} h_{ij} = s_{ij}(\bar{s}'_i \bar{s}_i^{-1} + h_{ij}) = s_{ij}(\beta_i + c_j)$
■

From lemmas 8, 9, and 6, we see that we have

[12]This paper does not discuss plan implementation. There are various implementations of these Natural Plans, but it doesn't work to simply add Kc_j to the availability s_{ij} of the current transition $i \to j$. That implements the plan $s'_{ij} = F_{ij}Kc_j$, which is not a Natural Plan.

Theorem 1
The following three conditions are equivalent. Either all three conditions hold or none of them do.
(1) $F'_{ij} = F_{ij}Kv_{ij}$ *for all* ij .
(2) $P'_{ij} = P_{ij}Kh_{ij}$ *for all* ij .
(3) There are numbers $\beta_1, \beta_2, \beta_3, ..., \beta_N$ *such that* $s'_{ij} = s_{ij}K(\beta_i + c_j)$ *for all* ij .

Theorem 1 gives us three equivalent definitions of Natural Plan.

Lemma 3 tells us that in any Natural Plan we have $\tilde{p}'_i = \tilde{p}_i K\bar{v}_i$.

Now let's look at string frequencies.

Suppose σ is the example string we gave in (21). There we gave the expressions for the string's frequency φ_σ and total value v_σ . We have the following.
$\varphi_\sigma = \tilde{p}_\alpha P_{\alpha\beta}P_{\beta i}P_{ij}P_{j\delta}P_{\delta\varepsilon}$
$\log(\varphi_\sigma) = \log(\tilde{p}_\alpha) + \log(P_{\alpha\beta}) + \log(P_{\beta i}) + \log(P_{ij}) + \log(P_{j\delta}) + \log(P_{\delta\varepsilon})$

$$\frac{\varphi'_\sigma}{\varphi_\sigma} = \frac{\tilde{p}'_\alpha}{\tilde{p}_\alpha} + \frac{P'_{\alpha\beta}}{P_{\alpha\beta}} + \frac{P'_{\beta i}}{P_{\beta i}} + \frac{P'_{ij}}{P_{ij}} + \frac{P'_{j\delta}}{P_{j\delta}} + \frac{P'_{\delta\varepsilon}}{P_{\delta\varepsilon}}$$

Now suppose we are using a Natural Plan. Then the last equation becomes
$\varphi_\sigma^{-1}\varphi'_\sigma = K\bar{v}_\alpha + Kh_{\alpha\beta} + Kh_{\beta i} + Kh_{ij} + Kh_{j\delta} + Kh_{\delta\varepsilon}$.
$K^{-1}\varphi_\sigma^{-1}\varphi'_\sigma = \bar{v}_\alpha + h_{\alpha\beta} + h_{\beta i} + h_{ij} + h_{j\delta} + h_{\delta\varepsilon}$
$= (b_\alpha - a_\alpha + c_\alpha) + (c_\beta - c_\alpha + a_\alpha) + (c_i - c_\beta + a_\beta) + (c_j - c_i + a_i) + (c_\delta - c_j + a_j) + (c_\varepsilon - c_\delta + a_\delta)$
$= b_\alpha + a_\beta + a_i + a_j + a_\delta + c_\varepsilon$
$= v_\sigma$

$$\varphi'_\sigma \;=\; \varphi_\sigma K v_\sigma \tag{24}$$

It is easy to see that if we are using a Natural Plan, then Equation (24) holds in general for any string σ of any length.

The converse is also true. If equation (24) holds in general for any string σ then the plan we are using must be Natural since equation $F'_{ij} = F_{ij}Kv_{ij}$ is a special case of equation (24). So we have:

Corollary 1
An adaptive plan is Natural if and only if
$\varphi'_\sigma = \varphi_\sigma K v_\sigma$ *for every transition string* σ .

Corollary 1 amounts to a fourth definition of Natural Plan.

Equation (24) is the same as equation (1).

Strings of transitions are strings in time, and strings of genes are strings in space, but their frequencies here are changing in the same way. Learning and evolution are similar processes.

5 Time Symmetry

This paper analyses only the cases where the Markov property holds. The results form the basis of analysis of more complicated cases, but that is beyond the scope of this paper. This paper can be thought of as a proof of theorem 1 and corollary 1.

In retrospect, theorem 1 looks almost obvious, but it long eluded us. To compare changes in transition frequencies with changes in gene frequencies we needed a usable formula for F'_{ij} , the change in *unconditional* transition frequency. All our formule for F'_{ij} were terribly complicated until we realized we should incude pre-values in the analysis. Then everything simplified. The time-symmetric approach works.

This paper shows that in the basic case, where the Markov property holds, string frequencies change according to the same simple formula whether the strings are strings of actions in time or strings of genes in space.

Learning and Evolution are similar processes.

References

[1] F. R. Gantmacher. *Applications of the Theory of Matrices.* Interscience Publishers, New York, 1959.

[2] J. H. Holland. *Adaptation in Natural and Artificial Systems.* Univ. of Michigan Press, Ann Arbor, 1975.

[3] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction, Second edition.* MIT Press, Cambridge, Mass., 2018.