J. theor. Biol. (1997) **184,** 31–39

The Design of Adaptive Systems: Optimal Parameters for Variation and Selection in Learning and Development

Steven A. Frank†

Department of Ecology and Evolutionary Biology, University of California, Irvine,
CA 92697-2525 *U.S.A.*

(*Received on* 30 *May* 1996, *Accepted in revised form on* 14 *August* 1996)

Some aspects of learning and development are based on evolutionary change within the organism. In trial and error learning, variant ideas or behaviors are generated and selective filters (learning rules) choose among the population of variants. Development may, in some cases, proceed by selection within a population of variant cellular lineages. This paper analyses abstract properties of selective systems to understand the evolutionary dynamics that occur within organisms. The Price Equation and Fisher's fundamental theorem of natural selection, two of the most powerful concepts in evolutionary genetics, are applied in a general way to internal selective systems in learning and development. This analysis emphasizes generative mechanisms and selective filters as genetically controlled phenotypes of individual organisms. Generative mechanisms create the variation on which selection acts. Selective filters determine the extent to which selection within the organism optimizes organismal performance. The methods of Price and Fisher provide a general way in which to partition evolutionary change into improvements caused by selection and the tendency of high performance variants to deteriorate because of competition or environmental change. This balance between selective improvement, at a rate equal to the variance in fitness, and a matching deterioration in performance, provides general insight into the common properties of adaptive systems in genetics, learning and development. These ideas are applied to a model of honey bee foraging. This example clarifies the relation between genes and phenotypes controlled by internal selective systems.

Introduction

Organisms use systems of variation and selection to develop their phenotype and learn about their environment. For example, the vertebrate immune system includes a complex adaptive system that generates a population of random molecular shapes and affinities. Those variants that bind invaders are selected, their frequency increasing to provide protection against pathogens (Golub & Green, 1991).

Honey bee colonies use a trial and error learning method for discovery and exploitation (Seeley *et al*., 1991). The colony finds a population of food sources by using scout bees that have a random component to their searching. Good sources are selectively amplified by recruiting other bees, poor sources attract declining numbers of bees and eventually "die out".

Certain pattern forming processes of development may be a special kind of adaptive system. According to this speculative idea, form is created by generating a population of alternative, competing developmental trajectories, and selecting those trajectories that meet innately determined design criteria (Sachs, 1988). The idea is attractive because it replaces the notion of a genetic blueprint for form with simpler generative rules for alternative trajectories and selective rules to promote or retard various pathways. Pattern formation of plant stomata (Kagan *et al*., 1992), fish fins (Wagner & Misof, 1993), and wiring of nervous systems (Changeaux & Danchin, 1976; Edelman,

---

†E-mail: safrank@uci.edu

0022–5193/97/010031 + 09 $25.00/0/jt960241

1987) have been explained by developmental selection.

Adaptive systems in learning and development are phenotypes of organisms subject to the usual processes of genetical evolution. In each case the evolving traits are the generative mechanisms that create a population of alternatives and the selective filters that channel phenotypes toward innate design criteria (Frank, 1996).

The idea that certain phenotypes can be analysed as adaptive systems has been around for a long time (Spiegelman, 1945; Gordon, 1966; Lewontin, 1970; Michaelson, 1987). Interestingly, the idea has neither died out nor gained widespread favor. One barrier to increase is the lack of a theoretical foundation for the analysis of selective systems. That claim may seem a bit surprising, since population genetics theory is highly developed and one of the great conceptual achievements of the biological sciences. But population genetics is very specific in its assumptions about generative rules, selection, and the forms of inheritance and reproduction.

The great success and specificity of population genetics has made it difficult to recognize the general properties of selective systems. The analogies between selection and various phenotypic processes is frequently acknowledged, but the analogy rarely leads directly to useful analysis. This failure favors the belief that the analogy is only a loose one, best abandoned when faced with a particular problem (Crick, 1989). The only useful efforts to develop selective theory beyond genetics have been the study of cultural evolution (Cavalli-Sforza & Feldman, 1981; Boyd & Richerson, 1985) and computer models of optimal design by simulated evolution (Holland, 1975).

I make a start on developing the formal theory of selective systems. My work begins with the Price Equation (Price, 1970), the most general and abstract description of selective systems. This equation has been used successfully in a wide variety of difficult problems in evolutionary genetics (Frank, 1995). But the extension to an abstract theory of selection, first suggested by Price (1995), has not been developed previously.

My development of abstract selective theory leads immediately to insights analogous to Fisher's (1958) famous fundamental theorem of natural selection. That theorem, which has been universally misunderstood in genetics and in the growing field of computer evolution (Holland, 1975), is illuminated by formal study in a new context.

In the second part of the paper I apply this new theory to a generalized version of the honey-bee foraging problem. This application shows that the concepts and the mathematics of selective theory can be useful in understanding the development of interesting phenotypes. In addition, the two-level nature of adaptive systems as phenotypes is emphasized: selection of genetic variants controls the generative rules and selective filters of the individual phenotype; the generative rules and selective filters control learning and development.

The Price Equation

The Price Equation is an exact, complete description of evolutionary change under all conditions (Price, 1970, 1972a). The equation adds considerable insight into many evolutionary problems by partitioning change into meaningful components.

Here is the derivation. Let there be a population (set) where each element is labeled by an index i. The frequency of elements with index i is q_i, and each element with index i has some character, b_i. One can think of elements with a common index as forming a subpopulation that makes up a fraction q_i of the total population. No restrictions are placed on how elements may be grouped.

A second (descendant) population has frequencies q_i' and characters b_i'. The change in the average character value, $\bar{b}$, between the two populations is

$$\Delta\bar{b} = \sum q_i' b_i' - \sum q_i b_i. \quad (1)$$

Note that this equation applies to anything that evolves, since b may be defined in any way. For example, b_i may be the gene frequency of entities i, and thus $\bar{b}$ is the average gene frequency in the population, or b_i may be the square of a quantitative character, so that one can study the evolution of variances of traits. Applications are not limited to population genetics. For example, b_i may be the value of resources collected by bees foraging in the ith flower patch in a region.

Both the power and the difficulty of the Price Equation come from the unusual way it associates entities from two populations, which are typically called the ancestral and descendant populations. The value of q_i' is not obtained from the frequency of elements with index i in the descendant population, but from the proportion of the descendant population that is derived from the elements with index i in the parent population. If we define the fitness of element i as w_i, the contribution to the descendant population from type i in the parent population, then $q_i' = q_i w_i/\bar{w}$, where $\bar{w}$ is the mean fitness of the parent population.

The assignment of character values b_i' also uses indices of the parent population. The value of b_i' is the average character value of the descendants of index i. Specifically, for an index i in the parent population, b_i' is obtained by weighting the character value of each entity in the descendant population by the fraction of the total fitness of i that it represents (see examples in later sections). The change in character value for descendants of i is defined as $\Delta b_i = b_i' - b_i$.

Equation (1) is true with these definitions for q_i' and b_i'. We can proceed with the derivation by a few substitutions and rearrangements:

$$\Delta\bar{b} = \sum q_i(w_i/\bar{w})(b_i + \Delta b_i) - \sum q_i b_i$$

$$= \sum q_i\ (w_i/\bar{w} - 1)b_i + \sum q_i(w_i/\bar{w})\Delta b_i$$

which, using standard definitions from statistics for covariance (Cov) and expectation (E), yields the Price Equation

$$\bar{w}\Delta\bar{b} = \mathrm{Cov}(w,b) + \mathrm{E}(w\Delta b). \qquad (2)$$

The two terms may be thought of as changes due to selection and transmission, respectively. The covariance between fitness and character value gives the change in the character caused by differential reproductive success. The expectation term is a fitness weighted measure of the change in character values between ancestor and descendant.

Fisher's Fundamental Theorem of Natural Selection

The rate of increase in the average fitness of a population is equal to the genetic variance in fitness. That is the usual interpretation of Fisher's (1958) fundamental theorem of natural selection. Indeed, Fisher did make that statement, but the idea has been taken out of context and interpreted as a way to describe how natural selection improves the average quality of a population.

Fisher was actually concerned with a differnt problem. He was interested in the specific role of natural selection in the context of evolutionary change. By his definitions, natural selection inevitably increases fitness, but environmental changes act simultaneously in a way that usually reduces fitness by approximately the same amount. This must be so because, as Fisher noted, if average reproductive rate (fitness) were continually increasing or decreasing, then populations would either overrun the earth or quickly disappear.

Price (1972b) was the first to recognize the true meaning of Fisher's theorem. The instantaneous change in average fitness caused by natural selection is an increase proportional to the variance in fitness. The full evolutionary change in average fitness is the sum of the "partial" change in fitness caused by selection and a second term that is the partial change in fitness caused by changes in the environment (Ewens, 1989).

Fisher's theorem and its interpretation in evolutionary genetics have been analysed extensively in the past several years (Ewens, 1989, 1992; Frank & Slatkin, 1992; Edwards, 1994). The particular details of the theorem depend on aspects of Mendelian genetics that may or may not be applied easily to abstract selective systems that lack explicit genetics. Rather than force the general selective systems into genetical language, in order to get an exact match to Fisher's theorem, I will pursue only a broad analogy with Fisher's theorem. I will show that Fisher's partitioning of evolutionary change into selective and environmental components is generally useful.

The Price Equation has a similar, although not identical, partitioning between selective and environmental effects on evolutionary change. If, for example, we take fitness as the character under study, $b \equiv w$, then

$$\begin{aligned}\bar{w}\Delta\bar{w} &= \mathrm{Cov}(w,w) + \mathrm{E}(w\Delta w)\\ &= \mathrm{Var}(w) + \mathrm{E}(w\Delta w), \qquad (3)\end{aligned}$$

where the first term is the variance in fitness and the second is the component of evolutionary change caused by changes in the environment.

This is all a bit abstract. I will show in the next section how these definitions can be useful in the context of a particular example. Before turning to that example, I conclude here with a few additional statements.

If we are interested in the evolution of a character, then we need a particular assumption about the relation between the character and fitness. A common assumption is that fitness depends linearly on character value, $w = 1 + sb$, yielding

$$\begin{aligned}\bar{w}\Delta\bar{b} &= \mathrm{Cov}(w,b) + \mathrm{E}(w\Delta b)\\ &= s\mathrm{Var}(b) + \mathrm{E}(w\Delta b). \qquad (4)\end{aligned}$$

When "parent" and "offspring" have identical character values because the environment does not influence characters, $\Delta b = 0$, then one obtains a standard equation from evolutionary genetics, $\bar{w}\Delta\bar{b} = s\mathrm{Var}(b)$, which is often cited incorrectly as Fisher's fundamental theorem.

Another result from eqn (3) is that, at equilibrium,

$$\mathrm{Var}(w) + \mathrm{E}(w\Delta w) = 0,$$

thus selective improvements in fitness, Var(w), must be exactly balanced by what Fisher called "deterioration of the environment", here represented by E($w\Delta w$). This is best explained by example.

An Allocation Problem in a Changing Environment

Many learning problems take the following form. The organism must allocate its limited resources among different strategies, for example the investment in exploiting energy reserves in different spatial locations. The value returned increases at a diminishing rate for each strategy or location. Changes in weather, season, or other extrinsic conditions change the potential value of each strategy or location.

Inherent in the problem is a trade-off between exploration, the search for new strategies, and optimally efficient exploitation of known, successful strategies (Holland, 1975). This type of learning problem is often discussed as a problem of variation and selection—the generation of variant strategies and the selection of the most successful ones (Staddon, 1983). However, such analyses have not been tied to the most powerful and general concepts of selection, the Price Equation and Fisher's fundamental theorem.

Foraging in honey bees provides a good example for analysis. The colony must find patches of flowers producing nectar and extract the energy in an efficient way. Seeley *et al.* (1991) described this system explicitly as learning by variation and selection. All bees, after return from a foraging trip, perform a dance that contains information about the location and quality of a patch of flowers. The colony allocates a certain portion of its foraging bees to act as searching scouts. These bees do not necessarily follow the information provided by dances, but instead search in a way that is partly random. These bees create new "variants" in the selective system. The remainder of the bees tend to follow the dances that indicate the best floral reserves. Thus foraging in successful patches increases by "reproduction," whereas trips to unsuccessful patches "die out". The dance and foraging decisions create a selective filter on the existing variation.

There are several formal models of honey bee foraging. Camazine & Sneyd (1991) and Seeley *et al.* (1991) discuss variation and selection in their theories, but their formal models do not relate general properties of selective systems to selective models of learning. I now introduce such a model in the context 1of the Price Equation.

GROUP SUCCESS AND THE FITNESS OF INDIVIDUAL UNITS

The problem is how to allocate foraging bees among flower patches. Label the patches $i = 1, \ldots, N$, and let the fraction of bees in each patch be q_i. The goal is to find values of q_i that satisfy certain criteria based on the harvesting of resources from flower patches. The total value of resources obtained from a patch in one time period is v_i, and the success per bee is $b_i = cv_i/q_i$, where c is a proportionality constant ($c = 1$ without loss of generality). The total resources obtained by the colony is

$$V_t = \textstyle\sum_{i=1}^{N} v_i.$$

Optimizing total colony success, V_t, is a standard allocation problem that depends on marginal values. Let the marginal rate of change in success per patch be $m_i = \mathrm{d}v_i/\mathrm{d}q_i$. If marginal returns increase at a diminishing rate, $m_i < 0$, then optimal allocation of bees occurs when the marginal values for all patches are the same.

A process of variation and selection does not necessarily lead to optimization of group success. As noted by Bartholdi *et al.* (1993), colony success is maximized by equalization of marginal values, but selective change in foraging behavior based on foraging dances leads to the equalization of per bee fitness for each patch, w_i. The tendency of selection to equilibrate fitnesses is a general property (Slatkin, 1978). This can be seen from the Price Equation, eqn (3), where, under the assumption that no force is changing the value of traits between "parent" and "offspring", $\Delta w = 0$, the condition for equilibrium is Var(w) = 0.

When fitness is linearly related to foraging success, $w_i = 1 + sb_i$, then selection equilibrates per bee foraging success, b_i, and at equilibrium Var(b) = 0, as in eqn (4).

When does selection optimize colony success? When the equilibration of individual fitnesses by selection causes the equilibration of per patch marginal values. This occurs when Var(w) = 0 implies that Var(m) = 0. Under linearity, $w_i = 1 + sb_i$, equilibration of fitnesses implies Var(b) = 0, so a sufficient condition for selection to maximize colony success is $b_i = km_i$, where k is a proportionality constant. Using the definitions for $b_i = v_i/q_i$ and $m_i = \mathrm{d}v_i/\mathrm{d}q_i$ above, this expands to

$$\frac{v_i}{q_i} = k\left(\frac{\mathrm{d}V_i}{\mathrm{d}q_i}\right).$$

This condition is satisfied when v is a power function of q, as in an example below. For most functional relations between v and q the condition is not satisfied, and learning by selection among foraging bees does not optimize colony success.

SPECIFIC ASSUMPTIONS

The fitness function w_i relates per bee foraging success, b_i, to the tendency of other foraging bees to copy or avoid the foraging behavior of an individual. This fitness function may, itself, be subject to modification by colony learning. But at some level the mechanisms by which the colony adjusts its behavior must be influenced by genetic (innate) factors subject to evolutionary modification. In a changing environment the optimal fitness function for the colony may depend on many details, and the solution to this general problem remains unknown.

I proceed by introducing an explicit, linear fitness function. I then turn my attention to a model in which the environment continuously changes. This allows study of the mechanisms by which variants are produced, in this case the scouts of honey bee colonies. In addition, the explicit model clarifies the Price Equation's unusual approach to separating selection from environment.

The two quantities that describe foraging success are

$$v_i = \alpha_i(q_i/\alpha_i)^a = q_i^a\alpha_i^{1-a}$$

$$b_i = v_i/q_i = (\alpha_i/q_i)^{1-a},$$

where v_i is the total value extracted by the colony from the ith flower patch, α_i is the quality of the patch, and q_i is the fraction of all foraging bees that go to the ith patch (Fig. 1). The fraction $(q_i/\alpha_i)^a$ is the relative proportion of resources extracted from the ith patch, where $0 < a < 1$ so that returns increase at a diminishing rate as the number of foragers in a patch increases. The per-bee success in patch i is b_i, which is the character value b that I will use in the Price Equation, matching eqn (2).

In a constant environment, selection equilibrates the b_i's, which implies that foragers are allocated in direct proportion to patch quality

$$q_i^* = \alpha_i/\sum\alpha_i.$$

When the environment varies, the colony is continuously adjusting its allocation of foragers by selection among per bee successes. The colony must also allocate some of its foragers to the scouting role, to locate in time and space patches that are increasing in quality.

I use the beta function (Lindgren, 1976, pp. 328–330) to create a distribution of patch qualities

$$\begin{aligned} \beta(x) &= cx^{g-1}\,(1-x)^{h-1} = 0 \qquad && 0 < x < 1 \\ &= 0 && \text{otherwise,} \end{aligned}$$

where c is chosen so that $\int_0^1 \beta(x)\mathrm{d}x = 1$. I use $g = h = 2$ for beta distribution parameters throughout my numerical analysis.

The patch array $i = 1, \ldots, N$ can be thought of as a circle with index N mapping to 0 to repeat the cycle. A fraction p of the patches have resources and $1 - p$ are empty, so that the number of patches with resources is $P = pN$. Specifically the patch qualities are given by

$$\alpha_i = \int_{(i-1)/P}^{i/P} \beta(x)\mathrm{d}x$$

which can be calculated by commonly available numerical methods for the incomplete beta function (Press *et al*., 1992). The environment changes by "moving" the distribution in each time period t according to

$$\alpha_i^{t+1} = \alpha_{i-1}^t.$$

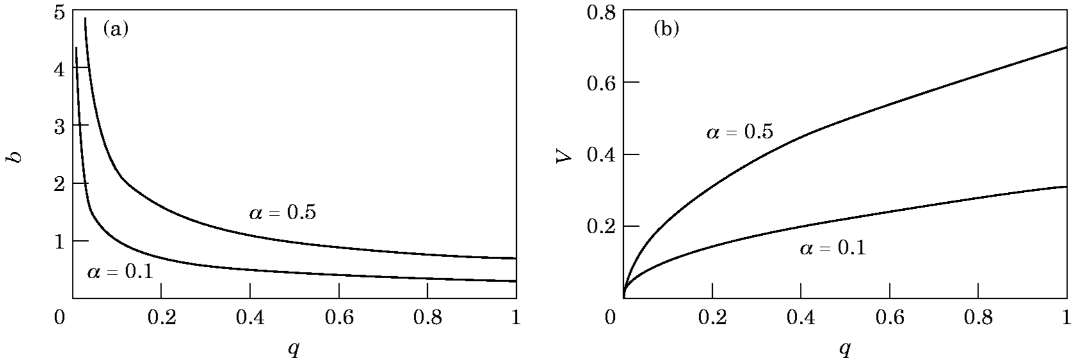

FIG. 1. Shapes of curves for per bee success, b, and per patch value, v, vs. the abundance of foragers in the patch, q. The shape parameter is $a = 0.5$ for these examples. The per bee success, b is proportional to the marginal value of v, that is, $b = k\mathrm{d}v/\mathrm{d}q$.

THREE WAYS IN WHICH THE ENVIRONMENT DETERIORATES

With this definition for change in patch quality we can now summarize properties of "environmental deterioration" in the Price Equation that must balance selection at equilibrium

$$s\mathrm{Var}\ (b) + \mathrm{E}(w\Delta b) = 0. \tag{5}$$

Selective improvement occurs according to variance in the foraging success of individual bees, Var(b). Those bees with higher success have more followers back to their patch in the next time period. This selective improvement must be balanced by deteriorating success among the followers of selectively favored bees, in other words, Δb must often be negative.

The difference in success between bees from the ith patch and their followers is $\Delta b_i = b_i' - b_i$, where the followers' success is

$$b_i' = (1-\mu)b_i^{t+1} + \mu\hat{b},$$

where b_i^{t+1} is the per bee success in the next time period for bees going to patch i. This can be expanded using the explicit definition of b_i given above

$$b_i^{t+1} = \left(\frac{\alpha_i^{t+1}}{q_i^{t+1}}\right)^{1-a}.$$

We can use these definitions to clarify three distinct aspects of environmental deterioration in Δb.

First μ is the scouting or "mutation" rate at which followers choose a patch at random rather than copy returning bees according to their recent foraging success. Random choice typically provides lower success than following a high quality forager, but is needed to explore the changing environment and maintain sufficient variation for effective selection. The success of a bee going to a random patch is $\hat{b}$, where

$$\hat{b} = \sum b_i^{t+1}/N$$

is the per bee success in the next time period taken as an unweighted average over patches.

Second, the quality of a particular patch is likely to vary over time, such that $\alpha_i^{t+1} \neq \alpha_i^t$.

Third, there is a density dependent effect. Successful bees will attract many followers, increasing the number of bees in the patch and decreasing the per bee success. This is accounted for by changing bee densities, $q_i^{t+1} \neq q_i^t$.

QUALITATIVE CONCLUSIONS

I first draw a few general insights from the Price Equation analysis before turning, in the next section, to the numerical properties of the honey bee problem.

It is useful to distinguish among equilibrium, perfection and optimality. The condition for equilibrium is given by eqn (5), which can also be written as $\mathrm{Cov}(b,b') = 0$. When $\mathrm{Cov}(b,b') > 0$ there is a positive correlation in the success of a bee and its followers, and selection will enhance the average value of b in the future by favoring "reproductively" those entities with high b in the present. When the correlation is negative, a positive selective coefficient s causes a decline in average performance by enhancing those bees with followers that are below average. A correlation of zero implies that selective improvements achieved by choosing high quality foragers are balanced by deterioration in the returns obtained by followers, as shown explicitly in eqn (5).

Perfection occurs when the marginal returns on all patches are equal. For the particular assumptions above, $\mathrm{Var}(b) = 0$ implies equality of marginal values. If the environment deteriorates, $E(w\Delta b) < 0$, then $\mathrm{Var}(b) > 0$ and perfection is not possible. In this model the parameters s and μ, selection and mutation in the internal colony learning system, are assumed to change by genetical evolution. Thus, the optimality problem is to find the values of s and μ that minimize environmental deterioration and Var(b).

Learning by this type of selective system cannot achieve perfection in a changing environment. This can be seen by noting that environmental change implies $\alpha_i^{t+1} \neq \alpha_i^t$ for some i. The perfect solution was derived above as

$$q_i^* = \alpha_i / \sum \alpha_i.$$

If, at time t, the distribution of q's is perfect, then $\mathrm{Var}(b) = 0$ and there will be no change in the q's except for random mutation. When, in the next time step, some of the α's change, the q's will no longer match perfectly. The problem is that a selective system necessarily lags the environment. A selective change can provide a closer match to the environment in the future only through the current mismatch. Thus, the optimality problem is not to match the environment perfectly but to track its changes as closely as possible.

Consider a situation at time t when the match is perfect. At time $t+1$ the environment changes, but there is no selective change in the q's because $\mathrm{Var}(b) = 0$ at time t. Suppose α increased in patch, i, that is, $\Delta\alpha_i^t = \alpha_i^{t+1} - \alpha_i^t > 0$. Positive selection, $s > 0$, will increase q_i^{t+2} based on the discrepancy during time $t + 1$. This will provide a closer match to the extent that the direction and magnitude of $\Delta\alpha_i^t$ is a good predictor of α_i^{t+2}. Too large an increase in q_i^{t+2} will cause severe environmental deterioration in that patch because of density dependent interactions.

No single predictor can be optimal for all types of environmental change. For a smoothly changing environment with a stationary distribution of α values, the $\mathrm{Corr}(\Delta\alpha^t, \Delta\alpha^{t+1})$ describes the extent to which an environmental change in one time step correlates with a change in the following time step. Positive selection, $s > 0$, can track environments with positive difference correlations, and negative selection, $s < 0$, can track environments in which increases regularly alternate with decreases. In the case of bees, negative selection would cause foragers to avoid patches that are currently providing better than average returns. Mutation may be required to track an environment to the extent that patch qualities near zero increase substantially, that is, exploration is required to accommodate unpredictable aspects of the environment. Better measures can probably be developed to match optimally the variation and selection parameters to the correlational structure of a changing environment. My main purpose here is simply to describe some general properties of learning and development in selective terms.

PROBLEMS OF DYNAMIC SUFFICIENCY AND A METHOD OF SOLUTION

The quantitative solution to our problem is an equilibrium distribution of bees, q_i^*, over the N patches. A solution requires $N - 1$ constraints plus the fact that

$$\sum q_i = 1.$$

The Price Equation, as we have used it, provides us with only a single equilibrium constraint, $\bar{w}\Delta\bar{b} = 0$, which is not sufficient for solution. The problem is that evolutionary dynamics depend on the full details of the distribution of trait values in the population (Barton and Turelli, 1987; Frank, 1995).

There are two basic methods of solution. The first is essentially a method of moments. The Price Equation gives us a way to use information about variance and transmission to calculate changes in character means—information about the second moment is sufficient for describing changes in the first moment. But in the next time-step the variance (second moment) will have changed, so we cannot calculate further changes in the first moment without dynamic equations for the second moment. For example, if we assume linear fitness, $w = 1 + sb$, then we can generalize the Price Equation to describe changes in the moments of the distribution of trait values as

$$\bar{w}\Delta\overline{b^n} = \mathrm{Cov}(b, b^n) + \mathrm{E}((1 + sb)\Delta b^n),$$

so that changes in the nth moment of b will typically depend on the $n + 1$st moment (Frank, 1995). Closure requires some way of relating higher moments to lower moments.

The second approach is to obtain $N - 1$ explicit recursions that set constraints at equilibrium. These recursions require many additional assumptions beyond the few simple conditions to establish that a system possesses properties of variation and selection. This fundamental fact of selective systems is, I think, why it is so easy to suggest that a process is controlled by selection within a population, yet such a suggestion inevitably appears to be rather loose and incomplete. On the formal side, a person setting out to model the dynamics of a system is inevitably so intent on the full set of assumptions required that one quickly loses sight of the basic selective properties that I have outlined above. Thus most models of variation and selection end up looking like physical models of dynamics because of the modeling process.

I use the recursion approach here to obtain a solution to the honey bee problem. It would have been possible to skip some of the general comments on selection vs. environmental deterioration and move directly to solution by recursion. I think that would be a mistake because formal models in biology must have a dual role. The first is to formalize and test a style of thinking by expressing the ideas rigorously. The second is the numerical consequences implied by particular parameter values. My claim is that too much modeling skips the first part to emphasize the second. General properties of selection are not easy to grasp, and the problem is exacerbated by the fact that so little work has focused on general selective systems outside the confines of mendelian genetics (Price, 1995).

OPTIMAL PARAMETERS FOR VARIATION AND SELECTION

Returning to the honey bee problem, our solution will follow from the constraint on environmental change, $\alpha_{i-1}^t = \alpha_i^{t+1}$. This is useful for calculation, but actually we only need the much weaker stationary condition that $\alpha_j^{t+1} = \alpha_i^t$ for any arbitrary one-to-one mapping of j to i, where the mapping can change in each time step. But for convenience I continue to map $i - 1$ in the current period to i in the next period.

With this steady shifting of patch quality in space

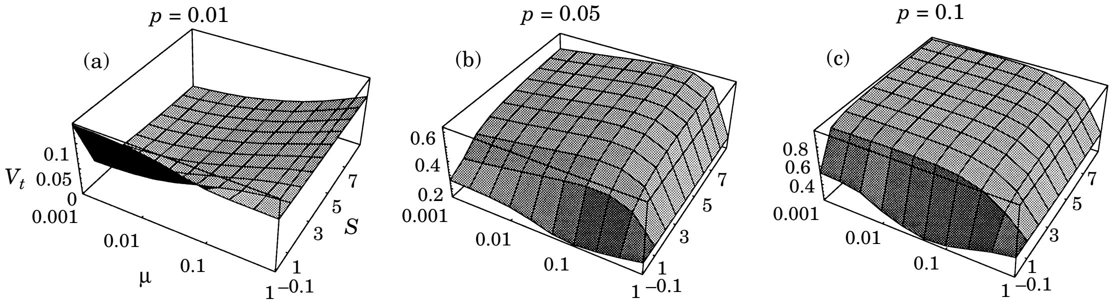

FIG. 2. Total value obtained by the colony for scouting rate, μ, and selection intensity, s. There are $N = 100$ patches; the fraction of patches with resources is p. The shape parameter a is 0.5. In panel (*a*), $\mathrm{Corr}(\Delta\alpha^t, \Delta\alpha^{t+1}) = -0.5$ and consequently a selective value of $s < 0$ is favored. For panel (*b*), $\mathrm{Corr}(\Delta\alpha^t, \Delta\alpha^{t+1}) = 0.2$, and for panel (*c*), $\mathrm{Corr}(\Delta\alpha^t, \Delta\alpha^{t+1}) = 0.1$.

and constancy of the distribution of quality at different points in time, the equilibrium conditions are clearly $q_{i-1}^t = q_i^{t+1}$, which can be expanded as

$$q_{\mathrm{i}-1} = (1 - \mu)q'_{\mathrm{i}} + \mu\bar{q}$$
$$= (1 - \mu)q_i w_i/\bar{w} + \mu\bar{q},$$

where $q'_i = q_i w_i/\bar{w}$ is the fraction of bees that follow foragers from the *i*th patch among those bees influenced by selection, μ is the fraction of bees that are scouts and choose patches at random, and $\bar{q} = 1/N$. Using $w_i = 1 + sb_i$ and the definition of b_i given above, we have a complete set of nonlinear constraints that must be solved numerically for the equilibrium distribution of bees in patches, q_i. With that solution we can calculate the per bee success, b_i, and the total value extracted by the colony, V_t.

The total value extracted by the colony is shown in Fig. 2 for various assumptions, where the value of V_t is the proportion of the maximum value that the colony could obtain by perfect allocation. The optimum values of selection, s, and scouting (mutation), μ, are given by the maximum value over the surface.

Figure 3 shows the distributions of patch quality and forager allocations for two sample parameter combinations. The values of μ and s are optimal for the parameters illustrated. Figure 3(a) is interpreted as follows. The dashed distribution shows the quality of patches, α_i, with

$$\sum \alpha_i = 1.$$

This distribution is shifted to the right by one patch in each time step, wrapping around so that patch 10 shifts to patch 1. The solid distribution shows the frequency of foragers in each patch at equilibrium. The forager distribution shifts to the right by one patch in each time step, keeping pace with the changing patch qualities. Maximum colony success, $V_t = 1$, occurs when the patch quality and forager distributions are identical. In Fig. 3(a), $V_t = 0.78$ and in Fig. 3(b), $V_t = 0.93$.

In Fig. 2(a) patch quality is ephemeral. At any time only one patch has resources and all others are empty, that is, $pN = 1$. The patch with resources will be empty in the following time step, thus recruiting foragers to a quality patch actually decreases colony

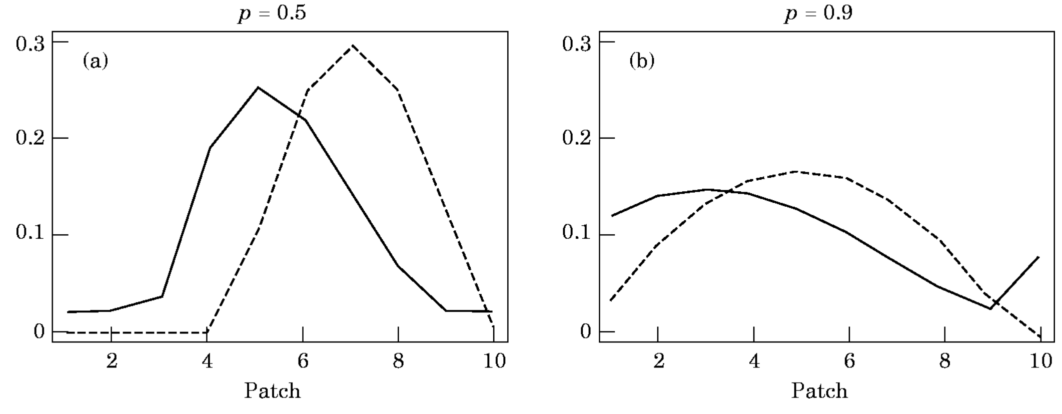

FIG. 3. Distributions of patch quality and forager frequency. Here $N = 10$, $a = 0.5$, scouting rate $\mu = 0.2$, and selection intensity $s = 10$. These parameters for μ and s are approximately optimal, that is, they yield the highest V_t obtainable. Key: --- patch quality, α_i; —— forager frequency, q_i.

success. The only information the colony can obtain is that a quality patch in one time step will have no resources in the following time step. The best fitness function has $s < 0$, with new foragers avoiding the patch from which successful foragers just returned.

Conclusion

I have emphasized learning with a particular example of honey bee behavior. But the same general problems apply to development, in which the relative success of different cellular lineages strongly influences the ontogeny of form. The cellular level of success and the organismal level of success will often conflict in the same way that per bee success differs from colony success. Organismal control depends on generative mechanisms that create variants and selective filters that assign fitnesses to lineages. Evolution at the lineage level affects organismal performance. To the extent that such internal selection occurs, genetic control will act directly on generative mechanisms and selective filters and only indirectly on final form. The performance, based on final form, then influences the selection of genetic variants among organisms (Plotkin & Odling-Smee, 1981; Frank, 1996).

This analysis of learning and development as selective systems makes explicit the idea that genes do not encode a complete blueprint for the organism. Rather, for many traits an organism may begin life with only broad rules for a dynamic system that adapts to its environment.

The general properties of selective systems have not received much attention. The Price Equation provides a natural way to explore such properties and to gain a deeper understanding of selective systems.

My research is supported by NSF grants DEB-9057331, and DEB-9627259 and NIH grant GM42403.

REFERENCES

Bartholdi III, J. J., Seeley, T. D., Tovey, C. A. & Vande Vate, J. H. (1993). The pattern and effectiveness of forager allocation among flower patches by honey bee colonies. *J. theor. Biol.* **160,** 23–40.

Barton, N. H. & Turelli, M. (1987). Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genet. Res.* **49,** 157–173.

Boyd, R. & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Camazine, S. & Sneyd, J. (1991). A model of collective nectar source selection by honey bees: self-organization through simple rules. *J. theor. Biol.* **149,** 547–571.

Cavalli-Sforza, L. L. & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton: Princeton University Press.

Changeux, J.-P. & Danchin, A. (1976). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature* **264,** 705–712.

Crick, F. (1989). Neural edelmanism. *Tr. Neurosci.* **12,** 240–248.

Edelman, G. M. (1987). Neural Darwinism: The Theory of Neuronal Group Selection. New York: Basic Books.

Edwards, A. W. F. (1994). The fundamental theorem of natural selection. *Biol. Rev.* **69,** 443–474.

Ewens, W. J. (1989). An interpretation and proof of the fundamental theorem of natural selection. *Theor. Popul. Biol.* **36,** 167–180.

Ewens, W. J. (1992). An optimizing principle of natural selection in evolutionary population genetics. *Theor. Popul. Biol.* **42,** 333–346.

Fisher, R. A. (1958). The Genetical Theory of Natural Selection, 2nd ed. New York: Dover.

Frank, S. A. (1995). George Price's contributions to evolutionary genetics. *J. theor. Biol.* **175,** 373–388.

Frank, S. A. (1996). The design of natural and artificial adaptive systems. In: Evolutionary Biology of Adaptation (Rose, M. R. & Lauder, G. V., eds) San Diego: Academic Press (in press).

Frank, S. A. & Slatkin, M. (1992). Fisher's fundamental theorem of natural selection. *Tr. Ecol. Evol.* **7,** 92–95.

Golub, E. S. & Green, D. R. (1991). Immunology: A Synthesis, 2nd Edn. Sunderland, MA: Sinauer Associates.

Gordon, R. (1966). On stochastic growth and form. *Proc. Natl. Acad. Sci. U.S.A.* **56,** 1497–1504.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press.

Kagan, M. L., Novoplansky, N. & Sachs, T. (1992). Variable cell lineages form the functional pea epidermis. *Ann. Bot.* **69,** 303–312.

Lewontin, R. C. (1970). The units of selection. *A. Rev. ecol. Syst.* **1,** 1–18.

Lindgren, B. W. (1976). *Statistical Theory*, 3rd Edn. New York: Macmillan.

Michaelson, J. (1987). Cell selection in development. *Biol. Rev.* **62,** 115–139.

Plotkin, H. C. & Odling-Smee, F. J. (1981). A multiple-level model of evolution and its implications for sociobiology. *Behav. Brain Sci.* **4,** 225–268.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992). *Numerical Recipes in C*, 2nd Edn. Cambridge: Cambridge University Press.

Price, G. R. (1970). Selection and covariance. *Nature* **227,** 520–521.

Price, G. R. (1972*a*). Extension of covariance selection mathematics. *Ann. hum. Genet.* **35,** 485–490.

Price, G. R. (1972*b*). Fisher's 'fundamental theorem' made clear. *Ann. hum. Genet.* **36,** 129–140.

Price, G. R. (1995). The nature of selection. *J. theor. Biol.* **175,** 389–396.

Sachs, T. (1988). Epigenetic selection: an alternative mechanism of pattern formation. *J. theor. Biol.* **134,** 547–559.

Seeley, T. D., Camazine, S. & Sneyd, J. (1991). Collective decision-making in honey bees: how colonies choose among nectar sources. *Behav. Ecol. Sociobiol.* **28,** 277–290.

Slatkin, M. (1978). On the equilibration of fitnesses by natural selection. *Am. Nat.* **112,** 845–859.

Spiegelman, S. (1945). Physiological competition as a regulatory mechanism in morphogenesis. *Q. Rev. Biol.* **20,** 121–146.

Staddon, J. E. R. (1983). *Adaptive Behavior and Learning*. Cambridge: Cambridge University Press.

Wagner, G. P. & Misof, B. Y. (1993). How can a character be developmentally constrained despite variation in developmental pathways? *J. Evol. Biol.* **6,** 449–455.