J. theor. Biol. (1995) **175,** 389–396

The Nature of Selection

George R. Price†

(*Written circa* 1971, *published posthumously*)

A model that unifies all types of selection (chemical, sociological, genetical, and every other kind of selection) may open the way to develop a general "Mathematical Theory of Selection" analogous to communication theory. [*Note added by S. A. Frank*: This previously unpublished manuscript was found among Dr. Price's papers when he died in 1975. In this paper Dr. Price did not provide a complete, general theory of selection. Rather, he argued why such a theory is needed and what some of its properties might be. The accompanying article provides commentary on this paper and describes Dr. Price's significant contributions to evolutionary genetics (S. A. Frank, 1995, *J. theor. Biol.* **175,** 373–388).]

Introduction

Selection has been studied mainly in genetics, but of course there is much more to selection than just genetical selection. In psychology, for example, trial-and-error learning is simply learning by selection. In chemistry, selection operates in a recrystallisation under equilibrium conditions, with impure and irregular crystals dissolving and pure, well-formed crystals growing. In palaeontology and archaeology, selection especially favours stones, pottery, and teeth, and greatly increases the frequency of mandibles among the bones of the hominid skeleton. In linguistics, selection unceasingly shapes and reshapes phonetics, grammar, and vocabulary. In history we see political selection in the rise of Macedonia, Rome, and Muscovy. Similarly, economic selection in private enterprise systems causes the rise and fall of firms and products. And science itself is shaped in part by selection, with experimental tests and other criteria selecting among rival hypotheses.

And yet, despite the pervading importance of selection in science and life, there has been no abstraction and generalization from genetical selection to obtain a general selection theory and general selection mathematics. Instead, particular selection problems are treated in ways appropriate to particular fields of science. Thus one might say that "selection theory" is a theory waiting to be born—much as communication theory was 50 years ago. Probably the main lack that has been holding back any development of a general selection theory is lack of a clear concept of the general nature or meaning of "selection". That is what this paper is about.

Let us pursue a little further the analogy with communication theory. Probably the single most important prerequisite for Shannon's famous 1948 paper on "A Mathematical Theory of Communication" was the definition of "information" given by Hartley in 1928, for it was impossible to have a successful mathematical theory of communication without having a clear concept of the commodity "information" that a communication system deals with. Hartley gave what he described as a "physical as contrasted with psychological" definition of *information*, which omitted all considerations of the meaningfulness of messages but measured attributes relevant to the design of communication systems. Similarly, for development of a useful mathematical theory of selection, one needs a physical rather than psychological definition of *selection*, which excludes psychological factors of preferences and decision making. It is my hope that the concept of *selection* proposed in this paper will contribute to the future development of "selection theory" as helpfully as Hartley's concept of *information* contributed to Shannon's communication theory.

---

† Address correspondence to: Steven A. Frank, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717, U.S.A.

0022–5193/95/150389+08 $12.00/0

In addition, I have been endeavouring to develop new mathematical treatments of selection appropriate to the generalized selection concept to be explained (Price, 1970, 1972). However, the present paper excludes almost all formal mathematics of that sort, since attention to mathematical details might distract from the main purpose of explaining the nature of selection.

Concepts of Selection

Two different main concepts of selection are employed in science. These are illustrated in Fig. 1. Historically and etymologically, the meaning of *select* (from *se*-aside + *legere* to gather or collect) was to pick out a subset from a set according to a criterion of preference or excellence. This we will call *subset selection* [Fig. 1(a)]. Darwin introduced a new meaning (as Wallace, 1916, pointed out to him), for offspring are not subsets of parents but new entities, and Darwinian natural selection [Fig. 1(b)] does not involve intelligent agents who pick out. (Darwin distinguished "natural selection" from "sexual selection". Of course the latter does involve preferences and decisions.)

These two concepts are seemingly discordant. What is needed, in order to make possible the development of a general selection theory, is to abstract the characteristics that Darwinian natural selection and the traditional subset selection have in common, and then generalize. This will lead us to a concept of selection as a certain type of transformation that may occur as the result of a natural force or as the result of preference, optimization calculation, and decision making by an intelligent agent. What we consider here is just the act or process of transformation, and not any preferring, calculating, or deciding that may have given rise to it. Inasmuch as Darwin's concept of selection exercised by a natural force is familiar to workers in all fields of science, no confusion should arise from this use of a "physical as contrasted with psychological" definition of *selection*. (In contrast, there has been confusion from the communication theory definition of "information" without reference to semantic content. We may say that Hartley defined the right concept, but chose a wrong name for it.)

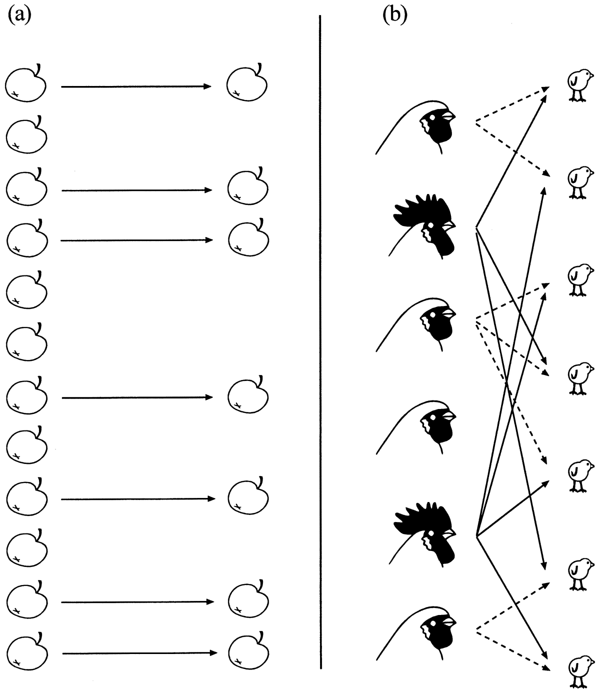

FIG. 1. Conventional concepts of selection. (a) Subset selection. (b) Darwinian selection.

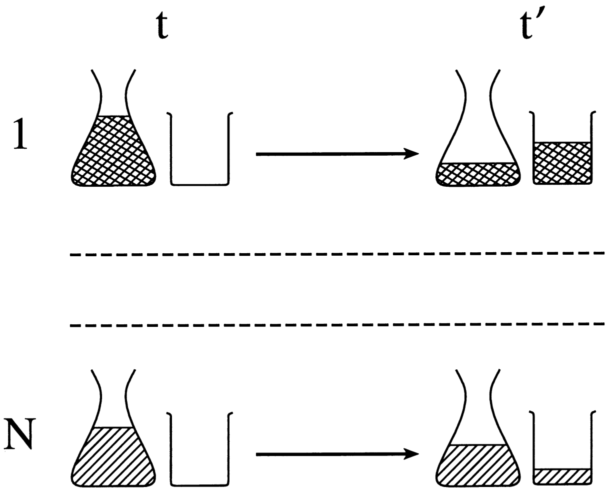

FIG. 2. A solution selection example.

We now proceed to describe a general model fitting all forms of selection, which will unify the seemingly discordant cases illustrated in Fig. 1.

A General Model for Selection

We begin with a very simple example. Line 1 of Fig. 2 shows that at a certain time (t) before selection there is an erlenmeyer flask containing a certain amount w of solution of concentration x. Selection involves pouring some of the solution from the flask into the beaker, so that at time t' (after selection) the beaker contains amount w' of solution of concentration x'. (In this example $x'=x$, but this restriction does not hold in all selection cases.)

Now suppose that we have N such flask and beaker pairs, each containing the same kind of solution though not all at the same concentration, as is indicated in Fig. 2 (where the darkness of shading indicates the solution concentrations). Let us suppose that we are interested in the change in (weighted) average concentration from the N flasks at time t to the N beakers at time t'. That is, if the initial solutions in all flasks at time t had been (we imagine) mixed together, the concentration of the total mixed solution would have been

$$X=(\Sigma\ w_i x_i)/\Sigma\ w_i$$

summed over $i=1$ to N (provided that amounts and concentrations are measured in mass units). Similarly, if the solutions in the beakers after selection were mixed together, the concentration in the mixture would be

$$X'=(\Sigma\ w_i' x_i')/\Sigma\ w_i'.$$

Let $\Delta X=X'-X$, the change in (weighted) average concentration. Suppose that large amounts of the more concentrated solutions and small amounts of the less concentrated solutions are poured into the beakers. Then we will ordinarily expect ΔX to be positive. Evidently "selection" means taking different amounts of the different solutions in relation to their concentrations. Therefore we might consider measuring the degree or amount of selection occurring in the Fig. 2 example by calculating the correlation or covariance between the concentrations x_i and the amounts w_i' of solution poured into the beakers.

Now we are ready to generalize. In Fig. 2 the important elements are the flasks at time t and the beakers at time t'. Therefore it suffices to show only these elements, as in Fig. 3(a). And Fig. 3(a) leads immediately to the idea of a schematic representation as in Fig. 4.

Figure 4 is the general selection model. It shows a pre-selection set or population P of 'packages' p_i, and a post-selection set or population P' of packages p_i'. Each pre-selection package has an amount w_i (symbolized by the area shaded) of something which has a quantitative property x, of numerical magnitude x_i (symbolized by the darkness of shading). Similarly, each post-selection package has an amount w_i' of something with numerical value x_i' for property x. The arrows from P to P' symbolize the selection process, with arrow thickness roughly representing the fractional amount selected,

$$s_i=w_i'/w_i,$$

which we will term the *selection coefficient*. The $X \longrightarrow X'$ at the bottom represents the change produced by the selection process in a *population property* X related to property x of individual set members. (For example, X might be the arithmetic mean of the x_i or their variance, and correspondingly for X' and the x_i' values.) *Selection* itself is illustrated in the figure by the association between darkness of shading of pre-selection packages and the thickness of the corresponding arrows.

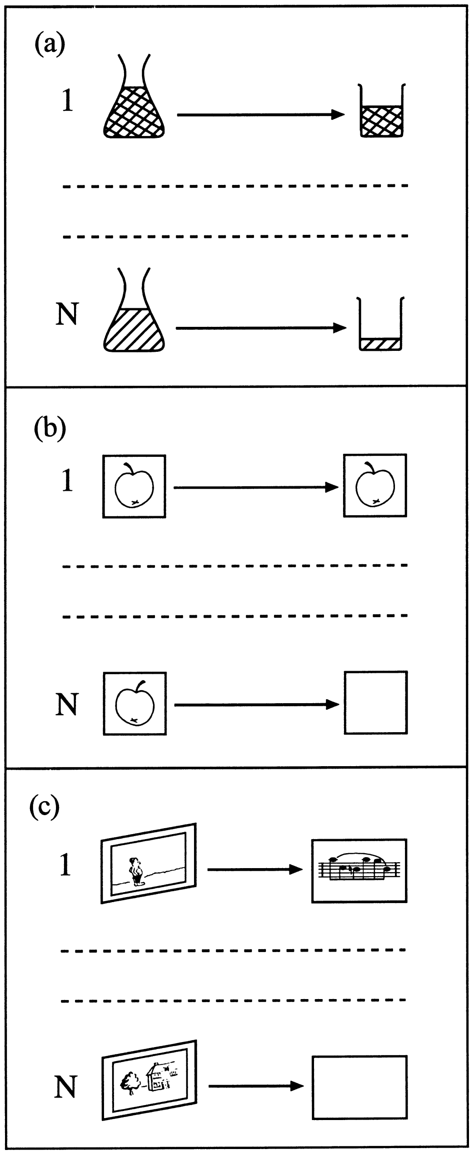

FIG. 3. Three selection examples arranged in the pattern of the general selection model. (a) The essential elements of the Fig. 2 example. (b) How the Fig. 1(a) example is fitted to the general model. (c) Moussorgsky's selection of "Pictures from an Exhibition".

Figure 4 illustrates the possibilities that $w_i'>w_i$ (see line N) and $x_i' \neq x_i$ (see line i). Both these possibilities result from Darwin's extension of the selection concept, since offspring can exceed parents in number and mass and can differ from them in many ways. We will speak of the change $x_i \longrightarrow x_i'$ in line i as an instance of *property change*. In general, the change

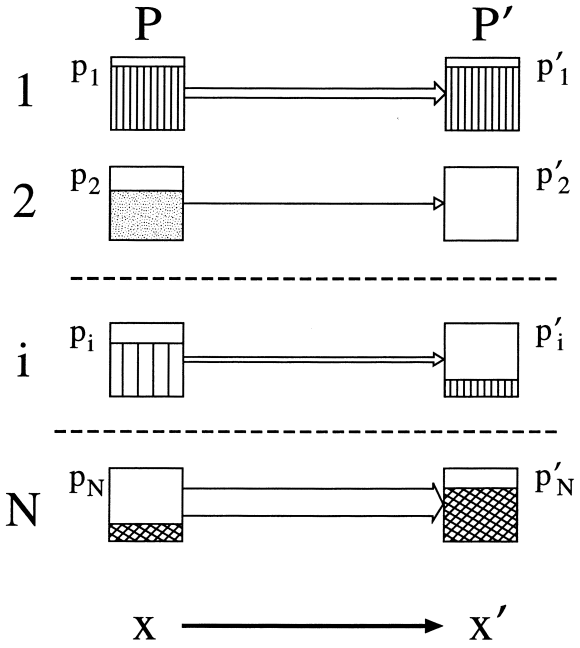

FIG. 4. The general selection model.

ΔX will be partly caused by selection and partly caused by such changes in individual x values.

Particularly to be noted is the one-to-one correspondence between P and P'. This should be contrasted with the 12-to-7 relation of the two sets in Fig. 1(a), and the complicated, criss-crossed 6-to-7 relation in Fig. 1(b). We will say that a set P' is a *corresponding set* to a set P if there exists a one-to-one correspondence such that, for each member p_i of P, there is a corresponding member p'_i of P' which (if not empty) is composed partly or wholly of the same material as p_i, or has been derived directly from p_i, or contains one or more replicas of p_i or some part of p_i, or has some other special close relation to p_i. Using this concept of a "corresponding set", we can now define *selection* as follows:

Selection on a set P in relation to property x is the act or process of producing a corresponding set P' in a way such that the amounts w'_i (or some function of them such as the ratios w'_i/w_i) are non-randomly related to the corresponding x_i values.

Accordingly, "random selection" is not selection—as in fact any housewife would insist if she had bought what were labelled as "Selected Eggs", and then learned that they had been "selected" at random.

If we compare the proposed "selection" definition with Hartley's definition of "information", it may appear that I have failed to accomplish something similar since the "selection" definition is quite broad whereas Hartley gave a precise mathematical definition (except that he left the logarithmic base unrestricted). However, Shannon somewhat extended Hartley's definition, indicating that the original definition was too narrow. Furthermore, both Hartley and Shannon defined "information" in terms of selection procedures. From this it is clear that selection is a much broader concept than "information" (in the communication theory sense). Therefore a fairly broad definition is desirable.

The meaning of the definition will become clearer as we consider some examples.

Unification of Selection Concepts

Now let us see how the Fig. 1 examples can be fitted to the general model. Figure 3(b) shows how the Fig. 1(a) subset selection example is handled by the device of showing an "empty package" with $w'_i = 0$ if apple i is not selected. Thus the 12-to-7 relation is converted into a one-to-one relation. We might define w to be apple mass. Alternatively, we might define w as the number of apples per package. In this case we would probably write n instead of w, with $n_i = 1$ for all i, and $n'_i = 0$ or 1 according as apple i is not or is selected. In either case, we would probably define x as some measure of apple quality, with X measuring the average quality.

In order to fit the Fig. 1(b) genetical selection example to the model, we expand it into the system shown in Fig. 5, in which in fact the pattern of the Fig. 4 model occurs twice. In Fig. 5, subscripts 1 and 2 label parent and offspring generations, respectively. Column P_1 is identical with the left column of Fig. 1(b) except for the addition of rectangles to symbolize the package concept, plus the portrayal of a "representative cell" beside each animal. Similarly, P'_2 is identical with the right column of Fig. 1(b) except for the addition of rectangles and representative cells. The two intermediate columns show sets of "successful gametes" in P'_1 and newly conceived offspring in P_2. The tiny circles in the representative cells, gametes and fertilized eggs portray the gene constitution at some gene locus, with the black circles representing type 'G' genes and the open circles representing any other genes (not of type G) present at that gene locus.

It should be noticed in Fig. 5 that the steps from P_1 to P'_1 and from P_2 to P'_2 are labelled as selection, whereas the middle step, from P'_1 to P_2, is described as redistribution. If one counts the two types of tiny circles, it will be found that gene numbers change in the selection steps, but do not change in the redistribution step. In fact, set P'_1 consists essentially of the genetical material of P_2 members sorted out according to the parents from which they came. The advantage from

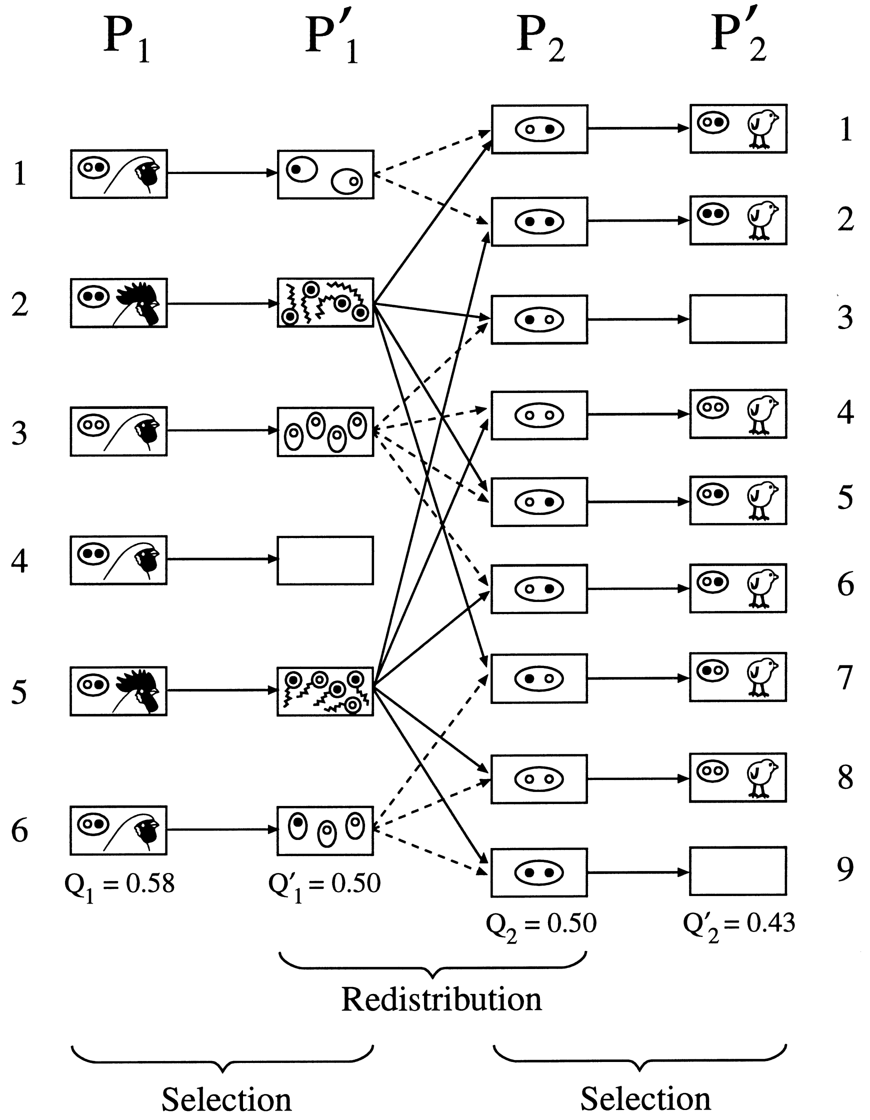

FIG. 5. A genetical selection example [showing how the Fig. 1(b) example is fitted to the general selection model].

this arrangement is that the selection steps are simple and orderly, and therefore convenient for analysis and calculation; while the complicated tangle of Fig. 1(b) is confined to the redistribution step, which for many purposes can be ignored since total gene numbers do not change there.

Suppose that we are interested in the effect of selection in changing the frequency of gene G in domestic fowl populations. As amount measure we use the number of tiny circles shown in a package, which we will represent by n. Thus $n=2$ for all P_1 and P_2 members. For P_1', $n'=2$, 4, 4, 0, 5, or 3. And for P_2', $n'=2$ for individuals that survive to hatching, but equals 0 for individuals 3 and 9, which die before hatching. As individual and population properties, we use the individual and population frequencies of gene G, which however we will represent by q and Q, respectively, instead of by x and X. To find these frequencies, we simply count the tiny black circles that represent type G genes, and divide by total numbers of tiny circles.

For example, $q_{1,5}$, the gene G frequency in member no. 5 of set P_1, $=1/2$; and $q_{1,5}'$, the gene G frequency in package no. 5 of set P_1', $=3/5$. Similarly, to find population frequencies, we count all tiny circles in a column. Thus $Q_1=7/12\approx 0.58$, and $Q_1'=9/18=0.50$. In generation 1, we can measure the fecundities of individuals by $s_{1i}=n_{1i}'/n_{1i}$, which is the number of offspring conceived by individual i if we credit a parent with half of each offspring conceived. It can be seen that Fig. 5 illustrates negative correlation between the frequencies q_{1i} and the fecundities s_{1i}, which constitutes selection against gene G, resulting in a decrease $\Delta Q_1 = Q_1' - Q_1 = -0.08$. Similarly, the right-hand side of Fig. 5 illustrates negative correlation between the frequencies q_{2i} and the survival coefficients $s_{2i} = n_{2i}'/n_{2i}$, resulting in a further decrease $\Delta Q_2 = Q_2' - Q_2 = -0.07$.

Thus the general selection model fits both Fig. 1 cases, thereby unifying the two seemingly discordant concepts of selection.

A further advantage resulting from the selection model and the point of view of selection that is associated with the model, is that the simple and orderly one-to-one pattern tends to clarify complex selection cases and in some cases can lead to perception of surprisingly simple mathematical relationships. For example, on the left below is a genetical selection equation recently published in *Nature* (O'Donald, 1969):

$$\Delta\bar{x} = \frac{2KV(\theta - \bar{x}) - K\mu_3}{1 - \alpha - K(\theta - \bar{x})^2 - KV} = \frac{\mathrm{Cov}(w, x)}{\bar{w}} = \Delta\bar{x}. \tag{1}$$

The Fig. 4 point of view leads to recognition of the much simpler relation shown on the right above, which is equivalent to equation (1) of Price (1970). (Here $w = 1 - \alpha - K(\theta - x)^2$ as defined in O'Donald (1969), $\mathrm{Cov}(w, x)$ is the covariance or first order central product moment of w and x, and the equivalence of the expressions given above can be verified by use of the standard equations relating central moments and moments around the origin.) It needs to be mentioned here that this covariance relationship has been widely misunderstood because it can be expressed also in terms of a regression coefficient, and this has resulted in confusion with conventional uses of regression coefficients to give linear approximations in totally dissimilar and much less accurate selection equations. But it appears that the covariance selection equations of Price (1970, 1972), despite their simplicity, have not been recognized earlier—yet it is almost inevitable to recognize them if one thinks of selection in terms of the Fig. 4 model.

Use of the General Selection Model

It is always possible to find the Fig. 4 pattern in any case of selection, but until one has gained some experience with this it can be puzzling to recognize the pattern and decide how to define amount and property variables. One difficulty is that some elements of the pattern may be present only implicitly or as mathematical constructs. Thus set P_1' in Fig. 5 is such a construct inasmuch as the "successful gametes" that give rise to a new generation are not likely all to be in existence at any given instant of time, let alone to be separated from other gametes and sorted out as in Fig. 5.

In some cases the very simplicity of a selection case may make it difficult to see that all the elements of the model are present. For example, suppose that we pick out a single book from a bookcase. Here P is the set of N books originally in the bookcase, and P' is one package containing the selected book plus $N - 1$ empty packages. The usual choice for amount measure would be number of books, so that $n_i = 1$ for all i, and n_1' and $s_i = 1$ for the selected book and $= 0$ for all other books. The appropriate choice for property x would depend on our purposes and interests. Thus it might be a measure of useful information or of the entertainment value in a book.

It is interesting to note that Hartley's way of defining "information" involved the same sort of selection pattern as in the book selection example, with selection of a single symbol (at any one time) from the set of all possible symbols for the given communication system. Shannon modified Hartley's definition by taking into account the probability $p_i(j)$ that an information source in state i would produce (or select) symbol j. We can represent these probabilities in our model as amounts w, for example writing $w_t(i)$ to represent the probability that symbol i will be selected under the conditions that exist at time t.

Next, we note that P' can contain elements quite different from those in P, as a consequence of some of the possibilities allowed in our definition of corresponding set. For example, in "Pictures at an Exhibition" [see Fig. 3(c)] Moussorgsky made a selection of pictures about which to compose music, and made further selection in respect to the amount of music composed in relation to a given picture. Here we can define P as the set of sketches at the 1874 Hartman exhibition, and define p_i' as the music (possibly none) related to picture i. Possible amount measures are $n_i = 1$ picture for all i, and $n_i' =$ the number of bars of music related to picture i. Then, if one could define interesting attributes of mood and subject matter that could be quantitatively evaluated in the paintings, one could measure (using definitions of "selection intensity" given in Price, 1972) the degree to which Moussorgsky selected for or against these.

Lastly, it may be helpful to explain some points about choosing appropriate properties to play the role of the "individual property" x. Property x can be a "primary selection property" or "primary criterion of selection" that directly influences the selection process, or it can be a property that has only an indirect or remote relation to the selection process. For example, if gold and silver are associated in some mineral, with the gold being easily apparent from its colour but the silver being difficult to recognize, then the gold content would be a primary selection property for a prospector who selected samples on the basis of appearance, while the silver content would be an associated property that would be indirectly influenced through the direct selection on a gold. Another possibility is that x can represent a composite property such as the sum or product of a number of more elementary properties. Also, in many selection cases there can be two or more independent selection properties, x_1, $x_2, \ldots$, with selection simultaneously occurring in respect to each of these. In selection cases of practical interest it will usually be obvious how to define property x (or properties x_1, $x_2, \ldots$), but one can easily be puzzled for a while if one seeks to invent odd selection cases to test whether the Fig. 4 model can be fitted to them.

Developing 'Selection Theory'

When Shannon's "Mathematical Theory of Communication" appeared in 1948, many scientists must have felt surprise to find that at so late a date there had still remained an opportunity to develop so fundamental a scientific area. Perhaps a similar opportunity exists today in respect to selection theory. If we compare the high level of communication technology reached 50 years ago with the very disappointing results usually reached nowadays in computer simulations of evolution (for example, as described by Bossert, 1967), and if we note the degree of understanding of communication systems shown in the 1928 papers of Nyquist (1928) and Hartley (1928) and then consider that it took another 20 years before Shannon's (1948) paper appeared, we can reasonably predict that much difficult work will be required before an interesting and useful "Mathematical Theory of Selection" can be developed. The remainder of this paper contains suggestions for readers who may wish to consider working on this problem themselves.

My main suggestion is that one should become well-acquainted with biological selection complexities, but work mathematically on abstract, generalized selection systems rather than on biological selection or other real selection systems. The reason for the first part of the suggestion is that unless one knows something about the range of remarkable things that happen in biological selection, one's ideas about what to investigate in mathematical models are almost certain to be too simple and unimaginative. (For example, who would ever think of inventing arrhenotokous reproduction if it did not actually occur? Yet is a highly successful system, characterizing those remarkably successful insects the ants, bees, and wasps.) The reason for the second part of the suggestion is obvious, so that it might seem to be unnecessary to make such a suggestion if it were not for the fact that mathematical biologists working on selection seem almost always to deal with actual rather than generalized systems. Therefore it may be helpful to say something further about abstract, generalised selection systems. Some possibilities are illustrated in Fig. 6.

Here an individual is represented as a point in a *property space* which may be of any desired dimensionality (x_1, $x_2, \ldots$), though only two dimensions are shown in the figure. Figure 6(a) shows successive stages in a multistep selection process involving alternate steps of property change without selection, and selection without property change. In

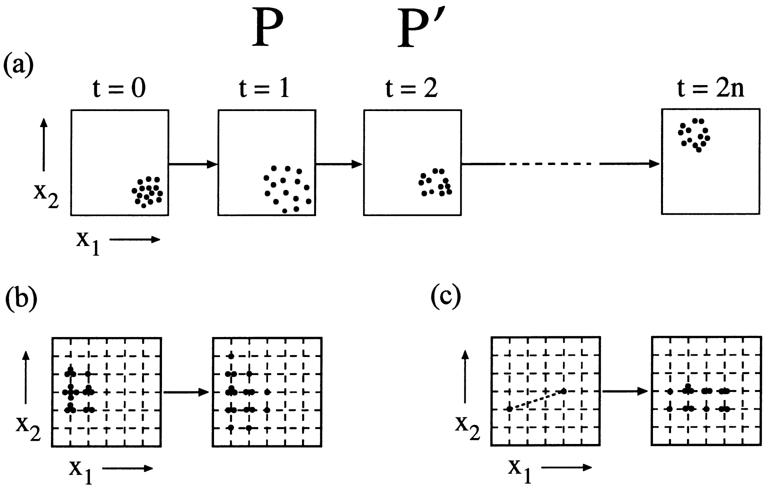

FIG. 6. Generalized "property space" treatment of selection.

the first step, first time 0 to 1, the 15 points (representing a population of 15 individuals) spread apart as though by a diffusion or repulsion process. The next step, from time 1 to 2, is a selection step, with $s_i = 0$ for most of the points, but $s_i = 2$ or 3 for five points with relatively low x_1 and high x_2 values. (The points at time 2 would be superimposed so that only five could be seen, but they have been drawn slightly displaced so that separate points can be distinguished.) The final view, at time $2n$, shows the result after n cycles of property change followed by selection.

There are a great number of possible ways in which the Fig. 6(a) system might be varied. One possibility that is suggested by Fig. 5 involves noticing that the individual gene frequencies in P_1, P_2 and P_2' are restricted to the three values 0, 1/2, and 1. This suggests that we quantize the property variables $x_1, x_2, \ldots$, to k possible values ($2 \leqslant k \leqslant \infty$), resulting in a lattice distribution as illustrated in Fig. 6(b) (where again some points are shown slightly displaced so that separate points can be distinguished). The change illustrated in Fig. 6(b) involves property change without selection. Lastly, Fig. 6(c) illustrates a transformation analogous to sexual reproduction, with a "mating pair" in the first diagram and their "offspring" in the second diagram, and with both property change and selection with $s > 1$. (Other population members besides those shown are supposed present.)

With such systems it should be possible to investigate many interesting and important problems. Here are two examples:

(i) Selection requires variation and acts to reduce variation: this results in a competing relationship between selection and entropy increase about which it would be interesting to have deeper understanding. What happens can be briefly described as follows. Since selection depends upon variation and uses up variation, multistage selection requires processes to produce new variation. When the selection intensity is high, such variation production enables selection to operate rapidly to create complexity and order; but if the selection intensity is low, these processes of variation production cause deterioration of complexity and order. Therefore selection is needed to maintain whatever order has been produced by selection. In terms of the Fig. 6 examples, it should be noted that the property change steps have been shown as increasing variation—which is why repulsion and diffusion processes were mentioned rather than attraction or coagulation. It may also be noted that lattice distributions with $k = 3$ (as in the great majority of multicellular species) result in low variance for genes with population frequencies near 0 to 1, and high variance for genes at intermediate frequencies. This gives rapid selection while a gene is changing from high to low or from low to high frequency, but slow selection, given much stabilization against entropy increase, once the gene has reached low or high frequency. It would be of great interest and value to have a general equation unifying and clarifying these and other matters of selection versus entropy increase.

(ii) A problem of great basic interest has been stated by Hamilton (1970): "How much does the evolved complexity of life owe to its physical setting? Cracked, contorted continents; tidal oceans; a blanket of capricious vapors—were these the necessary cradle for life? Are they a continuing incentive?" In other words, given a uniform world with constant weather, would the negentropy of solar radiation suffice to produce the necessary variation for natural selection to act upon? Or, stated more generally, are different types of variation in all ways interchangeable subject only to the limits set by the Second Law of Thermodynamics? Or are there different categories of variation such that two systems of equal negentropy can have very different capabilities in "fuelling" selection processes, and not be interconvertible even in theory?

Consideration of questions such as these, though in terms of abstract models rather than genes or continents, should lead to deepening understanding of selection much that in time someone will have the insight to take a very large step forward like that taken by Shannon in 1948.

I thank Professor Cedric A. B. Smith for scientific help, Mr. Alick Newman and Mr. Richard Bryant for the illustrations, and the Science Research Council for financial support.

REFERENCES

Bossert, W. (1967). Mathematical optimization: are there abstract limits on natural selection? In: *Mathematical Challenges to the Neo-Darwinian Interpretation of Evolution* (Moorhead, P. S. & Kaplan, M. M. eds) pp. 35–40. Philadelphia: Wistar Institute Press.

Hamilton, W. D. (1970). Ordering the phenomena of ecology. *Science* **167,** 1478–1480.

Hartley, R. V. L. (1928). Transmission of information. *Bell. Syst. Tech. J.* **7,** 617.

Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Trans. A.I.E.E.* **47,** 617–644.

O'Donald, P. (1969). "Haldane's dilemma" and the rate of natural selection. *Nature, Lond.* **221,** 815–816.

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

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

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* **27,** 379–423, 623–656.

Wallace, A. R. (1916). Letter of July 2, 1866 to Charles Darwin. In: *Alfred Russel Wallace: Letters and Reminescences* (Marchant, J., ed.) Vol. 1, p. 170. London: Cassell.