COMMENTARY

Diverse forms of selection in evolution and computer science

Nicholas H. Barton[1], Sebastian Novak, and Tiago Paixão
Institute of Science and Technology Austria (IST Austria), 3400 Klosterneuburg, Austria

The astonishing diversity of complex adaptations that we see in the living world has been produced by natural selection, over ~3.5 billion years of evolution. Information from the largely random survival and reproduction of past organisms has accumulated to produce genomes that are precisely fitted to diverse environments. In PNAS, Chastain et al. (1) show that natural selection on freely recombining populations is equivalent to the multiplicative weights update algorithm (MWUA), an efficient optimization algorithm that has been discovered many times in computer science, statistics, and economics. Whether this equivalence explains the extraordinary effectiveness of natural selection, or conversely, of artificial algorithms, depends on one's perspective. Perhaps surprisingly, theoretical results in population genetics and in computer science look quite different, even when they deal with essentially the same questions. Thus, the equivalence identified by Chastain et al. (1) allows these different results to be transferred between fields in both directions.

In each generation of selection, the frequency of each type is simply multiplied by its relative fitness, which corresponds precisely to MWUA. Imagine a panel of financial experts whose advice performs in an arbitrary way over time. In each step, the probability of choosing one or other expert is updated in proportion to their performance. Remarkably, provided that the updates are gradual, the ultimate performance of the algorithm is guaranteed to be close to that of the expert with the best total performance, or equivalently, to that of the allele with the highest total fitness, summed over generations. This can be seen as a justification for the Darwinian view that gradual evolution is most effective and as quantifying the performance of natural selection in a randomly fluctuating environment.

The MWUA can be represented by a simple maximization principle (Fig. 1). Probabilities are chosen to maximize the sum of two quantities: the expected total fitness of the chosen allele, summed over past generations, plus the entropy, which is a measure of the allelic diversity. Over time, differences between the total fitness of the different alleles almost always increase (Fig. 1B), and so the population converges toward the type with the best total performance; the entropy term delays this convergence, preventing immediate fixation of whichever type is currently best.

This maximization principle contrasts with one more familiar in population genetics (2, 3). In each generation, selection maximizes the gain in current fitness, subject to a constraint on the genetic distance moved, defined as $\sum_j [x_i^{t+1}(j) - x_i^t(j)]^2 / x_i^t(j)$, where $x_i^t(j)$ is the frequency of the jth type ("allele") at the ith gene, at time t. The net increase in mean fitness due to selection equals the additive variance in fitness (4), so that populations move toward peaks in mean fitness. Indeed, there are many ways to represent selection. Changes in allele frequencies and in the means of quantitative traits change in proportion to the gradient of mean fitness (2) and to the covariance with fitness (5, 6). Rather than thinking of the proportions of individuals in the current population, we can think of the proportions of ancestral lineages; Leibler and Kussell (7) have shown that these are selected according to their total historical fitness (as in Chastain et al.'s representation of the MWUA) and that the rate of increase of the mean historical fitness with respect to the strength of selection equals the variance of historical fitness. When random sampling drift is included, we see a yet wider variety of representations of selection—diffusion of allele frequencies through time (8, 9), probabilities of paths of allele frequency through time (10, 11), or distributions of genealogies in the ancestral selection graph (12). As in physics, where classical mechanics can be described either by Newton's laws or by the principle of least action, even the simplest systems can be understood in several mathematically equivalent ways.

The discovery of abundant molecular variation in the 1960s, confirmed by recent genomics, was surprising, and led to fierce argument over whether this diversity was maintained by selection (13). Similarly, it has long been known that most complex traits are highly heritable, due to many alleles of small effect, yet we still do not know whether this variation is primarily due to mutation or is positively maintained by selection.
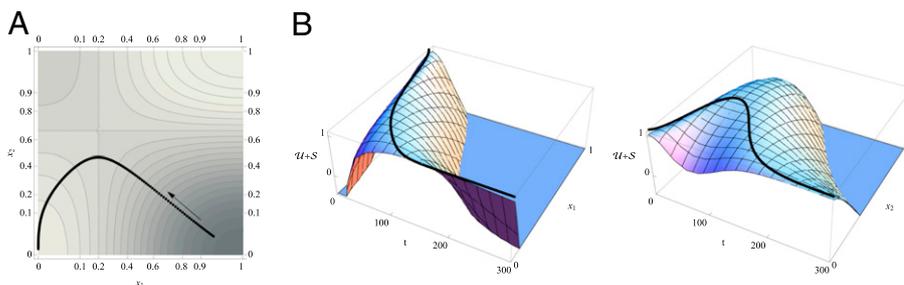
Fig. 1. (A) Allele frequencies change at a rate proportional to the gradient in mean fitness; on the arcsin-transformed scale used here ($x = \sin[\theta/2]^2$) (8), the increase in mean fitness is maximized, given the distance moved. In this example, fitnesses of genotypes {0,0}, {0,1}, {1,0}, and {1,1} are 1, 0.97, 0.90, and 1.02, respectively. Thus, there are two stable equilibria, corresponding to peaks in the fitness landscape, with either {0,0} or {1,1} fixed. Black dots show a population that starts at {0.95, 0.02} (*Lower Right*), and evolves toward fixation of the {0,0} alleles over 300 generations. Contours of mean fitness are spaced at intervals 0.005. (B) In any generation, allele frequencies (black curve) maximize the sum of the cumulative fitness, $\mathcal{U} = \sum_j x_i^t(j) U_i^t(j)$ and the entropy, $\mathcal{S} = -\sum_j x_i^t(j) \log[x_i^t(j)]$. As time proceeds, differences in $\mathcal{U}$ becomes larger, relative to the constant entropy, $\mathcal{S}$, thus forcing allele frequencies toward fixation. This figure shows trajectories for the two genes.

Chastain et al. (1) suggest that the representation of selection as (partially) maximizing entropy may help us understand how selection maintains diversity. However, it is widely believed that selection on haploids (the relevant case here) cannot maintain a stable polymorphic equilibrium. There seems to be no formal proof of this in the population genetic literature, where most effort has been devoted to the diploid case, when heterozygote advantage can maintain diversity (14). It is straightforward to prove that when haploids have constant fitnesses, selection does not allow stable polymorphism; correspondingly, coordination games with fixed payoff do not allow mixed strategies. However, whether time-varying selection can maintain diversity, when genotype fitnesses do not depend on frequency, remains an open problem. It is hard to see that the maximization principle identified by Chastain et al. (1) helps with this, because it extends to the diploid and frequency-dependent cases, as well as to haploids: although it determines the dynamics, it does not show whether trajectories are stable.

It has long been clear intuitively, in both evolutionary biology and computer science, that life without sex is difficult: asexual species do not persist, nonrecombining chromosomes degenerate, and in evolutionary computation, some form of recombination is almost always used in successful algorithms (15, 16). However, it remains unclear just how selection on individuals maintains high rates of recombination, at least in most eukaryotes. Chastain et al. (1) suggest that the tradeoff between increasing cumulative fitness and entropy helps us understand the role of sex in evolution. However, this is not so clear, because the MWUA is defined for a single "player," and so applies equally well to an asexual population. Moreover, Chastain et al.'s (1) analysis is entirely deterministic, and it is known that in the absence of random drift (i.e., in infinitely large populations), recombination is only favored if there are

Chastain et al. show that natural selection on freely recombining populations is equivalent to the multiplicative weights update algorithm.

special kinds of interaction between genes (15). Recombination is much more widely favored in finite populations, because it allows the reconstruction of combinations of genes that have been lost by chance; however, results on the deterministic MWUA do not directly address this issue, which is essentially stochastic.

Although computer science and evolutionary biology appear to be very different fields, there are some surprisingly close parallels, which are apparent if one thinks of natural selection as providing a general algorithm by which populations can efficiently learn about their environment (16–19). Although the questions in each field differ, there are common themes—not least, how rapidly can evolutionary algorithms act? In computer science terms, what is the complexity of an evolutionary algorithm, or in other words, how does the runtime scale with the dimensionality of the environment, genome size, population size, and so on? It is striking that, although these different fields deal with similar problems, their theoretical structures are quite different, giving an opportunity for fruitful transfer. Identification of the close analogy between MWUA and population genetics is a first step in this process.

1 Chastani E, Livnat A, Papadimitriou C, Vazirani U (2014) Algorithms, games, and evolution. *Proc Natl Acad Sci USA*, 10.1073/pnas.1406556111.
2 Wright S (1937) The distribution of gene frequencies in populations. *Proc Natl Acad Sci USA* 23(6):307–320.
3 Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30(2):314–334.
4 Fisher RA (1930) *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, UK).
5 Robertson A (1966) A mathematical model of the culling process in dairy cattle. *Anim Prod* 8(1):95–108.
6 Price GR (1970) Selection and covariance. *Nature* 227(5257):520–521.
7 Leibler S, Kussell E (2010) Individual histories and selection in heterogeneous populations. *Proc Natl Acad Sci USA* 107(29):13183–13188.
8 Fisher RA (1922) On the dominance ratio. *Proc R Soc Edinb* 42(1921-1922):321–341.
9 Kimura M (1955) Stochastic processes and distribution of gene frequencies under natural selection. *Cold Spring Harbor Symp Quant Biol* 20(1955):33–55.
10 Rouhani S, Barton NH (1987) Speciation and the 'shifting balance' in a continuous population. *Theor Pop Biol* 31(3):465–92.
11 Mustonen V, Lassig M (2010) Fitness flux and ubiquity of adaptive evolution. *Proc Natl Acad Sci USA* 107(9):4248–4253.

12 Krone SM, Neuhauser C (1997) Ancestral processes with selection. *Theor Popul Biol* 51(3):210–237.
13 Kimura M (1984) *The Neutral Theory of Molecular Evolution* (Cambridge Univ Press, Cambridge, UK).
14 Christiansen FB (1999) *Population Genetics of Multiple Loci* (Wiley, Chichester, UK).
15 Barton NH, Charlesworth B (1998) Why sex and recombination? *Science* 281(5385):1986–1990.
16 Watson RA (2006) *Compositional Evolution: The Impact of Sex, Symbiosis and Modularity on the Gradualist Framework of Evolution* (MIT Press, Boston).
17 Valiant L (2013) *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World* (Basic Books, New York).
18 Watkins C (2008) Selective breeding analysed as a communication channel: channel capacity as a fundamental limit on adaptive complexity. *Symbolic and Numeric Algorithms for Scientific Computing* (IEEE, New York), pp 514–518.
19 Lehre PK, Oliveto PS (2013) Runtime analysis of evolutionary algorithms: Basic introduction. *Proceeding of the Fifteenth Annual Conference on Genetic and Evolutionary Computation*, July 6–20, 2013, Amsterdam (Association for Computing Machinery, New York), pp 469–498.