Chapter 1

Critical states of fitness landscapes

Yukihiko Toquenaga
Institute of Biological Sciences
University of Tsukuba
toque@biol.tsukuba.ac.jp

The traffic cone hypothesis on evolving fitness landscapes is reconsidered here to incorporate aspects of self-organized criticality. This modification enables the model to explain coexistence of various kinds of genetic operations at multiple selection levels as well as conventional views of the distribution of mutational effects in evolutionary biology.

1.1 Introduction

A fitness landscape is a metaphor that casts evolutionary trajectories of phenotypes, genotypes, and gene frequencies of evolving populations. This has been extensively used in studies of evolutionary biology, yet there are several criticisms against it.

The most popular criticism is that a fitness landscape cannot be defined if interaction dominates selection regimes. This point of view has its root in the tautology argument against utility of the phrase, "survival of the fittest." This criticism, however, can be overcome by defining fitness landscapes as resultant snapshots rather than causes of evolution. Nevertheless, such a landscape can play a role of boundary condition for further evolutionary processes.

Another criticism is that there is no clear-cut definition of coordinates for projecting fitness values [14]. Fitness landscapes are mountainous images on which individuals can occupy locations, and means of a population travel as they evolve. People who discuss optimal evolutionary strategies often assign

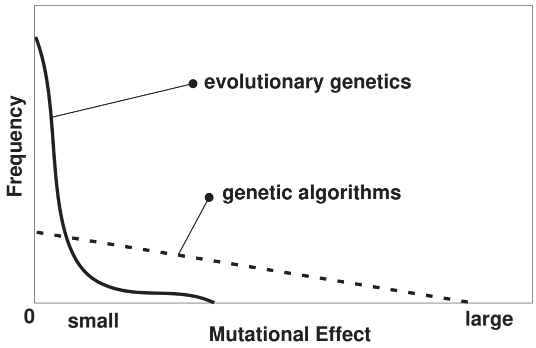

Figure 1.1: Mutational Effects. X- and y-axes represent absolute intensity of mutational effects and corresponding frequencies, respectively. In evolutionary biology, most mutational effects are believed to be neutral, and hence the distribution has its mode at the zero mutational effect. In genetic algorithms used in artificial life, a single mutation can have large as well as small effects. If the fitness values are randomly assigned, there is more instances with small mutational effects than large effects for a single mutation. The expected distribution of mutational effects decreases monotonically from small to large effects as shown in the dotted line in the figure.

phenotype to basal axes for plotting fitness values. Some geneticists use genotype, and others use allele frequency for the same purpose. Some evolutionary biologists even use a hyper cube of boolean network [6, 10].

Evolutionary trajectories on the landscape should be continuous. Otherwise, the mountainous landscapes are meaningless. In that sense, the basal axes should be additive. This prerequisite is often dismissed in the above theoretical arguments except for the boolean network approach. There are no specific guidelines for selecting appropriate basal axes for a fitness landscape.

In spite of this potential obscurity, the fitness landscape is a useful metaphor for evolutionary biology, and Wade and I proposed a hypothesis of evolution of fitness landscapes[16]. In the next section, I explain this hypothesis, called the "traffic cone hypothesis"(THC). In the following sections, I point out a weakness of the THC and propose a modification that incorporates the idea of self organized criticality. In the final section, I proposed the direction of future studies.

1.2 The Traffic Cone Hypothesis

This hypothesis was developed in 1996 while I was a visiting scholar with Michael J. Wade at the University of Chicago. We discussed the discrepancy between conventional assumptions of evolutionary biology and those being applied in artificial life studies.

The most prominent discrepancy between the two disciplines regards the

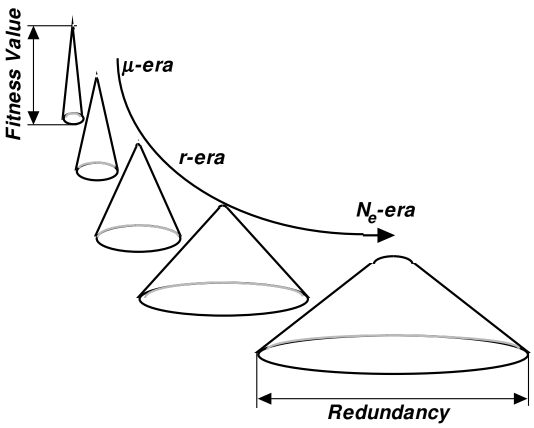

Figure 1.2: Traffic cone hypothesis. μ, *r, and* N_e *stand for mutation, recombination, and inter-demic selection, respectively.*

frequency distribution of mutational effects. In evolutionary biology, most mutations are assumed to be neutral or slightly deleterious. In contrast, artificial life simulations often assume that mutations of large effect are relatively common (Fig.1.1). Nevertheless the two disciplines share a common goal: to study evolution of life systems[15]. Based on the comparison between the two, we developed the following general scenario of biological evolution (the TCH) in terms of the fitness landscape [16];

At the very beginning of life, organisms faced very rugged fitness landscapes where point mutations in asexual reproduction frequently moved organisms or populations from one hill to another in a genotypic space (the mutation era). Lives begin adaptive walks, and seek successful ways of reproducing themselves. Most successful lives are those who can reproduce, or produce copies of themselves. Their reproductive way should be robust against drastic mutations. One way to achieve this robustness is to acquire genetic redundancy that maps several different genotypes to a single phenotype, or a fitness value. At the same time, the reproductive mechanism should keep a way for continuous changing not to be trapped in a certain domain for ever.

As time goes by the landscape becomes locally correlated by acquisition of genetic redundancy, or nonlinear relationships in mapping genotypes to phenotypes. Each domain of attraction in the genotypic space (i.e. hills) enlarges, and some smaller ones are engulfed by larger ones. The hight of merged peaks is shorten because of genetic trade-off among pooled genotypes. At the same time, the shape of peak becomes dull. This dynamical change of the fitness landscape causes the distance between domain centers to increase, and shorten the hight of the peaks on average. At this stage, a more powerful form of mutation, sexual recombination, is invoked to permit continued evolutionary exploration of the

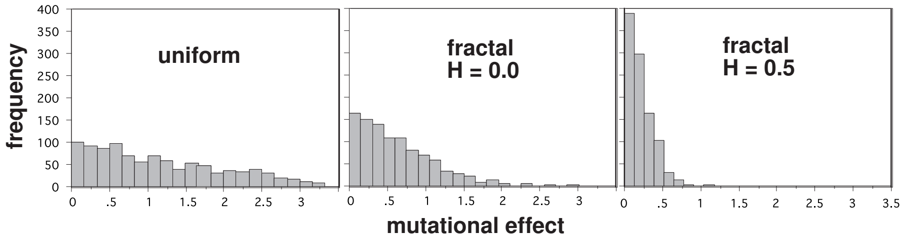

Figure 1.3: Mutational effects in uniform (upper most) and fractal (lower two) fitness landscapes. The fractal landscapes are generated by the midpoint displacement method [13]. The parameter H (Hurst exponent) describes roughness of the landscape at small scales.

fitness landscape (the recombination era).

Further evolution of fitness landscapes calls for more dynamic exchange of genotype configurations by group (inter-demic) selection described in the shifting balance theory[18]. At this stage each domain represents a species where a variety of genotypes have similar and suboptimal fitness values, and mutations often have little or no effect on fitness (the inter-demic selection era). Figure 1.2 schematically illustrates the whole scenario.

1.3 Power Balance of Genetic Operation at Different Selection Levels

The above scenario bridges the traditional model of evolutionary biology (most mutations are neutral) and assumptions of genetic algorithms in artificial life studies (any degree of effect can be achieved by a single mutation with equal probability). Resultant fitness landscapes consist of huge domains (very fat, short, and dull traffic cones) in the genotypic space at the late evolutionary phase (Fig.1.2). Low fitness domains can be easily invaded by higher ones. It remains unclear how the fatness of domains of attraction in the landscape would be maintained. There should be some mechanisms to prevent reverse evolution to the rugged fitness landscapes.

Another problem in evolutionary biology is that the highest level process (i.e. inter-demic selection) does not overwhelm the two at lower levels. Inter-demic selection is often thought to play a subordinate role in evolutionary biology [5]. Point mutation and recombination can explain much wider ranges of evolutionary phenomena without invoking inter-demic selection. The long-lasting controversy between the Fisherian and Wrightian schools of thought [17] is based on these distinctions. The TCH cannot explain such a power balance among the three kinds of genetic operation at the later evolutionary phases.

1.4 TCH at a Critical State

How can one achieve both genetic redundancy and maintenance of genetic operation at the lower selection levels? One possible explanation is that the fitness landscape might be at a self-organized critical state (SOC; [4, 2, 9]) rather than a simple cluster of smoothed and flattened cones.

Single-mutational effects can be defined as the absolute difference between the fitness value of a focal point on the landscape and a nearby point that can be achieved by single mutation from the focal point. If the fitness landscape is maximally rugged (i.e., a uniform random distribution of fitness across the landscape [10]), the frequency distribution of mutational effects becomes a monotonic decreasing function of the effect as shown in the left panel in Fig.1.3.

If the fitness landscape is fractal, the frequency distribution of mutational effects piles up at lower mutational effects (see the center and the right panels in Fig.1.3). This distribution corresponds to the prevalence of neutral mutations, which is widely accepted in evolutionary biology (see Fig.1.1).

Moreover, the fractal surface of landscape ensures that three different genetic operations (mutation, recombination, and inter-demic selection) can have the same or similar effects against fitness values. There is no domination of a genetic operation over the others. This is believed to be the essence of the shifting balance theory of Sewall Wright [18, 14].

1.5 Future Direction

It is very difficult to detect critical states in real biological data. Log-log plots are not enough to represent criticality, let alone self-organized criticality [9]. I would recommend two approaches to test the new TCH featuring SOC.

1.5.1 Strong Artificial Life

Strong Artificial Life (ALife) seeks to create and study theoretical representations of life forms apart from existing living things on the earth. One can examine snap-shots of fitness landscapes of such artificial creatures. However, the Bak and Sneppen Model [3] is hardly acceptable to biologists as an evolutionary model.

Alternatives are Avida [1], and other ALife systems. Lenski et al. [12] is a good example of collaboration between ALifers and evolutionary biologists. They address the evolution of the mutational effects (see Figure 2 in [12] and compare it with Fig.1.1 and Fig1.3). Evolutionary systems featured by the simple genetic algorithm might be the good starting points.

1.5.2 Mining in Genome Database

Data from the rapidly growing genomic databases are best suited for testing the modified TCH. Ubiquitous observations of fractal base sequences in real organisms may relate to the critical state [8, 11]. Prevailing duplication and palindrome structures in base sequences are believed to have some functions (e.g., folding and efficient replication) for the current as well as past usage. Introns rather than exons might be important targets if one searches for fossil records of the RNA world [7].

Acknowledgements

I am most grateful to Guy Hoelzer for his valuable and helpful comments on the manuscript.

Bibliography

[1] Adami, Chris, *Introduction to Artificial Life*, Springer-Verlag (1998).

[2] Bak, Per, *How nature works*, Copernicus (1996).

[3] Bak, Per, and K. Sneppen, "Punctuated equilibrium and criticality in a simple model of evolution", *Physical Review Letters* **24** (1993), 4083–4086.

[4] Bak, P., C. Tang, and K. Wiesenfeld, "Self-organized criticality: An explanation of 1/f noise", *Physical Review Letters* **59** (1987), 381–384.

[5] Coyne, J. A., N. H. Barton, and M. Turelli, "A critique of Sewall Wrightś shifting balance theory of evolution", *Evolution* **51** (1997), 643–671.

[6] Eigen, Manfred, *Steps towards life*, Oxford UP (1992).

[7] Forsdyke, D. R., "A stem-loop kissing for the initiation of recombination and the origin of introns", *Molecular Biology and Evolution* **12** (1995), 949–958.

[8] Jeffrey, H. Joel, "Chaos game representaion of gene structre", *Nucleic Acids Research* **18** (1990), 2163–2170.

[9] Jensen, Henrik J., *Self-organized criticality*, Cambridge UP (1998).

[10] Kauffman, Stuart A., *The origin of order*, Oxford UP (1993).

[11] Korolev, A. V., V. V. Solov'ev, and V. G. Tumanyan, "New method in the overall search for the functional regions of DNA using the fractal representaiton of nucleotide texts", *Biophysics* **37** (1992), 733–742.

[12] Lenski, Richard E., Charles Ofria, Travis C. Collier, and Christopher Adami, "Genome complexity, robustness and genetic interactions in digital organisms", *Nature* **400** (1999), 661–664.

[13] PEITGEN, Heinz-Otto, and Dietmar SAUPE, *The science of fractal images*, Springer-Verlag (1988).

[14] PROVINE, William B., *Sewall Wright and evolutinary biology*, Chicago UP (1986).

[15] TOQUENAGA, Y., and M. J. WADE, "Sewall Wright meets Artificial Life", *TREE* **11** (1996), 478–482.

[16] TOQUENAGA, Y., and M. J. WADE, "Evolution of fitness landscape", *Proceedings of the 3rd International Symposium on Artificial Life and Robotics, Vol. 1* (A. 3RD ed.), AROB (1998), 658–661.

[17] WADE, M. J., and C. J. GOODNIGHT, "The theories of Fisher and Wright in the context of metapopulations: When nature does many small experiments", *Evolution* **52** (1998), 1537–1553.

[18] WRIGHT, Sewall, *Evolution and genetics of populations, Vol. 4*, University of Chicago Press (1978).