Contingency and Memory in Evolution

High Dimensionality and Sparse Occupation Create History in Complex Systems

A universal, simple, and powerful principle for the erasure of history has been discovered in physics and consists in the approach to a stationary state or thermodynamic equilibrium. After stationarity has been achieved, any memory of the past is annulled. This fact can be visualized by a toy model considering trajectories on binary trees. As a matter of fact, the phenomenon is the same, whether we are considering a continuous or a discrete space for the embedding of the trajectories. For the purpose of illustration, however, the discrete space is much better suited and modeling by discrete steps is also closer to the biological application that will be considered later. Individual trajectories merge at confluences like rivers or streams do. We assume a single walker, who started somewhere in space and migrates step by step toward the equilibrium point, no external recording is taken that allows for a reconstruction of the past and as the walker passes confluences of trajectories one after the other, the closer his particular route comes to other trajectories which started at different points in space, and the smaller is the information that the position P_k carries. The last step toward equilibrium eventually erases all information on the past. The situation is illustrated in the Figure 1 by means of simple binary trees. Shannon's formula for the information content of a uniform probability distribution of n objects, $H = \sum_i x_i \ln x_i = \ln n$, is applied to compute the loss of the information in a walk progressing inwards from some point in the outmost shell, P_n, to a point P_k:

$$\Delta H(P_k) = (\ln n - \ln k)/\ln 2 [\text{bits}], \text{ with } k = 1, 2^1, 2^2 \ldots, n$$

The variable k represents the number of initial points from which trajectories lead to P_k. For binary tress, this number is given by: $k = n/2^\gamma$, where γ is the index of the shell around the equilibrium point. In other words at the beginning we need n bits to describe the position in shell $\gamma = \ln n/\ln 2$ and every step toward equilibrium erases one bit. As information is lost, retrieval of the walk is impossible per definition.

The inverse problem is to find an answer for the question how history is created in a system and appears to be more involved. Inverting the approach to a uniquely defined final state in our toy example is illustrative. The process starts at some given point on the binary tree, which could but need not be the equilibrium point (as assumed in the Fig. 1), and the walker migrates outwards passing one branching point after the other. A series of (binary) decisions determines the tra-

PETER SCHUSTER

Peter Schuster, Editor in Chief of Complexity, is at the Institut für Theortische Chemie der Universität Wien, A-1090 Wien, Austria. (email: pks@tbi.univie.ac.at)

FIGURE 1
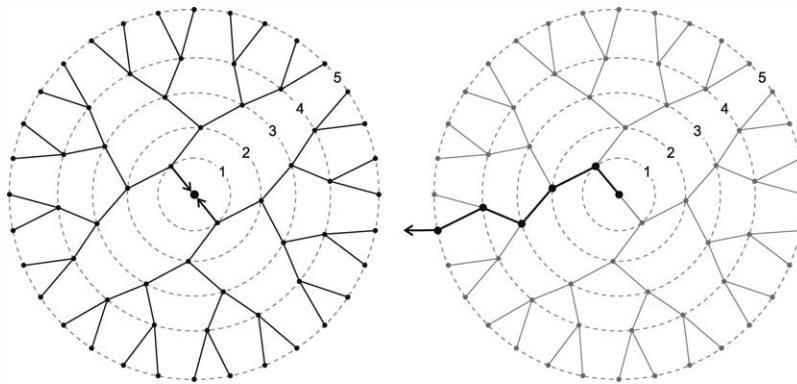
Memory-erasing and memory-creating walks on binary trees. The binary tree on the left hand side illustrates the approach toward a point attractor, equilibrium or asymptotically stable stationary state. Shells around equilibrium are numbered by γ = 0, 1, 2, 3, …. The number of points in a shell is 2^γ and hence γ bits of information are required to locate a point. The approach toward equilibrium erases the information of the location by one bit for every shell. On the right hand side, an outward going trajectory is shown on the same binary tree. Each step going outward involves a binary decision and increases the information content by one bit. The tree reconstruction problem in molecular evolution starts from the information stored in—as many as possible—sequences in the current, i.e., present day shell and retrieves the past by systematic sequence comparison.

jectory and constitutes thereby the history of the walk. For unbiased decisions, the series of (random) branching events will inevitably lead to a completely different trajectory even when the experiment is repeated under identical initial conditions. Applying Shannon's formula, each decision of the walker creates information and increases H by one bit for each step, $\Delta H(P_k) = \ln k / \ln 2$ [bits], where the initial condition $n = 1$ was applied. The history of the process may be stored by some external observer or the walker himself may carry a memory that allows for a reconstruction of the path he took. So far the diverging process creates history, but for nontrivial contingency, the choices at individual bifurcation points must introduce recognizable differences for the forthcoming walks. The simplest example is the introduction of a preference for certain decisions. For example, one might prefer to go left in the upper-

left part of the binary tree and to choose preferentially the right branch in the lower-right half. Then, the first decision—up or down—has an influence on all further decisions which, therefore, are dependent on the first step. History matters and a simple bias introduces contingency into the system.

The binary-tree toy model illustrates the basic causes for history and contingency, but it differs from reality in many aspects. The four most important points for the forthcoming discussion are: (i) The spaces in which the trajectories are embedded are high-dimensional (instead of the two dimensions suggested by the Figure), (ii) we are not dealing with a single walker but with an ensemble of walkers, (iii) the decisions to be taken commonly have more than two options, and (iv) the graph of the embedding space is not a tree but a network. High-dimensionality of spaces counteracts intuition as mentioned al-

ready by Sewall Wright, who is considered the "father" of fitness landscapes as a tool to visualize evolution [1]. Distances are generally smaller in high-dimensional spaces. For the purpose of illustration, we consider the largest distance in an n-dimensional array of N points. For $N = 1000$ points in one dimension, i.e. $n = 1$ and all points are situated on a line, the largest distance is $d_{max} = 999$, two-dimensional space with $n = 2$ and $N = 1000$ yields $d_{max} = 44.72$, and for $n = 3$ we find $d_{max} = 17.32$. The space, in which the evolution of genomes takes place, is commonly called sequence space:[1] Every sequence is represented by a point, and the distance between two points is the Hamming distance[1] between the corresponding two sequences [2, 3]. The sequence space of binary sequences is a hypercube of dimension n and consists of a network of closed loops rather than a tree, but for sufficiently long sequence lengths n, the sequence space is so large that a local approximation by a tree becomes meaningful. This condition is almost always fulfilled in nature where we are dealing with genome lengths of several thousand for viruses, several million for bacteria, and several billion for higher organisms.

Almost everything in nature carries some kind of memory on its past, but the underlying principle of storage is often not evident. Organisms are an exception because genetic information is always stored digitally in a DNA molecule that represents an excellent memory of the phylogenetic past. Here, the creation of contingency by mutation and selection in the sense of Charles Darwin will be considered.

---

[1]*The notion of sequence space has been introduced by Manfred Eigen for nucleic acid sequences [2] and by John Maynard Smith for proteins [3]. The appropriate metric for point mutations as source of variation is the Hamming distance. It counts the positions in which two aligned sequences differ.*

The currently available highly efficient techniques for sequencing DNA allowed for harvesting extensive information on genetic sequences and put the reconstruction of phylogenies from molecular data through comparison of present day genomes on a solid basis. In essence, the tree of life is a binary tree differing from the one shown in the figure by variable lengths of the edges–edges can bridge several shells. The current view that the majority of mutations is either deleterious or has no effect on selection at all—and hence, advantageous mutations are rare—is supported by a wealth of sequence data from all biological kingdoms. Research in molecular evolution led to the concept of a molecular clock, which states that the number of mutations recorded in the genome is approximately proportional to genome lengths and time [4, 5]. Molecular phylogeny reconstruction became the most important tool for timing and revealing history in biology.

Considering a "population of walkers" rather than a single individual brings the model close to evolution. Required is the assumption that walkers multiply along their trajectories and transmit their position on the tree to the progeny, which continues the walk of the parent. Some points on the tree—representing some genomes—provide more resources than others and progeny accumulates there. Variation is introduced at the branching points and selection is a result of finite population size—to compensate the increase in population size by birth, individuals have to die. A population has memory since walkers at less favorable points stay in the population for some generations before they are eliminated.

Migrating outwards in space faces "combinatorial explosion" of points—2^γ on a binary tree—and sooner or later populations cannot cover all accessible points. The size of populations is highly variable in nature—as illustrated

for mammals as an example where it ranges from several hundred for certain endangered species to several billion for men. Symbiotic bacterial colonies typically consist of several hundred thousand to 10^{14} cells, the latter observed, for example, in human guts. Even 10^{14} appears tiny in relation to the values of the exponential function at moderate chain lengths: $2^\gamma = 10^{14}$, for example, yields only $\gamma = 46.51$. Computer simulations have shown that in absence of differences in the properties relevant for selection, colonies that are unable to cover sequence space migrate and split into smaller clones [6–8]. If computer experiments are started several times under identical conditions,[2] the population migrates in different—almost orthogonal—directions in sequence space [9]. Clones created by replication and mutation form quasispecies consisting of a fittest genotype and a surrounding mutant distribution as predicted by the kinetic theory of evolution [2, 10] and later verified in virus populations [11]. The quasispecies concept gave rise to completely new insights into virus evolution [12] which, in essence, confirmed the concepts derived from computer simulations.

The origin of contingency in evolution has been recently demonstrated by an elegant experiment with *Escherichia coli* [13, 14]. In February 1988, Richard Lenski from Michigan State University in East Lansing started a long-time experiment with *Escherichia coli* [12]. Bacterial evolution is studied under controlled constant conditions giving rise to about six to seven generations or duplications per day and thus lead to about 45,000 generations in 20 years. The bacteria grow in a

minimal medium containing nutrients and citrate buffer to keep the pH constant. Every day a small sample is transferred to fresh medium, isolates are taken at regular intervals after every 100 generations and stored at a low temperature of $-80°$C ($-112°$F). Twelve subpopulations were prepared from a single ancestral clone; the subpopulations evolved independently and adapted to the new constant environment in the medium within the first few thousand generations [15]. After 31,500 generations, a dramatic change in growth was observed in one (and only one) of the 12 subpopulations, the "c-population," that manifested itself in a much higher concentration of bacteria and increased turbidity of the solution [11]. What had happened was reconstructed by careful investigations and biochemical analysis: The "c-population" had developed a membrane channel for the uptake of citrate into the cell.[3] The innovation created an enormous advantage as the buffer medium citrate was now available as a new nutrient and substantially more growth was the answer of the successful subpopulation. The next obvious question was whether the citrate channel arose by a single very rare mutation—and hence was not repeatable within reasonable time spans—or was a result of the historical development of the successful population. The answer to this question came from previous isolates that were used to initiate evolution of new strains. The innovation of a citrate channel was indeed reproducible with earlier—but not too early—isolates from the "c-population" but did never occur

---

[2]*In this context identical conditions means that everything was kept the same in the simulations except the seeds for the random number generator, which determine the sequence of random events.*

---

[3]*Under aerobic conditions, normal Escherichia coli bacteria have no means for the transport of citrate from the medium into the cell and hence cannot life on a medium containing only citrate when exposed to air. Rare special forms are known, which in contrast can incorporate citrate.*

in any of the other 11 subpopulations. Several explanations for the action of past mutations facilitating a mutation producing the citrate gate are possible: (i) epistatic effects favoring the citrate gate precursor gene, (ii) the preparation of a precursor by insertion of a sequence by mobile elements in the *Escherichia coli* genome, and (iii) migration of the population by neutral evolution into some region in sequence space from where the citrate gate sequence becomes accessible via an ordinary mutation. Deep sequencing of the "c-population" and the other subpopulation as planned by Richard Lenski and his group will very likely provide the answer. No matter which interpretation will eventually turn out to be correct, Richard Lenski's experiment with *Escherichia coli* has shown three important features:

i. The occurrence of contingency does not require different environmental influences as the evolutionary mechanism alone may create history by mutation and selection and as it was discussed already in the toy model.

ii. Rare events need not be rare per se since the history of a population may contribute to conditions that allow for a great innovation by means of an ordinary single mutation. This phenomenon has been characteristically named the "jackpot effect" [16].

iii. A memory carried by one particular population allowed for reproduction of events within a certain historical time span. If one went back too far on the way to the common ancestor of all 12 subpopulations, no advantage was left for the production of a successful mutant.

The Lenski experiment provides also the first hint from a laboratory experiment to answer the often raised question: "What would happen if the tape were played twice?" [17]. Most likely something else, as we have experimental evidence that contingency occurs under fairly simple conditions and within reasonable time spans. Differences between subpopulations and diversity will be created in repeated evolution events and accumulate as the trajectories diverge in sequence space. In short time intervals, as we have seen, for example, by the reproduction of the mutation toward the citrate channel, the second run of the tape would result in essentially the same or very similar phenotypic improvements. The genotypes giving rise to the same phenotypes, however, can be assumed to be different. In the long run also the phenotypes diverge as demonstrated by the fact that none of the other 11 subpopulations was able to produce a favorable mutation leading to citrate uptake within the time spans of the experiments. Hence, repeating evolution would give different results at least in all details.

REFERENCES

1. Wright, S. The surfaces of selective value revisited. Am Nat 1988, 131, 115–131.
2. Eigen, M. Self-organization of matter and the evolution of biological macromolecules. Naturwissenschaften 1971, 58, 465–523.
3. Maynard Smith, J. Natural selection and the concept of a protein space. Nature 1970, 225, 563–564.
4. Morgan, G.J. Emile Zuckerkandl, Linus Pauling, and the molecular evolutionary clock, 1959-1965. J Hist Biol 1998, 31, 155–178.
5. Kumar, S. Molecular clocks: Four decades of evolution. Nat Rev Genet 2005, 6, 654–662.
6. Derrida, B.; Peliti, L. Evolution in a flat fitness landscape. Bull Math Biol 1990, 53, 355–382.
7. Huynen, M.; Stadler, P.F.; Fontana, W. Smoothness within ruggedness: The role of neutrality in adaptation. Proc Natl Acad Sci USA 1996, 93, 397–401.
8. Fontana, W.; Schuster, P. Continuity in evolution. On the nature of transitions. Science 1998, 280, 1451–1455.
9. Fontana, W.; Schnabl, W.; Schuster, P. Physical aspects of evolutionary optimization and adaptation. Phys Rev A 1989, 40, 3301–3321.
10. Eigen, M.; Schuster, P. The hypercycle. A principle of natural self-organization. Part A. Emergence of the hypercycle. Naturwissenschaften 1977, 64, 541–565.
11. Domingo, E.; Sabo, D.; Taniguchi, T.; Weissmann, C. Nucleotide-sequence heterogeneity of an RNA phage population. Cell 1978, 13, 735–744.
12. Domingo, E.; Parrish, C.R.; Holland, J.J., Eds. Origin and Evolution of Viruses, 2nd ed.; Elsevier, Amsterdam, NL, 2008.
13. Blount, Z.D.; Borland, C.Z.; Lenski, R.E. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proc Natl Acad Sci USA 2008, 105, 7899–7906.
14. Lenski, R.E.; Rose, M.R.; Simpson, S. C.; Tadler, S.C. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am Nat 1991, 138, 1315–1341.
15. Elena, S.F.; Cooper, V.S.; Lenski, R.E. Punctuated evolution caused by selection of rare beneficial mutants. Science 1996, 272, 1802–1804.
16. Luria, S.E.; Delbrück, M. Mutations of bacteria from virus sensitivity to virus resistance. Genetics 1943, 102/103, 127–144.
17. Fontana, W.; Buss, L.W. What would be conserved if "the tape were played twice"? Proc Natl Acad Sci USA 1994, 91, 757–761.