

# The evolutionary ecology of technological innovations

Ricard V. Solé<sup>\*</sup>, <sup>1,2,3</sup> Sergi Valverde, <sup>1,2</sup> Martí Rosas Casals, <sup>4,1</sup> Stuart Kauffman, <sup>3</sup> Doyne Farmer, <sup>3</sup> and Niles Eldredge<sup>5</sup>

<sup>1</sup>*ICREA-Complex Systems Lab, Universitat Pompeu Fabra (GRIB), Dr Aiguader 80, 08003 Barcelona, Spain*

<sup>2</sup>*Institut de Biología Evolutiva, CSIC-UPF, Passeig Marítim de la Barceloneta, 37, 08003 Barcelona, Spain*

<sup>3</sup>*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe NM 87501, USA*

<sup>4</sup>*Sustainability Modeling and Measurement Lab, Universitat Politcnica de Catalunya, EET-Campus Terrassa, 08222, Barcelona, Spain*

<sup>5</sup>*Division of Paleontology, American Museum of Natural History, New York 10024, USA*

Technological evolution has been compared to biological evolution by many authors over the last two centuries. As a parallel experiment of innovation involving economic, historical and social components, artifacts define a universe of evolving properties that displays episodes of diversification and extinction. Here we critically review previous work comparing the two types of evolution. Like biological evolution, technological evolution is driven by descent with variation and selection, and includes tinkering, convergence and contingency. At the same time there are essential differences that make the two types of evolution quite distinct. Major distinctions are illustrated by current specific examples, including the evolution of cornets and the historical dynamics of information technologies. Due to their fast and rich development, the later provide a unique opportunity to study technological evolution at all scales with unprecedented resolution. Despite the presence of patterns suggesting convergent trends between man-made systems and biological ones, they provide examples of planned design that have no equivalent with natural evolution.

**Keywords:** Technology, evolution, tinkering, convergence, information technology, culturomics.

## I. INTRODUCTION

We can look at the history of technology as a human-driven, parallel experiment of evolution [1-4]. So far artifacts are not capable of self-reproduction, but the population-level dynamics of long-term technological innovation nonetheless resemble biological evolution in many ways. The design of new technologies is strongly influenced by existing technologies, and technological change can be viewed as a process of descent with variation and selection [5-7]. Both chance and the appropriate context are required for innovations to occur. Lineages of design often show rapid change and diversification as well as exaptation. The latter is illustrated by Gutenberg's printing press, when an existing technology (the screw press) that was co-opted to serve a completely novel purpose. Extinction and replacement are also common. As soon as a genuinely novel invention appears, it is typically followed by an enormous diversification, followed by the extinction (and turnover) of most competing inventions. Moreover, technological change also displays convergence: similar discoveries are made simultaneously by different inventors [4,9] such as the more than 20 different patents involving light bulb inventions prior to Edison's success [10]. The view that technological evolution follows similar rules to biological evolution has captured the interest of scientists, historians and engineers alike.

Despite the commonalities, technological evolution de-

parts from biological evolution in fundamental ways. For technological change long-term goals and expectations play a leading role in which the designers seek optimality, typically under explicit criteria such as efficiency, cost and speed. Moreover, as pointed by François Jacob [11-14], in contrast to artifacts, living structures are largely the result of tinkering, i.e. a widespread reuse and combination of available elements to build new structures. Technology is highly dependent on the combination of preexisting [2,4] inventions, but unlike biology, the introduction of new simple elements can completely reset the path of future technologies. In contrast, in biology, once established, solutions to problems are seldom replaced.

Both biological and technological innovations involve cost constraints. Thermodynamic efficiency can also help understanding the origin of some structures. Allometric scaling laws provide a good illustration of how a theory of biological distribution networks (including both vascular and respiratory ones) based on efficient energy dissipation on fractal trees [15]. Efficiency has also been driving technological improvements and marks the development of the steam engine and the bicycle [16]. The evolution of the latter can be traced as a succession of improvement steps towards increasing performance and lower metabolic cost. However, the coupling between energy costs and improvements is not a precondition for technological change to occur. On one hand, many examples illustrate a common pattern of development of a given invention: in early stages, inventions are often overly expensive and not perceived as economically relevant. The barrier to diffusion can only be overcome through the vision of individuals pursuing their views

---

\*corresponding author

and goals [4].

Is it possible to formulate a theory of technological evolution? How much can we take advantage from our theoretical understanding of biological evolution? Recent advances within network theory and a unique availability of the fossil record of human inventions might help in reaching that goal. Such theory needs to consider the existence of universal trends, the economic context [2,6,17-18] and history. We believe that a major effort in this direction would settle the debate on similarities versus differences.

## II. LOCK-IN AND CONTINGENCY

A fingerprint of biological evolution is the existence of frozen accidents in the spread and stabilization of a given trait. Similarly, economic constraints can enhance technological lock-in [2,6]. The QWERTY keyboard is an archetypal example, where a suboptimal configuration of keys was adopted due to a conflict with typing speed. Even if rational evaluation of a technology and its alternatives reveals that the dominant technological solution is suboptimal, issues of compatibility often make it impossible for better solutions to enter [2,3,7,19]. Video-cassette recorders, aircrafts or clocks all provide examples where choices were strongly affected by history. This can be explained theoretically by considering the positive feedback associated to the imitation dynamics characteristic of technological adoption of novelties [21]. A simple model illustrates the path-dependent nature of choices made in a system where preferential decisions are based of majorities.

Let us consider a population where users of a given technology can choose between two possibilities  $A$  and  $B$ . Imagine that we represent individuals choosing  $A$  and  $B$  as black and white squares on a lattice. At the beginning, both alternatives are equally chosen and we place  $A$  (black) and  $B$  (white) agents at random over the lattice. Assuming that agents follow the majority rule, a given agent will change its current state, say  $A \rightarrow B$ , if the majority of its neighbors is adopting the  $B$  state. For convenience, we have projected our simulations on a sphere. As a consequence of the contagion process, domains of agents adopting the same choice emerge and expand in time. If we start from a random lattice with  $A$  and  $B$  agents, small fluctuations are amplified towards either all- $A$  or all- $B$  equilibria (bottom spheres).

The lattice model can be easily formalized in mathematical terms. If  $a$  and  $b$  indicate the relative fraction of individuals choosing each strategy. Encounters among individuals can be described in a mean field approximation by using a simple mathematical model. Let  $a$  and  $b$  define the populations of agents adopting  $A$  and  $B$ , respectively. Agents using  $A$  ( $B$ ) will adopt technology  $B$  ( $A$ ) provided that they perceive that some minimal number of individuals are using the alternative. The dif-

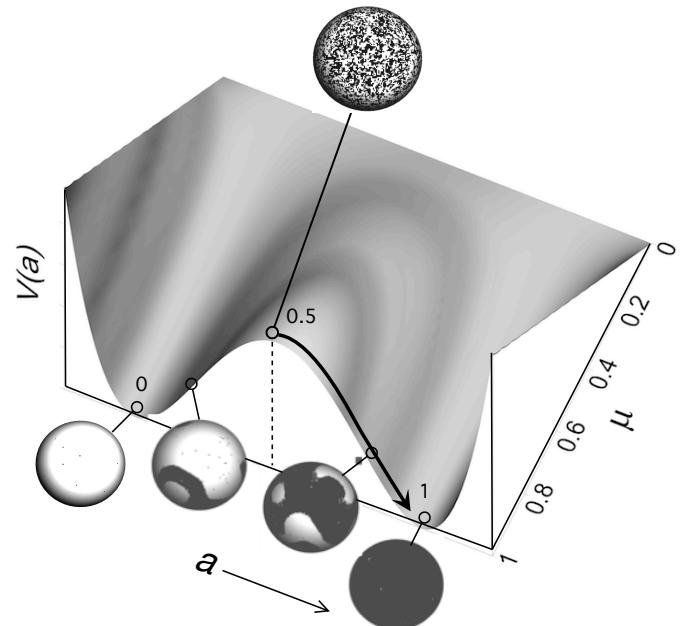


FIG. 1 Choosing between two given, initially equivalent choices can be represented by means of a two-well function. The top of this curve defines an unstable state. A marble located on this top will rapidly roll down towards one of the two possible valleys. Such choice is highly dependent on the small fluctuations happening at the beginning of the process. As a consequence, choices are highly path dependent: the original symmetry is broken.

ferential equation for  $a$  reads:

$$\frac{da}{dt} = \mu a(\eta - b) - a\phi(a, b) \quad (1)$$

The right-hand side of the equation described two terms. The first corresponds to an infection process: with an efficiency  $\mu$  and provided that the number of agents adopting  $B$  is below some threshold  $\eta$ . This can be interpreted as follows: there is a minimal number  $\eta$  of agents using  $B$  that define a critical mass able to promote the shift  $A \rightarrow B$ . Similarly, an equation for  $b$  is obtained as follows:  $db/dt = \mu b(\eta - a) - b\phi(a, b)$ . If a constant population constraint ( $a + b = 1$ ) is assumed, it is easy to show that  $\phi(a, b) = \mu(\eta(a + b) - 2ab)$ . The model becomes one-dimensional and we can write an equation for the time evolution of  $a$ :

$$\frac{da}{dt} = F_\mu(a) = 2\mu a(1 - a) \left( a - \frac{1}{2} \right) \quad (2)$$

The term  $a(1 - a)$  is just a logistic function (with saturation when  $a = 1$ ). Equation (4) actually corresponds to the population dynamics of one locus and two identical alleles due to selection alone [20]. The last term in the right-hand side captures the majority rule. If  $a < 1/2$ , the  $A$  choice is in a minority situation and a decrease in  $A$  use will occur. Instead, when a majority

has adopted A (namely  $a > 1/2$ ) growth will occur. The previous equation has three equilibrium points, obtained from  $da/dt = 0$ . These are  $a_0^* = 0$ ,  $a_1^* = 1$  and  $a_2^* = 1/2$ . Their stability is obtained by the sign of

$$\lambda_\mu(a_i^*) = \left( \frac{\partial F_\mu(a)}{\partial a} \right)_{a=a_i^*} \quad (3)$$

The first two correspond to winner-takes-all stable points, where A has failed to be adopted ( $a_0^*$ ) and has invaded the whole population. These are stable points, while the third one, associated to the coexistence between both technologies, is unstable.

A useful analog of the path-dependent nature of this model is obtained by computing the associated potential  $V_\mu(a)$  for the previous model [21]. The potential is linked to the dynamical equation through

$$\frac{da}{dt} = -\frac{\partial}{\partial a} \left[ - \int F_\mu(a) da \right] \quad (4)$$

where the integral within the parenthesis is the potential function. As defined, it will exhibit a maximum (minimum) when  $a^*$  is unstable (stable). For our example, this gives:

$$V_\mu(a) = -\frac{\mu}{2} (a^4 - 2a^3 + a^2) \quad (5)$$

The curve  $V_\mu(a)$  is displayed in figure 1. The two wells separated by a maximum are the accessible states and the symmetric, unstable one (since  $0 \leq a \leq 1$ , the values outside this domain are meaningless). The spheres shown as insets are lattice simulations at different time steps (the spherical projection is used for convenience). Here black and white represent A and B states. As we can see, the control parameter  $\mu$  only affects the depth of the alternative valleys but not their location. However, smaller  $\mu$  values will be more likely to be affected by stochastic fluctuations and thus the symmetry breaking process might be more difficult (or even impossible) to occur.

### III. THE GENOTYPE-PHENOTYPE MAPPING: TECHNOLOGY MEETS (SOMETIMES) BIOLOGY

The presence of sub-optimal choices, as illustrated in the previous section, reminds us that nonlinear dynamics can dominate over rational decisions. This effectively limits the ideal picture of man-made artifacts as optimal, purposeful designs. However, in some well known cases, optimal solutions are achieved and maintained. This is illustrated by the evolution of the fork [22]. A sequence of steps knife → two-tines → three-tines → four tines fork took place slowly and the final design has not changed since it was adopted in the seventeenth century. Four tines provide a good compromise between an object that performs its functions properly: It provides a broad surface, it fits in the mouth and can hold both large and

small pieces of food. A fitness function  $\Phi(N)$  could be measured using a population of users playing with different designs. Here  $N_t$  is the number of tines and we would like to know how  $N_t$  relates to  $\Phi(N_t)$ . We can imagine a simple experiment where different people use (repeatedly) different forks to different conditions and provide some score for each design. The average over different users provides an estimate for the  $N_t \rightarrow \Phi(N_t)$  mapping. In this simple example, a single-peak would be present at  $N_t = 4$ . If we add the cost of fabrication to the fitness (so that performance is computed over the number of tines) the peak would be sharper.

Cost constraints can strongly limit the potential solutions associated to a given problem. As a consequence, optimal designs sometimes lead to remarkable matching between biological and man-made solutions. As an example, the same laws associated to wiring very large scale integrated (VLSI) circuits under strong packing constraints [23] are followed by brains [24]. Convergence in design principles has also been reported from the study of brain circuits in the visual cortex, which obey the same basic architecture of parallel computer vision [25]. The close relationship between integrated circuits and neural systems is provided by the so called Rent's rule, which defines a power law in networks that exhibit hierarchical modularity. Assuming that we partition the system into sub-systems of size  $N$ , the rule establishes that the number of connections  $C$  linking the subset with the rest of the system scales as

$$C = \langle k \rangle N^p \quad (6)$$

where  $\langle k \rangle$  and  $p$  are characteristic parameters. Here  $\langle k \rangle$  gives the average number of links per node, whereas  $0 \leq \alpha \leq 1$  is the so called *Rent's exponent*. The latter is typically limited to the interval  $\alpha \in [0.45, 0.75]$ . Given the importance of the following result, it is worth to show how is this scaling law derived. Imagine a given system  $\Gamma$  (a network) involving  $N_0$  elements (gates in a circuit, neurons or groups of neurons in a brain) and consider a given sub-system of size  $1 \leq N \leq N_0$ . These components are connected forming a network. We indicate by  $\langle k \rangle$  the average number of links per node and we take a subset  $\Gamma_k(N)$  of elements within a given area, including  $N$  components. This subset will be connected to the rest of the system  $\Gamma - \Gamma_k$  by means of a number of links crossing its boundaries. Let us also indicate as  $C(N)$  the number of links connecting  $\Gamma_k$  with the rest of the system. If we add additional elements to  $\Gamma_k$ , expanding it in size, new links should be expected to connect the larger subset with the rest of the network. Indicating by  $\Delta C$  and  $\Delta N$  the increase in connectivity and size associated to the expansion, we can expect a relationship:

$$\Delta C = p(N) \left( \frac{C}{N} \right) \Delta N \quad (7)$$

The parameter  $p \in (0, 1)$  indicates a potential dependency between size and the degree of optimization in-

volved in the circuit design. VLSI designers use a number of strategies to reduce  $C$  and in particular a modular, hierarchical system is specially designed. For a random network where any element can connect to any other one with the same probability, we would observe  $p = 1$ . The last equation can be solved by using a continuous approximation in terms of a differential equation, which gives:

$$C(N) = \langle k \rangle \exp \left[ \int_1^N \frac{p(N)}{N} dN \right] \quad (8)$$

assuming that  $p(N) = p$ , we obtain the standard formulation of Rent's rule. In most known systems, this constant assumption seems the correct one.

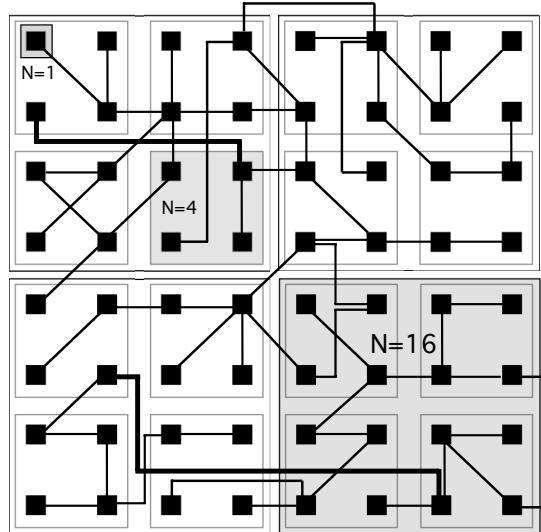


FIG. 2 Hierarchical modularity. In both circuits and neural systems, which experience cost constraints, a hierarchy of nested modules has been shown to exist, displaying some seemingly universal regularities, such as Rent's rule. This rule establishes that the amount of connections  $C$  between elements in a sub-system of size  $N$  (here three scales are shown) with the rest of the system scales as  $C = \langle k \rangle N^p$ .

The fact that the exponent is homogeneous tells us that, despite the scale change associated to increasing  $N$ , the relative fraction of links crossing boundaries scales with system's size in a constant way. This is characteristic of fractal objects (where the basic pattern is repeated at different scales) and thus an indication of the presence of hierarchical order. When this method was applied to neural networks, a striking convergence was found. Both the neural web of the nematode *C. elegans* and human cortical maps shared a Rent's exponent  $p$  close to what is the expected value for an optimally efficient hierarchical system. Such convergence, that shares many properties in common with VLSI circuits, illustrates the role played by cost constraints in promoting convergent designs.

A single peak defines the simplest landscape for optimization. But the presence of lock-in reminds us that fitness landscapes are, in general, much more complicated.

The ruggedness of the landscape is connected to the nature of the genotype-phenotype mapping [26-29]. The complexity of this mapping pervades the path-dependent character of evolution, how robust designs are and how evolvable they can be. It is in general far from trivial to know the landscape structure that pervades a given complex system. One exception is provided by those systems where a discrete sequence space and a well defined phenotype space can be defined and scanned. A success story in this context is the analysis of the genotype-phenotype mapping in both RNA and protein networks (refs). Here a mapping can be defined between a set of sequences  $\mathcal{H}(\mathbf{S})$  (nucleotide chains) and the set corresponding forms  $\mathcal{H}(\mathbf{F})$  resulting from the folding process:

$$\Omega : \mathcal{H}(\mathbf{S}) \longrightarrow \mathcal{H}(\mathbf{F}) \quad (9)$$

$$S_i \longrightarrow F_i \quad (10)$$

It was shown that, starting from a given sequence  $S_i \in \mathbf{S}$  it is possible to move through sequence space (figure 3a) through neutral mutations sharing the same phenotype  $F_i$  (a common fold). Sequence space is enormously vast. If the alphabet has size  $A$  and sequences have length  $L$  we have  $A^L$  possible sequences, an already hyperastronomic number. By introducing single mutations starting from a given sequence, it is possible to see if it has an impact on the shape (function) or instead is neutral. A key finding was that we can travel through this vast combinatorial space by spreading through very large neutral networks of sequences folding in the same shape (i. e. the same phenotype). In the simplified example shown here, three sequences are indicated. Two of them ( $S_1$  and  $S_2$ ) span finite neutral networks, whereas  $S_3$  percolates through sequence space.

The unexpected comes when we compare the previous results with the corresponding mapping associated to a seemingly disparate object: electronic circuits. These are now defined as sets of logic gates connected through a given wiring scheme [30]. The circuit performs a function defining the phenotype: for the given set of inputs (a fixed number) there is a specific output that defines the function performed. In other words, the wiring diagram is now the genotype and the Boolean function implemented the phenotype. The full list of outputs describes the Boolean function  $\Phi_i$ , i. e. the phenotype  $P \equiv \Phi_i$ . The genotype is defined as the ordered string of all weights  $w_{ij}$ . Two different sets,  $\Phi$  and  $\mathbf{W}$  describe the universe of possible functions and wirings i. e. the sets of possible phenotypes and genotypes. The set of (genotype-phenotype) mappings between wiring and function is defined as:

$$\Omega : \mathcal{H}(\mathbf{W}) \longrightarrow \mathcal{H}(\Phi) \quad (11)$$

where  $\mathcal{H}(\mathbf{W})$  is the space of all possible FFNs and  $\mathcal{H}(\Phi)$  the corresponding space of all Boolean functions. For each  $W_i \in \mathcal{H}(\mathbf{W})$  we have a genotype which maps into a given function described by a given  $\Phi_i \in \mathcal{H}(\Phi)$ .

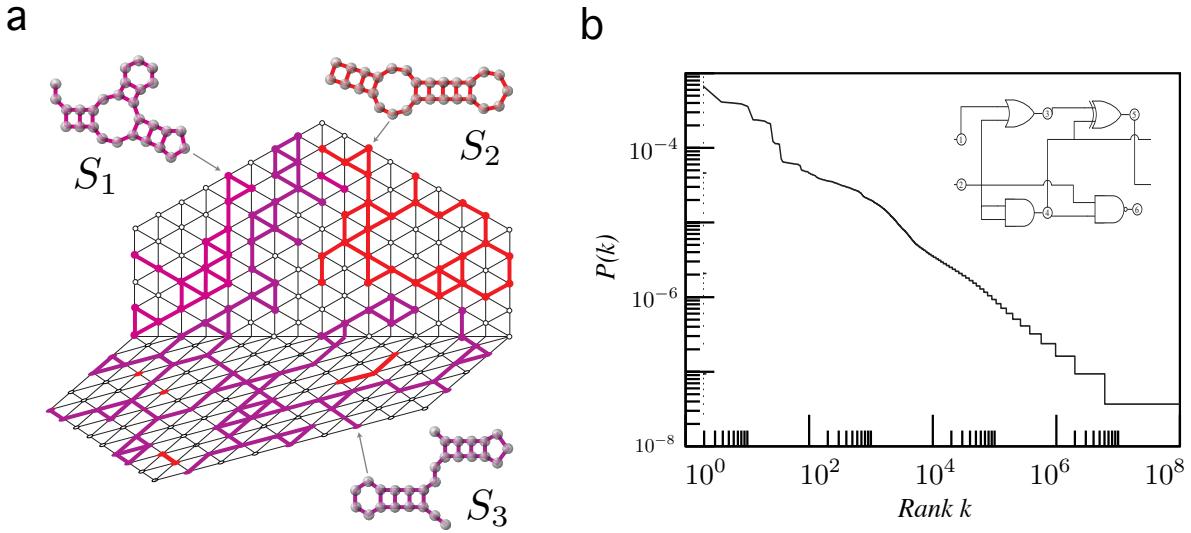


FIG. 3 Neutral landscapes are common in a number of biological systems. They have been shown to drive the evolution of RNA and proteins through sequence space (pulir). In (a) we display an idealized picture of a small piece of an RNA landscape (redrawn from Schuster 2001). Here each node represents a given sequence whereas links among nodes are possible paths connecting single-mutation events. Three different neutral networks are represented, associated to three different RNA folds. These landscapes are also present in technological systems. The analysis of field programmable gate arrays or FPGAs in short (Raman and Wagner 2010) has shown that a similar space can be defined, with given arrangements of gates as nodes and connections indicating possible transitions from one circuit to another (see text). In (b) a rank plot distribution of these circuits is shown, using a small array (inset, modified from Raman and Wagner 2010).

When the landscape associated to circuits performing computations [31,32] was analyzed, a very similar pattern was found. Here the genotype is identified with the network of elements and their connections, whereas the computation being performed describes the function. By "mutating" the wiring diagram of a given circuit (altering single connections) it can be shown that neutral networks are also widely present, the boundaries of which are occupied by circuits able to implement a great variety of novel functions. A common statistical regularity shared by both RNA and FPGA landscapes is the broad distribution of abundances, shown in figure 3b. Here we display the probability of finding a given circuit, against its rank (see [32] for details). The  $P(k)$  rank distribution provides the relative frequency of the  $k$ -th most common circuit. The distribution is highly skewed, and follows a law, the Zipf's law, namely:

$$p(k) = \frac{k^{-\gamma}}{\sum_{j=1}^n j^{-\gamma}} \quad (12)$$

with  $\gamma \approx 1$ . This is a highly skewed distribution, with a few sequences (or circuits) being highly common whereas most are rare.

A major consequence of remarkably common pattern is that key structural properties of the organization of man-made designs can have an unexpected potential for evolvability [32]. What these results do indicate [30-32] is that design principles in engineering could benefit from

understanding evolution on landscapes.

#### IV. DIVERSIFICATION, EXTINCTION AND TECHNOLOGICAL REGIMES

The rise of innovations often leads to both the decline of pre-existing solutions and the creation of new ones. An example is provided by the evolution of cornets [37,38] which diversified over a century until valve trumpets replaced them, after a radiation in styles and usages, somewhat similar to biotic replacements. Similarly, the invention of the car marked the decline and fall of carriages, harness shops and buggies, while favoring the expansion of the oil industry. Paved roads, gas stations or motels came later. Although a given innovation occurs in a well defined economic context, novel technologies can deeply transform how the economy is organized and how new economic regimes emerge. Events like the rise of electronics and personal computers represent a turning point from which a new *technological regime* is born [39-42]. Such regimes can be described in terms of the grammar or rule-set embedded in the way a given technology must be used. They include many facets such as product standards and design and engineering practices. The resetting of a regime and its replacement by a new one share some similarities with recovery from mass extinction [2-3,8].

Properly tracing the rise and fall of technologies is not always possible. As always in science, the availability of data is an important determinant of what we know, and the data for biological vs. technological evolution are very different. For many technologies we have good anecdotal evidence on performance and prevalence through time. Data on performance can take many forms. For example, for information technologies we have reasonably good historical data for properties such as processing speed, information storage, or transmission, both in terms of cost and physical characteristics [43-45]. There is no corresponding historical performance data in biology – we can only guess at the strength or running speed of dinosaurs.

In contrast, in biology we have a well-developed phylogeny and rich data on the history of phyla with no technological equivalent. This is not an accident – as discussed later, the phylogeny of technologies is not hierarchical, but rather is more similar to that of bacteria. While many studies have been done for specific families of technologies, there are no comprehensive databases on technological change. As a result, the data we have and the inferences that we can make about biological vs. technological change are very different. Two major sources of quantitative analysis help filling this gap.

## V. TEMPORAL TRENDS AND LEARNING CURVES

One of the biggest differences between technology and biology is that for technologies we have clear notions of progress. For example, in 1936 Theodore Wright conjectured that the unit cost of a given model of airplane drops as a well-defined power law in the cumulative number produced. This “law” has since been observed, at varying degrees of fidelity, for a large number of other technologies [43-48], and the general pattern goes under the name of an experience or learning curve. An alternative hypothesis about the rate of technological improvement is due to Gordon Moore, who proposed that the density of transistors for integrated circuits increases exponentially with a doubling time of about two years [49]. The rule of exponential change also holds quite well for many other aspects of information technology. An approximate exponential relationship between unit cost and other measures as a function of time has recently been shown to hold for many other technologies as well [43,45].

Moore’s law and Wright’s law are closely related: If production increases exponentially and unit cost drops exponentially, then Wright’s law holds. Scaling laws are commonly observed between cost  $C$  and production  $z$ , i.e.

$$C(z) \sim z^{-\alpha} \quad (13)$$

where the scaling exponent  $\alpha$  weights the ratio of improvement. Some general trends displayed by learning curves are consistent with expected trends displayed by

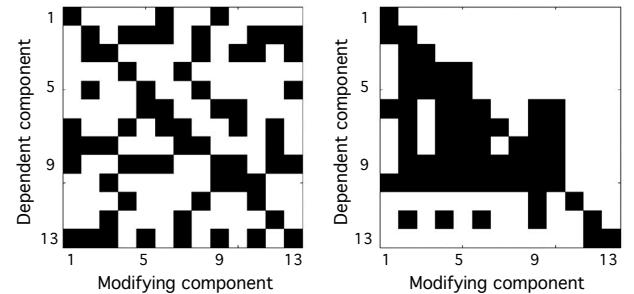


FIG. 4 Example design structure matrices (DSMs) with  $n = 13$  components. Black squares represent links. The DSM on the left was randomly generated to have fixed out-degree for each component. The DSM on the right represents the design of an automobile brake system (31). All diagonal elements are present because a component always affects its own cost.

populations climbing rugged fitness landscapes [47,50-53]. In this context, *design structure matrices* (DSMs) offer a promising link between technological evolution and its underlying fitness landscape [54,55]. As illustrated in figure 4, DSM are described by an  $n \times n$  matrix  $M = (M_{ij})$  with  $M_{ij} \neq 0$  if the component  $i$  has some cost impact on component  $j$  (and is zero otherwise). They capture the internal constraints derived from the interaction of different components, providing a representation of the genotype-phenotype mapping and how it is affected by epistatic interactions among units and how they affect scaling laws.

A simple example illustrates the approach followed in [55]. Imagine an idealized scenario where a given technology has a cost  $c(t)$  which can change over time as new innovations occur. For simplicity, we consider a finite range of values, to be normalized with uniform distribution  $f(c)$  on  $[0, 1]$ . In other words we have

$$\int_0^1 f(c)dc = 1 \quad (14)$$

Cost will be reduced if chance events allow a new design to be used at a smaller cost  $c$ . It can be shown that the equation describing the change of cost follows:

$$\frac{dc}{dt} = -\frac{\langle c \rangle}{2} P[c' < c] \quad (15)$$

$$= -\frac{\langle c \rangle}{2} \int_0^c f(c)dc \quad (16)$$

$$= -\frac{1}{2} \langle c \rangle^2 \quad (17)$$

which can be easily solved, giving a power law decay:

$$c(t) = \frac{c(0)}{1 + c(0)t/2} \approx \frac{1}{t} \quad (18)$$

where  $c(0)$  indicates initial cost and the last term is the approximation for long times. As we can see, this simple

model provides a correct prediction for the time decay of cost in technological evolution. More detailed models can be built by considering the internal structure of the technological artifact and how different parts affect others performance [55]. By using DSMs, an effective mapping between the network of hardware interactions and the function being performed is defined. Because of this, the DSM captures pleiotropic effects and how they influence the rates of improvement and the existence of internal limitations. It has been shown [55] that the matrix connectivity pattern (the density of links or the presence of modularity) largely determines the speed of improvement, which can be continuous or punctuated.

## VI. INNOVATION NETWORKS

The assumption that combinatorics is at the heart of innovation provides a possible rationale for the accelerating growth of innovations. If pairs of components can be combined with some probability, the number of new designs will increase with the number of already present designs. This picture has led to the suggestion [56] that technological improvement leads inevitably to singularities, with the number of inventions exploding on a finite time scale.

The idea that innovations happen mainly by combination [2] predicts (assuming unlimited resources) that the number of inventions will grow in time following a hyperbolic curve [33,34]. Specifically, if  $x$  indicates the number of inventions and we assume that combining previous designs creates new inventions, the time evolution of  $x$  would follow, in its simplest form, the equation

$$\frac{dx}{dt} = \mu(t)x^2 \quad (19)$$

which says that a reaction between two given inventions is needed to create new ones. Here  $\mu(t)$  is the rate at which two inventions are used (by an inventor or group of inventors) to generate new ones. In general, this would be a time-dependent parameter, since it might depend on the number of inventors, economic factors or aging phenomena. The model can be easily solved, leading to

$$x(t) = \frac{x(0)}{1 - x(0) \int_0^t \mu(\tau) d\tau} \quad (20)$$

where  $x(0)$  would be the initial number of inventions. For the simplest scenario where  $\mu$  can be considered constant, a singularity exists at a finite time

$$\tau = \frac{1}{\mu x(0)} \quad (21)$$

when  $x$  becomes infinite.

One consequence of such pattern is that a system where combination is easy to happen will rapidly explore the *adjacent possible* [4,35] defining the next layer of potential designs obtainable through combination of existing ones.

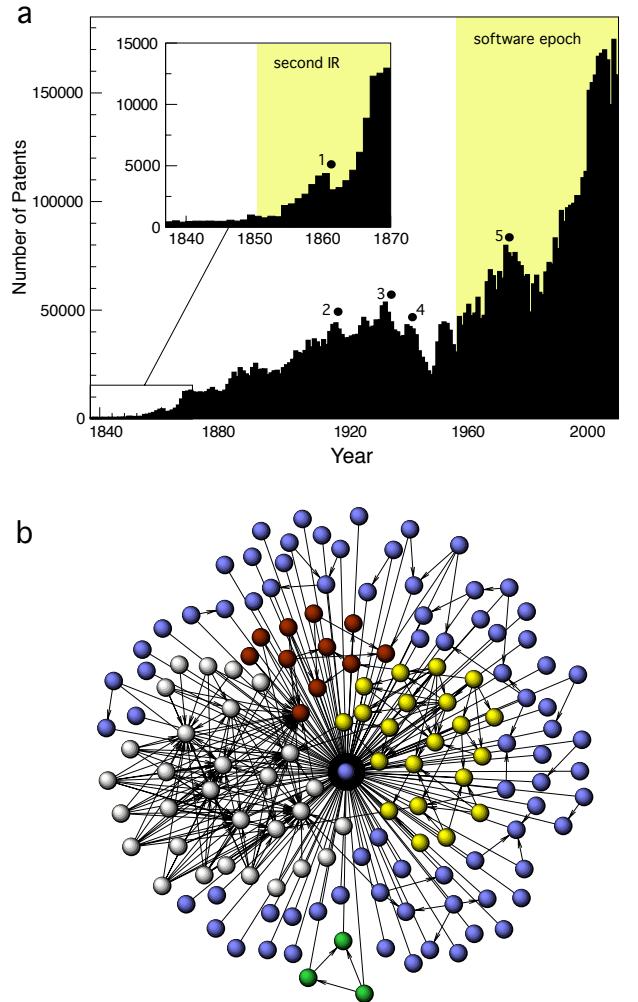


FIG. 5 Patent citations and innovation by combination. In (a) we display the time series of USA patents from 1837 to 2009. Drops in citation records are often associated to crashes and wars 1-4) but also due to legislation (5). A closer picture is obtained by looking at the actual pattern of cross-citations among patents. This defines a network (b) where nodes are filed patents and the (directed) links are citations among them. The network shown here corresponds to early inventions in tomography (see [59]) and colors indicate topological modules, i. e. subsets of patents that share a larger number of citations among themselves compared with the rest of the dataset. Topological modules are usually related to functional subsets with specialized traits.

Due to the immense potential for combination, the diversity of potential novel inventions could make the outcome essentially unpredictable [35]. However, even when very large search spaces need to be explored, solutions are nevertheless found. This is illustrated by drug search in chemical space [36] of which only an extremely small fraction is screened. This suggests that the combinatorial potential might be strongly reduced due to the underlying syntax constraining what objects can be linked together.

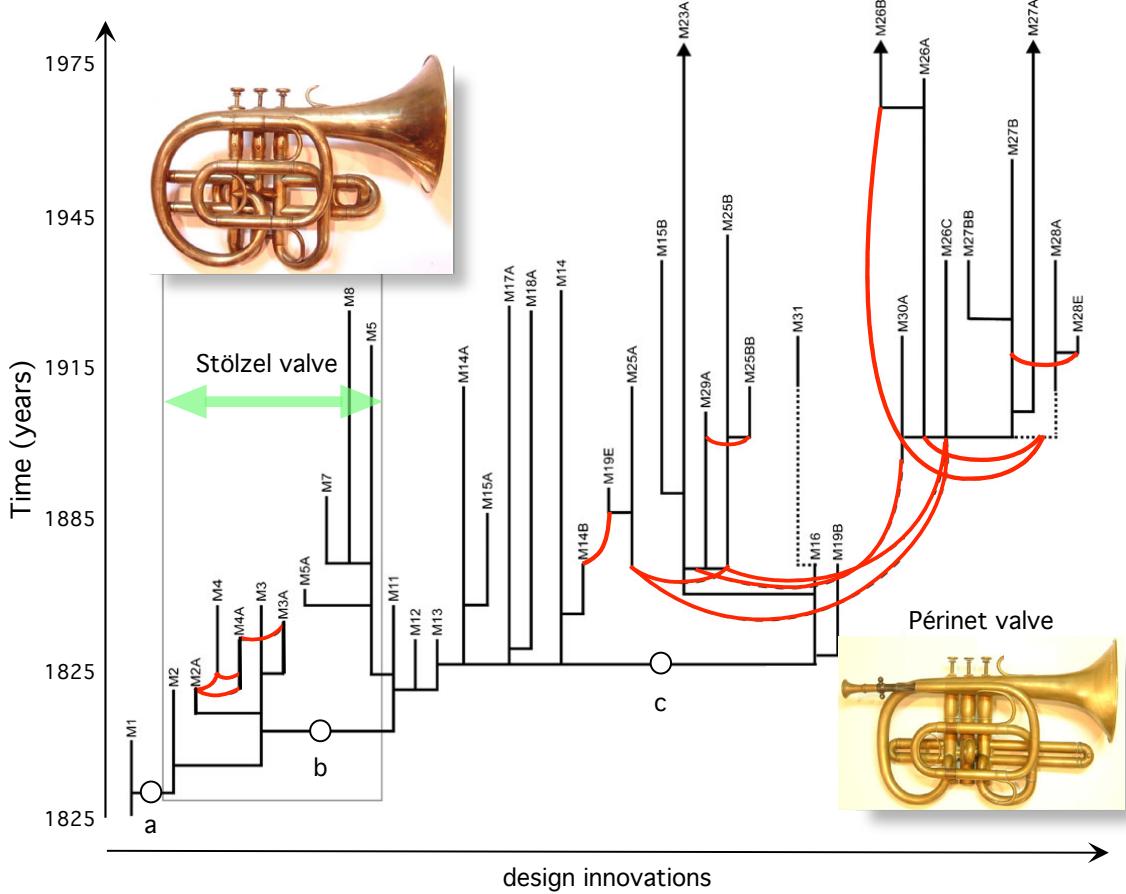


FIG. 6 Lateral transfer muddies phylogeny in material cultural evolution, as illustrated here by the evolutionary tree of cornets. Vertical lines represent the periods of manufacture of distinct cornet models, here plotted against time. Red arcs indicate so called *reticulations*, i. e. potential non-vertical instances of information transfer among makers. Here two main groups are indicated, corresponding to the so called Stölzel (left rectangle) and Périnet (unmarked right domain) cornets (Redrawn after [37]).

Despite the appeal of the previous picture, dedicated efforts using different measurements and databases suggest a more complex scenario [43]. Let us take for example the number of patents (figure 1a) over time. The inset shows a transition from the industrial revolution to the so called second industrial revolution and characterized by the rise of mass production and the rapid growth of chemical, electrical and automotive industries and the availability of energy resources. Similarly, the global record of citations has a broad accelerated trend of growth, particularly obvious in the last decades, strongly influenced by the rise of software (colored area) which started in the 1960s and rapidly accelerated during the 1980s. As a written description of innovations, patents provide a surrogate of technological dynamics [57-59]. The rapid growth in patent diversity have been compared to Phanerozoic diversity time series [3] but once again the similarity is only partially significant. During periods of crisis or war their numbers fall. This is shown in figure

5a, where we indicate the starting point of only five of them using black dots. Here we pointed: (1) American Civil War, (2) USA enter World War I, (3) great depression and (4) USA enters World War II. However, a different type of external trigger can also provoke a decay in patenting, as shown by (5) which indicates the point where a much more strict intellectual property legislation and higher costs for patenting rights. Such top-down control has obviously no biological counterpart.

Patent inventories give us a kinetic view of innovation, but much more can be uncovered by looking at how patents are related to each other. Since a given patent must cite others from which the novel technology gets inspiration, a network linking a given patent with past ones can be defined. The number of outward links starting from a patent relates it to previous inventions. Similarly, the number of incoming links provides a rough estimate of its importance. Using available data sets, a network of patent citations can be constructed [59] (figure 5b).

Each node represents a patent and arrows indicate that a citation is made to a previous invention. The central node in this case is the oldest invention. Colors indicate modules: groups of patents that have more connections among them than with others within the network. These modules correlate with groups of related inventions, suggesting that there is considerable structure in the way new technologies emerge.

Patent networks are highly heterogeneous, with some hubs involving many citations, whereas most nodes have only one or two [59]. This pattern can be accounted for using a rich-gets-richer mechanism favoring highly connected nodes (the popular patents) [59-62]. By following the time evolution of citations, it is actually possible to detect two major components affecting the relative importance of an invention. The first is the rapid growth of citations often associated with the combinatorial potential of the invention. The second is an aging process: as patents get older, they tend to be less cited. This reflects the changing landscape of technology and the inevitable replacement of old by new designs.

Can phylogenetic analyses help untangling the origins of technological change? Unfortunately, the differences are again important. As noted in [38] the post-industrial pace of innovation is highly labyrinthine in terms of how information is exchanged between inventions and their makers. It is now accepted that an appropriate approach to cultural evolution requires a network perspective: trees are clearly not enough. The dichotomous picture of branching is being replaced by reticulate phylogenies (see below). In figure 6 the problem is illustrated with the reconstructed historical evolution of cornets [37] where the vertical branches correspond to periods of manufacture of particular models. Instruments equipped with so called Stölzel and Périnet valve systems respectively. Nonvertical links (indicated by curved red lines) are reticulations, i. e. instances of information transfer among makers. The open circles at the bottom mark key innovations in cornet design, including: (1) valve number, (2) shifting of the second valve slide and valve alignment, (3) changing of bell exit position and bell placement, and (4) alteration of bell shape (trumpetization). The reticulated nature of the diagram is largely due to the rapid and large information exchange between makers, who introduced different types of innovations and changes in the number and location of valves and other features. The data set provides clear illustrations of gradual evolution, stasis followed by punctuation, extinction and selection, but also examples of a special feature of cultural evolution: the Lazarus effect. In contrast to living species, artifacts (such as the so called shepherd's crook design) can "resurrect" based on available information about them.

## VII. EVOLVING INFORMATION TECHNOLOGY

Software systems have a relatively recent role within technological evolution and are currently the most widespread and influential systems in our society and economy. They were introduced in the 1950s (fig 7, colored areas) as soon as computer hardware and a theory on computing architectures emerged [63]. Hardware was clearly a precondition for software to emerge, as it might have been in early life [64] but, once in place, a rapid coevolution unfolded. In spite of its relevance, software has been largely ignored in the literature on technological evolution. In many ways, software allows us to get closer to the picture of biological evolution where a genomic description is present, since software deals mainly with performing functions and is coded [9,65].

One of the greatest advantages of software is that, on all scales, its fossil record is extremely well preserved, involving multiple levels of detail, from written code to the social networks of engineers that designed it. At a coarse grained scale, a phylogeny of programming languages can be defined. Some of them are displayed in Figure 7a. After the invention of the first programming language, new versions came up in order to facilitate the interactions between man and machine. In this context, software could run on different hardware thus making it much more adaptive [65]. FORTRAN in particular became a standard for engineering and physics and still survives, mainly because its simplicity and its first-move character: libraries developed to solve general calculations are constantly reused. Many other languages failed and got extinct [66]. The diagram reveals both radiations and merging. We can appreciate this in the branching from Algol-60, which solved some syntax problems associated to FORTRAN and was as much hardware-independent and portable as possible.

Later on, other major transitions took place, including the emergence of so-called Object-Oriented Programming where instructions and data are both included within well-defined structures called *objects* [65,66]. This implies a modularization of the program organization that enhances their evolvability. As it happens with cornets, engineers were able to merge different parts of previous systems and incorporate them into new languages. But here the scale of such events is much larger and conceptually more complex. The carefully planned merging of different structures used in different languages and their correct integration have no match in any other evolutionary trend. In this respect, the evolution of programming languages deeply differentiates from any biological counterpart, including the evolution of language.

Beyond the language programming scale, a hierarchy of levels of software evolution is defined by the long-term development of complex networks of interacting pieces of code and data. Although software engineering is obviously an intentional, purpose-driven activity, it is strongly affected and shaped by reuse. The evolution of large programs has been shown to occur in parallel to an

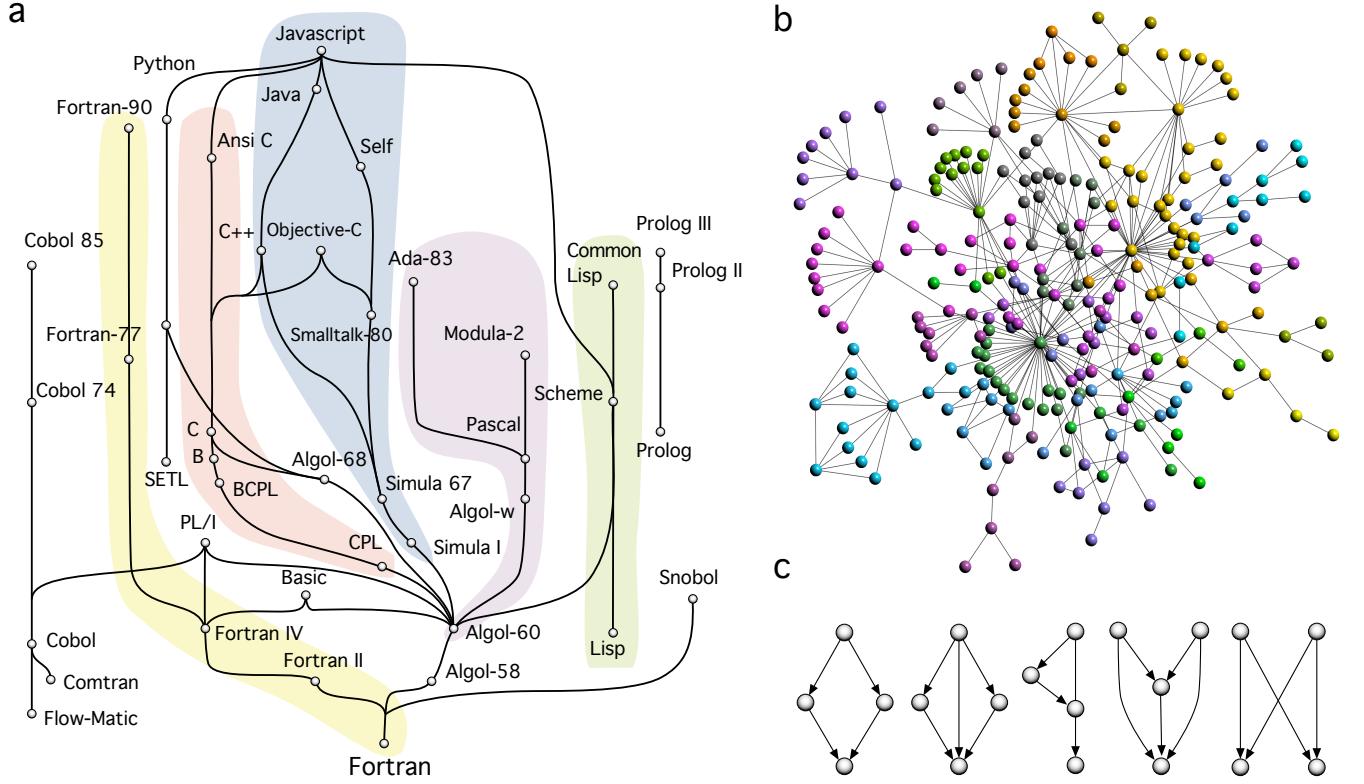


FIG. 7 The software hierarchy. At the largest scale involving the actual origins of programming languages (a) we can appreciate a pattern of diversification and multiple fusion that reveals not a tree but a network of innovations. Here we highlight the different branches associated to special classes of languages (yellow and violet represent procedural languages, blue depicts object-oriented, green indicates an important class of languages used in artificial intelligence, and red represents system programming languages). Multiple branches can fuse on a single, novel programming language (such as Javascript) through a planned process of large-scale integration of different components. Within each language, large-scale projects can be described in terms of a network (b, Aztec software) connecting different subsets of functionally well-defined blocks (b). Here each node in the graph for one of these large scale programs represents one such functional block, which depend (are linked to) on others. A directed arrow thus indicates that a given block needs to "call" another block. Such network is scale free and modular, being dominated by a handful of hubs and having most elements just one or two links with others. Such a structure seems to be a consequence of the widespread copy-paste process of growth. The rules of reuse also explain the biased statistical presence of some special, small subgraphs known as network motifs. In (c) we show some of the most common network motifs found in the graph displayed in (b). The frequency of these motifs has been shown to be a byproduct of a duplication-rewiring process [67,68] thus not (necessarily) related to special advantages of given structures at the small scale.

ever-increasing frequency of copy-paste [67-70]. An unintended consequence of this tinkering is that, as it occurs with patent networks, it spontaneously generates complex modular graphs (figure 7b). Their patterns of organization seem to be explained through extensive cloning [58,59] thus questioning Jacob's separation between technology and biological evolution in terms of the presence or absence of reuse. Even at the smallest scale the statistical pattern of network motifs, defined as small subgraphs whose frequency strongly departs from random expectations can be accounted for by reuse (figure 2c). As it occurs with other systems [13,71-74] a non-adaptive interpretation of their biased frequency is obtained.

## VIII. DISCUSSION

It seems undeniable that technological change displays numerous life-like features, suggesting a deep connection with biological evolution. This is specially well illustrated by common convergent patterns, particularly when optimization due to cost/energy constraints is at stake. But some "disanalogies" [75] are also noticeable. The co-existence of both similarities and dissimilarities occurs because technological and biological change share universal traits and causal mechanisms that pervade their long-term trends, no matter how different they appear at a lower scale. What is at the root of these potential universals? Statistical physics has been successful in showing that disparate systems often behave in the same way (once they are properly scaled) and that this uni-

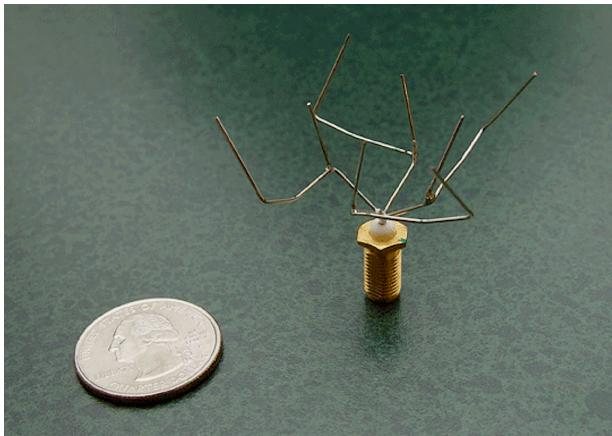


FIG. 8 A pure engineering design can be obtained by following a Darwinian selection process. Using computer simulations, NASA Ames researchers evolved a population of antennas with a fitness function defined in terms of a signal-processing problem. The best fit candidates were rather organic in their shape, as the one shown here. Image from NASA Ames Center

versal behavior can be explained by considering only a very reduced number of properties [76-78]. It has been found for example that a magnet, a flock of birds or even the stock market behave similarly in terms of displaying tipping points, transient dynamics or response to external stresses. This universal behavior is present in the structure of neutral landscapes, the existence of frozen accidents or the punctuated equilibria.

Any theory of technological change will require new metaphors and tools, particularly when dealing with information systems. Reticulate networks, instead of trees [79-81] provide the adequate test to the role played by pure combination. Networks allow defining modules and their relations and the success within language evolution [82,83] indicates that they are more appropriate when dealing with cultural dynamics. Additionally, technological design can be obtained through *in silico* Darwinian selection [84-87] thus defining a new level of complexity. As the distance between technology and biology shrinks, some unexpected results emerge as hardware is itself evolved.

One important message is that the results of Darwinian selection often escape from our intuition and expertise, suggesting that part of the design landscape is not accessible to our minds. This is illustrated by hardware design involving Darwinian selection. Genetic algorithms in particular [73-75] allow a population-level exploration of vast parameter spaces. In this approach, each member of the population is more or less fit in solving the given problem, and its survival and number of offspring directly depends on its functional success. Surviving solutions are copied with mutations: small changes are introduced with some given probability.

Evolved circuits, engines, game strategies or market

forecasts can be obtained by using evolutionary programming. Often these algorithms provide improved modifications that make the previous designs more efficient. However, some challenge our intuition. An example is shown in figure 8, where an evolved antenna is displayed. It was obtained under evolutionary dynamics to solve a difficult task on a NASA mission satellite to study the Earth magnetosphere. Using a population of candidate structures and evolving them over many generations, each antenna was tested against a number of technical requirements that were simulated on a supercomputer. The evolved structure is far from what an engineer would ever had imagined. This example offers a very interesting situation where the peaks in the technological landscape that solve the problem cannot be captured by the mental structure of the human designer.

Another differential trait of technology is the role played by social and economic factors [88] on innovation. Networks of inventors act as ecosystem engineers [89] able to transform their own innovation landscape (see [90]) thus changing the ways other technologies will interact. Moreover, as early noticed by Robert Solow, economic growth requires the presence of innovation [91] and thus coevolution among economy and technology must be included. Similarly, we need to identify the scales at which technological hierarchies operate. In biology, such hierarchy can be described [92] including different levels, from population dynamics to genotype-phenotypic maps.

In technology, population changes can be estimated using *culturomics* [93], i. e. exploiting the abundance of given strings of words (the so called *n*-grams) extracted from millions of digitalized books. The relative frequency of these n-grams within the culturome provides a measure of their importance over time. A kinetic theory of n-grams could take advantage of useful ecological metaphors used within other domains of cultural evolution, such as language change [94]. Defining appropriate mappings between genotype and phenotype is a much more difficult task [95]. Once again, information technology, with all its richness and multiplicity of scales, offers our best bet to achieve this goal. There is a whole fossil record there to be unearthed.

### Acknowledgments

We thank Peter Schuster and Ricardo Haussman for useful comments. This work has been supported by research grant FIS2004-05422 (RS,SV), by the James McDonnell Foundation (RS), the Marcelino Botin Foundation (SV,RS), the TEKES Foundation of Finland (SK), the National Science Foundation (NSF) Grant SBE0738187 (DF) and by the Santa Fe institute.

## IX. REFERENCES

1. Basalla, (1989) The evolution of technology. Cambridge U. Press.
2. Arthur, B. (2009) The nature of technology. What it is and how it evolves. Free Press.
3. Kelly, K (2010) What technology wants. Viking.
4. Johnson S (2010) Where good ideas come from: a natural history of innovation. Riverhead Press.
5. Sahal, D. (1981). Patterns of technological innovation. Addison-Wesley.
6. Schuster, P. (2006) Untamable curiosity, innovation, discovery and bricolage. Complexity 11, 9-11.
7. Arthur, B. (1997) Increasing returns and path dependence in the economy. Michigan U. Press.
8. Arthur, B. and Polak, W. (2006). The evolution of technology within a simple computer model. Complexity 11, 23-31.
9. Ogburn, WF and Thomas D (1922) Are inventions inevitable? A note on social evolution. Polit. Sci. Quat. 37, 83-98.
10. Friedel, R. and Israel, P. (1986) Edison's electric light: biography of an invention. Rutgers U. Press. pp. 115 - 117.
11. Jacob, F. (1977). Evolution as tinkering. Science 196, 1161-1166.
12. Wilkins, A. (2007) Between design and bricolage: genetic networks, levels of selection and adaptive evolution. Proc. Natl. Acad. Sci. USA 105, 8590-8596.
13. Solé, R.V., Ferrer, R., Montoya, J. M. and Valverde, S. (2002). Selection, Tinkering, and Emergence in Complex Networks. Complexity 8, 20-33.
14. Schuster, P. (1996). How does complexity arise in evolution? Complexity 2, 22-30.
15. Brown, J.H. and West, G. B. (2000). Scaling in biology. Oxford U. Press.
16. Minetti, A, E., Pinkerton, J. and Zamparo, P. (2001) From bipedalism to bicycism: evolution in energetics and biomechanics of historic bicycles. Proc. R. Soc. Lond. B 268, 1351-1360.
17. Farmer D (1999) Market force, ecology and evolution. Ind. Corp. Change 11, 895-953.
18. Eldredge, N. (1997) Evolution in the marketplace. Struct. Change Econ. Dyn. 8, 385-398.
19. Gould, SJ (1987) The Pandas thumb of technology. Natural History 96, 14-23.
20. Kimura, M. (1964). Diffusion models in population genetics. J. Appl. Prob. 1, 177-232.
21. Solé, R.V. (2011) Phase Transitions. Princeton U. Press.
22. Petrovski, H (1992) The evolution of useful things. Vintage Books.
23. Christie, P. (2000). The interpretation and application of Rent's rule. IEEE Trans. VLSI Syst. 8, 639-648.
24. Bassett, D.S., Greenfield, D.L., Meyer-Lindenberg, A. et al. (2010). Efficient physical embedding of Topologically Complex Information Processing Networks in Brains and Computer Circuits. PLoS Comp. Biol. 6, e1000748.
25. Nelson, ME and Bower, JM (1990) Brain maps and parallel computers. Trends Neurosci. 13, 403-408.
26. Schuster, P., Fontana, W., Stadler, P. F. and Hockecker, I. L. (1994). From sequences to shapes and back: a case study in RNA secondary structures. Proc. R. Soc. Lond. B 255, 279284.
27. Huynen, M. A., Stadler, P. F. and Fontana, W. (1996). Smoothness within ruggedness: The role of neutrality in adaptation. Proc. Natl. Acad. Sci. USA 93, 397-401.
28. Nimwegen, E., Crutchfield, J. P. and Huynen, M. (1999). Neutral evolution of mutational robustness. Proc. Natl. Acad. Sci. USA 96, 9716-9720.
29. Wilke, C.O. (2001). Adaptive evolution on neutral networks. Bull. Math. Biol. 63, 715-730.
30. Ferrer, R., Janssen, C. and Solé, R. V. (2001). Topology of Technology Graphs: Small World Patterns in Electronic Circuits. Phys. Rev. E 64, 32767
31. Macia, J. and Solé, R. V. (2009) Distributed robustness in cellular networks: insights from synthetic evolved circuits. J. R. Soc. Interface 6, 393400.
32. Raman, K. and Wagner, A. (2012). The evolvability of programmable hardware. J. R. Soc. Interface. doi: 10.1098/rsif.2010.0212.
33. Hanel, R, Kauffman SA and Thurner S (2005) Phase transitions in random catalytic networks. Phys. Rev. E 72, 036117.
34. Hanel, R, Kauffman SA and Thurner S (2007) Towards a physics of evolution: critical diversity dynamics at the edges of collapse and bursts of diversification. Phys. Rev. E 76, 036110.

35. Kauffman SA (2000) *Investigations*. Oxford U. Press.
36. Lipinski, C. and Hopkins, A. (2004). Navigating chemical space for biology and medicine. *Nature* 432, 855-861.
37. Témkin, I. and Eldredge, N. (2007). Phylogenetics and Material Cultural Evolution. *Curr. Antrop.* 48, 146-153.
38. Eldredge, N. (2011) Paleontology and cornets: Thoughts of material cultural evolution. *Evo. Edu. Outreach* 4, 364-373.
39. Nelson RR and Winter SG (1982) An evolutionary theory of economic change. Belknap Press.
40. Nelson RR and Winter SG (1977) In search of a useful theory of innovation. *Res. Policy* 6, 36-76.
41. Rip, a and Kemp R. (1988) Technological change. In: Rainer S and Malone L (Eds) *Human choice and climate change*, Vol 2, pp. 327-400. Batelle Press.
42. Van den Ende, J and Kemp, R. (1999) Technological transformations in history: how the computer regime grew out of existing computing regimes. *Res. Policy* 28, 833-851.
43. Nagi, B., Farmer, J.D., Trancik, J. E. and Gonzales, J.P. (2011). Superexponential Long-Term Trends in Information Technology. *J. Tech. Forecast. Soc. Change* 73, 1061-1083.
44. H. Koh and C. L. Magee, (2006) Technological Forecasting and Social Change 73, 1061 - 1083.
45. Farmer, J. D. and Trancik, J. (2007). Dynamics of Technological Development in the Energy Sector. In: London Accord Final Publication. Eds. J-P Onstwedder and M Mainelli. pp. 1-24.
46. Argote, L and Epple E (1990) Learning curves in manufacturing. *Science* 247, 920-924.
47. Auerswald P, Kauffman, SA, Lobo, J and Shell K (2000) The production recipes approach to modeling technological innovation: An application to learning by doing. *J. Econ. Dyn. Control* 24, 389-450.
48. McDonald A and Schrattenholzer L (2001) Learning rates for energy technologies. *Energ. Policy* 29, 255-261.
49. Moore GE (1965) Cramming more components onto integrated circuits. *Electronics* 38, 114-117.
50. Kauffman, S. A. and Levin, S. (1987) Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.* 128, 11-45.
51. Kauffman, S. A. (1995) At home in the universe. Chapter 9. Oxford U. Press.
52. Frenken K (2000) A complexity approach to innovation networks. The case of the aircraft industry. *Res. Policy* 29, 257-272.
53. Marengo L and Valente M (2010) Industry dynamics in complex products spaces: An evolutionary model. *Struct. Change Econ. Dyn.* 21, 5-16.
54. Baldwin CY and Clark KB (2000) *Design rules*. MIT Press.
55. McNerney, J, Farmer JD, Redner S and Trancik, JE (2011) Role of design complexity in technology improvement. *Proc. Natl. Acad. Sci. USA* 108, 90089013.
56. Kurzweil, R. (2005) *The singularity is near*. Viking.
57. Tuomi, I. (2006) *Networks of innovation*. Oxford U. Press.
58. Jaffe, AB and Trajtenberg, M (2003) Patents, Citations and Innovations. MIT Press.
59. Valverde, S, Solé RV, Bedau M. and Packard, N (2007) Topology and evolution of technology innovation networks. *Phys. Rev. E* 76, 032767.
60. Gress, B. (2010). Properties of the USPTO patent citation network: 1963-2002. *World Patent Info.* 32, 3-21.
61. Bommarito II M.J., Katz D.M., Zelner J.L., Fowler J.H. (2010). Distance measures for dynamic citation networks. 389, 42014208
62. Verspagen, B. (2009). Mapping technological trajectories as complex networks. *Adv. Complex Syst.* 10, 93-115.
63. Williams, MR (1985) *A history of computer architecture*. Prentice Hall.
64. Dyson G (1997) *Darwin among the machines*. Addison-Wesley.
65. Sammet JE (1969) *Programming languages: History and fundamentals*. Prentice-Hall.
66. Mens T. and Demeyer, S (2008) *Software evolution*. Springer-Verlag.
67. Valverde, S, Ferrer, R and Solé RV (2002) Scale-free networks from optimal design. *Europhys. Lett.* 60, 512-518.
68. Valverde, S. and Solé, R.V., (2005). Network Motifs in Computational Networks: A Case Study in Software Architecture, *Phys. Rev. E* 72, 026107.

69. Myers, C.R. (2003) Software systems as complex networks: structure, function, and evolvability of software collaboration graphs. *Phys. Rev. E* 68, 048116.
70. Concas G, Marchesi M, Pinna S and Serra N (2007) Power-Laws in a Large Object-Oriented Software System. *IEEE Trans. Soft. Eng.* 33, 687-708.
71. Solé, R.V. and Valverde, S. (2006) Are Network Motifs The Spandrels of Cellular Complexity?. *Trends Ecol Evol.* 21, 419-22.
72. Solé, R.V. and Valverde, S. (2007) Spontaneous emergence of modularity in cellular networks. *J. Roy. Soc. Interface* 5, 129-133.
73. Lynch M (2007) The evolution of genetic networks by non-adaptive processes. *Nat. Rev. Genet.* 8, 803-813.
74. Wagner, G. Pavlicev, M., Cheverud, J. M. (2007). The road to modularity. *Nat. Rev. Genet.* 8, 921-931.
75. Ziman, J. editor. (2000). Technological innovation as an evolutionary process. Cambridge U. Press.
76. Kauffman, S. A. (1993). The Origins of Order: Self-Organization and Selection in Evolution. Oxford U. Press
77. Cowan, G. A., Pines, D. and Meltzer, D. (1996). Complexity: metaphors, models and reality. Perseus Books.
78. Mitchell, M. (2009). Complexity: a guided tour. Oxford U. Press.
79. Makarenkov, V. and Legendre, P. (2004). From phylogenetic tree to a reticulated network. *J. Comput. Biol.* 11, 195-212.
80. Lipo, C.P., O'Brien, M.J., Collard, M. and Shennan, S.J. , editors (2009) Mapping our ancestors. Aldine Transaction Publishers.
81. Kunin, V., Goldovsky, L., Darzentas, N., and Ouzounis, C. A. (2005). The net of life: reconstructing the microbial phylogenetic network. *Genome Res.* 15, 954959.
82. Dagan T, Martin W (2009) Getting a better picture of microbial evolution en route to a network of genomes. *Phil. Trans. R Soc. Lond B Biol Sci* 364, 2187-96.
83. Nelson-Sathi S. et al., (2011) Networks uncover hidden lexical borrowing in Indo-European language evolution. *Proc. Royal Soc. London B* 278, 1794-1803.
84. Mitchell M (1998) An introduction to genetic algorithms. Bradford Books.
85. Forrest S (1993) Genetic algorithms: principles of natural selection applied to computation. *Science* 261, 872-878.
86. Koza, J.R., Keane, M.A. and Streeter, M.J. (2003) Evolving inventions. *Sci Am.* February 52-59.
87. Long, J. (2012). Darwin's devices. Basic Books.
88. Jones, C. G., Lawton, J. H. and Shachak, M. (1994). Organisms as ecosystem engineers. *Oikos* 69, 373-386.
89. Erwin, D.H. (2008). Macroevolution of ecosystem engineering, niche construction and diversity. *Trends Ecol Evol.* 23, 304-310.
90. Hidalgo, C.A. , Klinger B. , Barabási, A.-L. and Hausmann, R. (2007). The product space conditions the development of nations. *Science* 317, 482-487.
91. Solow, R. (1952) A contribution to the theory of economic growth. *Q. J. Econ.* 76, 65-94.
92. Schuster, P. (1996) How does complexity arise in evolution? *Complexity* 2, 22-30.
93. Michel, Shen, J.-B., Aiden, Y.K., Veres, A.P, et al. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, 33, 176-182.
94. Solé, R. V., Corominas-Murtra, B. and Fortuny, J. 2010. Diversity, competition, extinction: the eco-physics of language change. *J. Royal Soc. Interface* 7, 1647-1664.
95. Lobo, J., Miller, J.H. and Fontana, W. 2004. Neutrality in technological landscapes. Santa Fe Institute Working Paper