Complexity, Organization, and Stuart Kauffman's *The Origins of Order*

Frank H. Westhoff, Beth V. Yarbrough, Robert M. Yarbrough *

*Department of Economics, Amherst College, Amherst, MA 01002 USA*

Received 30 September 1993; revised 24 January 1994

**Abstract**

The economics of organization asks why we observe the organizational variety we do, why that observed variety is only a subset of all possible organization types, and how organizational forms change. This review essay evaluates Stuart Kauffman's *NK* models from the complexity sciences for investigating these questions. Such models focus on the statistical properties of complex systems that allow us to examine the total ensemble of organization types, which types are 'neighbors' of which other types, and the pattern of change in organizational forms, as well as situations where the interactions among the relevant elements or the dynamic characteristics of the system are unknown.

**JEL classification:** L2; D23; C69; N60

**Keywords:** Complexity, Organization; Industrial enterprise; Fitness landscape

1. **Introduction**

In *Scale and Scope*, Alfred Chandler depicts a historical pattern of industrial development in which the United States and Britain exhibited distinctive patterns of firm management during the Second Industrial Revolution. * U.S. firms moved
quickly toward innovative professional management of large-scale multidivision structures, integrating upstream into research and development and downstream into distribution. British firms, in contrast, clung to traditional family-based management techniques and eschewed both massive changes in scale and vertical integration. One can infer from Chandler’s account that Britain failed to take the optimal path – the innovative one taken in the United States. This essay examines new techniques from the complexity sciences, in particular Stuart A. Kauffman’s *The Origins of Order*, to begin investigating such issues.

Herbert Simon was among the first economists to consider organizational questions in the light of computational complexity (Simon, 1962; Simon, 1972; Simon, 1978a, Simon, 1978b). A key concern involved what he called the ‘needle in the haystack’ problem:

‘If needles are distributed randomly in a haystack..., then to find the sharpest needle in the stack, we have to search the entire stack, and the search time will ... be linear with size, which does not seem too bad until we remember that the haystack of life is essentially infinite. If we are satisfied with any needle (after all, they are all sharp enough to sew with), then the amount of search to find one will ... be independent of the size of the stack. Complexity independent of the size of the problem domain is a property we badly need in algorithms designed to face the problems of the real world.’

As Simon stressed, choices among decision methods frequently exhibit a trade-off between computational costs and improvement in the decision reached, and both quantities tend to vary with the size of the problem at hand. In the theory of computational complexity, problems of a given size, \( N \), are considered ‘tractable’ if their solution times \( (T_1) \), measured as the number of simple computational steps required) rise only linearly \( (T_1 = aN) \) or polynomially \( (T_1 = aN + bN^2 + \ldots) \), although, as Simon notes, even these so-called tractable problems can prove daunting within real world time frames. Computationally ‘complex’ problems, on the other hand, are those whose solution times increase exponentially in \( N \) \( (T_1 = a^N) \), a task of a prohibitive order of magnitude for large \( N \).

In terms of Simon’s needle-in-a-haystack example, the relevant haystack may be too large to complete a search for the sharpest needle within a feasible time horizon. Then economists’ typical functional argument — the needle observed in use must be the sharpest one available — loses much of its force, because its

---

2 Chandler also investigates the history of German firms, which exhibited a combination of the characteristics of British and U.S. firms.

3 See, for example, Chandler (1990), p. 294. Closely related literatures include work on path dependence (for example, David, 1975; David, 1985; Arthur, 1988a; Arthur, 1988b; Arthur, 1990) and on the role of culture (for example, Hamilton and Biggart, 1988, and Boyd and Richerson, 1993).

reliance on convexity conditions that equate local and global optima may be misleading or without operational significance. However, approximate solutions, which are more numerous than optimal ones, generally can be found within an acceptable time horizon even when the problem is complex. And as Simon's needle example highlights, enlarging the number of solutions for which we are willing to settle also reduces the expected solution time for problems with linear or polynomial search times.

If British firms in Chandler's study settled for less than the sharpest needle in Simon's sense, was it due to differences in the British and American environments or to the computational complexity of the problem? Of course, British firms developed earlier than those in the United States and faced different market sizes, transportation and communication technologies, geography, growth rates, government policies, and random events. These factors may have channeled British and American firms toward different management techniques and organizational structures. Or, maybe not.

2. The \( N \)-dimensional haystack

Chandler specifies many margins or characteristics on which U.S. and British firms differed. Of those many \((N)\) possibilities, Fig. 1 lists just four: management technique, scale of operation, extent of vertical integration, and organizational structure. Suppose each margin takes one of only two values, which can be 'coded' as 0 or 1 as in Fig. 1, for example, management is either family (0) or professional (1), and scale either small (0) or large (1).

Even considering only four of \(N\) potential margins, and restricting each to only two values, the total ensemble or set of all possible combinations contains \(2^4 = 16\) organizational types, each of which can be denoted by a four-bit string. For example, 1111 is one type, and 0000 is another. Each type can be changed to another by altering one bit – from 0 to 1 or from 1 to 0 – so each type has four 'one-margin' neighbors. For example, 1110 and 0111 are two of 1111's neighbors, each one Hamming distance away. 5 In a four-dimensional space, we can represent the entire ensemble of \(2^4 = 16\) possible combinations, with each next to its 4 one-margin neighbors. Panel (a) in Fig. 2 illustrates all 16 possible combinations of four firm characteristics as vertices on a four-dimensional Boolean hypercube, representing a type space. 6 Each vertex, labeled with the corresponding bit string,

---

5 The Hamming distance between two \(N\) element bit strings can be defined as the number of bits by which the two strings differ.

6 This section utilizes a simplified example based on Chandler to present the basic concepts central to Kauffman's framework: binary coding, the ensemble or type space, and the fitness landscape; see Kauffman's Chapter 2. More generally, the ensemble contains \(A^N\) members where \(A\) represents the number of possible values for each margin. The number of one-margin neighbors then becomes \((A - 1)N\).
represents one of the combinations of characteristics, and each is linked to the four other types that differ from it by a single bit being switched from 0 to 1 or from 1 to 0. Therefore the four-dimensional type space both represents the entire ensemble of $2^4$ organizational types and records which four types comprise the one-margin neighbors of each.

Suppose the capacity of each organizational type to accomplish a particular task under specified conditions is a well-defined property. For example, Chandler argues that firms differed by organizational type in their abilities to manage production and distribution ('throughput') in a way that maintained capacity utilization even during downturns of the business cycle. We could then define for each type the capacity to maintain throughput as the type's 'fitness'. Such fitness would be distributed across the various members of the ensemble; in other words, firm types could be ranked according to that capacity from best (1) to worst (16). Panel (b) of Fig. 2 arbitrarily assigns such a ranking to each of the 4-bit types. The distribution of capacities across the type space constitutes the fitness landscape with respect to that particular task. If the firm searches for fitter types by alteration of single bits at each of the margins, the 'hill-climbing' process can be thought of as an adaptive walk between adjacent vertices on the fitness landscape. If

---

7 Fitness relative to a particular task need not necessarily imply overall fitness of the firm.
additionally each step must improve function (that is, flow with the arrows in the figure), a walk started from any vertex will reach a local optimum, fitter than all its one-margin neighbors. Given the arbitrarily assigned fitness ranking in Fig. 2(b), there are three local optima (the global optimum 1111, the second-fittest type 0000, and the third-fittest type 1010). If a walk is restricted to climbing to fitter neighbors and begins on any one of these local-optimum types, the process stops immediately. If not, hill climbing proceeds for one or a few steps until it reaches such an optimum. In neither case will the process necessarily reach the global optimum.

3. Self-organization and selection

If organizations use only contiguous hill climbing to seek their optimal type, some may become stranded at local, not global, optima. Most economists believe selective forces eventually will eliminate at least vastly suboptimal organizational types. If so, only a subset of the ensemble of possible types will be observed, from which most economists conclude that the selection process favors the characteristics exhibited by types in the observed subset.

Stuart Kauffman, on the other hand, asks whether the observed types are necessarily the products of selection. Could they instead be the products of self organization? If certain properties are widespread throughout the ensemble, perhaps selection cannot avoid them. Suppose most but not all entities in the ensemble exhibit some spontaneously ordered property that may or may not affect fitness. If that property is selected for, we will observe it in existing entities. But Kauffman argues we might also observe the property if it is not selected for and, in fact, even if it is mildly selected against, depending on the extent to which selection can move a population to particular areas of its fitness landscape. If selection is weak and unable to move the population to particular regions of the landscape, observed types will exhibit any properties widespread throughout the landscape, not because of selection but despite it. If selection is powerful enough...

8 Walks constrained to pass via one-mutant neighbors correspond to one plausible limiting case of a population's adaptive flow under mutation, selection, and recombination. Gillespie (1983, 1984) shows that such constrained walks correspond to a population in which the rate of finding fitter variants is very low compared with the fitness differentials between the less and more fit alleles. Gillespie shows that, in the limit, the adaptive process can be treated as a continuous-time, discrete-state Markov process, with each state corresponding to one genotype. The population as a whole jumps with different probabilities to one or another of the fitter neighboring genotypes if the product of population size and mutation rate is low compared with the rate of finding fitter variants.

9 Economists' interest in evolutionary selection predates Darwin; see, for example, Jones (1986). Seminal contributions by Alchian (1950) and Nelson and Winter (1982) are prominent among contemporary studies following an explicitly evolutionary approach. See Dosi et al. (1988) and the citations therein for the current state of the art.
to move the population to virtually any area of the landscape, the selection process may be able to avoid a self-ordered property exhibited by most but not all entities, but it may not. Kauffman argues that the extent to which we should expect to observe types with characteristics typical or atypical of the ensemble as a whole depends on the shape of the rugged fitness landscape, and his \( NK \) model is a theory of the statistical properties of such landscapes.

4. Kauffman's \( NK \) Model

\( N \) denotes the number of elements in the system, for example, the number of characteristic margins along which organizational types differ. As we have seen, if organizational types differ on four margins, each of which can take one of two values, the total number of organizational types is \( 2^4 \). More generally, with \( N \) characteristics, each of which can take \( A \) values, the total ensemble numbers \( A^N \) members. Each is a one-characteristic neighbor of \( N(A - 1) \) other types, which we can represent in an \( N(A - 1) \) dimensional space analogous to Fig. 2.

Each characteristic makes a contribution to the type's fitness. The fitness of a particular type depends on the contribution of each of the type's characteristics. That contribution depends on the characteristic itself and on \( K \) other characteristics from among the \( N \). In other words, \( K \) denotes how interdependent the characteristics are, or the number of 'epistatic interactions' among them, in determining their fitness contributions. If \( K = 0 \), each characteristic's fitness contribution is independent of all other characteristics. In terms of our example, \( K = 0 \) would imply \textit{inter alia} that the fitness contribution of professional management would be independent of whether the organization was large or small, integrated or not, and uni- or multidivisional. At the other extreme, if \( K = N - 1 \), each characteristic's fitness contribution depends on all the other characteristics.

Generally, we expect \( (N - 1) \geq K > 0 \). For example, we might expect the fitness benefit of professional management to be greater for large, integrated, multidivisional firms than for small, nonintegrated, unidivisional ones. However, we often do not know the precise pattern of the epistatic interactions. This raises obvious questions of both measurement and modeling, whether in the economics of organization or in Kauffman's primary field of research, biophysics.

'In general, we truly have almost no idea what those mutual influences on overall fitness might be. Take Mendel's peas. He found two alleles for seed color, yellow and green, and two alleles of a second gene for seed texture, rough and smooth. \textit{A priori} we have no idea which of the four combinations of these traits will be of highest fitness, nor how changing from any one combination of traits to any other

---

\(^{10}\) This section outlines the model from Kauffman's Chapters 2 and 3.
will affect fitness. If the fitness contribution of each gene is epistatically affected by a large number of other genes, the possible conflicting constraints among the complex web of epistatically interacting genes are both unknown and likely to be extremely complex. This complexity suggests that it might be useful to confess our total ignorance and admit that, for different genes and those which epistatically affect them, essentially arbitrary interactions are possible. Then we might attempt to capture the statistical features of such webs of epistatic interactions by assuming that the interactions are so complex that we can model the statistical features of their consequences with a random fitness function. This leads to the NK model.\(^\text{11}\)

Consider our simple example based on Chandler's firms, letting \( K = N - 1 = 3 \) to produce the maximum interaction among characteristics in determining their fitness contributions. The fitness contribution of each characteristic depends on \( K + 1 = 4 \) characteristics (\( K = 3 \) others and itself). Fig. 3(a) illustrates the pattern of these epistatic interactions.\(^\text{12}\) The number of possible combinations of these characteristics is \( 2^{K+1} = 16 \). Kauffman's model 'confess[es] our total ignorance' by assigning each of those combinations a randomly drawn fitness contribution from the uniform distribution over the unit interval. Therefore the possible values of the fitness contribution of the \( i \)'th characteristic, \( w_i \), are given by a \( 2^{K+1} \)-long list of random numbers between 0 and 1. Panel (b) reports an arbitrary assignment of fitness contributions for each characteristic given its epistatic inputs, as well as a fitness value for each possible organizational type (measured as the mean value of the fitness contributions of the four characteristics, or \( W = \Sigma w_i / N \)). This specification of the simplest NK model defines the fitness landscape in panel (c), where the values in parentheses report the fitness values from panel (b).

Even with a random assignment of fitness values, knowledge of just two parameters, \( N \) and \( K \), can still produce a surprising amount of information about the statistical properties of the fitness landscape. Most notably, as \( K \) increases, conflicting constraints result in a more and more rugged and multipeaked landscape, with obvious implications for the prevalence of local optima and for the efficacy of hill climbing.\(^\text{13}\)

4.1. \( K = 0 \) yields a smooth, correlated, single-peaked landscape

If each characteristic's fitness contribution is independent of all others (\( K = 0 \), so there are no epistatic interactions among the \( N \) characteristics), the fitness

\(^{11}\) Kauffman (1993), p. 41. Unless noted, all page citations refer to Kauffman's Origins.

\(^{12}\) See Kauffman (1993), Fig. 2.2, p. 42 for the simpler \( N = 3 \) case.

\(^{13}\) The range of fitness values exhibits sensitivity to the distribution from which the underlying fitness values are drawn; see Kauffman (1993), p. 44. At the same time, many landscape features of interest appear robust to parameters of the model other than \( N \) and \( K \), allowing a focus on those as the primary determinants of the structure of rugged fitness landscapes.
Fig. 3. (a) Assignment of $K = 3$ epistatic interactions, (b) arbitrary assignment of random fitness values, and (c) corresponding fitness landscape on a Boolean hypercube.

A smooth landscape is one in which neighboring types have similar fitnesses. When one characteristic changes, the fitness contributions of the other characteristics remain the same, so the one-characteristic change cannot alter fitness by more than $1/N$. The larger is $N$, the smoother or more highly correlated the landscape.

---

Such fitness landscapes have a single global optimal type, and all other types are suboptimal but can climb to the global optimum via fitter neighbors. For each characteristic, one of the two possible values, 0 or 1, makes the higher fitness contribution, and the global optimal type exhibits the higher fitness value for each characteristic. Any suboptimal type can be changed to the optimal type by one-by-one switching of each low-fitness-value characteristic to the corresponding high-fitness value. Therefore, any type can reach the optimum via fitter neighbors. In addition, on any climb toward the optimum, the number of fitter neighbors falls by one at each upward step, and the expected number of steps to the optimum (given by \( N/2 \)) increases linearly with \( N \).

In terms of our firm-type example, \( K = 0 \) would imply that each margin or characteristic—management, scale, degree of vertical integration, and organization structure—would have an optimal value, either 0 or 1, independent of all other margins. Therefore, the optimal firm type would be the one exhibiting the superior value at each of the four margins. A firm of any type could climb to the single, global optimal type simply by switching its value on each margin to correspond to that margin’s superior value. Each one-margin change would improve the firm’s fitness, so the firm could hill-climb to the optimum from any point via fitter one-margin neighbors.

4.2. \( K = N - 1 \) yields a jagged, uncorrelated, multipeaked landscape

\( K \) acts as a parameter that tunes the ruggedness of the fitness landscape.\(^{15} \) As \( K \) increases from zero toward its maximum value of \( N - 1 \), the fitness landscape changes from smooth, correlated, with a single global optimum to increasingly jagged, uncorrelated, and multipeaked. Increasing the number of epistatic interactions presents conflicting constraints on the system, reflected in the ruggedness of the landscape.

When \( K = N - 1 \), the landscape is completely uncorrelated; in other words, the fitness value of one type conveys no information about the fitness of its neighbors. The fitness contribution of each characteristic depends on all other characteristics; so beginning from any type, alteration of any one characteristic changes the fitness contribution of all the others to new random values. The fitness value of a neighboring type is, therefore, a sum of \( N \) new random values and completely uncorrelated with the fitness of the original type.

In addition to being completely uncorrelated, \( K = N - 1 \) landscapes contain a large number of local optima, as can be seen by noting that the probability of a type being a local optimum equals the probability that its fitness exceeds that of all its neighbors. With \( 2^N \) types, the expected number of local optima with respect to one-characteristic moves is \( 2^N/(N + 1) \). As \( N \) rises, the number of local optima

rises exponentially. Hill climbing on landscapes with so many local optima almost inevitably leads to trapping and failure to attain the global optimum.

The jagged structure of $K = N - 1$ landscapes also limits the number of local optima attainable from any starting point. Even beginning from the least-fit type, the upper bound on the expected number of accessible local optima is $N^{(\log N - 1)/2}$. Combining this result with the expected number of local optima implies that the upper bound on the share of local optima reachable from a given starting point declines with $N$. In our $N = 4$ example of firm characteristics, the expected number of local optima is just over 3, of which two (or approximately 63 percent) are expected to be accessible from the least-fit type. If $N$ doubled to 8, the expected number of local optima would rise to over 28, but only 8 of those (or approximately 28 percent) would be expected to be accessible from the least-fit type.

Another perspective on the likelihood of reaching the global optimum involves calculating an upper bound on the fraction of total types that can reach that optimum via fitter one-characteristic neighbors. That bound is $\sum_{j=0}^{N} N^{(\log N - 1)/2}$. If $N = 4$, this upper bound is 12, which corresponds to 75 percent of the total number of types. But when $N$ rises to 8, the bound rises to 80, or only 31 percent of the 256 total types. Therefore, as $N$ increases, a rapidly rising share of adaptive walks via fitter neighbors end on local optima that fall short of the global one.

Just as important, and perhaps more surprising, is that as $N$ increases, the fitness peaks dwindle in height and approach the mean fitness of the entire ensemble of types. This result implies a limit on attainable fitness regardless of the strength of selection and regardless of the type of hill-climbing procedures used, because, by definition, the peaks of the fitness landscape limit what can be attained. Conflicting constraints drive this result. If characteristics $i$ and $j$ epistatically interact, the value taken by each affects the fitness contribution of the other. The value for characteristic $i$ that maximizes the fitness contribution of characteristic $j$ typically will differ from the value that maximizes $i$'s own contribution, or that of any of the other characteristics with which $i$ interacts. Therefore, the best mutual choices become increasingly constrained both as $N$ rises (given $K = N - 1$) and as $K$ rises as a share of $N$.

This is the case for which Fig. 3 represents one highly stylized example. With

---

16 This result can be generalized in two ways. If each characteristic can take $A$ values rather than only two and walks can alter as many as $r$ characteristics per step, the expected number of local optima becomes $A^{N/\sum_{j=0}^{r} N^{(\log N - 1)/2}}$ (Kauffman, 1993, pp. 47–48).

17 Generalizing the number of possible values taken by each characteristic from 2 to $A$, this becomes $N(A - 1)^{\log A - (A - 1)^{1/2}}$ (Kauffman, 1993, pp. 51–52).

18 However, in our particular example in Fig. 3, all three local optima are accessible from the least-fit type 0111.

$K = 3$, the landscape is uncorrelated. A firm altering any margin experiences a change in the fitness contribution of all other margins. Our example has three local optima – Chandler's prototype U.S. firm (1111 or professionally managed, large, integrated, and multidivisional), his prototype British firm (0000 or family managed, small, nonintegrated, and unidivisional), and a type (1010) corresponding to professionally managed, small, integrated, unidivisional firms. Firms hill climbing to either of the latter two types via one-margin neighbors become trapped, unable to reach the global optimal type (1111).

But this is not the end of Kauffman's story. Many questions of interest in his field of biophysics, as well as puzzles in the economics of organization, revolve around dynamic systems. Kauffman extends the framework of his $NK$ model to random Boolean network models of dynamic systems, yielding results in situations that sometimes prove intractable using continuous nonlinear differential equations.

5. Dynamic random Boolean $NK$ networks

A Boolean network consists of $N$ binary (0,1) variables, each regulated by a set of elements in the network that act as its inputs. Specifying a random Boolean $NK$ network involves specifying randomly the $K$ inputs and randomly assigning a Boolean function for each element. The Boolean function or logical switching rule then determines whether each variable takes its 0 or its 1 value in the next period – based on the values of its input elements in the current period.

Fig. 4 presents perhaps the simplest example of a Boolean function, 'EQUIVALENCE', where element 2 receives input from element 1. If element 1 takes its 0 value at a given moment, then element 2 will take its equivalent value (0) at the next moment, and if the input element takes its 1 value, the regulated element will take its (equivalent) value of 1 the next moment. If each element receives input from only one element ($K = 1$), there are only four possible Boolean functions: an element can be always 0, always 1, the same as its input's value, or the opposite of its input’s value.

More generally, with $K$ binary inputs there are $2^K$ possible combinations of those inputs. For each combination, the Boolean function must specify whether the regulated element is to take its 0 or its 1 value. Therefore, there are $2^{2^K}$ possible Boolean functions regulating each of the $N$ elements in the network. Fig. 5 represents the 16 Boolean functions of $K = 2$. For each function, the figure reports its name or symbol, a tabular representation, and a two-state automata representa-

---

20 Kauffman's Chapter 5 introduces his theory of dynamic random Boolean networks. The Boolean idealization captures the essential features of continuous dynamic systems in which processes are characterized by S-shaped or sigmoidal curves. Such processes typically have two stable steady-state extreme values (corresponding to 0 and 1 in the Boolean idealization) and an unstable intermediate value from which any perturbation drives the value to one of the two extreme values.
Fig. 4. Boolean function ‘EQUIVALENCE’, where element 2 receives input from element 1.

Fig. 5. Sixteen Boolean functions for \( K = 2 \) and their automata representations

In the tabular representations, values along the left-hand side and top of the table denote values of the \( K = 2 \) inputs, and values in the table denote the resulting output value. Consider the ‘OR’ function, for example. Each output

\[ 0 \]

\[ 1 \]

element has two inputs (its own current value and that of one other element) and takes a value of 1 in the next period if either of those elements or both have 1 as their current value, and takes a value of 0 otherwise. In the ‘AND’ function, the output element takes a value of 1 in the next period only if both its input elements’ current values are 1, and takes a value of 0 otherwise.

In the automata representations, circles denoted by 0 or 1 represent the element’s own current value, and labels on the arrows denote the value taken currently by the other input element. Arrows represent the Boolean function’s transition rules that carry an element from one value to another in response to its own value and that of the other input. The automata representations are familiar from their applications in game theory, where the automata conveniently denote strategies. For example, if we allow a value of 0 to correspond to the action of defecting and a value of 1 to correspond to the action of cooperating, the Boolean function in the right-hand column of Fig. 5 depicts the TIT-FOR-TAT strategy made famous by Axelrod (1984, Axelrod (1989): cooperate as long as the other player does, defect if the other player does, and revert to cooperation as soon as the other player does. Similarly, the Boolean ‘AND’ function depicts a GRIM strategy in game theory: cooperate as long as the other player does, but if the other player defects, defect permanently.

Fig. 6 summarizes two Boolean networks consisting of the four elements listed in Fig. 1 as representative of the many characteristics of Chandler’s British and U.S. firms. In each network, each element is assigned an arbitrary Boolean function. For reasons that will become clear below, we label the left panel as a stylized ‘British’ case and the right panel as a stylized ‘U.S.’ case. In the British case, element 1 is regulated by the ‘AND’ function, which sets the element’s value at 1 in the next period only if elements 2, 3, and 4 currently all take their 1 values. Elements 2, 3, and 4 are each governed by the Boolean ‘EQUIVALENCE’ function that switches the regulated element’s value in the next period to match element 1’s current value. Each period, each element is updated based on the values taken by its input elements and its assigned Boolean switching function. For example, state 0001 switches to state 0000 in the next period because element 4’s ‘EQUIVALENCE’ rule changes element 4 from its 1 value to its 0 value to match the 0 at element 1. The left panel’s state-transition graph reports which

---

22 On automata theory, see Hopcroft and Ullman (1979); Gelfand and Walker (1984); Bienenstock et al. (1986); Robert (1986); Soulié et al. (1987); Carroll and Long (1989); Weisbuch (1990); Wolfram (1983) and Wolfram (1984).

23 See, for example, Aumann (1981); Ncyman (1985); Rubinstein (1986); Abreu and Rubinstein (1988); Ben-Porath (1988); Gilboa (1988); Gilboa and Samet (1989); Banks and Sundaram (1990); Young and Foster (1991); Binmore and Samuelson (1992); Ellison (1992); Nachbar (1992); Papadimitriou (1992); Kandori et al. (1993) and Young (1993).
Fig. 6. (left: stylized “British firm” case) State-transition graph illustrating state transitions to successor states with element 1 governed by ‘AND’ and elements 2, 3, and 4 governed by ‘EQUIVALENCE’ with input from element 1, and (right: stylized “U.S. firm” case) state-transition graph illustrating state transitions to successor states with element 1 governed by ‘OR’ and elements 2, 3, and 4 governed by ‘EQUIVALENCE’ with element 1.

states’ trajectories fall on each of the network’s three state cycles, A, B, and C. The network contains three attractors: two steady states (0000, or state cycle A, and 1111, or state cycle C) plus state cycle B which consists of a limit cycle, where the system cycles between 1000 and 0111, rather than a single steady-state type. The left panel illustrates the attractors’ respective basins of attraction (that is, the set of states flowing into each attractor). Once in one of the steady states, the network will remain in that state if left undisturbed.

Boolean networks are subject to two kinds of perturbations: minimal and structural. A minimal perturbation consists of a flip of one binary element to its opposite state. In the left panel, none of the 7 states in state cycle A is stable to all possible minimal perturbations. Flipping one element from 0 to 1 (or from 1 to 0) may move these states out of 0000’s basin of attraction, given the elements’ Boolean functions; for example, perturbing state 0001 to 1001 moves it from state cycle A to state cycle B. However, the states are stable to other minimal perturbations; for example, altering state 0001 to 0011 by flipping its third element does not move it from state cycle A. Similarly, the states on state cycle B are stable to some minimal perturbations, but not to others that shift the states to state

---

24 A trajectory is just the succession of points on the state space as the system changes through time.
25 Attractors include points and limit cycles onto which the system flows.
26 The concept of path dependency in economics (see the literature cited in footnotes 3 and 38) asks when the attractor a system reaches varies with the system’s initial condition. This corresponds to the concept of instability to minimal perturbations.
cycle A or C. State cycle C consists of a single state, obviously unstable to any minimal perturbation.

The second type of stability in a dynamic Boolean network – structural stability – refers to the impact of changes in the input connections or in the Boolean functions governing the elements. As an example, consider the effect of changing the Boolean function governing element 1 from ‘AND’ to ‘OR’ as in the right panel of Fig. 6, the stylized U.S. case. Now any one of elements 2, 3, or 4 taking its 1 value is sufficient to switch element 1 to its 1 value. Comparing the right and left panels makes apparent the instability of the state cycles to this structural perturbation. The change in the Boolean function governing element 1 from ‘AND’ to ‘OR’ shifts six of the seven states that had been in state cycle A (that is, all but the steady state 0000 itself) onto state cycle B. In addition, six of the eight states previously associated with state cycle B now fall on cycle C. The change renders the state 0000 unstable to all minimal perturbations.

This is an example of a dynamic Boolean network in which a small structural change dramatically affects the relative sizes of the basins of attraction. In the left panel, firm type 0000, corresponding to Chandler’s ‘British’ firm – family managed, small, nonintegrated, and unidivisional – drains a relatively large basin, while ‘U.S.’ firm type 1111 exists as an isolated steady state unstable to any minimal perturbation. Because element 1 (management) is governed by the ‘AND’ function in the ‘British’ scenario, firms’ innovations on one or two of the other margins (represented here by elements 2, 3, and 4) fail to switch the management technique to professional; and any switch to professional management gets canceled unless all the other margins take their 1 values.

But a simple change in the Boolean function governing element 1 suffices to alter radically the left panel system’s dynamic behavior. In the right panel’s ‘U.S.’ scenario, the management element now is governed by the ‘OR’ function, making it more responsive to moves on other margins, that is, moves to larger firms, vertical integration, or multidivisional structure. Chandler’s ‘U.S.’ firm type (1111) – professionally managed, large, integrated, and multidivisional – now drains a large basin of attraction, while the ‘British’ firm type (0000) becomes an isolated steady state vulnerable to any minimal perturbation.

Obviously, our assignment of input linkages and Boolean functions in the example is arbitrary, chosen to suggest a greater responsiveness of management technique to innovations on other margins in the United States than in Britain – as indicated by Chandler’s historical case study. However, the differential responsiveness of management technique in the two countries does not suffice to dictate

\[27\] In dynamical systems theory more generally, structural stability refers to systems in which most parameter changes do not cause the system to cross a bifurcation surface into another volume of the space. We are grateful to a referee for noting, under that more general definition of structural stability, our example in Figs. 6(a) and (b) represents a system in which minimal perturbations can cause structural instabilities.
that the modern industrial enterprise would arise in the United States and not in Britain, but merely to imply that a broader range of circumstances were consistent with its rise in America than in Britain. This important distinction can be illustrated in Fig. 6. Chandler's 'modern industrial enterprise' or firm type 1111 could have occurred in Britain, in the United States, or both. But type 1111 drains a much larger basin of attraction in the right panel than in the left panel, making the enterprise's appearance in the stylized U.S. environment more likely.

The further promise of Kauffman's framework lies in its next step. Thus far, we have focused on a single $N = 4$ dynamic Boolean network by arbitrarily assigning particular inputs and a specific Boolean switching rule to each of the four elements, producing results that obviously depend on those arbitrary assignments. But one can construct an entire ensemble of such networks. Discovering the statistical properties of random dynamic Boolean networks involves random assignment of the $K$ inputs to each element and random assignment of the Boolean function to each element. To derive these properties, Kauffman samples randomly from the entire ensemble of possible networks (representing all possible input and function assignments), examines the behavior of the sample networks, and documents that behavior as a function of $N$ and $K$. This new statistical mechanics identifies the average features of all the networks in the ensemble. While these averages may differ substantially from the behavior of any arbitrarily selected single network, they nevertheless represent the ensemble's statistically typical behavior and provide a basis of prediction for the behavior of the system.

5.1. $K = N$: The grand ensemble

This constitutes the largest possible ensemble of Boolean networks containing $N$ variables. Each element acts as input to all others, and each is governed by one of $2^{2N}$ possible Boolean functions. All other ensembles are subsets of this one. The system's behavior is 'chaotic' in three senses. First, the median length of its state cycles $(0.5 \times 2^{N/2})$ increases exponentially in $N$. For large $N$, this implies prohibitively long periods before the system settles into a steady state. Second, neighboring initial states tend to diverge over time. Finally, because the successor to each state is totally random (determined by randomly chosen input elements and a randomly chosen Boolean function), perturbations in the form of random flipping of a single element send the system onto totally different dynamic paths than would have occurred without the perturbation. Nevertheless, these

---

29 We are grateful to a referee for pointing out a problem with Kauffman's use of chaotic to describe behavior in Boolean networks. While truly chaotic systems admit chaotic or strange attractors (see, for example, Kauffman (1993), pp. 178–179), all attractors in Boolean networks exhibit limited periodicity.
"chaotic" systems contain only a small number of attractors \((N/e)\), a few of which typically drain large basins of attraction while the others drain small basins.

5.2. \(K = 2\): Sudden order

When \(K\) falls from \(N\) to 2, the dynamic behavior of the system suddenly becomes orderly. Expected mean state-cycle length falls to about \(N^{1/2}\), a strong contrast to the astronomical lengths of state cycles when \(K = N\). A system of \(N = 10000\) and \(K = 2\), for example, would exhibit expected state-cycle length of 100 states, implying the system restricts itself to only \(1/10^{2998}\) of its entire state space. Even such a large system would contain only approximately \(N^{1/2} = 100\) steady-state attractors. Eighty to ninety percent of all states are stable to minimal perturbations, or flipping of a single element's value. Most of the system's elements (70 percent or more) reach a fixed 0 or 1 state. Structural perturbations cause only small changes in the dynamic behavior of the system.

5.3. Sources of order in dynamic random Boolean networks

There are two basic sources of order that restrict 'chaotic' behavior in dynamic random Boolean networks: canalyzing Boolean functions that create forcing structures, and biased Boolean functions that create homogeneity clusters. Both come into play as \(K\) declines from \(N\) to 2. Both act to create what Kauffman calls a connected mesh or frozen core of elements, each stuck in its 0 or 1 state. This core represents a span of constancy that divides the dynamic system into functionally isolated islands prevented by the frozen core from influencing one another.

Forcing structures arise from canalyzing Boolean functions, that is, those having at least one input with at least one value that suffices to guarantee that the regulated element takes a specific value regardless of other inputs. The functions we have used in our firm-type examples are canalyzing. In the 'OR' function, an input element taking the value 1 is sufficient to force the regulated element to take a value of 1. In 'AND,' an input exhibiting its 0 value forces the regulated element to 0. For 'EQUIVALENCE,' either input value forces the regulated element to the same value. For the \(K = 2\) case, the only noncanalyzing Boolean functions are 'XOR' and 'IFF'. In 'XOR,' the output element takes a value of 1 if one and only 1 of its input elements takes a value of 0. Therefore, neither input possesses a value that suffices to determine the output element's value; the outcome always is sensitive to the value from both inputs, making the function noncanalyzing. Similarly, in 'IFF', the output takes a value of 0 if one

---


and only one of its inputs elements takes a value of 1; so it, too, is noncanalyzing. When all the Boolean functions in a system are canalyzing, once one input element takes its forcing value (say, 1), it forces all the elements that it serves as an input to move to the 1 state. Once a loop of such influences has been forced, it becomes immune to outside influence and constitutes a wall through which disturbances cannot pass. Only islands of unforced elements remain free to move, and they are functionally isolated by the mesh of forcing structures.

The proportion of Boolean functions that are canalyzing decreases rapidly as $K$ increases above 2 (for example, dropping from 87.5 percent for $K = 2$ to less than 5 percent for $K = 4$). Therefore, there are two means of achieving order based on canalyzing functions: restrict $K$ to 2 (which guarantees that a large proportion of the possible Boolean functions will be canalyzing) or restrict the set of Boolean functions to those that are canalyzing.

The two stylized networks of firm types depicted in Fig. 6 rely exclusively on canalyzing functions (‘AND’ and ‘EQUIVALENCE’ in the left panel and ‘OR’ and ‘EQUIVALENCE’ in the right panel). In the left panel, element 1 taking a value of 0 forces elements 2, 3, and 4 to their 0 values, and element 1 taking a value of 1 forces elements 2, 3, and 4 to their 1 values. This forcing structure freezes elements and builds a wall between the basins of attraction that is impenetrable by many disturbances. As a result, the system contains only a few state cycles, each of which is short.

The second source of order in random dynamic Boolean networks is homogeneity clusters based on a subset of Boolean functions – those biased toward producing a particular value. The internal homogeneity of a Boolean function $(P)$ is just the fraction of 0 or 1 values of the $2^K$ states, whichever exceeds 50 percent. Consider the ‘OR’, ‘AND’, and ‘EQUIVALENCE’ functions used in Fig. 6. The ‘OR’ function has three inputs and therefore $2^3 = 8$ possible input combinations. Seven of the eight combinations (all except 000) produce a value of 1; therefore, the ‘OR’ function’s internal homogeneity with $K = 3$ is 87.5 percent. Similarly, the ‘AND’ function with $K = 3$ exhibits an internal homogeneity of 87.5 percent since seven of its eight input combinations (all but 111) produce a 0 value. ‘EQUIVALENCE’ is unbiased because it results in 50 percent 0 values and 50 percent 1 values.

Obviously, if all Boolean functions in a network are biased toward a value, either 0 or 1, the network’s states will tend to converge to the one exhibiting that value at all its elements. Fig. 6 provides an illustration. In the left panel’s network, three of the four Boolean functions are ‘AND’, which we have seen has a

---

33 Fig. 5’s tabular representation of the 16 $K = 2$ Boolean functions provides a useful framework for recognizing the two noncanalyzing functions. ‘XOR’ and ‘IFF’ are the only functions with 0 values on the main diagonal and 1 values on the off-diagonal, or vice versa, reflecting the fact that every outcome is responsive to both input values.

homogeneity bias of 87.5 percent toward its 0 value. As a result, seven of the sixteen states converge to 0000. In the right panel’s network, the bias runs in the opposite direction. Three of the four Boolean functions are ‘OR’, which exhibits a homogeneity bias of 87.5 percent in favor of its 1 value. Therefore, a much larger number of states fall onto 1111’s basin of attraction in the right panel’s network than in the left panel’s network.

Homogeneity bias reduces the expected median state-cycle length from $0.5x2^{N/2}$ to $0.5xP^{-N/2}$, but this still increases exponentially in $N$. The implication is that homogeneity bias alone cannot eliminate ‘chaotic’ behavior in large dynamic random Boolean networks. Like canalizing functions and their forcing structures, the role of biased Boolean functions is not independent of the role of $K$, since the average internal homogeneity of all $2^K$ Boolean functions is at a maximum when $K = 2$ and decreases with $K$.

Thus far, we have assumed that a group’s organization types are insulated on their own fitness landscapes characterized by local optima attainable through hill climbing. However, in many circumstances, choices made by one set of actors reshape or deform the fitness landscape facing others, complicating the process of seeking even local, much less global, optima. This is the situation in Kauffman’s third model, the $N(K + C)$ model of coevolving systems.

6. Coevolution in the $N(K + C)$ model

Coevolution links or couples the fitness landscapes of various groups so that actions by one group alter both the fitness and the fitness landscape of the other. Hill climbing by one group on its fitness landscape may increase or decrease the fitness of other groups as well as change the adaptive walks available to them. Kauffman captures this idea by extending his concept of epistatic interactions. Now, each characteristic’s fitness contribution to one group depends not only on $K$ other intra-group characteristics, but on $C$ characteristics from another group or groups. For each of $N$ characteristics, we must now assign a random fitness contribution for each combination of the $K$ internal characteristics and the $C$ external characteristics. This process implies that the shape of one group’s fitness landscape reflects in part the current location of the other group or groups on their landscapes. Changes in one group’s characteristics not only alter that group’s fitness, but also both the fitness value and fitness landscape for the other group(s). For example, British firms and U.S. firms during the Second Industrial Revolution competed for existing markets and for technological advances creating new ones.

---

35 See Kauffman (1993), Chapter 6, although he does not refer to this extension as the $N(K + C)$ model.
As firms from one country filled market niches or opened new markets, other firms were forced to compete, adapt, or die.

The $N(K + C)$ model of coevolution shares obvious characteristics with game theory: the payoff (or fitness) of each player (characteristic) in one group depends not only on the group’s own other characteristics and how they interact, but on the characteristics of the other group. However, unlike more traditional game-theoretic models, Kauffman’s coevolution framework allows simultaneous consideration of the structure of the individual fitness landscapes, the degree of coupling among different groups’ landscapes, and the number of coupled groups. In general, the model produces two possible behavior patterns. In one, groups evolve continually as their landscapes deform in response to one another’s actions. In the other, groups reach a steady state roughly analogous to a Nash equilibrium. 36

From simulations of randomly connected coevolving systems, Kauffman finds that such systems do reach steady states in which neither group can improve its fitness, given the other, by altering a single characteristic. The larger is $K$ (number of intra-group epistatic interactions) relative to $C$ (number of inter-group epistatic interactions), the shorter the expected waiting time to a Nash equilibrium, reflecting the greater number of local optima on each landscape. More complex systems (larger $N$) experience longer expected times to reach a steady state. As $C$ increases relative to $N$ and $K$, the proportion of systems attaining a Nash equilibrium in a set period declines, as does the fitness of both parties during the pre-Nash phase of the system. In addition, overall average fitness for both parties is higher when $K$ and $C$ are both high or both low, rather than highly unequal. When $C$ is high, high-$K$ parties seem to outperform low-$K$ ones and to allow their coevolving parties to fare better. But when $C$ is low, low-$K$ players gain an advantage and enjoy higher mean fitness.

Consider a simple example based on Chandler’s firms. He argues that the rise of the ‘modern industrial enterprise’ (type 1111) in the United States gave U.S. firms first-mover advantages and foreclosed many important markets to British firms. Although Chandler does not directly address the counterfactual, presumably had the enterprise arisen in Britain, U.S. firms would have found themselves similarly disadvantaged as second movers. The Kauffman framework sharpens our insight into the problem considered by Chandler. Suppose the fitness contribution of each characteristic for both British and American firms depended not only on $K = 3$ intra-group epistatic interactions (as we have assumed throughout), but on $C = 4$ additional epistatic connections with characteristics from the other group of firms. Chandler’s story suggests, in particular, that once U.S. firms took the

36 As Kauffman (1993), p. 245, notes such a steady state differs in one important respect from a Nash equilibrium. The steady state requires only that neither party can improve its fitness by altering a single characteristic, given the characteristics of the other party. A true Nash equilibrium, on the other hand, would require that a party not be able to improve its fitness by moving to any of the $2^N$ other types. We follow Kauffman in referring to the weaker condition loosely as a Nash equilibrium.
'modern industrial enterprise' (1111) form, the fitness contributions of all British-firm characteristics fell, as did the overall fitness of those firms. If so, was the fall uniform across all firm types (that is, a uniform 'sinking' of British firms' fitness landscape), or did the landscape deform, relocating local and global optima and altering hill-climbing routes?

Kauffman also considers the number of groups coevolving. As that number rises, waiting time to a Nash equilibrium increases, mean fitness decreases, and fitness fluctuations increase dramatically. If sudden fluctuations to low fitness result in extinction, then large numbers of interacting groups may set the stage for such events. For example, during Chandler’s Second Industrial Revolution, the growth of international markets brought firms based in different countries into increasingly close contact and competition (that is, $C$ increased). Kauffman’s results suggest that poor performance and even extinction of some firm types, organizational change in response to poor fitness (which, after all, increases the number of uphill adaptive changes available), and stellar fitness by other types might be a predictable result.

7. Complexity and organization

Stuart Kauffman’s work represents just one contribution in a rapidly growing literature in the complexity sciences. Contributions have come mainly from computer scientists, physicists, evolutionary biologists, and biophysicists. 37 A few works by economists have appeared, most concerned with path dependence and technological change. 38 This paper has tried to suggest how insights from the complexity literature might be applied to the economics of organization more broadly. Several characteristics of complexity-based approaches hold promise. Of course, transporting theoretical models and empirical techniques – even ones developed through a highly multidisciplinary process – never is straightforward, easy, or without its pitfalls. Applying complexity theory to the economics of organization is no exception to this general rule.

7.1. Limitations of complexity approaches for the economics of organization

Complexity models provide an approach to systems ‘made up of a large number of parts that interact in a nonsimple way’. 39 Models such as Kauffman’s


focus on systems' statistical properties in order to achieve tractability for analyzing such systems, rather than on providing 'definitive' answers. Kauffman's framework cannot, for example, produce a simple yes or no answer to whether British culture during the Second Industrial Revolution prevented the rise of modern industrial enterprises in Britain. But, of course, neither can alternative approaches, in particular equilibrium-selection models that use (sometimes arbitrary) assumptions to choose among multiple equilibria.

Opportunities for laboratory manipulation in the economics of organization are limited compared to opportunities in Kauffman's biophysics. The precision of the answers Kauffman's framework can provide depends on the amount and quality of information we can bring to bear in a particular application. The model's output will be most revealing in cases where we know the pattern of epistatic interactions and can narrow the subset of Boolean functions to fit the situation. In cases where our ignorance requires totally random assignment of epistatic interactions and Boolean functions, the output will necessarily be only broadly suggestive. But, again, this limitation applies to alternative approaches as well.

Finally, complexity models will have to be enhanced to incorporate the learning, conscious imitation, and elaborate search techniques that individuals and groups utilize in shaping their organizational behavior. The random mutations and one-margin hill climbing of Kauffman's models oriented toward biological evolution obviously represent only a portion of the behavior important in the economics of organization.

Despite these cautions, Kauffman's work and complexity models more generally have properties that deserve the attention of economists interested in organizational issues.

7.2. Strengths of complexity approaches for the economics of organization

Perhaps the fundamental set of questions defining the economics of organization concerns why we observe the organizational variety that we do, why that variety consists of only a subset of all possible organization types, and how and why organizational forms change through time with their environments and other organizations. Despite the long-standing centrality of these questions, many traditional economic models are poorly suited to answering them. Few such models focus on organizational form, and even fewer encompass organizational variety. Kauffman's NK model, in contrast, explicitly draws our attention to the total ensemble of possibilities, embodied in the type space and fitness landscape. In addition, the N-dimensional characterization of organization types invites thought about the relationships among different forms (that is, which types are usefully thought of as 'one-characteristic neighbors' of other types?). The dynamic version

---

40 Dow (1990) provides an interesting example.
of Kauffman's framework provides a nondeterministic vehicle to explore the pattern of change in organizational forms.

A second attractive feature of complexity models, as their very name suggests, is their ability to handle complex systems or multi-part, multi-connection situations in which more traditional approaches quickly become intractable. This applies to large-\(N\) cases of the simple \(NK\) model, as well as to dynamic random Boolean networks as characterizations of nonlinear dynamic systems.

Finally, complexity models take an ensemble approach, to borrow Kauffman's term. They focus on the statistical properties of ensembles of systems. This characteristic provides an opportunity for modeling situations in which we may not know the pattern of interactions among the relevant elements, or may not know the dynamic characteristics of the system. Kauffman's work demonstrates that, even when all epistatic interactions and Boolean functions are modeled randomly, we can still learn a good deal from the statistical behavior of the random system, based on a few simple parameters. At the same time, whatever knowledge we do have of a particular system can be applied – through appropriate nonrandom assignment of the epistatic interactions as well as choice of the Boolean functions from among a subset deemed appropriate given our information. Given such a specification, the model can reveal statistical patterns, indicating what is possible or likely. On many questions central to the economics of organization, such information would represent a significant advance.

Acknowledgements

The authors gratefully acknowledge Amherst College's financial support of this research. We also appreciate an anonymous referee's helpful comments which allowed us to improve the paper.

References


Ellison, Glenn, 1992, Learning, local interaction, and coordination, Harvard University working paper.


Gelfand, Alan E. and Crayton C. Walker, 1984, Ensemble modeling: Inference from small-scale properties to large-scale systems, Marcel Dekker, New York.


Gillespie, J.H., 1984, Molecular evolution over the mutational landscape, Evolution 38, 1116.


Hopcroft, John E. and Jeffrey D. Ullman, 1979, Introduction to automata theory, languages, and computation, Addison-Wesley, Reading, Mass.


Prigogine, Ilya and Isabelle Stengers, 1984, Order out of chaos, Bantam, Toronto.


Young, H. Peyton and Dean Foster, 1991, Cooperation in the short and in the long run, Games and Economic Behavior 3, 145–156.