COMPLEXITY IN ORGANIZATIONS

by

Stephen J. DeCanio*
William E. Watkins**
Glenn Mitchell*
Keyvan Amir-Atefi*
Catherine Dibble***

For information contact:

Stephen J. DeCanio
Department of Economics
University of California
Santa Barbara, CA 93106
805-893-3130
fax: 805-893-8830
E-mail: decanio@econ.ucsb.edu

October 16, 1998

* Department of Economics, University of California, Santa Barbara
** Board of Governors of the Federal Reserve System
*** Department of Geography, University of California, Santa Barbara

The authors are members of the **Computational Laboratories Group, University of California, Santa Barbara**. The research was supported in part by a grant from the United States Environmental Protection Agency, and time on the Cray T3E supercomputer at the Lawrence Berkeley National Laboratory was provided by the National Energy Research Scientific Computing Center (NERSC) funded by the U. S. Department of Energy, Office of Energy Research. We acknowledge helpful comments from Jeffrey Williams, Skip Laitner, John Gliedman, and participants in seminars at the Energy and Resources Group at the University of California, Berkeley, the Lawrence Berkeley National Laboratory, the Environmental Sciences Department at the University of California, Santa Cruz, and the Tellus Institute, Boston. The views expressed in the paper are those of the authors alone, and are not those of the U.S. Environmental Protection Agency, the U.S. Department of Energy, or the Board of Governors of the Federal Reserve System.

ABSTRACT

Algorithmic models specifying the kinds of computations carried out by economic organizations have the potential to account for the serious discrepancies between the real-world behavior of firms and the predictions of conventional maximization models. The algorithmic approach uncovers a surprising degree of complexity in organizational structure and performance. The fact that firms are composed of *networks* of individual agents drastically raises the complexity of the firm's optimization problem. Even in very simple network models, a large number of organizational characteristics, including some whose computation cannot be carried out in polynomial time, appear to influence economic performance. We explore these effects using regression analysis, and through application of standard search heuristics. The calculations show that discovering optimal network structures can be extraordinarily difficult, even when a single clear organizational objective exists and the agents belonging to the firms are homogeneous. One implication is that firms are likely to operate at local rather than global optima. In addition, if organizational fitness is a function of the ability to solve multiple problems, the structure that evolves may not solve any of the individual problems optimally. These results raise the possibility that externally-driven objectives, such as for energy efficiency or pollution control, may shift the firm to a new structural compromise that improves other objectives of the firm as well, rather than necessarily imposing economic losses.

JEL L2, Q4

Keywords: Organization theory, computational complexity, networks.

COMPLEXITY IN ORGANIZATIONS

> *As we shall see…, the parallels between branching evolution in the tree of life and branching evolution in the tree of technology bespeak a common theme: both the evolution of complex organisms and the evolution of complex artifacts confront conflicting "design criteria." Heavier bones are stronger, but may make agile flight harder to achieve. Heavier beams are stronger, but may make agile fighter aircraft harder to achieve as well. Conflicting design criteria, in organism or artifact, create extremely difficult "optimization" problems—juggling acts in which the aim is to find the best array of compromises* (Kauffman 1995, p. 14).

I. Introduction

Computational or algorithmic models of organizational behavior offer an alternative to traditional economic thinking about organizations. The standard theory specifies an "objective function" for the firm[1] and proceeds to maximize that objective function subject to various constraints embodying technological possibilities, availability of information, and the firm's expectations about the state of the world and the reactions of other actors. However, the very act of writing down the objective function and the constraints subsumes a great deal of implicit theorizing. In particular, it abstracts from the ubiquitous *structural interconnections* that characterize all human organizations.[2]

---

1. We will use the "firm" as the typical example of an economic organization. Firms are central to the study of market economies, and profit maximization is the archetype of the traditional economic approach. Nevertheless, it should be kept in mind that our approach carries over to other human organizations (government bureaucracies, non-profits, etc.) with modification only of the criteria by which the organizations survive and evolve.

2. The specialized literature on the theory of the firm frequently goes beyond the simple specification and maximization of an objective function for the firm, and focuses on principal-agent problems and other manifestations of the non-unitary nature of organizations (Williamson 1985; Alchian and Woodward 1988). Considerations of this type are not our central concern, however, because they do not involve modeling the network structure of the firm.

Requiring that organizational behavior be described in terms of algorithmic computation is a natural alternative approach for modeling organizations. Defining the processes and rules of procedure followed by individuals who are members of an organization, together with their patterns of communications and lines of authority, is equivalent in a very real sense to defining the organization. Making explicit the "task" facing the organization, as well as how the agents interact to accomplish the task, amounts to specification of an algorithm that describes the functioning of the organization. Instead of assuming *ab initio* that the organization sets out to maximize a particular objective function, this algorithmic approach allows a richer description of the process by which the choices and behavior of individuals results in collective action (see Olson 1971 for a classic treatment of this distinction).

In the models developed in this paper, we will focus on two kinds of organizational tasks: (1) the building up of shareholder value through adoption of a profitable innovation; and (2) performance of an "associative task," specifically, adding up a set of numbers.[3] In both cases, the key enhancement to the modeling framework is addition of the requirement that information can flow only through defined channels of communication. The problem of optimization is then one of finding the organizational structure that most effectively carries out the organization's task or tasks. The information-processing capability of the individual agents is defined very simply, to maintain focus on the structural issues that are usually ignored. Explicit modeling of the algorithm by which organizations carry out their tasks also enables us to perform experiments using search heuristics characterized by simulated

3. The associative task is akin to production in which a number of parts have to be brought together to produce a finished whole but in which the assembly can take place in any order, hence the task is "associative."

selection, mutation, and reproduction of model organizations in a population over time. Our approach allows us to examine quantitatively such questions as (a) how different organizations compare in their fitness to carry out particular tasks, (b) the computational complexity of the firm's optimization problem, and (c) how market forces and selection pressures shape the development of organizations over time.

It bears emphasis that we do not claim that our stylized models are in any sense "descriptive." Our goal is not to mimic the behavior of real-world organizations, but rather to illustrate *by example* the kinds of complexities that appear once the importance of a firm's network characteristics is acknowledged. Nor do we argue that our computational models are in any sense "better" than others that have been proposed in the literature. The point is that the structural features that enhance or detract from performance depend on the organization's rules of procedure and its environment, and that finding optimal structures is difficult even in very simple cases. This can most clearly be shown with models that are as sparse as possible, so that the conclusions about rule-dependence and complexity follow *a fortiori* for more elaborate and realistic cases.

II. The Model

Mathematically, the organization is defined as a *digraph*, with the agents as vertices and the actual or potential interactions as edges.[4] The edges may be thought of as channels of communication between the members of the organization. Formally, an organization will

---

4. The type of model to be developed here was introduced by DeCanio and Watkins (1998) with a fitness function based on the time taken for an organization to adopt an innovation completely. Hägerstrand (1967) developed a Monte Carlo agent-based model of innovation diffusion within a population. In that model, diffusion was a stochastic process structured by the *a priori* spatial locations of agents. We use a similar approach to *evaluate* network structures as a function of their effective diffusion properties.

be defined as a digraph G consisting of n vertices (members, agents, or individuals) denoted by v_i, $i = 1, \dots, n$, and m edges, (v_i, v_j), where the order of the vertices in an edge means that agent i can receive information from agent j, or that agent i "sees" agent j. These directed edges distinguish a digraph (directed graph) from an ordinary undirected graph in which the edge $\{v_i, v_j\}$ is the same as the edge $\{v_j, v_i\}$. Of course, it is possible for both (v_i,v_j) and (v_j,v_i) to be edges in a particular digraph.[5] It is often convenient to represent the digraph G by its *adjacency matrix* $\mathbf{A}_G = \{a_{ij}\}$, where $a_{ij} = 1$ if agent i sees agent j, and $a_{ij} = 0$ if agent i does not see agent j. It is known that the number of non-isomorphic digraphs with a given number of vertices n grows very rapidly with n; no simple formula exists for the relationship, but the number of digraphs for any given number of vertices can be calculated using the Pólya Enumeration Theorem (Trudeau 1976; Pólya and Read 1987 (1937); Harary 1955). For the remainder of this paper, we will sometimes refer to digraphs simply as "graphs" with the understanding that we are discussing only directed graphs unless we specify otherwise.

A. Adopting a Profitable Innovation.

The first of the tasks modeled is the adoption of a profitable innovation. It is assumed that at any given time, each agent can be in one of two states, designated by 0 and 1, representing non-adoption and adoption of the innovation. The first innovator is selected at random from the members of the organization at time zero. In each successive time period, the agents "see" the others to whom they are connected ("seeing" being the directional linkage) and adopt the innovation with probability P_i, where

---

5. For reasons that will become apparent below, we rule out (v_i, v_i) "loops" in which agents see themselves.

$$P_i\,(1\,|\,0) = f\,[(\Sigma_j\ Y_{ij})/H_i)] \qquad (1)$$

Here, $Y_{ij} = 1$ if agent j is connected to agent i (in the sense that i sees j's state) and agent j is in state 1, while $Y_{ij} = 0$ if agent j is seen by agent i and agent j is in state 0. H_i is the total number of agents seen by agent i and does not change over time. (H_i is the sum of the elements of row i of the adjacency matrix, that is, $H_i = \Sigma_j\ a_{ij}$.) The time subscript is suppressed to simplify the notation, but it should be borne in mind that P_i is recalculated at each discrete time step, and that the state of non-adoption or adoption of the innovation is evolving over time for each agent. Thus, $\Sigma_j\ Y_{ij}$ is the number of agents seen by agent i that are in state 1 at the time step for which P_i is being evaluated. Once an agent switches to state 1, the agent remains in that state permanently.

When an agent adopts the innovation, the agent receives a payoff A. The dollar value of this reward is the same for all agents, and can be thought of as the value of a perpetual annuity beginning at the point in time the innovation is adopted. From the standpoint of the organization, the shareholder value associated with a particular adoption sequence is the present value of the rewards summed over all the members of the organization. That is,

$$\text{🏳}_1 = \sum_i^n\ A/(1 + r)^{\blacklozenge_i} \qquad (2)$$

where $\blacklozenge_i$ is the time period in which agent i adopts the innovation, n is the number of agents in the organization, and r is the discount rate. Thus, the sooner the agents adopt the innovation, the greater the shareholder value of the organization.

The last element of the model is specification of the probability function *f*. If this function is S-shaped and depends on the information processing capability of the agents, then there will be a tension between having a high degree of connectedness within the organization (which would tend to reduce the number of “steps” the innovation must go through as it diffuses within the organization, but which has the downside of exposing the agents to “information overload”) versus having only sparse connectedness (which reduces the information overload problem but increases the number of steps). A suitable functional form for *f* is

$$f(x_i) = 1 / \{1 + e^{-[x_i - (a/c)]/(b/c)}\} - 1 / [1 + e^{(a/b)}] \quad (3)$$

where x_i is the same as the argument of *f* on the right-hand side of equation (1). Here, *a* and *b* are scaling parameters, and *c* is the parameter that specifies the information processing capability of the agents in the organization, given exogenously. The *c* parameter could be related to agents’ human capital, or to the amount of computing power they have available to them. For low *c*, the probability of an agent’s adopting the innovation is low unless a relatively large proportion of the others seen have already adopted the innovation; for high *c*, the agents will adopt with high probability even if only a relatively small fraction of those seen have adopted. Given *c*, the optimization problem for the firm reduces to finding the best structure of linkages between agents. Thus, for high *c*, a completely connected organization will be best, while for very low *c*, the best organization will be one in which each agent sees only one other. The interesting cases are the intermediate ones in which it is not clear how best to connect the members of the organization to maximize the speed of diffusion of the innovation.[6]

B. Solving an Associative Problem.

The second kind of task we set for the model organizations is to find the sum of a set of n numbers (where n is the size of the organization), denoted by $\{z_1, z_2, \ldots, z_n\}$. The numbers are assigned at random to the members of the organization, and the process by which they find the sum is as follows:

(i) At each discrete time step t, pick one of the nodes of the digraph at random and denote it by i. (The time subscript is suppressed in describing the algorithm to simplify the notation.) Let $\{z_{i1}, z_{i2},\ldots,z_{ik}\}$ denote the set of numbers assigned to the k nodes that are seen by node i. (Note that k is the same as H_i in the notation used to describe the adoptive task.)

(ii) Replace z_i with $z_i + \sum_{j=1}^{k} z_{ij}$.

(iii) Reset all the elements of $\{z_{i1}, z_{i2},\ldots,z_{ik}\}$ to zero.

(iv) Pick another node at random and repeat the process (for time step $t + 1$). (The random selection of nodes is done with replacement, so that a node can be picked more than once.)

(v) Stop when only one node has a non-zero value assigned to it.

To give an economic value to the solution of this adding-up problem, it is necessary to specify a cost to each step and to assign a reward for solution of the problem. (If steps

6. For a more extensive discussion of this probability function and its properties, see DeCanio and Watkins (1998).

(i)-(v) could be performed costlessly, the optimal organization would simply be the completely connected digraph and the first iteration would yield the sum.) The cost function for the t'th iteration (the iteration occurring at time t) is defined as

$$C_t = \mu^{k(t)}/(1 + r)^t \qquad (4)$$

where μ is the cost parameter, $k(t)$ is the number of nodes seen by the node being sampled at the t'th iteration, and r is the discount rate. The total cost of solving the problem is the sum of the C_t over all time periods. The payoff for solving the problem is B, so if the problem is solved at time τ^*, the present value for finding the solution is given by

$$\Pi_2 = B/(1 + r)^{\tau^*} - \sum_{t}^{\tau^*} C_t \qquad (5)$$

There is no guarantee that any particular organization will eventually solve the associative problem, so we arbitrarily give the organizations in the simulations 100 time periods or iterations to succeed.[7] With organizations of size 8, this is a sufficient number of periods for the process described by steps (i)-(v) to solve the problem for most trials, provided the organizational network structure is capable of solving it. Because of the discounting, organizations that can solve the problem relatively quickly have an advantage relative to those that take longer. The organizational tension here is between a high degree of connectedness (which tends to lead to a quicker solution) versus the cost of each iteration

7. Indeed, for high enough values of the cost parameter relative to the reward, it can be optimal for the organization to do nothing. If there are no connections (so that nothing ever happens), the problem will never be solved, but cost will be low. This situation might be characterized as the "procrastination equilibrium."

(which goes up non-linearly with the number of other agents seen by the individual sampled).[8]

The same kind of associative problem (viz., adding up a string of numbers) has been explored by Radner (1992, 1993), Van Zandt (1997), and Miller (1996), who examine tree-like structures for solving the problem. The primary difference between our setup and these is the *rules of procedure* by which the organization performs the task. In Radner's model, for example, processing takes place in parallel by all the agents in the organization, and optimality is achieved by minimizing the amount of "idle time" experienced by the members of the organization. In our model, processing takes place sequentially (with the sequence determined at random) and costs are incurred at each step depending on the computational "load," that is, how many numbers the particular individual selected has to add up at once. All of these models are highly stylized representations of an algorithm that will accomplish the task; our algorithm for solving the associative task is not better, just different. Our objective is to show how the structural characteristics of well-adapted organizations differ depending on the rules of procedure, not to design efficient procedures. In the real world, the computational routines followed by organizations will be far more complex and multifaceted than the rules of any of these model organizations, but the simple models of the type we are examining can yield insights into the more general cases.

8. To facilitate comparisons, we set a time limit of 100 iterations for organizations to solve the adoption of technology problem as well. Discounting the rewards for individuals' adoption of the technology provides the positive incentive for early adoption in the solution of that task, as it does in the case of the associative problem.

III. Performance Depends on Structure

To examine the ways in which organizational structure influences performance, we first generated a sample of small organizations (size 8) and regressed the profitability or "fitness" measures π_1 and π_2 on a variety of characteristics of the network structure. A separate regression was performed for each of a number of values of the c parameter (corresponding to different information-processing capabilities of the agents) for the adoption task, and for various values of the μ parameter (representing different levels of cost) for the associative task. Table 1 presents the network characteristics used as regressors, brief descriptions of these variables, and their abbreviated names. This table also contains a brief description of the graph characteristics. More detailed definitions can be found in Garey and Johnson (1979). The Combinatorica (Wolfram Research 1996) add-on package for Mathematica (Wolfram 1996) was used to compute the graph characteristics.

Some discussion of the population from which the sample of digraphs was drawn is required. It would be possible to pick a random sample from the population of all *labeled* digraphs of size 8 simply by allowing each possible edge to be present or absent with equal probability. (A labeled digraph affixes a different "name" to each vertex.) Thus, if there are eight vertices, there would be $2 \times {}_8C_2 = 56$ possible edges, and each digraph in the random sample would be constructed by giving each of these 56 possible edges a 0.5 probability of being present. This would yield 2^{56} possible structures. However, the labeling of vertices has no intrinsic significance, and most graph characteristics are independent of any particular assignment of names to vertices. Two digraphs of the same size are *isomorphic* if one can be transformed to the other (they have the same adjacency

Table 1 - Variable Names and Definitions

Variable Name	Definition
avec	*Average eccentricity*. The eccentricity of a vertex is defined as follows: Find the shortest path from the given vertex to each other vertex in the graph. The eccentricity of the vertex is the longest of these shortest paths. The average eccentricity for a graph is the average over the vertices of the eccentricities of the vertices.
diam	*Diameter*. The diameter of a graph is the maximum eccentricity.
rad	*Radius*. The radius of a graph is the minimum eccentricity.
aved	*Average number of edges*. This is the number of other vertices each agent sees, averaged over all the vertices of the graph.
arvr	*Number of articulation vertices*. An articulation vertex of a graph G is a vertex whose deletion disconnects G.
brdg	*Bridges*. A bridge is an edge such that removing it increases the number of disconnected components in a graph.
edcn	*Edge connectivity*. This is the minimum number of edges whose deletion would disconnect the graph.
vrcn	*Vertex connectivity*. This is the minimum number of vertices whose deletion would disconnect the graph.
mxcl	*Maximum clique*. A clique is a subgraph of G that is completely connected. The maximum clique is the number of vertices in the largest clique.
mnvr	*Minimum vertex cover*. A vertex cover X of size K is a subset of vertices in G such that for each edge in G, at least one of its endpoints is in X. The minimum vertex cover is the number of vertices in the smallest vertex cover.

Sources: See text.

matrix) simply by renaming the vertices while preserving adjacency. If the complete population of digraphs of size 8 is thought of as the population of non-isomorphic digraphs (that is, structures that are intrinsically different), then the population size is smaller than the 2^{56} ($\approx 7 \times 10^{16}$) labeled digraphs; it is known that there are approximately 2×10^{12} non-isomorphic digraphs of size 8 (Wilson 1985). If we were to sample from the population of labeled digraphs, a large proportion of the picks would be isomorphic, so that there would be little variation in the graph characteristics that differ only for non-isomorphic digraphs. This lack of variation in the independent variables would make it difficult to discern any effects of the graph characteristics on organizational performance.

For this reason, we constructed what amounts to a stratified sample as follows. For each digraph in the sample, we first chose a random variable p from the uniform distribution over [0,1]. Then for each of the 56 possible edges, another random variable q from [0,1] was picked, and the edge was set to being present if $q > p$ and was set to being absent if $q \leq p$. The effect of this stratification was to “spread out” the variety of digraphs in the sample. (No two digraphs can be isomorphic if they have a different number of edges.) It is still possible that some isomorphic digraphs would be selected using this procedure, but the fraction of isomorphisms would be smaller than if the edges had been inserted at random with equal probability for all digraphs picked.[9] This procedure was repeated until a sample of digraphs of size 2,000 was drawn. Then for each graph, the average fitness (profitability) measures π_1 and

9. This procedure is similar to what might be done if one were interested in assessing the influence of peoples’ heights on some personal characteristic such as income. Rather than selecting individuals at random from the population (which would yield a high proportion of individuals close to the mean height) and regressing income on height, one would select a stratified sample having equal numbers of individuals of each different height (or height interval). This would increase the variation in the explanatory variable and result in a better test of the null hypothesis.

⚐$_2$ were computed over a series of 2,500 simulations. For the adoption task, the innovation was introduced to a vertex randomly chosen in each run. The fitness measures were averaged to control for random variation due to the initial site of the innovation and the realizations of f (in the case of the adoption model), or the randomness of the sequence of nodes chosen to perform the associative task. Next, the graph characteristics of Table 1 were computed for each of the digraphs. Finally, average fitness was regressed on the set of graph characteristics, including linear and quadratic terms in the regression. Thus, the regression equation estimated (with distinct observations denoted by the subscript s) was

$$y_s = \beta_0 + \Sigma_j \beta_j x_{js} + \Sigma_j \gamma_j x_{js}^2 + u_s, \qquad (6)$$

where y_s is the average profitability of organization (graph) s, the β's and γ's are the coefficients of the independent variables x_j and their squares, and u is the error term. Equation (6) was estimated by ordinary least squares.

It should be noted before presenting the results that computation of two of the variables, *maximum clique* and *minimum vertex cover*, are NP-complete problems. This means that the time required for any known algorithm to compute either of these two variables rises faster than any polynomial function of the number of vertices.[10] It is possible to calculate the *maximum clique* and *minimum vertex cover* values for the sample of digraphs we selected only because the digraphs were small. Even problems whose computation time increases exponentially with problem size can be solved in finite time, and the finite time computation can feasibly be

---

10. For a full discussion of NP-completeness and computational complexity, see Garey and Johnson (1979) and Papadimitriou (1995). The decision problem of determining whether a particular graph has a clique larger than a particular number (of vertices) is known to be NP-complete, and for problems of this type (in which the cost function is relatively easy to evaluate), "the decision problem can be no harder than the corresponding optimization problem" (Garey and Johnson 1979, p. 19). Whether polynomial time algorithms will some day be found that solve problems in the NP-complete class is considered to be one of the most important open questions in theoretical computer science.

implemented on existing hardware if the problem is small enough. Restricting ourselves to organizations of size 8 enables us to test for the effect of these “NP-complete variables” on the performance of the organizations.

Two versions of the regressions were performed in addition to the results reported in Table 2. Three of the variables whose coefficients are shown in Table 2—the radius, diameter, and average eccentricity—can take on infinite values for randomly generated digraphs. (If there are two vertices in the digraph such that there is no path from one to the other, then the eccentricity of the “starting point” vertex is infinite.) The regression results reported in Table 2 exclude all such digraphs, so the sample sizes are less than 2,000. We ran the same regressions without these three variables, with essentially the same results as those reported in Table 2. Similar results were also obtained when the regressions were run for organizations of size 16.

The first notable feature of Table 2 is that the graph characteristics explain a very large fraction of the variance in profitability across the organizations in our sample. The adjusted R^2 is above .94 in all the regressions. Second, it is clear that the coefficients vary across the regressions. That is, the coefficients describing the influence of graph characteristics on performance are different depending on the task, and depending on the capabilities of the agents (as measured either by the agents’ processing capacity c or by the cost parameter ♏).[11] For example, in the regression for associative fitness, the linear coefficient of average eccentricity is positive and statistically significant for the high-cost case (♏ = 10) but negative and significant for the low-cost case (♏ = 3). In the regressions with adoptive fitness as the dependent variable, the

---

Table 2a- Regressions Explaining Adoptive Task Fitness, Various Processing Capability Parameters, Dependent Variable ⚐$_1$ of Equation (2)

c = 0.6

	Explanatory Variable	Coefficient	Std. Error	t	P > |t|
R^2 = .962	avec	- 1.2812	0.1609	- 7.961	0.000
N = 1235	avec2	0.2031	0.0342	5.939	0.000
	diam	0.0482	0.0500	0.964	0.335
	diam2	- 0.0200	0.0082	- 2.428	0.015
	rad	- 0.2936	0.0959	- 3.061	0.002
	rad2	0.0964	0.0294	3.277	0.001
	aved	0.1386	0.0746	1.857	0.064
	aved2	- 0.1108	0.0082	- 13.510	0.000
	arvr	- 0.0236	0.0768	- 0.307	0.759
	arvr2	- 0.0176	0.0317	- 0.555	0.579
	brdg	0.0031	0.0799	0.039	0.969
	brdg2	- 0.0030	0.0330	- 0.092	0.927
	edcn	0.1410	0.0237	5.948	0.000
	edcn2	- 0.0146	0.0034	- 4.298	0.000
	vrcn	0.0281	0.0157	1.794	0.073
	vrcn2	- 0.0023	0.0019	- 1.250	0.211
	mxcl	- 0.0797	0.0305	- 2.613	0.009
	mxcl2	0.0061	0.0035	1.740	0.082
	mnvr	0.7950	0.1227	6.481	0.000
	mnvr2	- 0.0600	0.0104	- 5.718	0.000
	constant	6.4073	0.3627	17.667	0.000

c = 3

	Explanatory Variable	Coefficient	Std. Error	t	P > |t|
R^2 = .995	avec	0.0677	0.0104	6.511	0.000
N = 1235	avec2	- 0.0301	0.0022	- 3.618	0.000
	diam	0.0184	0.0032	5.722	0.000
	diam2	- 0.0041	0.0005	- 7.768	0.000
	rad	- 0.0155	0.0062	- 2.510	0.012
	rad2	0.0079	0.0019	4.157	0.000
	aved	0.1907	0.0048	39.558	0.000
	aved2	- 0.0111	0.0005	- 21.005	0.000
	arvr	- 0.0003	0.0050	- 0.061	0.951
	arvr2	- 0.0011	0.0020	- 0.521	0.602
	brdg	0.0042	0.0052	0.817	0.414
	brdg2	- 0.0021	0.0021	- 1.001	0.317
	edcn	- 0.0048	0.0015	- 3.169	0.002
	edcn2	0.0008	0.0002	3.799	0.000
	vrcn	0.0004	0.0010	0.409	0.683
	vrcn2	0.00001	0.0001	0.094	0.925
	mxcl	- 0.0107	0.0020	- 5.439	0.000

11. Note that Combinatorica computes *arvr* and *brdg* for *un*directed graphs only. The software assumes that every edge is undirected when computing these two variables. Thus, it is not surprising that these variables do not perform well in the regressions.

	mxcl2	0.0014	0.0002	6.121	0.000
	mnvr	- 0.0219	0.0079	- 2.762	0.006
	mnvr2	0.0018	0.0007	2.721	0.007
	constant	6.5807	0.0234	280.957	0.000

Table 2b - Regressions Explaining Associative Task Fitness, Various Cost Parameters, Dependent Variable 🏱$_2$ of Equation (5)

♏ = 10

	Explanatory Variable	Coefficient	Std. Error	t	P > ✹ t ✹
R^2 = .975 N=1235 ♌s ⬦ 1000	avec	10.3043	0.5183	20.037	0.000
	avec2	- 1.5985	0.1102	-14.512	0.000
	diam	- 2.1328	0.1611	-13.241	0.000
	diam2	0.2413	0.0265	9.100	0.000
	rad	7.1397	0.3088	23.120	0.000
	rad2	- 1.3227	0.0948	-13.960	0.000
	aved	2.5292	0.2404	10.521	0.000
	aved2	- 0.3510	0.0264	-13.287	0.000
	arvr	- 0.3103	0.2474	- 1.254	0.210
	arvr2	0.1941	0.1019	1.904	0.057
	brdg	0.2483	0.2572	0.965	0.335
	brdg2	- 0.2233	0.1061	- 2.104	0.036
	edcn	0.2530	0.0763	3.316	0.001
	edcn2	- 0.0093	0.0110	- 0.845	0.398
	vrcn	- 0.0268	0.0550	- 0.531	0.596
	vrcn2	0.0132	0.0060	2.186	0.029
	mxcl	- 0.5264	0.0983	- 5.355	0.000
	mxcl2	0.0596	0.0113	5.292	0.000
	mnvr	- 0.3134	0.3950	- 0.793	0.428
	mnvr2	0.0240	0.0338	0.711	0.477
	constant	-23.6204	1.1679	-20.224	0.000

♏ = 3

R^2 = .940	avec	- 1.0769	0.2528	- 4.261	0.000
N = 1235	avec2	0.0099	0.0537	0.184	0.854
	diam	- 0.3489	0.0786	- 4.441	0.000
	diam2	0.0676	0.0129	5.228	0.000
	rad	1.0484	0.1506	6.961	0.000
	rad2	- 0.2749	0.0462	- 5.949	0.000
	aved	1.8300	0.1172	15.610	0.000
	aved2	- 0.1691	0.0129	-13.127	0.000
	arvr	- 0.0118	0.1207	- 0.098	0.922
	arvr2	0.0065	0.0497	0.131	0.896
	brdg	- 0.0030	0.1255	- 0.024	0.981
	brdg2	- 0.0093	0.1518	- 0.180	0.857
	edcn	0.5836	0.0372	15.679	0.000
	edcn2	- 0.0507	0.0053	- 9.483	0.000
	vrcn	- 0.0454	0.0246	- 1.843	0.066
	vrcn2	0.0053	0.0029	1.809	0.071
	mxcl	- 0.2763	0.0479	- 5.763	0.000

	mxcl2	0.0222	0.0055	4.035	0.000
	mnvr	- 0.2798	0.1926	- 1.452	0.147
	mnvr2	0.0263	0.0165	1.594	0.111
	constant	2.8122	0.5696	4.937	0.000

coefficients of the linear terms for average eccentricity, edge connectivity, and minimum vertex cover all reverse signs from the case of $c = 0.6$ (low processing capacity) to the case of $c = 3$ (higher processing capacity).

Next, it is clear that the NP-complete variables, maximum clique and minimum vertex cover, influence performance of both the adoptive and associative tasks. Most of the individual coefficients of the NP-complete variables and their squares are statistically significant in Table 2, as well as in other regressions (not reported here) for different values of ♏ and c. The linear hypotheses that the coefficients of the NP-complete variables were all zero could be rejected by an F-test in every regression we ran.

These results are not the same thing as a proof that the problem of finding the optimally performing firm is NP-complete. Indeed, for some values of the parameters, the optimum structure for each of the tasks can be derived deductively. If ♏ = 1 for the associative task, for example, the cost of each time step's operation is a constant no matter how many other agents are seen, so a completely connected organization will have the highest present value. (Similarly, for very high c in the adoptive task, the completely connected organization will be optimal.) For low c or high ♏, a sparse and/or hierarchical organization must be best (if ♏ is high enough, the "do nothing" organizational form will minimize cost, because all structures that can solve the problem have larger negative net present values). The total number of possible organizational configurations is finite, so for intermediate values of c or ♏, some degree of connectedness between complete and sparse will work best. There does not appear to be any continuous mapping of the alternative graph structures onto fitness values, however; the regression results show that graphs with the same degree of "connectedness" can have very different performance (otherwise, the only variable that would show up with a statistically significant coefficient in the regressions would be the average number of edges).[12] Finding that optimum may be computationally complex. We conjecture that both tasks are NP-complete or harder for *some range of parameter values*, but we do not know how to delimit those ranges. It appears that the computational complexity of both tasks exhibits phase transitions depending on the values of the c and ♏ parameters, much as other NP-complete problems show phase transitions depending on the value of an "order parameter" (Cheeseman, Kanefsky, and Taylor 1991).[13]

We do know that brute-force search methods will not solve the optimization problem in polynomial time, because the number of non-isomorphic digraphs grows faster than polynomially in the number of members of the organization. An asymptotic approximation for the number of non-isomorphic digraphs is known from published and unpublished work of Pólya (see Harary and Palmer 1973, Chapter 9). In particular, the number d_n of digraphs with n vertices satisfies

$$d_n = [2^{n^2 - n}/n!][1 + 4n(n-1)/2^{2n} + ☺(n^3/2^{3n})] \qquad (7)$$

12. Recall that in all the models we are considering the agents making up the organization are *identical*. This imposes a great deal of symmetry on the problem that is unlikely to correspond to real-world situations. The presence of explanatory variables other than average edges is indicative of the existence of deeper structural determinants of performance. Comparison runs not reported here also show that simple connected graphs (such as the hypercube or k'th order rings (in which each agent i sees agents $i+1$, $i+2$, …, $i+k$), are not optimal performers for ranges of values of the cost and processing power parameters.

13. There is a growing body of literature that finds problems of computational complexity at the heart of standard economic models. See, for example, Deng and Papadimitriou (1994) for the solution concepts of game theory, Papadimitriou and Tsitsiklis (1986) for problems in control theory, Spear (1989) for rational expectations, and Rust (1997) and DeCanio (1998b) for general surveys.

The terms in the second pair of brackets on the right side of (7) are additive, so it is necessary only to show that part arising from the first one grows faster than polynomially. Applying Stirling's formula to the $n!$ in the denominator of (7), one sees that

$$d_n \approx (1/\sqrt{2\pi})2^{n^2 - n - (n + \star)\log_2 n} e^n \qquad (8)$$

from which it is clear that d_n grows faster than polynomially in n.

The substantive economic point is that there may be no practical way to find the "best" organizational form. The performance of any particular structure depends on the task the organization is trying to perform, the cost of its computational procedures, and the processing capabilities of the agents comprising it. In addition, it does not appear to be possible to approximate an optimal solution by continuous techniques. It may be that a structure well-suited to carry out either task is one in which the organization is segmented into subgroups that are connected by a small number of edges. Removal of these connections would split the organization into non-communicating groups, making it impossible to adopt the innovation fully or to solve the associative problem at all. In the regression, the variables measuring edge connectivity and vertex connectivity are statistically significant for some values of c and ♏. The problem of designing the optimal organizational structure is intrinsically a "discrete" problem and, as such, its solution may be subject to the limits imposed by computational complexity.[14]

Exogenous changes in the environment (or new demands on the firm) also will change the optimal structure. Furthermore, if the firm is required to perform multiple tasks, the structure that emerges may not be ideally suited for any of the tasks individually. If a

14. Combinatorial optimization problems are not the only ones subject to the limits of computational complexity. In addition to the paper by Papadimitriou and Tsitsiklis (1986) cited above, see also Ko (1991).

firm is focused on its performance of the associative task, it may adopt a structure that is less suited for adopting profitable innovations. These results all follow from the extremely simplified and stylized computations expressed in our model; it can safely be inferred that the same considerations will extend to real-world organizations of much larger size, which are faced with far more difficult tasks, and which have potential structures of great variety.[15]

IV. Finding Better Organizational Structures

If the problem of optimizing organizational structure is indeed NP-complete or harder, then no known algorithm can guarantee finding the structure with maximal fitness in polynomial time (as a function of organization size), and the actual processes of organizational change can only be expected to lead to improvements, not a global optimum. In the real world, both selection pressures and conscious efforts to restructure organizations contribute to the evolution of performance. Without claiming that we can describe the historical course of organizational change, it is possible to specify computable methods that produce organizations with improved fitness within our model framework. These search heuristics are sometimes called "evolutionary algorithms," although it should be kept in mind that they are not meant to describe the actual mechanisms of organizational evolution.[16] We employ both a Hill Climbing Algorithm (HCA)[17] and a Genetic Algorithm (GA)[18] to compute the results reported below for organizations of size 8.

---

15. Only the simplest forms of communication and information processing represented. In the real world, additional layers of complexity are present. For example, instead of the network linkages being simple "on" or "off" (and represented as 0's or 1's), real linkages could be functions with the range [0,1], where the value of the function indicates the frequency or intensity of communication.

16. Biological or evolutionary descriptions of economic dynamics have been offered byAlchian (1950), Rothschild (1990) and perhaps most extensively by Nelson and Winter (1982) and their followers. De Vany (1997) describes a model of evolution leading to emergent order that has similarities to ours. It is worth noting that the biological/evolutionary metaphor for economic life has long been an alternative to the

A. Search by a Hill-Climbing Algorithm

Hill-climbing algorithms can be set up in various ways. We selected a particularly direct method. Starting with a random graph (constructed with each potential edge being "on" or "off"), we simply picked an edge at random, changed it from its current status, then compared the fitness of the new graph to the previous one. If the average fitness of the new graph over a large number of trials was greater than the fitness of the original graph, the new graph was substituted, and the random change of an edge was repeated. This process was continued for a specified number of "generations," and the graph remaining after that process was considered to the product of that particular run of the HCA.

In order to determine whether this search process was effective, the entire algorithm was repeated, thereby generating a sample of "adapted" graphs. The characteristics of graphs in this sample were then compared to the graph characteristics of the original random graphs. The results of these comparisons are presented in Table 4. Several inferences can be drawn. Consider first the associative task: for all three cost parameters represented, the graphs resulting from 100 steps of the hill climb show clear differences from the sample of random graphs representing the starting points. The low value of the cost parameter (♏ = 1) corresponds to the case in which we know that the optimal graph is completely connected, and indeed the search process moves in that direction. The average

---

rational choice models that are most frequently invoked by economic theory. Krugman (1996) provides a very nice discussion of the similarities between *conventional* economic models and biological models.

17. Simulated Annealing, an algorithm having some similarities to both Hill Climbing and Genetic Algorithms, is not used in this paper. For a recent application of Simulated Annealing to a question of political economy, see Kollman, Miller, and Page (1997).

18. See Goldberg (1989) and Holland (1992) for an introduction to Genetic Algorithms.

number of edges in the adapted sample equals 5.84, versus 3.58 in the initial random sample. It is interesting to note, however, that even after 100 generations, the HCA has not (on average) reached the optimal completely connected graph with an average number of edges equal to 7.

It is clear, however, that the HCA search process tends to move in the right direction. For the low value of ♏, the organizations evolve towards greater connectedness, while for the high-cost value of ♏, evolution moves towards sparser connectedness. But more is involved in determining organizational fitness than the average number of edges. The other graph characteristics listed in Table 4 are also different in the adapted samples than in the original samples of randomly constructed graphs. The signs of the Z-statistics corresponding to the tests of equality of means are almost always opposite when ♏ = 1 and when ♏ = 10, especially in the cases in which the difference in means between the original and final populations is statistically significant. An exception is that in all cases, the number of bridges falls in the adapted population. (This is likely due to the fact that the *brdg* variable is computed as if the graphs were undirected. See footnote 11.) Graphs with improved performance are characterized by features other than simply the average number of edges.

Except for two cases (both involving the minimumvertex cover), the variances of the characteristics of the adapted graphs are lower than the variances of the characteristics of the randomly selected initial graphs. This indicates that not only are the adapted graphs different from the initial population, but that the “good” organizational forms produced by the HCA are more alike than those in the general population.

Table 4 - Comparison of Sample Statistics for Random and Adapted Populations, Hill Climbing Algorithm

	Associative Task					***Adoptive Task***				
	Random Population		Adapted Population			Random Population		Adapted Population		
	Smpl. Mean	Smpl Std. Dev.	Smpl Mean	Smpl. Std. Dev.	Z for H_0: $\mu_R=\mu_A$	Smpl. Mean	Smpl. Std. Dev.	Smpl. Mean	Smpl. Std. Dev.	Z for H_0: $\mu_R=\mu_A$
	♏ = 1					***c* = 9**				
arvr	0.55	0.87	0.02	0.14	6.01**	0.54	0.96	0.00	0.00	5.63**
aved	3.58	1.98	5.84	0.58	-11.0**	3.74	2.11	6.45	0.41	-12.6**
brdg	1.10	2.04	0.04	0.28	5.15**	1.10	2.21	0.00	0.00	4.98**
mxcl	2.93	1.49	4.60	1.07	-9.10**	3.16	1.79	5.80	1.23	-12.2**
mnvr	5.14	1.58	6.48	0.56	-7.99**	5.15	1.72	6.82	0.39	-9.47**
edcn	2.22	1.96	4.66	1.09	-10.9**	2.35	2.15	5.35	1.06	-12.5**
vrcn	2.26	2.12	4.44	1.39	-8.60**	2.54	2.31	5.73	1.43	-11.7**
	♏ = 3					***c* = 3**				
arvr	0.58	0.89	0.35	0.76	1.97*	0.68	0.97	0.00	0.00	7.01**
aved	3.36	2.16	4.00	2.11	-2.12*	3.55	2.07	6.41	0.46	-13.5**
brdg	1.26	2.02	0.10	0.44	5.61**	1.62	2.47	0.00	0.00	6.56**
mxcl	2.91	1.66	2.88	1.49	0.13	3.13	0.69	5.77	1.36	-17.3**
mnvr	4.87	1.86	4.91	2.42	-0.13	5.12	1.49	6.81	0.39	-11.0**
edcn	1.93	2.07	3.29	1.85	-4.90**	2.16	2.11	5.42	1.04	-13.9**
vrcn	2.09	2.18	2.96	1.76	-3.11**	2.17	2.31	5.89	1.43	-13.7**
	♏ = 10					***c* = 0.6**				
arvr	0.61	0.90	0.59	0.82	0.16	0.56	0.92	0.26	0.46	2.92**
aved	3.42	2.22	2.03	1.46	5.23*	3.54	1.95	2.29	0.29	6.34**
brdg	1.62	2.14	0.82	1.78	2.87**	1.40	2.55	0.44	0.83	3.58**
mxcl	2.97	1.91	1.96	1.47	4.19**	2.97	1.71	1.93	0.29	6.00**
mnvr	4.85	1.79	4.22	1.93	2.39*	5.15	1.46	4.99	0.36	1.06
edcn	2.06	2.24	1.41	1.49	2.42*	2.13	2.06	1.27	0.45	4.08**
vrcn	2.25	2.37	1.09	1.73	3.95**	2.18	2.18	1.01	0.66	5.14**

* Probably-value < 0.05 under H_0 of no difference in population means.
** Probably-value < 0.01 under H_0 of no difference in population means.

The same sort of results hold true for the adoptive task. When the agents have high processing capacity (c = 9), the search process finds structures of greater connectivity (an increase in average edges), and when agents' processing capacity is poor (c = 0.6) adaptation proceeds in the opposite direction. In this case, however, the number of articulation vertices and bridges moves in the same direction for the c = 9 case and the c = 0.6 case.

For most parameter values in our model, it seems unlikely that the HCA will be able to find a globally optimal structure. In models exhibiting rugged fitness landscapes, the HCA is quite likely to get stuck on a local maximum. The search algorithm can only switch one edge "on" or "off" at a time, and hence will never be able to make a "jump" to a higher local maximum that differs in structure by more than two edges. Thus, this search procedure can in general only be expected to find structural improvements, not the globally optimal organizational form.

B. Search by a Genetic Algorithm.

Genetic Algorithms operate in a manner that imitates sexual reproduction in biological populations. In addition to allowing for the possibility of random mutations (which are akin to the bit flips occurring in the HCA), a GA allows entire sets of genes to be exchanged between members of a population, so that offspring are created having a mixture of the genes of their parents. If the fitter members of the population are differentially selected for reproduction, it is possible to search over a wider range of the space of genotypes than can be reached by relying on random mutation alone. GAs are preferred to

HCAs in situations where non-local jumps are required to reach higher points on the fitness landscape (Goldberg 1989).

In our model, the structure of a particular organization is naturally encoded by specifying its chromosome simply as the row vector obtained by concatenating the successive rows of its adjacency matrix. Fitness is calculated as in Section II above, depending on whether the adoption problem or the associative problem is being solved. We used a GA with settings for reproduction, crossover, and mutation in the range of magnitudes suggested by Grefenstette (1986).[19] The GA behaved normally, with rapid improvement in the early generations followed by a leveling off of both the population average fitness and the fitness of the best-performing organization. Organizational innovations after the early rapid increases in fitness appear as discrete "jumps" in fitness levels that are otherwise largely flat over time.

We applied the GA repeatedly for each of the values of the cost or processing capacity parameters that were used in Table 4. Each GA run was begun with a size-100 "uniformly distributed" stratified random population of the type described in Section III. In each run, we let the GA operate for 100 generations, then selected the "champion" structure with the highest fitness as the outcome of that search. This procedure was carried out 100 times for each separate value of c or ♏, to obtain a population of 100 champion organizations for each of the parameter values. We then computed statistics on the fitness scores of the members of the adapted populations of champions.

Table 5 compares the fitness statistics for the original and adapted populations for both search procedures. Several features of these results are interesting. First, application of either search algorithm increases population fitness considerably. For example, for low processing capability in the adoptive task ($c = 0.6$), the GA increases average population fitness by 53%, while the HCA increases average population fitness by 45%. Second, both the GA and the HCA "converge" in the sense that the variance in fitness of the adapted populations is much smaller than the variance of the original populations. Third, the search algorithms are in some cases able to find the structure yielding the theoretical maximum present value of the firm. For example, the minimum number of steps for complete adoption of the profitable innovation is 2, with an associated present value of 7.3636. For $c = 3$ and $c = 9$, the best organizations are at or quite near this maximum value after 100 generations, although the population averages fall slightly short of this maximum. In the case of $c = 0.6$, the best of the GA champions is 7.2% below the theoretical maximum and the average of the champions is 7.5% below the theoretical maximum. The best organization from 100 runs of the HCA is 7.5% below the theoretical optimum, while the population average is 8.6% below the optimum. For the associative task, the theoretical maximum present value if ♏ = 1 is 9,090, which is achieved by the best organizations in both the random and adapted populations. However, the theoretical optimum depends on the value of the cost parameter if ♏ > 1. This optimum is not known exactly, but Table 5 shows that, for example, in the case of the high cost parameter (♏ = 10) the GA champions are consistently able to show a positive present value, despite the fact that the average fitness for the initial set of populations is quite negative.

---

19. Details are available from the authors.

Table 5 - Fitness Statistics for Random and Adapted Populations

	Associative Task				**Adoptive Task**			
Fitness Statistic	Genetic Algorithm		Hill Climbing Algorithm		Genetic Algorithm		Hill Climbing Algorithm	
	Gen. 1	*Gen. 100*	*Gen. 1*	*Gen. 100*	*Gen. 1*	*Gen. 100*	*Gen. 1*	*Gen. 100*
	♏ = 1				**c = 9**			
mean	4,563	9,079	4,801	8,357	5.8202	7.3635	5.9358	7.3182
median	5,370	9,090	5,926	8,394	7.0351	7.3636	7.0620	7.3220
std. dev.	3,317	37.26	3,256	420	2.0737	0.000045	2.0678	0.0341
max	9,090	9,090	9,090	9,090	7.3636	7.3636	7.3636	7.3636
min	-9.274	8,851	-9.274	7,283	1.0000	7.3634	1.0000	7.2083
	♏ = 3				**c = 3**			
mean	3,760	7,083	3,496	5,354	5.7755	7.3635	5.8573	7.3146
median	4,750	7,103	4,428	6,750	7.0274	7.3635	7.0090	7.3219
std. dev.	2,680	39.93	2,825	2,785	2.1102	0.000055	1.9380	0.0380
max	7,103	7,103	7,103	7,148	7.3636	7.3636	7.3635	7.3635
min	-890	6,923	-165.7	-16.76	1.0000	7.3634	1.0000	7.1647
	♏ = 10				**c = 0.6**			
mean	-2,133,431	2,063	-2,368,339	- 453,780	4.4496	6.8105	4.6316	6.7289
median	-166,244	2,204	- 102,130	624	4.8021	6.8122	4.7535	6.7345
std. dev.	3,056,876	600	3,209,361	1,989,975	1.5043	0.0123	1.3370	0.0569
max	1,242	2898	- 9	2,232	6.7105	6.8346	6.4944	6.8109
min	-9,130,480	-31.85	-9,093,000	-9,081,820	1.0000	6.7762	1.0000	6.4998

The substantive economic point is that adapted populations will in general show a *range* of fitness values, even if selection pressure has been operating for a considerable period of time in a stable environment. Not all or even many of the surviving firms achieve the highest attainable level of fitness (nor is there any reason to expect them to). The GA champions seem to generally have a higher average level of fitness than the population of HCA survivors (although the best HCA result for the associative task with ♏ = 3 was better than any of the GA champions), but this has no clear economic meaning because much more computation is involved in the GA searches than in application of the HCA. (The total "initial population" for all the GA runs for one of the parameter values consists of 10,000 organizations, while the HCA initial population has only 100.) The performance of the GA relative to the HCA may be a reflection of the fact that, as in biological populations, diversity of genotypes provides resiliency and the "genetic raw material" for a variety of new organizational forms that may have fitness advantages.

V. Conclusions and Potential Policy Implications

Organizations exhibit a degree of complexity that is not incorporated into standard economic models. Recognizing the network structure of communication within a firm, and requiring that performance of the firm's productive tasks be carried out by well-specified computational algorithms, projects the theory of the firm into the realm of combinatorial optimization and computational complexity. This conclusion is important in itself, but it also raises a set of issues related to the way economic phenomena are conceptualized.

Economic theorizing and policy analysis are informed by what characteristics one does or does not expect idealized model firms to exhibit. In the conventional neoclassical approach, firms are expected to maximize profits subject to their market and technological constraints. The firms' productive configurations are presumed to be efficient in the sense that one kind of output can be increased only if another is reduced. Firms should be identical unless specific differences in endowments or access to technology can be identified. In contrast, models whose firms' optimal organizational structures are too complex to be determined exactly will have different properties. In such models, the heterogeneity of firms and measurable differences in their performance are a natural consequence of whatever imperfect search heuristics are used to improve performance. Organizational structure will influence profitability, sometimes in non-intuitive ways. Path dependence will be the rule, as firms' computational algorithms are built up from components that have been developed and tested in past circumstances.

As in the case of biological organisms, selection, not optimization, would be the driving force behind the emergence of order in this view. Selection (at least in market economies) creates powerful pressures for improvements in organizational performance, but even if optimal network structures could be found for coping with the economic environment at a particular moment in time, that environment shifts so rapidly that the useful adaptations of yesterday can become liabilities today. From this perspective, it is possible to make theoretical sense of phenomena such as the "efficiency paradox" (DeCanio 1998a and the literature cited therein; DeCanio and Watkins 1998) or the organizational change required to realize the productivity benefits of computerization (Brynjolfsson and Hitt 1996). If firms are faced with multiple tasks that are computationally complex, there is no reason to expect that a perfectly profit maximizing set of investments (in energy efficiency or any other technologies) will be made.

Giving greater explanatory weight to the structural characteristics of organizations also has policy implications. The efficiency paradox is important because it matters a great deal whether, for example, substantial reductions in greenhouse gas emissions can be accomplished without large cost (in terms of conventionally measured output) to the economy. Charging a price for the emissions through a carbon tax or emissions permits would internalize the greenhouse externality, but might do so only at the expense of a reduction in profits or income as conventionally measured. But what if maximization of profits or utility with respect to energy usage or technology choices is not to be *expected?* In that case, a whole range of unconventional policy options to reduce greenhouse gas emissions could become feasible and attractive. Policies that would influence the network structures of firms and markets or that would induce jumps from the current points on the firms' fitness landscapes (possibly at local optima) to other local optima (perhaps ones that are more energy efficient) might be possible. Moving to a local optimum that is environmentally more benign may have no impact or even a positive impact on the other dimensions of a firm's performance, as suggested by the Porter Hypothesis (Porter 1990, 1991; Porter and van der Linde, 1995a, 1995b).

Such possibilities remain conjectural. What has been demonstrated is that the introduction of network structure adds a relatively unexplored layer of complexity to the theory of firms and other economic organizations. Search heuristics that are beginning to see widespread application in economics and other fields offer one promising approach to examining how the performance of complex organizational structures might be improved. Further characterization of the properties of more or less efficient structural forms may be possible through application of measures that have been developed in mathematical graph theory and the study of networks. But even if better static descriptions of the functioning of organizations can be developed with such methods, an important area for further research is to develop algorithms that model the actual processes of organizational evolution and change.

REFERENCES

Alchian, Armen A., 1950 "Uncertainty, Evolution, and Economic Theory." *The Journal of Political Economy*, Vol. LVIII, No. 3 (June): 211-221.

Alchian, Armen A., and Susan Woodward, 1988. "The Firm is Dead; Long Live the Firm: A Review of Oliver E. Williamson's *The Economic Institutions of Capitalism*," *Journal of Economic Literature*, March 1988, Vol. 26, No. 1: 65-79.

Brynjolfsson, Erik, and Lorin Hitt, 1996. "Paradox Lost? Firm-level Evidence on the Returns to Information Systems Spending," *Management Science*, Vol. 42, No. 4 (April): 541-558.

Buckles, Bill P., and Frederick E. Petry, 1992. *Genetic Algorithms*. Los Alamitos, CA: IEEE Computer Society Press.

Cheeseman, Peter, Bob Kanefsky, and William M. Taylor, 1991. "Where the *Really* Hard Problems Are," in J. Mylopoulos and R. Reiter, eds., *Proceedings of IJCAI-91*. San Mateo, CA: Morgan Kaufmann, pp. 331-337.

DeCanio, Stephen J., 1997. "Economic Modeling and the False Tradeoff Between Environmental Protection and Economic Growth," *Contemporary Economic Policy*, Vol. XV, No. 4 (October): 10-27.

______________, 1998a. "The Efficiency Paradox: Bureaucratic and Organizational Barriers to Profitable Energy-Saving Investments," *Energy Policy*, Vol. 26, No. 5 (April): 441-454.

______________, 1998b. "Estimating the Non-Environmental Consequences of Greenhouse Gas Reductions is Harder Than You Think," unpublished manuscript, University of California, Santa Barbara.

DeCanio, Stephen J., and William E. Watkins, 1998. "Investment in Energy Efficiency: Do the Characteristics of Firms Matter?" *The Review of Economics and Statistics*, Vol. 80, No. 1: 1-13.

______________, 1998. "Information Processing and Organizational Structure," *Journal of Economic Behavior and Organization*, Vol. 36, No. 3 (30 August): 275-294.

De Vany, Arthur, 1997. "Information, Chance, and Evolution," in *Uncertainty and Economic Evolution: Essays in honor of Armen A. Alchian*, ed. John R. Lott, Jr. London: Routledge.

Garey, Michael R., and David S. Johnson, 1991. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York: W. H. Freeman and Company.

Goldberg, David E., 1989. *Genetic Algorithms in search, optimization, and machine learning*. Reading, MA: Addison-Wesley.

Grefenstette, John J., 1986. "Optimization of Control Parameters for Genetic Algorithms," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 16, No. 1: 122-128. Reprinted in Buckles and Petry (1992).

Hägerstrand, Torsten, 1967. *Innovation Diffusion as a Spatial Process*. Chicago: University of Chicago Press.

Harary, Frank, 1955. "The number of linear, directed, rooted, and connected graphs," *Transactions of the American Mathematical Society*, Vol. 78, No. 2 (March): 445-463.

Harary, Frank, and Edgar M. Palmer, 1973. *Graphical Enumeration*. New York: Academic Press.

Holland, John H., 1992. *Adaptation in Natural and Artificial Systems*. Cambridge, MA: MIT Press.

Kauffman, Stuart A., 1995. *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. New York: Oxford University Press.

Ko, Ker-I, 1991. Complexity Theory of Real Functions. Boston: Birkhäuser.

Kollman, Ken, John H. Miller, and Scott E. Page, 1997. "Political Institutions and Sorting in a Tiebout Model," *American Economic Review*, Vol. 87, No. 5 (December): 977-992.

Krugman, Paul, 1996. "What Economists Can Learn from Evolutionary Theorists," talk given to the European Association for Evolutionary Political Economy (November). Available at http://web.mit.edu/krugman/www/evolute.html.

Miller, John H., 1996. "Evolving Information Processing Organizations," manuscript, Department of Decision and Social Sciences, Carnegie Mellon University.

Nelson, Richard R., and Sidney G. Winter, 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press, 1982.

Olson, Mancur, 1971. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge: Harvard University Press.

Papadimitriou, Christos H., 1995. *Computational Complexity*. Reading, MA: Addison-Wesley Publishing Company.

Papadimitriou, Christos H., and Xiaotie Deng, 1994. "On the Complexity of Cooperative Solution Concepts," *Mathematics of Operations Research*, Vol. 19, No. 2 (May): 257-266.

Papadimitriou, Christos H., and John Tsitsiklis, 1986. "Intractable Problems in Control Theory," *SIAM Journal of Control and Optimization*, Vol. 24, No. 4 (July): 639-654.

Pólya, George, and Ronald C. Read, 1987 (1937). *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*. New York: Springer-Verlag.

Porter, Michael E., 1990. *The Competitive Advantage of Nations*, Free Press, New York.

___________, 1991. "America's Green Strategy: Environmental Standards and Competitiveness," *Scientific American*, Vol. 264, No. 4 (April): 168.

Porter, Michael E., and Claas van der Linde, 1995a. "Toward a New Conception of the Environment-Competitiveness Relationship," *Journal of Economic Perspectives*, Vol. 9, No. 4 (Fall): 97-118.

___________, 1995b. "Green and Competitive: Breaking the Stalemate," *Harvard Business Review*, Vol. 73, No. 5 (September-October): 120-134.

Radner, Roy, 1992. "Hierarchy: The Economics of Managing," *Journal of Economic Literature*, Vol. 30 (September): 1382-1415.

______________, 1993. "The Organization of Decentralized Information Processing," *Econometrica*, Vol. 61, No. 5 (September): 1109-1146.

Rothschild, Michael, 1990. *Bionomics: Economy as Ecosystem*. New York: Henry Holt and Company.

Rust, John, 1997. "Dealing with the Complexity of Economic Calculations," paper for workshop on "Fundamental Limits to Knowledge in Economics," Santa Fe Institute, August 3, 1996.

Spear, Stephen E., 1989. "Learning Rational Expectations Under Computability Constraints," *Econometrica*, Vol. 57, No. 4 (July): 889-910.

Trudeau, Richard J., 1976. *Dots and Lines*. The Kent State University Press.

Van Zandt, Timothy, 1997. "The Scheduling and Organization of Periodic Associative Computation," unpublished manuscript, Princeton University (January 28).

Williamson, Oliver E., 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press.

Wilson, Robin J., 1985. *Introduction to Graph Theory* (3rd ed.). New York: Longman Group Limited.

Wolfram, Stephen, 1996. *The Mathematica Book* (3rd ed.). Champaign, IL: Wolfram Media/Cambridge University Press.

Wolfram Research, 1990. *Mathematica 3.0 Standard Add-on Packages*. Champaign, IL: Wolfram Media/Cambridge University Press.