

# Near Decomposability and the Speed of Evolution

Herbert A. Simon  
Department of Psychology  
Carnegie Mellon University

Complex Information Processing  
Working Paper #549

# Near Decomposability and the Speed of Evolution

Herbert A. Simon  
4 April 2000

This paper proposes an explanation for a fundamental property that appears to be shared by all multi-celled organisms. Such organisms consist of a hierarchy of components, such that, at any level of the hierarchy, the rates of interaction within components at that level are much higher than the rates of interaction between different components. Systems with this property are called nearly completely decomposable, or more briefly, nearly decomposable (ND). The explanation for the ubiquity of the ND property is that, under the usual conditions of mutation and/or crossover and natural selection, ND systems will increase in fitness, and therefore reproduce, at a much faster rate than systems that do not possess the ND property.

## *Introduction*

In contemporary molecular biology, the evolutionary process is expressed in its most general form by a fitness function,  $F$ , of the genome and the environment:  $F(G, E)$ . Fitness is measured by the relative rates of growth of the populations of different genomes, which are determined, in turn, by natural selection.

Genomes, which can be represented by vectors of genes, change, from generation to generation, through the workings of such stochastic mechanisms as mutation and crossover; and there is also continual change in the environment (the *fitness landscape*), defined by the ecology of co-existing evolving plant and animal species, and by the geological, atmospheric, oceanic and astrophysical environment.

Describing the environment as a fitness landscape calls attention to the fact that, for survival and proliferation, an organism must obtain enough energy from the environment to sustain its metabolic processes and to reproduce; the organism's metabolic and reproductive requirements are in turn determined by its processes for energy acquisition and use. The wolf must hunt successfully in order to metabolize and breed, and it must metabolize efficiently in order to maintain sensory and motor processes that enable it to hunt. The law of conservation of energy rules here as elsewhere in nature.

Just as certain environments are friendly to the survival and reproduction of a great variety of organisms (e.g., tropical jungles and warm shallow seas), and other environments are austere (deserts, polar wastes and ocean trenches), so certain organismic architectures may be more conducive to survival and proliferation, over a wide range of environments, than others. Quantities of water and solar energy received (and their seasonal fluctuations) might characterize the fecundity of fitness landscapes; while efficiency in energy use and capability for energy acquisition (in relation to size) might characterize the generalized propensity of organisms for fitness. A number of evolutionary speculations, e.g., about the sudden emergence of new phyla in the "Cambrian Explosion" (and the emergence of almost no new phyla since), suggest that a limited number of Ur-forms are exceptionally well designed for high fitness. Recently there has been an extension of theories of "optimal" body form, and empirical testing of these theories, with considerable success in fitting the empirical data for a wide range of species.

### *Nearly Decomposable Complex Systems*

In examining the general conceptual framework just sketched, I will focus on one specific architectural feature of organisms as central to determining speed of adaptation and increase in fitness. I will argue that an important and very basic property is shared by almost all complex systems observed in nature (inorganic as well as organic, stretching from elementary particles to social systems), and including the whole range of Ur-forms for multi-celled organisms that the Cambrian "revolution" produced. This common architectural feature both confers greater fitness on individual organisms; and, most important, it accelerates the rate at which the fitness of organisms possessing it increases over time through the standard processes of genome change and natural selection.

The property in question is called *nearly complete decomposability* *NCD*, which I will abbreviate to *near decomposability* (*ND*) (Simon & Ando, 1961). To characterize it, let me propose a metaphor: We imagine a large building, with very many rooms (I'll call it the Mellon Institute). The rooms are large with thick exterior and interior walls, and most of them are divided into smaller cubicles, with thinner walls. By some violent magic (a tornado, perhaps?) we force it suddenly into a condition of extreme temperature disequilibrium, so that each cubic centimeter of air has a distinctly different temperature from each adjoining cubic centimeter, each cubicle a sizeable temperature difference from each adjoining cubicle, each room from each adjoining room, and the whole set

of rooms from the outdoors. We hold the outdoor temperature constant, shut off the heating and air conditioning equipment, close all doors and observe what happens.

Very rapidly, of course — perhaps in a matter of minutes — the temperatures of all the cc's of air in any single cubicle will become essentially equal. By the end of an hour, the temperature of all the cubicles in a given room will be the same. By the end of eight hours, the temperatures of all the rooms will be about the same. And by the end of a day, all the rooms will be at the same temperature as the outside air. The exact times of equilibration of the Mellon Institute would depend on the Newton coefficients of heat transmission through the walls and ceilings, but the sequence is clear.

The system I have just described is the archetype of a nearly decomposable system. It can be thought of as a boxes-within-boxes hierarchy with an arbitrary number of levels. Its special characteristic is that equilibrating interactions *within* boxes at any level take place much more rapidly than the interactions *between* boxes at that same level, and similarly all the way to the top of the hierarchy (Table 1).

Table 1 illustrates a ND system of three layers, with two cells at the top level, each divided into two sub-cells, and each of these subdivided into three sub-sub cells. The numbers of the rows and columns designate the cells. cells 1-6 and 7-12 constitute the two top-level subsystems, cells 1-3, 4-6, 7-9 and 10-12 the four-second level subsystems. The interactions within the latter subsystems have intensity  $\alpha$ , those within the former two subsystems, intensity  $\epsilon_1$ , and those between components of the largest subsystems, intensity  $\epsilon_2$ .

## A Nearly-Decomposable System

	1	2	3	4	5	6	7	8	9	10	11	12
1	-1	a	a	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$
2	a	-1	a	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$
3	a	a	-1	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$
4	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	-1	a	a	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$
5	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	a	-1	a	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$
6	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	a	a	-1	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$
7	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	-1	a	a	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$
8	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	a	-1	a	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$
9	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	a	a	-1	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$
10	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	-1	a	a
11	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	a	-1	a
12	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_2$	$\epsilon_1$	$\epsilon_1$	$\epsilon_1$	a	a	-1

**Table 1:** In a mathematical model of the steady state of a dynamic system, each row would represent the left side of an equation that is equal to a constant, and each column a variable. The diagonal element is the dependent variable in the equation of the corresponding row. Typically, the dependent variable and other variables with negative coefficients will represent the outputs of a process defined by the row, and the other elements in the row (the inputs), will be positive.

The dynamic, non-equilibrium, behavior of the Mellon Institute example can be represented by a simple set of differential equations (Simon & Ando, 1961, Iwasaki & Simon, 1994). I will consider the special case where the equations are linear (but the qualitative results can be extended to more general cases):

$$dT_i/dt = \sum_j (a_{ij}T_j - T_i) ,$$

where  $i$  and  $j$  designate the elements at the lowest level, and the  $a_{ij}$ 's (the  $a$ 's and  $\epsilon$ 's of the table), with  $a_{ij} = a_{ji}$ , are, in the Mellon Institute example, the coefficients of heat transmission.

In Figure 1, the matrix of coefficients of the equations is arranged in near-diagonal form. The larger coefficients will soon bring about equilibrium of the subsystems that lie in the small blocks along the diagonal, and the epsilons will bring about, more slowly and successively, equilibrium at the two layers of larger blocks that are composed of the smaller ones. For the example, this approximating procedure first finds the equilibrium temperature for each cubicle, then averages these to find the equilibrium temperature for each room, then averages again to find the equilibrium temperature for the whole building before equilibrating with the outside temperature.

The corresponding computational consequence is that, to a first approximation, we can solve each of the static equations ( $dT/dt = 0$ ) corresponding to one of the innermost squares along the diagonal (nearly) independently of the others. We can then aggregate these results, replacing each square by a single average coefficient, and repeat the process at the next level above. If the numbers of squares at the various levels are  $n_0, n_1, n_2, n_3$ , respectively, then to solve the whole system of equations simultaneously, we would have to solve  $(n_0 \times n_1 \times n_2 \times n_3)$  simultaneous equations. By making use of the property of near decomposability, we need to solve just the four separate sets of  $n_0, n_1, n_2$ , and  $n_3$  equations each, a very much simpler task.

In the example of Figure 1, the exact solution would require solving 12 simultaneous equations, the approximation, only three sets of 3 and three sets of 2 simultaneous equations. Because solution times for simultaneous equations increase quadratically with the number of equations, a system of 12 equations will require about 16 times as long to solve as a system of 3, and 36 times as long as a system of 2, the approximation, even in this simple case, will save more than a quarter of the computational time. The savings in computation time rise very rapidly with the size of the matrix and the number of levels.

The theory of near-decomposability has been independently discovered several times and is widely used in engineering and science to facilitate the solutions of large systems of equations, especially those involving a wide range of temporal frequencies: for example in designing large electrical power grids (Korn's method of "tearing") and in so-called

"renormalization" in quantum electrodynamics. ND systems are close relatives of fractals.

But our aim here is not just to solve equations. As the example shows, ND systems have very special dynamic behavior. When disturbed from equilibrium, the subsets at the lowest level of the system return to equilibrium while the sets at the next level above are still changing dynamically (relatively slowly), and the same is true (and even more decisively) for the still higher levels, which are essentially stationary (but not in equilibrium) on this time scale. In fact — and this property is used continually by the sciences — we can study the equilibrium at any given level without much concern for the slow dynamics at the levels above, and we can study the dynamics at any level without much concern for the rapid return to equilibrium at the levels below, or for the (very slowly moving) disequilibrium states at the levels above.

What, if any, implications can we draw from this architecture that are relevant for biological evolution? I wish to propose two such implications. The first has been known since at least 1969; the second perhaps not until about 1996, although, like most ideas, it has earlier harbingers.

### *The Evolution of Complex Systems by Composition*

The first implication is summarized by the following claim:

If we begin with a set of simple elements that, when they meet at random, are capable sometimes of forming stable combinations, and if the combined systems thus formed are themselves capable of combining into still larger systems, then the complex systems we will observe after this process has proceeded for some time will almost all be ND systems.

In the *Sciences of the Artificial* (3rd ed., 1996), this claim is illustrated by a parable of two watchmakers who make nearly identical watches of 1,000 parts each. However, the watches of the one watchmaker are composed of ten stable subassemblies of 10 parts each, and each subassembly of ten sub-subassemblies of 10 parts each. The other watchmaker's watches have no stable subassemblies. Then it is easy to show that if the watchmakers are frequently interrupted during the assembly process, the first watchmaker, who needs to assemble only 10 parts between interruptions, will complete vastly more watches than the second, who must assemble 1,000 parts with no interruption. That is to

say, watches built as ND structures will be fitter than those built on other architectures and will win out in evolutionary competition.

To formalize the argument, we need simply to compute the probabilities that a watch will be completed in a given period of time, with a given rate of interruption, in the two conditions. We can use the argument not only to discuss the assembly of mythical watches, but also to examine the time required for complex systems to evolve by combination and natural selection. Applying realistic parameters (with respect to numbers of elements, probabilities of meeting and joining, and assuming that bonds among small components, being shorter than those among large components, will consequently be much stronger and faster-acting), it has been shown that, if the organisms and their subsystems are ND, then the time available for evolution of living organisms on the Earth is sufficient to produce organisms of the complexity that is actually observed (say, bacterial complexity), but the time is probably insufficient if they must be completed by an uninterrupted sequence of unions of elementary-level components (Simon 1996, p. 189).

The gradual evolution of the elements from primeval fundamental particles and hydrogen — then the evolution of successively complex molecules — has observably produced ND systems with clearly defined particle, nuclear, and atomic levels, and a whole sequence of molecular levels above the atomic (small inorganic molecules, amino acids, proteins). Moreover the frequencies of interaction and binding energies among the systems decline with successive levels exactly as required by the property of near-decomposability.

#### *Acceleration of Evolution by ND Architectures at All Levels of Complexity*

In the previous argument, I chose "bacterial complexity" as the upper limit of synthesis of complex systems, for although inorganic evolution, and possibly even the evolution of single-cell organisms, fits the postulated pattern, it is highly unlikely that existing living multi-celled organisms were formed by the combination of single-celled organisms. Therefore, if we are to account for the rapid appearance of multi-celled organisms, especially during and after the Cambrian era, we must discover some other implication of the nearly-decomposable architecture. This second implication is summarized by a claim, initially set forth briefly in Simon (1996, pp. 204-205), which I would now restate in slightly modified terms:



If we begin with a population of systems of comparable complexity, some of which are ND and some of which are not, the ND systems will, on average, increase their fitness through evolutionary processes much faster than the remaining systems, and will soon come to dominate the entire population.

Notice that the claim is *not* that more complex systems will evolve more rapidly than less complex systems but that, at any level of complexity, ND systems will evolve much faster than systems of comparable complexity that are not ND. Thus, this second claim, unlike the first one, does not predict a steady increase in the complexity of the systems that are evolving; what it predicts is a growing predominance, at all levels of complexity, of systems that are nearly decomposable. Architectures that are not ND will simply be unable to compete successfully in rate of adaptation, hence ultimately in fitness, with ND architectures.

Again, we can preface our argument for the claim with a parable: this time, the parable of the clicking safe. Suppose that we have a safe with ten dials, and each dial has settings from 1 to 1,000. Then, on average, a safecracker will have to try  $5 \times 10^{29}$  settings in order to open the safe. Suppose now that the safe is defective, so that when any dial is set to the correct setting it emits an audible click. In this case, the safecracker will have to try, on average, only  $5 \times 10^3$  settings. In a third variation, each dial also emits a click, but the click-emitting setting changes with the changes in the settings of the remaining dials, so that the click is not indicative of whether the dial has been set correctly unless all the other dials are already at their correct settings. Then, this safe offers no advantage to the safecracker over the safe that does not click.

In a multi-celled organism, each cell is believed, on the basis of strong evidence, to contain the same genome as every other cell, but different subsets of the genes are activated in different types of cells (e.g., cells of different tissues and organs). Consider a multi-celled organism in a fitness landscape where the rise or fall of fitness when a particular type of cell's active genome is changed by a specific mutation or crossover depends on the pattern of active genes in each of the remaining cells. Then any improvement in fitness by change in this particular cell may be negated at any time by mutations in other cells that, even if they increase the fitness of the latter, again reduce the fitness of the former. We are in the position of the safecracker with the silent safe, or one in which the

correct position of any one dial depends on the positions of the other dials.

When the active genomes of all of the cells are subject to independent change, and when the efficiency of the active genome of a particular cell depends on the alleles of genomes of the other cells, fitness in the environment provides little or no information about whether a change in design of a particular cell or kind of cell is permanently advantageous or not; for the momentarily advantageous change may be disadvantageous in the context of the independent changes that will soon affect other cells. With interdependence among cells of the characteristics that affect fitness, a single fitness measure for the entire organism, which is all that nature provides, is highly inefficient in guiding organismic movement through the fitness landscape.

On the other hand, consider the case of an organism that is ND. Here, a more efficient set of processes in one set of cells will remain more efficient for those cells, and for the organism, over some range of changes in the processes that are active in the other cells. The more nearly complete the decomposability, the greater the range of independence of efficient design in different cells, and the greater the probability that a change in alleles favorable under current conditions will continue to be favorable after changes in other cells. Because the only feedback the organism receives is the net fitness of the whole organism, it is only under the condition of near-independence that the usual selective processes can work successfully for complex systems.

We can use Kauffman's NK model (Kauffman, 1993; Frenken, Marengo & Valente, 1999) to illustrate the effect of interdependence of component design upon movement through the fitness landscape. Consider a system with  $N$  genes each with  $A$  alleles, where the optimal allele in any gene for the organism depends on which allele is present in  $K < N$  other genes. For the simple case of 3 binary genes (2 alleles each), we have the examples of the fitness landscapes, Table 2a if  $K=0$  and of Table 2b if  $K=N-1$  ( $=2$ ).

	f1	f2	f3	F
000	0.2	0.6	0.8	0.53
001	0.2	0.6	0.5	0.43
010	0.2	0.9	0.8	0.63
011	0.2	0.9	0.5	0.53
100	0.7	0.6	0.8	0.70
101	0.7	0.6	0.5	0.60
110	0.7	0.9	0.8	0.80
111	0.7	0.9	0.5	0.70

Figure 2a  
Fitness Landscape  
with N=3, K=0

	f1	f2	f3	F
000	0.6	0.3	0.5	0.47
001	0.1	0.5	0.9	0.50
010	0.7	0.9	0.5	0.70
011	0.3	0.5	0.8	0.53
100	0.9	0.9	0.7	0.83
101	0.7	0.2	0.3	0.40
110	0.6	0.7	0.6	0.63
111	0.4	0.8	0.1	0.43

Figure 2b  
Fitness Landscape  
with N=3, K=2

Notice that in Figure 2a, allele 1-- is always superior in fitness to 0--, independently of the values of f<sub>2</sub> and f<sub>3</sub>; -1- is always superior to -0-, independently of f<sub>1</sub> and f<sub>3</sub>; and --0 is always superior to --1, independently of f<sub>1</sub> and f<sub>2</sub>. In contrast, in Figure 2b, where the fitness, F, of the organism is not simply an additive function of the independent contributions of the three genes, 100 is superior to 110, whereas 000 is inferior to 010. Which allele of the first gene produces greater fitness depends, in the case of Figure 2b, on which allele of the second gene is present.

The conditions of near-independence do not have to extend to individual cells. In most cases, a genetic change will affect an entire tissue or an entire organ, for the some of the same portions of the genome will be active throughout that tissue or organ. In the case of those parts of the genome that regulate general intracellular processes, and which are largely common to all the cells and to bacteria as well, local independence also becomes irrelevant.

Again, the argument we have made is readily mathematized, although, under realistic assumptions, it will not be possible to solve the systems of equations in closed form. Fortunately, the model has recently been formulated and simulated in a very ingenious way by Frenken, Marengo and Valente at the University of Trento, Italy (1999). By combining the ND model with Kauffman's model of the fitness landscape, the authors have demonstrated most convincingly over a series of

experiments that "in an evolutionary environment, the decomposition of the system into sub-systems, even if it is only an approximate decomposition, allows agents to improve their fitness in very short time, even though they usually are not able to find the global optimum. Thus, satisficing strategies aiming at a high rate of improvement outcompete optimizing strategies which aim at the maximum end result (p. 18)" The Frenken, Marengo and Valente simulation provides strong validation for the second implication of ND stated earlier.

Notice that there is no claim that the ND systems reach, or even approach, the optimal point in the evolutionary landscape. What is claimed is that they reach higher levels of fitness than the systems that are not ND, hence ultimately dominate the biota. Evolution by natural selection does not guarantee "the survival of the fittest"; what it guarantees is "the survival of the fittest among those alternatives that are actually generated." The space of possible organisms is unimaginably vast, and even over long evolutionary ages, there is no possibility of generating, so as to test their fitness, more than a tiny fraction of them.

The situation is like that in the games of chess and Go, where the winner is not the player who finds the optimal choice of moves, but the player who finds a better sequence of moves than his or her opponent. There are many organisms that would replace various of those that now exist if mutation and crossover at some time generated them. The pests, introduced into a biota by migration, like the hares of Australia, provide numerous empirical demonstrations of this point.

### *The Scaling of ND Systems: Allometric Regulation*

Of course, near-decomposability is not the whole story of organismic evolution. At least since the appearance of D'Arcy Thompson's *Life and Form*, the forces that determine the general shape of organisms and the stable allometric relations among their parts has been an important item on the biologist's agenda. Allometry continues to be an important and fruitful object of study, that has yielded impressive regularities, and sometimes convincing explanations in terms of energetics. In emphasizing the near-independence of the components of ND systems, I have been largely ignoring the need for compatibility of the scales of these parts. To keep inputs and outputs for the entire organism efficient and in balance, the size and pumping capacity of the heart, for example, must obviously maintain a close relation to the rate of metabolism of the entire organism, and vice versa. We clearly require genetic control of some general parameters that can regulate the relative

sizes and rates of activities of the various components of the whole organism.

It is not hard to provide for such coordination within the scheme that we have been considering. There are at least two non-exclusive possible kinds of coordinating mechanisms that are wholly compatible with a high degree of independence in the evolution of the components of the total organism; one kind of mechanism is entirely genetic, the other involves adaptation through use (but not, of course, Lamarckian inheritance of that adaptation).

To consider adaptation first, it is commonplace that muscles grow (though not without limits) when they are exercised. What we suppose here are feedback mechanisms, which can be highly decentralized, that cause increases and decreases of inputs and outputs of various processes on demand for the outputs, or as a function of the inventory of unused inputs. The lactose control mechanisms discovered by Jacob and Monod are a classical example of such mechanisms. Other mechanisms can accelerate cell division in tissues or absorb existing cells.

The feedback mechanisms can be an integral part of the decentralized processes in components of the ND structures at all levels. The genetic controls of the rates of processes may be nearly orthogonal to the genetic determinants of the chemistry of the processes themselves, for the former can often depend, for example, on gross adjustments of numbers of cells.

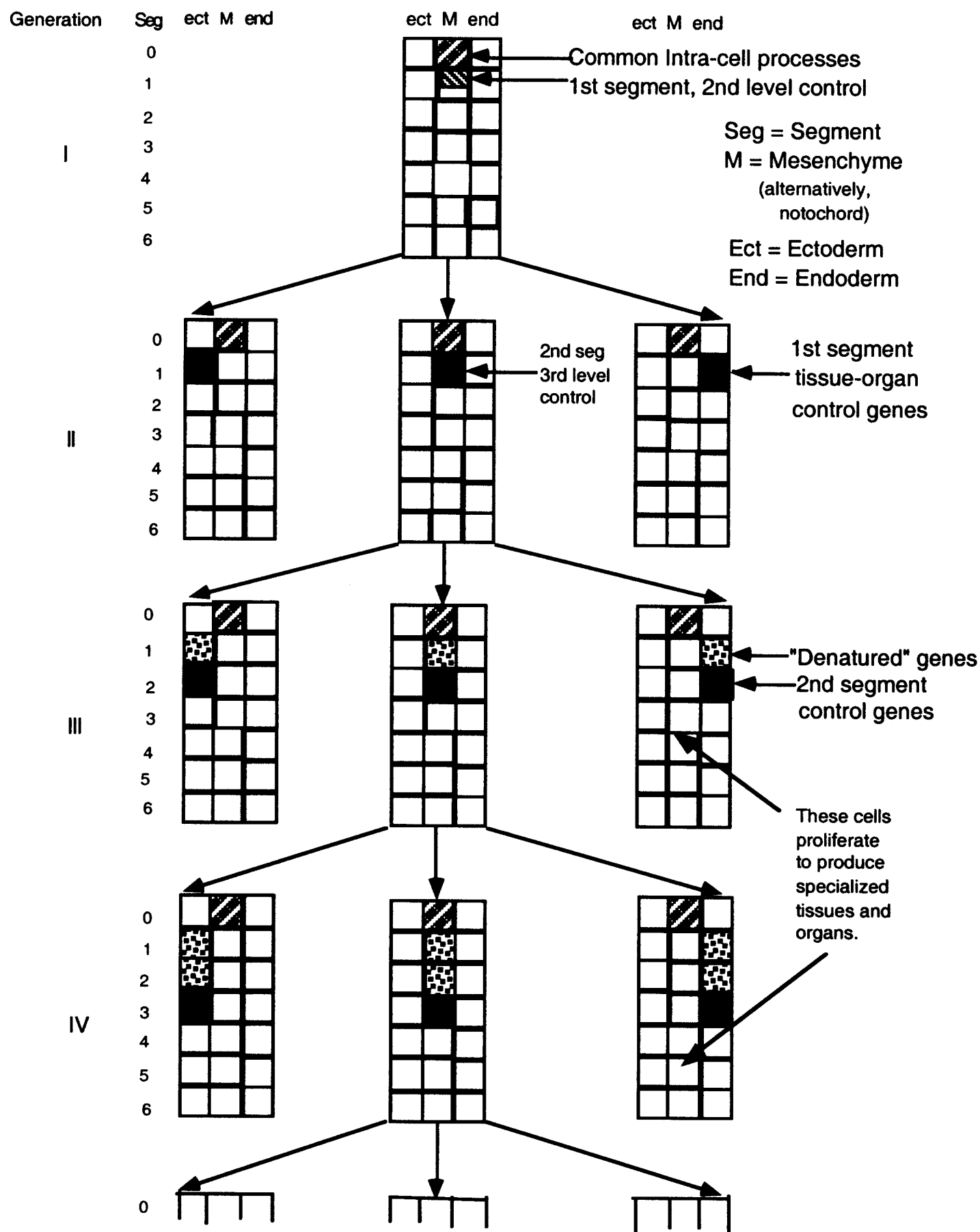
Almost the same thing can be said of genetic controls. Genetic changes in the general sizes and shapes of organs and tissues can take place almost independently of changes in the chemical processes and energetic flows in these same organs and tissues. The general architectural rule, to preserve near-decomposability, is akin to the basic rule of top-down computer programming: in order to retain the ability of the processes themselves to adapt nearly independently in the interest of efficiency, the controls that maintain the balance of the larger system must focus on inputs and outputs of processes, and must not "micro-manage" the detailed structure of the processes themselves. This suggests that when we come to understand the organization of the genome, we will almost certainly find in it a corresponding hierarchical structure (of which a number of signs have already been noticed in recent research).

It might even prove to be the case that an important part of the controls that coordinate the entire organism are to be found in those

portions of the genome that it shares with the (possibly ancestral) bacteria or other micro-organisms. But whether the controls be old or new, the important architectural consideration, and the important consideration for rate of evolution, is the extent to which they leave the locally active genes that are mainly responsible for the detailed structure and processes of components free to adapt nearly independently of the genes in other components, except for preserving relative scale proportions.

*Sketch of a Control Structure  
for Embryonic and Developmental Growth*

Figure 1 is a crude diagram of the architecture of a development process whereby a combination of centralized and decentralized controls could guide the specialization of cells during embryological development. In at least broad outline, I believe that it is consistent with current knowledge about these controls in vertebrates, utilizing the notochord and mesenchyme as the chief vehicles for the allocation of control and specialization.



Hypothetical mechanism for mesenchyme control of cell differentiation

Figure 1

The basic idea of this scheme is that the mesoderm (e.g., notochord and mesenchyme) at each cell division sets the controls differently for the next cell generation of endoderm and ectoderm cells, respectively, revising their own descendants' control structures at the same time. The same process can be repeated not only by the mesoderm cells, but by the

descendants of the endoderm and ectoderm cells that are spun off in this way. Thus there is a transmission of constantly changing and specializing control structures that bring about structural and functional specialization of the cells that become particular tissues and organs. We can think of the development of the control structure in the genes of the embryo and young organism as itself a process of hierarchical proliferation, with each element operating as a subroutine that calls on a set of sub-routines.

One might speculate that this control of controls, a meta-control which functions primarily during the period of cell proliferation early in the life of the organism, can account for some of the large amounts of apparently functionless DNA that are found in the genomes of multicellular organisms. A further step of speculation would suggest that it was the "invention" of this use of DNA for higher-level control functions that allowed multi-celled organisms to evolve.

While the crude scheme depicted here and the speculations elaborating on it are certainly wrong in detail, they do offer a sketch of a control system that requires only mechanisms of known kinds. There is already in the experimental literature of the past decade a substantial body of evidence for control mechanisms of these general sorts.

### *Conclusion*

Observation shows that the complex systems in our world today, and in particular, the living systems, all share the property of nearly complete decomposability. This fact calls for an explanation that is consistent with the mechanisms of Darwinian evolution that we believe have produced these systems. The explanation has two parts. One, chiefly applicable to inorganic nature and perhaps to single-celled organisms, derives from the fact that, when systems are assembled probabilistically from components, systems that are produced by successive assemblies of small numbers of components will emerge much more rapidly than systems that are assembled in one step by uniting large numbers of components. In the former case, the subsystems are themselves relatively stable systems, in the latter case there are not stable component systems in the larger system.

The second part of the explanation is applicable to complex living organisms. It derives from the fact that systems in which the efficiency of design of each component is relatively independent of the designs of



other components will increase their fitness through mutation, crossover and natural selection much more rapidly than systems that do not have this kind of near-decomposability. There must be limits, of course to the independence of components in order to assure their compatibility, especially their allometric compatibility, with the over-all organizational design. These limits can be provided by control structures that operate largely to regulate the inputs and outputs of the components without regulating the processes that convert inputs into outputs.

Near-decomposability is an exceedingly powerful architecture for effective organization. That being the case, it should occasion no great surprise that it appears with such regularity in the systems studied by physics, chemistry, and biology alike. In fact it appears with regularity also in human social organizations — for example, business firms and government agencies — with their many-layered hierarchies of divisions, departments, and sections; and even in complex human artifacts like computer programs with their hierarchies of subroutines. Perhaps there are other forms of system architecture with comparable versatility and power, but near decomposability appears to have been a central feature of the forms that have been produced in evolution by natural selection, and the formal dynamic properties of ND systems explain why this is so.

## *References*

- Frenken, K., Marengo, L., & Valente, M. (1999). Interdependencies, near-decomposability and adaptation. Working Paper 1999-03, Department of Economics, University of Trento, Trento, Italy.
- Iwasaki, Y., and Simon, H.A. (1994). Causality and model abstraction. *Artificial Intelligence*, 67 , 143-194.
- Kauffman, S.A. (1993). *The origins of order: Self-organization and selection in evolution*. New York, NY: Oxford University Press.
- Simon, H.A. (1996). *The sciences of the artificial (3<sup>rd</sup> ed.)*. Boston, MA: MIT Press.
- Simon, H.A., & Ando, A. (1961). Aggregation of variables in dynamic systems. *Econometrica* 29, 111-138.