

# **THE SANTA FE INSTITUTE**

Talk given January 9, 1990 in Santa Fe, New Mexico

Murray Gell-Mann

Robert A. Millikan Professor of Theoretical Physics,  
California Institute of Technology  
&  
Co-Chairman of the Science Board,  
Santa Fe Institute

The Santa Fe Institute is devoted to the study of complex systems, including their relation to the simple laws that underlie them, but emphasizing the behavior of the complex systems themselves.

Our work involves the collaboration of mathematicians, computer scientists, physicists, chemists, molecular biologists, immunologists, evolutionary biologists, ecologists, psychologists, economists, anthropologists, archaeologists, linguists, and historians, among others!

Traditional universities and institutes of technology are severely hampered in dealing with this kind of research because the departmental boundaries are so jealously defended. Sometimes an interdisciplinary center is created, housed in a dilapidated building, headed by a professor who has distinguished himself in some traditional subject and is quite secure. Most of the young researchers, however, are on soft money, have little connection with teaching and none with university policy-making, and are not on tenure track. Their work is not the central interest of any one department.

We are synergistic with the great research universities and not antagonistic toward them. Some day we may take away a few bright students and a handful of professors, but we serve them by

1. challenging them by our example and
2. giving some of their most creative faculty, post-doctoral fellows, and even graduate students an arena in which to satisfy their longing for contact with one another in order to engage in interdisciplinary work on complex systems.

We have formed and are forming research networks, somewhat similar to those of the J. D. and C. T. MacArthur Foundation Health Program, each consisting of scientists and scholars from many fields and many institutions from across the U.S. and abroad. They meet from time to time in Santa Fe; they communicate by FAX, electronic mail, visits, exchange of graduate students, and so forth. Some individuals come to the Institute for several months or even a year or more.

There is also a Science Board, of which David Pines and I are co-chairs, and which includes a stellar collection of scientists and scholars. It meets once or twice a year to determine the research program, and its steering committee keeps the work of the Board going during the rest of the year. We are beginning to recruit members from the humanities, including Frederic Wakeman, the historian of the Far East who has just finished a term as head of the Social Science Research Council.

The members of the Science Board, the other members of the various networks, together with visitors and participants in particular meetings, constitute the SFI family, a remarkably diverse and talented group of people. I have recruited many of them, often very prestigious figures, in many cases from fields very distant from my own. Each time I expect to be told, "Yes, it sounds interesting, but I already have a great many

commitments—don't call me, I'll call you.” Instead, the reply is nearly always more like: “When can I come? I've been waiting for this all my life.” When they do come, they find that they can have productive discussions of research that for various reasons they cannot have at home, even if it is those very colleagues from home with whom they are talking! They are allowed to say things at SFI that they are somehow not allowed to say at their own institutions.

SFI has a working budget of about 1.5 million dollars per year from NSF, DOE, the MacArthur Foundation, Citicorp, other firms, and individual donors, but so far no capital for a core of resident senior or junior professors or for a campus. We rent a convent from the Catholic Church, and our lease expires next year!

We have a successful and well-attended summer school on complex systems each year, in collaboration with some other institutions, and visitors sometimes bring or send their graduate students, but we do not yet have any formal program of graduate instruction—when we do have one, by the way, it will not be very formal.

So far, the research is theoretical and computational in nature, although we intend to have some experimental work in the future. (There is no need of “big science,” however, especially with Los Alamos nearby and radio and optical observatories in Arizona and New Mexico, as well as the huge accelerator being built in Texas.) Computation, using modern facilities, with the cooperation of Los Alamos, is crucial, because it permits simulation by computer of complicated systems that would otherwise be very difficult to study mathematically, and also because many kinds of computational devices and procedures are themselves complex systems of great interest.

Now, what is SFI studying? There are, of course, complex systems the study of which lies entirely within the physical sciences, like turbulence in fluid flow, certain meteorological systems, and so forth. Here, in the nonlinear dynamics of complex physical systems, the phenomenon of “chaos” was first discovered, in which the outcome of a process is infinitely sensitive to changes in initial conditions. Some fascinating work is done at SFI on such problems, but it is a small proportion of the total. Chaos, by the way, is not at all confined to situations of that kind.

Most of the work of SFI is on complex adaptive systems, systems that evolve or learn, such as:

1. prebiotic chemical evolution, including the chemical processes that gave rise to terrestrial life around four billion years ago, others that must have given rise to various lifelike assemblages elsewhere in the universe, and related chemical processes that can be studied in the laboratory, with the potential for developing products of great utility;
2. biological evolution, leading through mutation and selection to the enormous variety of life forms on earth and also the existence and evolution of ecosystems;

3. the behavior of mammalian immune systems, in which specialized cells undergo mutation at a very rapid rate accompanied by selection processes that facilitate attacks on invaders of the mammal's body;
4. individual learning and thinking in animals, including human beings;
5. human cultural evolution, in which information is transmitted between individuals and to succeeding generations, so that the whole society evolves (as an example, we have begun, together with the School of American Research, a famous archaeological institution in Santa Fe, to study the evolution of human society in the prehistoric Southwest);
6. the evolution of the human languages;
7. the global economy as a complex, evolving system—that is the subject of our largest and most successful program so far, one that has succeeded in luring a number of distinguished neo-classical economists away from their preoccupations with equilibrium and perfect information—recall the story of the elderly neo-classical economist and his polite granddaughter, who asks his permission to pick up a \$20 bill from the sidewalk—he says “Don't pick it up; if it were real, someone would already have taken it”;
8. the programming of computers to evolve, by mutation and selection, new strategies that no human has designed, for example for playing games.

In Fig. 1, we see how one kind of evolution or learning spawns another as an alternative way of “solving problems.” For example, biological evolution can “solve a problem” by producing a hard-wired pattern of behavior in an organism or else by producing enough intelligence so that the organism can arrive at the necessary behavior by a learning process.

In analyzing any evolutionary process, we must unpick three kinds of strands:

1. basic rules (like the rule that all DNA in the genetic material of terrestrial life is composed of four fundamental nucleotides);
2. frozen historical accidents(for example, the fact that terrestrial life utilizes certain right-handed molecules and not the corresponding left-handed ones is probably such an accident);
3. the emergence of what is adaptive (as exemplified by convergent biological evolution—note how the cacti of arid regions in the Americas resemble plants of the euphorbia family in arid parts of Africa although they are not closely related).

What appear to be fundamental rules on one scale may, on a much larger scale, reveal themselves to be frozen accidents—for example, while surviving terrestrial life has genetic material based on the four famous nucleotides, that circumstance may be nearly an accident on the cosmic scale, since lifelike complex adaptive systems of very many kinds may exist, on planets near distant stars, that are based on quite different chemical mechanisms. The theory of complex adaptive systems, if properly developed, should apply to such systems wherever they occur in the universe.

Now how does a complex adaptive system operate? How does it engage in passive learning about its environment, in prediction of the future impacts of the environment, and in prediction of how the environment will react to its behavior?

We may ask as well how it differs from a system like turbulent flow in a fluid, a complex phenomenon but not one that is likely to be called adaptive. Yet in turbulent flow there are eddies that give rise to smaller eddies and so on, and certain eddies have properties that allow them to survive in the flow and have offspring, while others do not and die out. Why is such turbulent flow not regarded as an evolutionary system?

The answer lies in the way the information about the environment is recorded. In complex adaptive systems, it is not merely listed, recorded in what computer scientists would call a look-up table, but rather encapsulated in highly compressed form as a model or theory or schema. Such a schema is usually approximate, sometimes wrong, but it may be adaptive if it can make useful predictions including interpolation and extrapolation and sometimes generalization to situations very different from those previously encountered.

In the presence of new information from the environment, the compressed schema unfolds to give prediction or behavior or both.

For example, in biological evolution, experience of the past is condensed in the genetic message that each organism carries in its DNA. The DNA of a baby organism unfolds, in the presence of the environment, in the sense that development from embryo to adult takes place, determined by the DNA and by the fresh information from the environment.

In the scientific enterprise, the compressed schema is a theory. In a given set of circumstances, the theory unfolds, as the result of calculation, to give predictions that can be compared with experiment.

Now in each example the schema is relatively robust and persistent, but it is subject to change, typically by random processes, and to competition with other schemata. The outcome of the unfolding leads to events in the real world that affect the survival of the schema or of related schemata. That feedback loop is the essential feature of adaptation, evolution, or learning. It is illustrated in Fig. 2.

In the case of a scientific theory, the feedback comes from the comparison of the results of calculation with experiment. In biological evolution, the success in producing offspring of an organism with particular DNA affects the survival of that or a similar pattern of DNA.

For individual learning and thinking, schemata can be ideas, including creative ideas, or patterns of thought that are ways of interpreting the world. The behavior generated by those ideas or patterns in particular situations can influence how those ideas or patterns fare in competition with others.

The adaptive process need not always be extremely effective in achieving success at the level of prediction and behavior, what corresponds in biological evolution to the “phenotypic” level of the organism as opposed to the “genotypic” level of the DNA.

Before modern science, for example, the theories that prevailed included, for example, sympathetic magic. In some places, sprinkling water from a perforated gourd was supposed to bring rain, and in others eating the heart of a lion was thought to produce courage. The first might occasionally succeed by chance and the second through psychological effects, but the feedback loop did not include careful testing of theories by observation and experiment and comparison with chance effects. Instead, the feedback loops that kept sympathetic magic alive for so long included the power it conferred on individuals who then had a stake in its survival. They also included the fact that complex adaptive systems like us keep searching for patterns in such a way that finding them is to some extent its own reward. Thus, in the case of superstitious beliefs, the schema is ensuring its own survival—the selfish scheme analogous to the selfish gene!

In individual human psychology there are, to take another example, “person schemata” formed by interaction with particular individuals (usually important figures in one’s early life) and then applied later to interaction with others in similar (or vaguely similar) situations. These “person schemata” may be rather maladaptive on the “phenotypic” or behavioral level, but they are often highly resistant to change. A “neurotic” pattern of repeated maladaptive behavior is established. If the person can find a way (or can be helped to find a way) to make such a schema more open to modification, then more adaptive ones may emerge.

Likewise, getting a valuable creative idea may involve escaping from some well-established way of thinking that has entrenched itself without being completely correct. Such a pattern tends to dig itself in and attract the thinking process to itself getting a better idea involves breaking out of that “basin of attraction” into others that may be more suitable. Here again, a schema is being sustained by a feedback loop.

In the case of human societies, the schemata are institutions, customs, traditions, and myths. They are, in effect, kinds of cultural DNA. In the presence of the circumstances in which the society finds itself at a given time (environmental change, interaction with neighbors, etc.), the schemata lead to societal behavior, the consequences of which affect the viability of those schemata. But again some features may be persistent but maladaptive on the behavioral level, even if they were adaptive a long time ago. Take, for instance, our human habit of forming groups, based on arbitrary small differences, that often dislike one another and are inclined to violent competition with one another. That habit may have been adaptive in earlier stages of human and anthropoid development; clearly, in a world of dangerous weapons, such a tendency is no longer adaptive on the behavioral level. Whether it will be possible to modify it in time to prevent a large scale disaster may depend on the extent to which the tendency is hard-wired and the extent to which it is a product of culture.

In other cases, the feedback loop is fairly straightforward, with the viability of a schema depending rather directly on its performance of a given task. In other words, there is a well-defined “fitness function” or “pay-off.” The scientific enterprise has this character in great measure, with the fitness of a theory depending on its ability to account for observed facts more accurately and more coherently than its rivals.

A computer can be equipped with software that allows it to evolve strategies for dealing with a situation with a clear pay-off (playing a game, for example). The “genetic algorithm” method of John Holland et al. has the individual instructions in a program undergoing random changes; it introduces a tendency for those changes that improve the play to be retained and those that make it worse to be eliminated. The computer then evolves strategies that may never have occurred to a human being but are highly effective.

Those who work with genetic algorithms are now employing conventional computer hardware. Other researchers are studying learning in so-called “neural nets,” mathematical networks of interconnected nodes that are supposed to resemble neurons, and some of those researchers are building new kinds of computer hardware on the “neural net” principle. We need not necessarily believe the claims of a close correspondence with the functioning of actual brains to acknowledge the utility of this kind of complex adaptive system.

It is often true of complex adaptive systems that a model or schema in relatively few dimensions is exploring a gigantic strategy space, far from any optimum or equilibrium. Think, for example, of a machine learning to play chess. Now some time in the future, if and when chess is a solved problem, as tic-tac-toe is today, the situation will be different, and adaptive learning will no longer be necessary for chess, the adaptation will be completed.

In biology, a somewhat analogous case is, say, that of the thiobacteria—sulfur-producing bacteria, among the oldest known living things, which have changed very little over billions of years. They have fully adapted to their stable ecological niche. When the pay-off or “fitness function” is clearly defined, then it may also be true that the adaptation is very efficient, and one of the challenges to SFI is to understand the circumstances under which that is the case. When the adaptation proceeds efficiently, the whole process can be represented as a search for relative optima over a “rugged landscape” of many variables. As we discussed in the case of creative thinking, the trick is then to have enough “noise” around to jar the system out of relative optima with only modest pay-offs so as to attain relative optima with much better pay-offs or even absolute optima. That can give rise to something like the traditional picture of evolutionary “progress” toward greater and greater fitness.

One of the most important characteristics of complex nonlinear systems is that they cannot be successfully analyzed by determining in advance a set of properties or aspects that are studied separately and then combining those partial approaches in an attempt to form a picture of the whole. Instead, it is necessary to look at the whole system, even if

that means taking a crude look, and then allow possible simplifications to emerge from the work.

In academic and bureaucratic circles today, one is not encouraged to take a “crude look at the whole.” Instead, careers are made by specialization and it is taken for granted that serious work can be done only by looking at one or a few aspects of a complex system. Yet anyone who becomes a CEO or a president finds it necessary to make decisions as if all aspects of a situation have been considered, including the interactions among them. If there are only specialists to consult, then the collation of their results will not result in a fair picture of the whole.

SFI stands for taking crude looks at whole systems. It is therefore important that the SFI family includes not only people with an interdisciplinary outlook who can see broad principles that apply to a variety of fields, but also scientists and scholars who are responsibly familiar with the facts and successful theories in those fields—that way we are not discussing fake biology, fake psychology, fake linguistics, or fake economics. Often the specialists and the generalists are the same people, but sometimes they are not. One of the most ambitious programs being contemplated by SFI is to explore, in cooperation with other institutions, the whole question of the approach to sustainability of human activity on this planet during the coming decades, including not only considerations of population and environmental impact, but a crude look at the whole complex of economic, technological, institutional, political, military, diplomatic, social, and ideological forces, all in interaction with one another, that are involved. As Jonas Salk has been pointing out for decades, the steeply rising curves of population and resource depletion have to turn over some time during the next century. The question is whether human foresight can be used to cause them to turn smoothly into the S-curves associated with an approach to sustainability or whether they will undergo sharp fluctuations associated with wide-spread famine, pestilence, and war. The issue of freedom versus tyranny is, of course, intimately related to the outcome. (Environmental and demographic issues are important, but the other elements of a global security have not gone away.)

It may be that such a crude integrated study of all the factors in global security will turn out to be impractical, but it is worth the effort of exploring the possibility.

The Santa Fe Institute is making progress on new approaches to such diverse subjects as the global economy, theoretical immunology, the study of learning and creative thinking, the links between the brain and the mind, and theoretical ecology. But the principal contribution of SFI lies in encouraging people to break down artificial barriers to understanding by making connections.

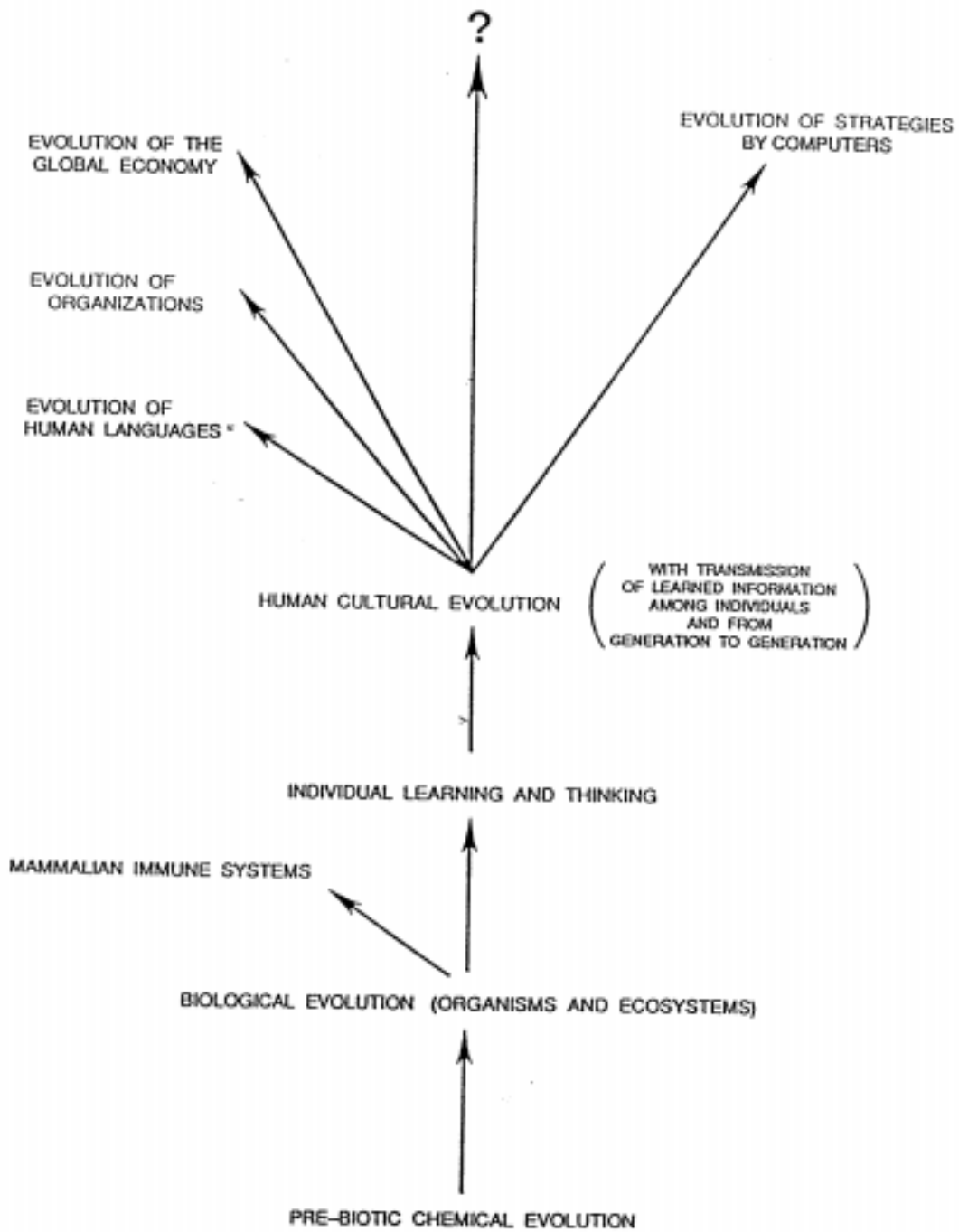


Figure 1

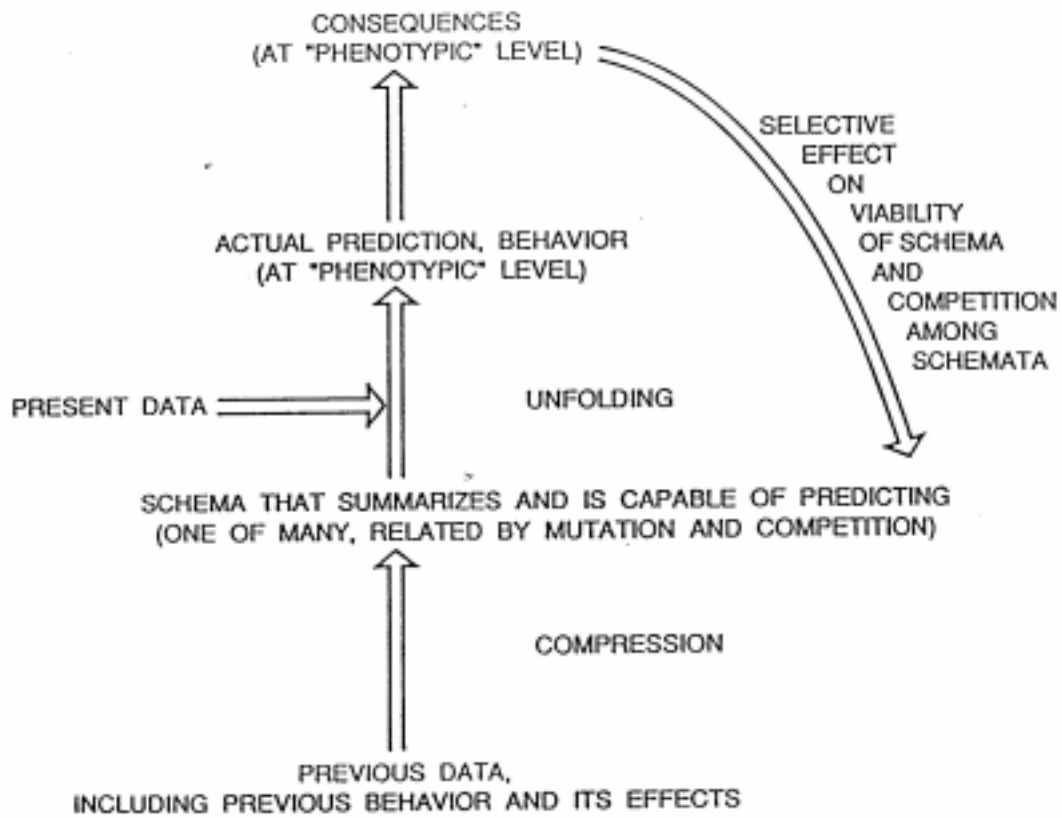


Figure 2