

Plectics: The study of simplicity and complexity

Murray Gell-Mann, Santa Fe Institute, Santa Fe, USA

The subject that I call plectics is the study of simplicity and complexity. At the Santa Fe Institute, which I helped to start, we deal to a great extent with matters of simplicity and complexity. I arrived at the name, plectics, in the following way. The word “complex” comes from *plexus*, originally meaning braided, and *com-*, meaning together, hence braided together. “Simple” comes in a similar way from roots meaning once folded; and the Latin words for “braided” and “folded” both owe their ultimate origin to the Indo-European root **plek-*. In Greek, that root gives rise to *plektos*, meaning braided. So, in using the word plectics we are describing the subject of simplicity and complexity without committing ourselves as to whether we are talking about something simple (once folded) or something complex (braided together).

What do we mean by complexity and its opposite, simplicity? It would take a great many concepts, a great many quantities to capture all the various meanings implicit in our use of the word complexity. But there is one concept — what I call effective complexity — that represents most closely what we usually mean in everyday conversation and also in scientific discourse when we use the word. A non-technical definition of effective complexity would be the length of a highly compressed description of the regularities of the entity under consideration. Compression — the elimination of redundancy — is very important; otherwise the length of the message would be of very little concern to us.

In my book “The Quark and the Jaguar” I describe the elementary school teacher who assigned to her class a 300-word essay to be turned in on Monday. One pupil, who spent the weekend playing around outdoors (as I would have done as a child), hastily scribbled the following on Monday morning: “Yesterday the neighbours had a fire in their kitchen, so I leaned out of the window and yelled “FIRE, FIRE, FIRE, FIRE,…” Of course he could have compressed that description by saying “I leaned out of the window and yelled FIRE 282 times”, but the teacher insisted on a 300-word essay. Now, one way to discuss compression is to make use of the concept of algorithmic information content. It is defined for a string of bits, that is a string of zeros and ones, or an entity that is described by that string of zeros and ones. The algo-

rithmic information content is the length of the shortest program that will cause a given universal computer U to print out that bit string and then stop computing.

Any kind of definition that utilises the length of description, even the length of a very concise description, will still involve a certain amount of arbitrariness or context dependence. We are describing a certain entity and the description is coded into a bit string. The level of detail at which we are describing the entity obviously matters. In physics that is called the *coarse graining*. The language in which the original description is expressed may also matter, and certainly the assumed knowledge and understanding of the world are important. All of these things help to determine the length of description. When the description is then coded into a string of bits to be printed out by a computer, we have additional context dependence, through the coding convention and also the choice of the universal computer. But if we put up with all that context dependence, we can give a more technical definition of effective complexity: the algorithmic information content of the regularities of an entity. That means that the entire algorithmic information content of the entity is split into two terms — one describing the regularities and the other describing the remaining features, which are regarded as incidental or random

Ties and bits

Let me use neckties as an example. The length of description of regular ties is trivial — you just give the colours, widths, and spacings of the stripes and the background colour of the tie, and you have it. But, of course, we are making a distinction here between the regularities in the pattern and various other features of the tie that we are treating as random or incidental, for example soup stains or little irregularities in the weave, and so on and so forth... Those are not included in the description.

If we are just concentrating on the regularities in the pattern of other ties, this by contrast, is quite complex. For a hand-painted tie from Austin, Texas, describing the regularities of its pattern would take a long time. It has a high effective complexity.

Let us consider a bit string consisting entirely of ones. It is obviously simple, since describing its regularity is so easy. It has a very low effective complexity as well as a very low algorithmic information content.

At the other end of the scale of algorithmic information content we would have a long bit string with almost no regularities — an incompressible or “random” string of zeros and ones, with no regularities except the length. Its algorithmic information content is very high, in fact maximal for the length of the string. The shortest program to describe it would be one that says “Print” followed by the string. Even though it has the greatest possible algorithmic information content, its effective complexity is again very small, because there are no regularities except the length.

So at both extremes, for the very simple string and for the very messy one, we have low effective complexity. Effective complexity can be large only in between, for something that lies between the extremes of order and disorder, something that has lots of different regularities.

Now, there can be no well-defined mathematical procedure that is guaranteed to find all the regularities of a bit string or of an entity described by it. In general we identify regularities by means of shared information, known technically as mutual information. If we process a bit string in certain ways and find that after processing it we can divide it into two or more parts and those parts share a lot of information, then we conclude that there is regularity present. While mutual information is diagnostic of regularity, it does not supply a measure of the algorithmic information content of the regularities that are present, and so mutual information is quite distinct from effective complexity.

Although there are mechanisms for identifying regularity, there is, as I indicated already, no mathematical procedure that will guarantee finding all the regularities. Thus the effective complexity, the algorithmic information content of the regularities, depends to some extent on who or what is describing the entity.

Let’s go back to the necktie, when we were discussing the effective complexity of the pattern and ignoring the various kinds of stains. However, suppose that we are dry-cleaners. We wouldn’t care so much about the pattern; we would concentrate instead on the soup stains, the blood stains, the wine stains, and so on. Those are the regularities that are most important for the dry-cleaner.

What are the systems that identify perceived regularities and then compress their description into a brief message? They are what I call complex adaptive systems, ones that learn or adapt or evolve as living things do. Each of us is such a system, capable of identifying perceived regularities in the data stream reaching us and distinguishing them from what is perceived as random or incidental, and then compressing the description of those regularities into a brief message.

We can give a description of how a complex adaptive system functions. It takes in a stream of data about the world, including itself. Certain regularities are identified in data that it has already taken in, and those regularities are then compressed into a very brief message, which I call a schema. This schema can be used, in combination with further data from the stream, to describe features of the world, predict the behaviour of things in the world, or prescribe behaviour for the complex adaptive system in the world. But then there are consequences in the real world: the description is better or worse; the predictions come true or they

fail; the prescription for behaviour in the real world results in success or failure, survival or disappearance. Those consequences in the real world then feed back to exert selection pressures on the competition among various candidate schemata. (A schema, although it has to have some degree of stability or robustness, must also be capable of changing, of undergoing mutation or replacement by another schema.)

Complex adaptive systems and life

All of the examples on Earth of complex adaptive systems, at least all the ones we know of, are connected in some way with life, although they are not necessarily alive or even parts or aggregations of living systems. Also, they have a tendency to give rise to other complex adaptive systems.

Among the complex adaptive systems on our planet were the prebiotic chemical reactions that led to life in the first place. Then biological evolution is a complex adaptive system, and so is the behaviour of each individual organism resulting from biological evolution. Parts of organisms can also function as complex adaptive systems — for example our immune system. The functioning of the human brain,

giving individual learning and thinking, is also a complex adaptive system. Then we may look at the behaviour of organised groups of people: human cultural evolution in general is a complex adaptive system, and human organisations, such as business firms, evolve as complex adaptive systems.

There are also non-living complex adaptive systems. Our computers are now sophisticated enough so that complex adaptive systems can be established on computers, usually by the use of software. Where is the connection with life? It’s generally agreed that the nerds who perfect the

software for these complex adaptive systems on computers are actually living things.

Now, what are the schemata? Well, let’s take a very familiar complex adaptive system involving many human beings, namely the scientific enterprise in which most of us are engaged. The schemata are theories. The theories are robust and when they have success in predicting properties of the real world they generally survive. When observations that are done carefully and are repeated over and over again disagree with the theory, then the theory is likely to be modified or have another one substituted for it.

For a hand-painted tie... describing the regularities of its pattern would take a long time. It has a high effective complexity.



Besides the scientific enterprise, where the schemata are theories, we may consider biological evolution, where the schemata are genotypes, and the evolution of a human society, where the schemata are laws, traditions, myths, customs, and so forth. Those schemata are composed of units that Richard Dawkins has christened memes, analogous to the genes in biological evolution. Sets of memes are the cultural DNA for societal evolution.

Let's return to the complex adaptive systems on computers. There are genetic algorithms, based on a very crude analogy with biological evolution. There are the so-called neural nets on computers, which are based on a very crude analogy with the way that the human nervous system – in particular the brain — is thought to operate. But there could be many more. We might have dozens of different kinds of adaptive computation methods, and they don't have to be based on analogies with ideas about the brain or biological evolution. There must be a wide class of possible complex adaptive systems on computers. What is that class? Which members of that class are suitable for solving which problems? We know that there are some problems for which genetic algorithms are suitable, while for others they are not. The same for neural nets — there are certain optimisation problems, for example, for which neural nets work very well, while there are others for which they don't. It can be shown that there is no complex adaptive system on computers that would be good for all optimisation problems. Each one has its domain of applicability, and it is a great challenge to theory to understand the whole class of complex adaptive systems on computers and figure out which one is good for which kind of problem.

Now I should point out, as a warning, that not everyone uses my notation. John Holland, who was the original inventor of genetic algorithms and a colleague and friend, uses a different terminology. What I call a complex adaptive system is something like what he calls an adaptive agent. What I call a schema he calls an internal model. Also, he uses the term complex adaptive system to mean what I would call a loose aggregation of complex adaptive systems that are modelling one another. (Examples include a market composed of investors and an ecological system composed of organisms.) By using different terminologies, we are both illustrating the famous adage that a scientist would rather use another person's toothbrush than another scientist's nomenclature.

Sometimes apparent complexity does not reflect high effective complexity. Besides the length of the shortest program that will cause a fixed universal computer to print out a description of the regularities of the entity in question, we must also consider how long it takes that computer to go from a brief program to printing out the description. That is called the logical depth of the regularities, as discussed by Charles Bennett.

Apparent and effective complexities

Take, for example, the energy levels of atomic nuclei. The rules for those energy levels look, at first sight, as if they are very complicated, but we now believe that they are obtained from a couple of simple physical theories: quantum electrodynamics (the quantum field theory of electromagnetic interactions) and quantum chromodynamics (the quantum field theory of quarks and gluons). We believe that if you put these two together you will get a description of atomic nuclei in great detail, including the positions of all their energy levels. But the computations are extremely long and difficult on our existing computers, using known methods, and most of them haven't even been done yet. So here is a case where we are looking at something apparently complex that has in fact low effective complexity, but a lot of log-

ical depth. In other words, a short program is involved, but that program is associated with a very long computation time.

Consider a case where we are still not sure whether apparent complexity reflects effective complexity or logical depth. Do the universal features of biochemistry on Earth constitute a unique system? Or are there many different kinds of possible biochemistries for systems resembling life that may flourish on other planets, orbiting other stars, in other parts of the universe? It doesn't look as if there is anything particularly special about our solar system or our planet or, for that matter, life, which appeared soon after the heavy bombardment of the Earth came to an end. It is most likely, therefore, there are very many planets in the universe with something like life, that is, complex adaptive systems with a chemistry resembling in some way the biochemistry on Earth. But do those chemistries have to be the same? Or is there a wide class of possible biochemistries? We don't really know and the experts disagree.

Some of them think that the fundamental constraints of physics limit biochemistries to just a very few possibilities. Other theorists believe that there are many possible biochemistries, of which we happen to have one example here on Earth. If we accept the ideas of the first set of theorists, then biochemistry has low effective complexity because it is derivable from the laws of physics. That derivation might be lengthy, however, implying a good deal of logical depth. If the other theorists are correct, then biochemistry on Earth could have some appreciable effective complexity, since it would depend as much on accidents of history as on fundamental physics.

This brings up the matter of how complexity arises in the universe. Where does effective complexity come from? We who work on the fundamental laws of physics mostly believe, as I do, that those laws are extremely simple. There are two of them. The first is a unified theory of all the elementary particles and all the forces of nature. It may be that we already have that theory, in the form of the wonderful candidate that has evolved from superstring theory into what is now called "M Theory."

The other fundamental principle of physics is the initial condition of the universe near the beginning of its expansion, some ten billion years ago. That may also be simple. In fact some specific ideas have been proposed about ways in which it could be simple.

A hundred years ago, we would have said that, given the fundamental theory and the initial condition, we could in principle predict the history of the universe. But today we know that is not the case. Our theories are probabilistic rather than fully deterministic. Thus the history of the universe is co-determined by these two fundamental principles and by an inconceivably long sequence of accidents — chance events — with various possible outcomes. In advance of each event, only probabilities for the different outcomes are available. A very simple example from the laboratory is the disintegration of a radioactive nucleus, emitting, for example, an alpha particle. The direction in which that alpha particle will come out is completely unknowable before it emerges; all directions are equally probable. When it does come out, then it is possible to discover the direction of emission.

We can think of the alternative, suitably coarse-grained histories of the universe as forming a branching tree, with probabilities at each branching. As time goes on and a given branching is reached, one of its branches is selected. Before the branching occurs, however, there are only probabilities for the different alternatives.

Frozen accidents and the history of the universe

The wonderful Argentinian writer, Jorge Luís Borges, wrote a short story called "The Garden of Forking Paths", in which somebody built a model of the branching histories of the universe in the form of such a garden.

Just think of all the accidents that have given rise to us in this discussion group: the little fluctuation that produced our galaxy; the accidents that were responsible for the formation of the solar system; the accidents that determined the character of the Earth; the accidents early in the history of the Earth that gave rise to life; all the accidents of biological evolution, which, together with natural selection, have produced the life forms on Earth today, including human beings; and then the accidents of sperm and egg, of sexual selection, of development in the womb and in childhood that led to the specific adults in our group.

Now, out of all the accidents in the history of the universe, some of them are more productive of mutual information, more productive of regularities than others. The fluctuation that gave rise to our galaxy, for example, might not be considered very important on a cosmic scale, but to everything in this galaxy it is of enormous significance that the galaxy came into being. Similarly, many events in human history represent branchings with huge effects on the future history of humanity on the Earth.

Historians nowadays like to talk about Annie Oakley, the famous female sharpshooter in Buffalo Bill Cody's Wild West Show, which travelled across Europe in 1889. One of the features of the show was that Annie Oakley offered to shoot part of a cigar out of the mouth of a volunteer from the audience. Typically, no man volunteered for this dangerous job and her husband, himself a famous marksman, would step forward with his cigar. Annie Oakley would shoot the ash off her husband's cigar and the audience would applaud. But in Berlin, in 1889, there was a volunteer from the audience — the Kaiser. He had been on the throne for just one year when he volunteered to be the victim of Annie Oakley's shooting stunt. She was worried about her aim — she had been drinking heavily the night before — but she had no choice. The Kaiser removed the band from his expensive Havana cigar, clipped off the end with his silver cigar cutter, put the cigar into his mouth and lit it. Annie aimed at the end of it and shot off the ash. She did not kill the Kaiser, but what if she had? History would probably have been quite different. The First World War might have been very different — indeed, it might never have happened — and so on and so forth. Here we have an example of a frozen accident, one that produces a great deal of future mutual information in a portion of the universe that is of interest.

We can now answer the question of why there is, in so many domains of experience, a tendency for entities of greater and greater complexity to come into being as time goes on. The fundamental laws of nature, we have said, are very simple, including the initial condition of the universe, but those laws are probabilistic. So we have a branching tree with probabilities, with accidents at the branchings. Some of these accidents — the frozen accidents — are more important than others. Now, as time goes on, more and more frozen accidents can accumulate, making possible the emergence of more and more regularities. If the accumulation of the results of frozen accidents outstrips the erasure or disappearance of the results of such accidents, then things of greater and greater complexity can come into being as time goes on. It is of course, not true that every individual thing becomes more complex. Far from it; for instance, organisms and civilisations die and become, naturally, much less complex in the process. But what we can say is that in many cases the envelope of complexity keeps getting pushed out as time goes on, so that

more and more complex entities come into being.

Now, the appearance of more and more complex entities as time goes on is in no way incompatible with the famous second law of thermodynamics, which states that the average disorder, the average entropy, of a closed system has a tendency to increase with time. But that is *average* disorder — there is nothing to stop mechanisms of self-organisation from producing local order at the expense of greater disorder elsewhere. We know of many mechanisms of self-organisation, for example gravitational attraction, which has produced galaxies, stars, planets, rocks, etc. Likewise, low temperatures give rise to the beautiful, regular shapes of crystals or snowflakes.

Let me end by asking about the future. Will things of greater and greater complexity keep on appearing in the universe? Well, we don't know for sure, but we can speculate about it. Many theoretical physicists believe, although so far the experiments have not proved it, that the proton will ultimately be found to be unstable. If so, then atomic nuclei are all unstable, with lifetimes perhaps something like 10^{35} years (that's a one with 35 zeros after it — a very large number of years). After some such time, nuclei would mostly be gone, and atoms and molecules would disappear too. Most of the regularities with which we are familiar would disappear, and it may be that then entities of greater and greater complexity would no longer keep appearing. In fact, the envelope of complexity might start to shrink when most of the nuclei are gone.

However, that is not of immediate concern. We have many more pressing worries.

(talk delivered to the symposium "Frontiers of Science" held in Coimbra, Portugal, 1999)

European Group for Optics and Photonics (EGOP)

The Quantum Electronics and Optics Division of the European Physical Society (QEOD of EPS) and the European Optical Society (EOS) share a strong desire to establish one unique voice to speak for the optical science and engineering community in Europe.

As a first step towards greater integration, the Prof. Dr. Theo Tschudi (President of EOS) and Prof. Dr. Guenter Huber (President of QEOD-EPS) signed an agreement in November 2001 to create a European Group for Optics and Photonics (EGOP), as a coordination mechanism between the two societies.

EGOP is concerned particularly with those topics that benefit from a close collaboration between physicists and optical engineers. The QEOD and EOS will use EGOP to coordinate their relevant activities in optics and photonics, and to represent the two societies in discussions with outside partners when both societies wish to be represented as one. Especially, there is a strong wish of both societies to coordinate and collaborate in conferences. EGOP is an informal structure and has no consequence on the membership and structure of QEOD-EPS and EOS, or on their rules of operation.