10

RICHARD DAWKINS

Selfish Genes and Selfish Memes

Selfish Genes

In the beginning was simplicity. It is difficult enough explaining how even a simple universe began. I take it as agreed that it would be even harder to explain the sudden springing up, fully armed, of complex order—life, or a being capable of creating life. Darwin's theory of evolution by natural selection is satisfying because it shows us a way in which simplicity could change into complexity, how unordered atoms could group themselves into ever more complex patterns until they ended up manufacturing people. Darwin provides a solution, the only feasible one so far suggested, to the deep problem of our existence. I will try to explain the great theory in a more general way than is customary, beginning with the time before evolution itself began.

Darwin's 'survival of the fittest' is really a special case of a more general law of *survival of the stable.* The universe is populated by stable things. A stable thing is a collection of atoms which is permanent enough or common enough to deserve a name. It may be a unique collection of atoms, such as the Matterhorn, which lasts long enough to be worth naming. Or it may be a *class* of entities, such as rain drops, which come

Excerpt from *The Selfish Gene* by Richard Dawkins. Copyright © Oxford University Press 1976. Reprinted by permission of Oxford University Press.

into existence at a sufficiently high rate to deserve a collective name, even if any one of them is short-lived. The things which we see around us, and which we think of as needing explanation—rocks, galaxies, ocean waves —are all, to a greater or lesser extent, stable patterns of atoms. Soap bubbles tend to be spherical because this is a stable configuration for thin films filled with gas. In a spacecraft, water is also stable in spherical globules, but on earth, where there is gravity, the stable surface for standing water is flat and horizontal. Salt crystals tend to be cubes because this is a stable way of packing sodium and chloride ions together. In the sun the simplest atoms of all, hydrogen atoms, are fusing to form helium atoms, because in the conditions which prevail there the helium configuration is more stable. Other even more complex atoms are being formed in stars all over the universe, and were formed in the "big bang" which, according to the prevailing theory, initiated the universe. This is originally where the elements on our world came from.

Sometimes when atoms meet they link up together in chemical reaction to form molecules, which may be more or less stable. Such molecules can be very large. A crystal such as a diamond can be regarded as a single molecule, a proverbially stable one in this case, but also a very simple one since its internal atomic structure is endlessly repeated. In modern living organisms there are other large molecules which are highly complex, and their complexity shows itself on several levels. The hemoglobin of our blood is a typical protein molecule. It is built up from chains of smaller molecules, amino acids, each containing a few dozen atoms arranged in a precise pattern. In the hemoglobin molecule there are 574 amino acid molecules. These are arranged in four chains, which twist around each other to form a globular three-dimensional structure of bewildering complexity. A model of a hemoglobin molecule looks rather like a dense thornbush. But unlike a real thornbush it is not a haphazard approximate pattern but a definite invariant structure, identically repeated, with not a twig nor a twist out of place, over six thousand million million million times in an average human body. The precise thornbush shape of a protein molecule such as hemoglobin is stable in the sense that two chains consisting of the same sequences of amino acids will tend, like two springs, to come to rest in exactly the same three-dimensional coiled pattern. Hemoglobin thornbushes are springing into their "preferred" shape in your body at a rate of about four hundred million million per second, and others are being destroyed at the same rate.

Hemoglobin is a modern molecule, used to illustrate the principle that atoms tend to fall into stable patterns. The point that is relevant here is that, before the coming of life on earth, some rudimentary evolution fo molecules could have occurred by ordinary processes of physics and

chemistry. There is no need to think of design or purpose or directedness. If a group of atoms in the presence of energy falls into a stable pattern it will tend to stay that way. The earliest form of natural selection was simply a selection of stable forms and a rejection of unstable ones. There is no mystery about this. It had to happen by definition.

From this, of course, it does not follow that you can explain the existence of entities as complex as man by exactly the same principles on their own. It is no good taking the right number of atoms and shaking them together with some external energy till they happen to fall into the right pattern, and out drops Adam! You may make a molecule consisting of a few dozen atoms like that, but a man consists of over a thousand million million million million atoms. To try to make a man, you would have to work at your biochemical cocktail-shaker for a period so long that the entire age of the universe would seem like an eye-blink, and even then you would not succeed. This is where Darwin's theory, in its most general form, comes to the rescue. Darwin's theory takes over from where the story of the slow building up of molecules leaves off.

The account of the origin of life which I shall give is necessarily speculative; by definition, nobody was around to see what happened. There are a number of rival theories, but they all have certain features in common. The simplified account I shall give is probably not too far from the truth.

We do not know what chemical raw materials were abundant on earth before the coming of life, but among the plausible possibilities are water, carbon dioxide, methane, and ammonia: all simple compounds known to be present on at least some of the other planets in our solar system. Chemists have tried to imitate the chemical conditions of the young earth. They have put these simple substances in a flask and supplied a source of energy such as ultraviolet light or electric sparks—artificial simulation of primordial lightning. After a few weeks of this, something interesting is usually found inside the flask: a weak brown soup containing a large number of molecules more complex than the ones originally put in. In particular, amino acids have been found—the building blocks of proteins, one of the two great classes of biological molecules. Before these experiments were done, naturally occurring amino acids would have been thought of as diagnostic of the presence of life. If they had been detected on, say, Mars, life on that planet would have seemed a near certainty. Now, however, their existence need imply only the presence of a few simple gases in the atmosphere and some volcanoes, sunlight, or thundery weather. More recently, laboratory simulations of the chemical conditions of earth before the coming of life have

yielded organic substances called purines and pyrimidines. These are building blocks of the genetic molecule, DNA itself.

Processes analogous to these must have given rise to the "primeval soup" which biologists and chemists believe constituted the seas some three to four thousand million years ago. The organic substances became locally concentrated, perhaps in drying scum round the shores, or in tiny suspended droplets. Under the further influence of energy such as ultraviolet light from the sun, they combined into larger molecules. Nowadays large organic molecules would not last long enough to be noticed: they would be quickly absorbed and broken down by bacteria or other living creatures. But bacteria and the rest of us are late-comers, and in those days large organic molecules could drift unmolested through the thickening broth.

At some point a particularly remarkable molecule was formed by accident. We will call it the *Replicator*. It may not necessarily have been the biggest or the most complex molecule around, but it had the extraordinary property of being able to create copies of itself. This may seem a very unlikely sort of accident to happen. So it was. It was exceedingly improbable. In the lifetime of a man, things which are that improbable can be treated for practical purposes as impossible. That is why you will never win a big prize on the football pools. But in our human estimates of what is probable and what is not, we are not used to dealing in hundreds of millions of years. If you filled in pools coupons every week for a hundred million years you would very likely win several jackpots.

Actually a molecule which makes copies of itself is not as difficult to imagine as it seems at first, and it only had to arise once. Think of the replicator as a mold or template. Imagine it as a large molecule consisting of a complex chain of various sorts of building block molecules. The small building blocks were abundantly available in the soup surrounding the replicator. Now suppose that each building block has an affinity for its own kind. Then whenever a building block from out in the soup lands up next to a part of the replicator for which it has an affinity, it will tend to stick there. The building blocks which attach themselves in this way will automatically be arranged in a sequence which mimics that of the replicator itself. It is easy then to think of them joining up to form a stable chain just as in the formation of the original replicator. This process could continue as a progressive stacking up, layer upon layer. This is how crystals are formed. On the other hand, the two chains might split apart, in which case we have two replicators, each of which can go on to make further copies.

A more complex possibility is that each building block has affinity not

lasted a long time, or they replicated rapidly, or they replicated accurately. Evolutionary trends toward these three kinds of stability took place in the following sense: If you had sampled the soup at two different times, the later sample would have contained a higher proportion of varieties with high longevity/fecundity/copying-fidelity. This is essentially what a biologist means by evolution when he is speaking of living creatures, and the mechanism is the same—natural selection.

Should we then call the original replicator molecules "living"? Who cares? I might say to you "Darwin was the greatest man who has ever lived," and you might say, "No, Newton was," but I hope we would not prolong the argument. The point is that no conclusion of substance would be affected whichever way our argument was resolved. The facts of the lives and achievements of Newton and Darwin remain totally unchanged whether we label them "great" or not. Similarly, the story of the replicator molecules probably happened something like the way I am telling it, regardless of whether we choose to call them "living." Human suffering has been caused because too many of us cannot grasp that words are only tools for our use, and that the mere presence in the dictionary of a word like "living" does not mean it necessarily has to refer to something definite in the real world. Whether we call the early replicators living or not, they were the ancestors of life; they were our founding fathers.

The next important link in the argument, one which Darwin himself laid stress on (although he was talking about animals and plants, not molecules) is *competition*. The primeval soup was not capable of supporting an infinite number of replicator molecules. For one thing, the earth's size is finite, but other limiting factors must also have been important. In our picture of the replicator acting as a template or mold, we supposed it to be bathed in a soup rich in the small building block molecules necessary to make copies. But when the replicators became numerous, building blocks must have been used up at such a rate that they became a scarce and precious resource. Different varieties or strains of replicator must have competed for them. We have considered the factors which would have increased the numbers of favored kinds of replicator. We can now see that less-favored varieties must actually have become *less* numerous because of competition, and ultimately many of their lines must have gone extinct. There was a struggle for existence among replicator varieties. They did not know they were struggling, or worry about it; the struggle was conducted without any hard feelings, indeed without feelings of any kind. But they were struggling, in the sense that any miscopying which resulted in a new higher level of stability, or a new way of reducing the stability of rivals, was automatically preserved and multi-

plied. The process of improvement was cumulative. Ways of increasing stability and of decreasing rivals' stability became more elaborate and more efficient. Some of them may even have "discovered" how to break up molecules of rival varieties chemically, and to use the building blocks so released for making their own copies. These proto-carnivores simultaneously obtained food and removed competing rivals. Other replicators perhaps discovered how to protect themselves, either chemically or by building a physical wall of protein around themselves. This may have been how the first living cells appeared. Replicators began not merely to exist, but to construct for themselves containers, vehicles for their continued existence. The replicators which survived were the ones which built *survival machines* for themselves to live in. The first survival machines probably consisted of nothing more than a protective coat. But making a living got steadily harder as new rivals arose with better and more effective survival machines. Survival machines got bigger and more elaborate, and the process was cumulative and progressive.

Was there to be any end to the gradual improvement in the techniques and artifices used by the replicators to ensure their own continuance in the world? There would be plenty of time for improvement. What weird engines of self-preservation would the millennia bring forth? Four thousand million years on, what was to be the fate of the ancient replicators? They did not die out, for they are past masters of the survival arts. But do not look for them floating loose in the sea; they gave up that cavalier freedom long ago. Now they swarm in huge colonies, safe inside gigantic lumbering robots, sealed off from the outside world, communicating with it by tortuous indirect routes, manipulating it by remote control. They are in you and in me; they created us, body and mind; and their preservation is the ultimate rationale for our existence. They have come a long way, those replicators. Now they go by the name of genes, and we are their survival machines.

* * *

Once upon a time, natural selection consisted of the differential survival of replicators floating free in the primeval soup. Now natural selection favors replicators which are good at building survival machines, genes which are skilled in the art of controlling embryonic development. In this, the replicators are no more conscious or purposeful than they ever were. The same old processes of automatic selection between rival molecules by reason of their longevity, fecundity, and copying-fidelity, still go on as blindly and as inevitably as they did in the far-off days. Genes have no foresight. They do not plan ahead. Genes just *are,* some genes more so

than others, and that is all there is to it. But the qualities which determine a gene's longevity and fecundity are not so simple as they were. Not by a long way.

In recent years—the last six hundred million or so—the replicators have achieved notable triumphs of survival-machine technology such as the muscle, the heart, and the eye (evolved several times independently). Before that, they radically altered fundamental features of their way of life as replicators, which must be understood if we are to proceed with the argument.

The first thing to grasp about a modern replicator is that it is highly gregarious. A survival machine is a vehicle containing not just one gene but many thousands. The manufacture of a body is a cooperative venture of such intricacy that it is almost impossible to disentangle the contribution of one gene from that of another. A given gene will have many different effects on quite different parts of the body. A given part of the body will be influenced by many genes, and the effect of any one gene depends on interaction with many others. Some genes act as master genes controlling the operation of a cluster of other genes. In terms of the analogy, any given page of the plans makes reference to many different parts of the building; and each page makes sense only in terms of cross-references to numerous other pages.

This intricate interdependence of genes may make you wonder why we use the word "gene" at all. Why not use a collective noun like "gene complex"? The answer is that for many purposes that is indeed quite a good idea. But if we look at things in another way, it does make sense too to think of the gene complex as being divided up into discrete replicators or genes. This arises because of the phenomenon of sex. Sexual reproduction has the effect of mixing and shuffling genes. This means that any one individual body is just a temporary vehicle for a short-lived combination of genes. The *combination* of genes that is any one individual may be short-lived, but the genes themselves are potentially very long-lived. Their paths constantly cross and recross down the generations. One gene may be regarded as a unit which survives through a large number of successive individual bodies.

* * *

Natural selection in its most general form means the differential survival of entities. Some entities live and others die but, in order for this selective death to have any impact on the world, an additional condition must be met. Each entity must exist in the form of lots of copies, and at least some of the entities must be *potentially* capable of surviving—in the form of

copies—for a significant period of evolutionary time. Small genetic units have these properties; individuals, groups, and species do not. It was the great achievement of Gregor Mendel to show that hereditary units can be treated in practice as indivisible and independent particles. Nowadays we know that this is a little too simple. Even a cistron is occasionally divisible and any two genes on the same chromosome are not wholly independent. What I have done is to define a gene as a unit which, to a high degree, *approaches* the ideal of indivisible particulateness. A gene is not indivisible, but it is seldom divided. It is either definitely present or definitely absent in the body of any given individual. A gene travels intact from grandparent to grandchild, passing straight through the intermediate generation without being merged with other genes. If genes continually blended with each other, natural selection as we now understand it would be impossible. Incidentally, this was proved in Darwin's lifetime, and it caused Darwin great worry since in those days it was assumed that heredity was a blending process. Mendel's discovery had already been published, and it could have rescued Darwin, but alas he never knew about it: nobody seems to have read it until years after Darwin and Mendel had both died. Mendel perhaps did not realize the significance of his findings, otherwise he might have written to Darwin.

Another aspect of the particulateness of the gene is that it does not grow senile; it is no more likely to die when it is a million years old than when it is only a hundred. It leaps from body to body down the generations, manipulating body after body in its own way and for its own ends, abandoning a succession of mortal bodies before they sink in senility and death.

The genes are the immortals, or rather, they are defined as genetic entities which come close to deserving the title. We, the individual survival machines in the world, can expect to live a few more decades. But the genes in the world have an expectation of life which must be measured not in decades but in thousands and millions of years.

* * *

Survival machines began as passive receptacles for the genes, providing little more than walls to protect them from the chemical warfare of their rivals and the ravages of accidental molecular bombardment. In the early days they "fed" on organic molecules freely available in the soup. This easy life came to an end when the organic food in the soup, which had been slowly built up under the energetic influence of centuries of sunlight, was all used up. A major branch of survival machines, now called plants, started to use sunlight directly themselves to build up complex

molecules from simple ones, reenacting at much higher speed the synthetic processes of the original soup. Another branch, now known as animals, "discovered" how to exploit the chemical labors of the plants, either by eating them, or by eating other animals. Both main branches of survival machines evolved more and more ingenious tricks to increase their efficiency in their various ways of life, and new ways of life were continually being opened up. Subbranches and sub-subbranches evolved, each one excelling in a particular specialized way of making a living: in the sea, on the ground, in the air, underground, up trees, inside other living bodies. This subbranching has given rise to the immense diversity of animals and plants which so impresses us today.

Both animals and plants evolved into many-celled bodies, complete copies of all the genes being distributed to every cell. We do not know when, why, or how many times independently, this happened. Some people use the metaphor of a colony, describing a body as a colony of cells. I prefer to think of the body as a colony of *genes*, and of the cell as a convenient working unit for the chemical industries of the genes.

Colonies of genes they may be but, in their behavior, bodies have undeniably acquired an individuality of their own. An animal moves as a coordinated whole, as a unit. Subjectively I feel like a unit, not a colony. This is to be expected. Selection has favored genes which cooperate with others. In the fierce competition for scarce resources, in the relentless struggle to eat other survival machines, and to avoid being eaten, there must have been a premium on central coordination rather than anarchy within the communal body. Nowadays the intricate mutual coevolution of genes has proceeded to such an extent that the communal nature of an individual survival machine is virtually unrecognizable. Indeed many biologists do not recognize it, and will disagree with me.

* * *

One of the most striking properties of survival-machine behavior is its apparent purposiveness. By this I do not just mean that it seems to be well calculated to help the animal's genes to survive, although of course it is. I am talking about a closer analogy to human purposeful behavior. When we watch an animal "searching" for food, or for a mate, or for a lost child, we can hardly help imputing to it some of the subjective feelings we ourselves experience when we search. These may include "desire" for some object, a "mental picture" of the desired object, an "aim" or "end in view." Each one of us knows, from the evidence of his own introspection, that, at least in one modern survival machine, this purposiveness has evolved the property we call "consciousness." I am not philosopher

enough to discuss what this means, but fortunately it does not matter for our present purposes because it is easy to talk about machines which behave *as if* motivated by a purpose, and to leave open the question whether they actually are conscious. These machines are basically very simple, and the principles of unconscious purposive behavior are among the commonplaces of engineering science. The classic example is the Watt steam governor.

The fundamental principle involved is called negative feedback, of which there are various different forms. In general what happens is this. The "purpose machine," the machine or thing that behaves as if it had a conscious purpose, is equipped with some kind of measuring device which measures the discrepancy between the current state of things and the "desired" state. It is built in such a way that the larger this discrepancy is, the harder the machine works. In this way the machine will automatically tend to reduce the discrepancy—this is why it is called *negative* feedback—and it may actually come to rest if the "desired" state is reached. The Watt governor consists of a pair of balls which are whirled round by a steam engine. Each ball is on the end of a hinged arm. The faster the balls fly round, the more does centrifugal force push the arms toward a horizontal position, this tendency being resisted by gravity. The arms are connected to the steam valve feeding the engine, in such a way that the steam tends to be shut off when the arms approach the horizontal position. So, if the engine goes too fast, some of its steam will be shut off, and it will tend to slow down. If it slows down too much, more steam will automatically be fed to it by the valve, and it will speed up again. Such purpose machines often oscillate due to overshooting and time-lags, and it is part of the engineer's art to build in supplementary devices to reduce the oscillations.

The "desired" state of the Watt governor is a particular speed of rotation. Obviously it does not consciously desire it. The "goal" of a machine is simply defined as that state to which it tends to return. Modern purpose machines use extensions of basic principles like negative feedback to achieve much more complex "lifelike" behavior. Guided missiles, for example, appear to search actively for their target, and when they have it in range they seem to pursue it, taking account of its evasive twists and turns, and sometimes even "predicting" or "anticipating" them. The details of how this is done are not worth going into. They involve negative feedback of various kinds, "feed-forward," and other principles well understood by engineers and now known to be extensively involved in the working of living bodies. Nothing remotely approaching consciousness needs to be postulated, even though a layman, watching its apparently deliberate and purposeful behavior, finds it hard to

believe that the missile is not under the direct control of a human pilot.

It is a common misconception that because a machine such as a guided missile was originally designed and built by conscious man, then it must be truly under the immediate control of conscious man. Another variant of this fallacy is "computers do not really play chess, because they can only do what a human operator tells them." It is important that we understand why this is fallacious, because it affects our understanding of the sense in which genes can be said to "control" behavior. Computer chess is quite a good example for making the point, so I will discuss it briefly.

Computers do not yet play chess as well as human grand masters, but they have reached the standard of a good amateur. More strictly, one should say *programs* have reached the standard of a good amateur, for a chess-playing program is not fussy which physical computer it uses to act out its skills. Now, what is the role of the human programmer? First, he is definitely not manipulating the computer from moment to moment, like a puppeteer pulling strings. That would be just cheating. He writes the program, puts it in the computer, and then the computer is on its own: there is no further human intervention, except for the opponent typing in his moves. Does the programmer perhaps anticipate all possible chess positions and provide the computer with a long list of good moves, one for each possible contingency? Most certainly not, because the number of possible positions in chess is so great that the world would come to an end before the list had been completed. For the same reason, the computer cannot possibly be programmed to try out "in its head" all possible moves, and all possible follow-ups, until it finds a winning strategy. There are more possible games of chess than there are atoms in the galaxy. So much for the trivial nonsolutions to the problem of programming a computer to play chess. It is in fact an exceedingly difficult problem, and it is hardly surprising that the best programs have still not achieved grand master status.

The programmer's actual role is rather more like that of a father teaching his son to play chess. He tells the computer the basic moves of the game, not separately for every possible starting position, but in terms of more economically expressed rules. He does not literally say in plain English "bishops move in a diagonal," but he does say something mathematically equivalent, such as, though more briefly: "New coordinates of bishop are obtained from old coordinates, by adding the same constant, though not necessarily with the same sign, to both old x coordinate and old y coordinate." Then he might program in some "advice," written in the same sort of mathematical or logical language, but amounting in human terms to hints such as "don't leave your king unguarded," or

useful tricks such as "forking" with the knight. The details are intriguing, but they would take us too far afield. The important point is this: When it is actually playing, the computer is on its own and can expect no help from its master. All the programmer can do is to set the computer up *beforehand* in the best way possible, with a proper balance between lists of specific knowledge and hints about strategies and techniques.

The genes too control the behavior of their survival machines, not directly with their fingers on puppet strings, but indirectly like the computer programmer. All they can do is to set it up beforehand; then the survival machine is on its own, and the genes can only sit passively inside. Why are they so passive? Why don't they grab the reins and take charge from moment to moment? The answer is that they cannot because of time-lag problems. This is best shown by another analogy, taken from science fiction. *A for Andromeda* by Fred Hoyle and John Elliot is an exciting story, and, like all good science fiction, it has some interesting scientific points lying behind it. Strangely, the book seems to lack explicit mention of the most important of these underlying points. It is left to the reader's imagination. I hope the authors will not mind if I spell it out here.

There is a civilization two hundred light years away, in the constellation of Andromeda.* They want to spread their culture to distant worlds. How best to do it? Direct travel is out of the question. The speed of light imposes a theoretical upper limit to the rate at which you can get from one place to another in the universe, and mechanical considerations impose a much lower limit in practice. Besides, there may not be all that many worlds worth going to, and how do you know which direction to go in? Radio is a better way of communicating with the rest of the universe, since, if you have enough power to broadcast your signals in all directions rather than beam them in one direction, you can reach a very large number of worlds (the number increasing as the square of the distance the signal travels). Radio waves travel at the speed of light, which means the signal takes two hundred years to reach Earth from Andromeda. The trouble with this sort of distance is that you can never hold a conversation. Even if you discount the fact that each successive message from Earth would be transmitted by people separated from each other by twelve generations or so, it would be just plain wasteful to attempt to converse over such distances.

This problem will soon arise in earnest for us: it takes about four minutes for radio waves to travel between Earth and Mars. There can be no doubt that spacemen will have to get out of the habit of conversing

*Not to be confused with the Andromeda galaxy, which is two million light years away. —Eds.

in short alternating sentences, and will have to use long soliloquies or monologues, more like letters than conversations. As another example, Roger Payne has pointed out that the acoustics of the sea have certain peculiar properties, which mean that the exceedingly loud "song" of the humpback whale could theoretically be heard all the way round the world, provided the whales swim at a certain depth. It is not known whether they actually do communicate with each other over very great distances, but if they do they must be in much the same predicament as an astronaut on Mars. The speed of sound in water is such that it would take nearly two hours for the song to travel across the Atlantic Ocean and for a reply to return. I suggest this as an explanation for the fact that the whales deliver a continuous soliloquy, without repeating themselves, for a full eight minutes. They then go back to the beginning of the song and repeat it all over again, many times over, each complete cycle lasting about eight minutes.

The Andromedans of the story did the same thing. Since there was no point in waiting for a reply, they assembled everything they wanted to say into one huge unbroken message, and then they broadcast it out into space, over and over again, with a cycle time of several months. Their message was very different from that of the whales, however. It consisted of coded instructions for the building and programming of a giant computer. Of course the instructions were in no human language, but almost any code can be broken by a skilled cryptographer, especially if the designers of the code intended it to be easily broken. Picked up by the Jodrell Bank radio telescope, the message was eventually decoded, the computer built, and the program run. The results were nearly disastrous for mankind, for the intentions of the Andromedans were not universally altruistic, and the computer was well on the way to dictatorship over the world before the hero eventually finished it off with an axe.

From our point of view, the interesting question is in what sense the Andromedans could be said to be manipulating events on Earth. They had no direct control over what the computer did from moment to moment; indeed they had no possible way of even knowing the computer had been built, since the information would have taken two hundred years to get back to them. The decisions and actions of the computer were entirely its own. It could not even refer back to its masters for general policy instructions. All its instructions had to be built-in in advance, because of the inviolable two-hundred-year barrier. In principle, it must have been programmed very much like a chess-playing computer, but with greater flexibility and capacity for absorbing local information. This was because the program had to be designed to work not just on earth, but on any world possessing an advanced technology, any of a set of worlds

whose detailed conditions the Andromedans had no way of knowing.

Just as the Andromedans had to have a computer on earth to take day-to-day decisions for them, our genes have to build a brain. But the genes are not only the Andromedans who sent the coded instructions; they are also the instructions themselves. The reason why they cannot manipulate our puppet strings directly is the same: time-lags. Genes work by controlling protein synthesis. This is a powerful way of manipulating the world, but it is slow. It takes months of patiently pulling protein strings to build an embryo. The whole point about behavior, on the other hand, is that it is fast. It works on a time scale not of months but of seconds and fractions of seconds. Something happens in the world, an owl flashes overhead, a rustle in the long grass betrays prey, and in milliseconds nervous systems crackle into action, muscles leap, and someone's life is saved—or lost. Genes don't have reaction times like that. Like the Andromedans, the genes can do only their best *in advance* by building a fast executive computer for themselves, and programming it in advance with rules and "advice" to cope with as many eventualities as they can "anticipate." But life, like the game of chess, offers too many different possible eventualities for all of them to be anticipated. Like the chess programmer, the genes have to "instruct" their survival machines not in specifics, but in the general strategies and tricks of the living trade.

As J. Z. Young has pointed out, the genes have to perform a task analogous to prediction. When an embryo survival machine is being built, the dangers and problems of its life lie in the future. Who can say what carnivores crouch waiting for it behind what bushes, or what fleet-footed prey will dart and zigzag across its path? No human prophet, nor any gene. But some general predictions can be made. Polar bear genes can safely predict that the future of their unborn survival machine is going to be a cold one. They do not think of it as a prophecy, they do not think at all: they just build in a thick coat of hair, because that is what they have always done before in previous bodies, and that is why they still exist in the gene pool. They also predict that the ground is going to be snowy, and their prediction takes the form of making the coat of hair white and therefore camouflaged. If the climate of the Arctic changed so rapidly that the baby bear found itself born into a tropical desert, the predictions of the genes would be wrong, and they would pay the penalty. The young bear would die, and they inside it.

* * *

One of the most interesting methods of predicting the future is simulation. If a general wishes to know whether a particular military plan will

be better than alternatives, he has a problem in prediction. There are unknown quantities in the weather, in the morale of his own troops, and in the possible countermeasures of the enemy. One way of discovering whether it is a good plan is to try it and see, but it is undesirable to use this test for all the tentative plans dreamed up, if only because the supply of young men prepared to die "for their country" is exhaustible and the supply of possible plans is very large. It is better to try the various plans out in dummy runs rather than in deadly earnest. This may take the form of full-scale exercises with "Northland" fighting "Southland" using blank ammunition, but even this is expensive in time and materials. Less wastefully, war games may be played, with tin soldiers and little toy tanks being shuffled around a large map.

Recently, computers have taken over large parts of the simulation function, not only in military strategy, but in all fields where prediction of the future is necessary, fields like economics, ecology, sociology, and many others. The technique works like this. A model of some aspect of the world is set up in the computer. This does not mean that if you unscrewed the lid you would see a little miniature dummy inside with the same shape as the object simulated. In the chess-playing computer there is no "mental picture" inside the memory banks recognizable as a chess board with knights and pawns sitting on it. The chess board and its current position would be represented by lists of electronically coded numbers. To us a map is a miniature scale model of a part of the world, compressed into two dimensions. In a computer, a map would more probably be represented as a list of towns and other spots, each with two numbers—its latitude and longitude. But it does not matter how the computer actually holds its model of the world in its head, provided that it holds it in a form in which it can operate on it, manipulate it, do experiments with it, and report back to the human operators in terms which they can understand. Through the technique of simulation, model battles can be won or lost, simulated airliners fly or crash, economic policies lead to prosperity or to ruin. In each case the whole process goes on inside the computer in a tiny fraction of the time it would take in real life. Of course there are good models of the world and bad ones, and even the good ones are only approximations. No amount of simulation can predict exactly what will happen in reality, but a good simulation is enormously preferable to blind trial and error. Simulation could be called vicarious trial and error, a term unfortunately preempted long ago by rat psychologists.

If simulation is such a good idea, we might expect that survival machines would have discovered it first. After all, they invented many of the other techniques of human engineering long before we came on the

scene: the focusing lens and the parabolic reflector, frequency analysis of sound waves, servo-control, sonar, buffer storage of incoming information, and countless others with long names, whose details don't matter. What about simulation? Well, when you yourself have a difficult decision to make involving unknown quantities in the future, you do go in for a form of simulation. You *imagine* what would happen if you did each of the alternatives open to you. You set up a model in your head, not of everything in the world, but of the restricted set of entities which you think may be relevant. You may see them vividly in your mind's eye, or you may see and manipulate stylized abstractions of them. In either case it is unlikely that somewhere laid out in your brain is an actual spatial model of the events you are imagining. But, just as in the computer, the details of how your brain represents its model of the world are less important than the fact that it is able to use it to predict possible events. Survival machines which can simulate the future are one jump ahead of survival machines who can only learn on the basis of overt trial and error. The trouble with overt trial is that it takes time and energy. The trouble with overt error is that it is often fatal. Simulation is both safer and faster.

The evolution of the capacity to simulate seems to have culminated in subjective consciousness. Why this should have happened is, to me, the most profound mystery facing modern biology. There is no reason to suppose that electronic computers are conscious when they simulate, although we have to admit that in the future they may become so. Perhaps consciousness arises when the brain's simulation of the world becomes so complete that it must include a model of itself. Obviously the limbs and body of a survival machine must constitute an important part of its simulated world; presumably for the same kind of reason, the simulation itself could be regarded as part of the world to be simulated. Another word for this might indeed be "self-awareness," but I don't find this a fully satisfying explanation of the evolution of consciousness, and this is only partly because it involves an infinite regress—if there is a model of the model, why not a model of the model of the model? . . .

Whatever the philosophical problems raised by consciousness, for the purpose of this story it can be thought of as the culmination of an evolutionary trend towards the emancipation of survival machines as executive decision-takers from their ultimate masters, the genes. Not only are brains in charge of the day-to-day running of survival-machine affairs, they have also acquired the ability to predict the future and act accordingly. They even have the power to rebel against the dictates of the genes, for instance in refusing to have as many children as they are able to. But in this respect man is a very special case, as we shall see.

What has all this to do with altruism and selfishness? I am trying to build up the idea that animal behavior, altruistic or selfish, is under the control of genes in only an indirect, but still very powerful, sense. By dictating the way survival machines and their nervous systems are built, genes exert ultimate power over behavior. But the moment-to-moment decisions about what to do next are taken by the nervous system. Genes are the primary policy-makers; brains are the executives. But as brains became more highly developed, they took over more and more of the actual policy decisions, using tricks like learning and simulation in doing so. The logical conclusion to this trend, not yet reached in any species, would be for the genes to give the survival machine a single overall policy instruction: do whatever you think best to keep us alive.

Selfish Memes

The laws of physics are supposed to be true all over the accessible universe. Are there any principles of biology which are likely to have similar universal validity? When astronauts voyage to distant planets and look for life, they can expect to find creatures too strange and unearthly for us to imagine. But is there anything which must be true of all life, wherever it is found, and whatever the basis of its chemistry? If forms of life exist whose chemistry is based on silicon rather than carbon, or ammonia rather than water, if creatures are discovered which boil to death at −100 degrees centigrade, if a form of life is found which is not based on chemistry at all but on electronic reverberating circuits, will there still be any general principle which is true of all life? Obviously I do not know but, if I had to bet, I would put my money on one fundamental principle. This is the law that all life evolves by the differential survival of replicating entities. The gene, the DNA molecule, happens to be the replicating entity which prevails on our own planet. There may be others. If there are, provided certain other conditions are met, they will almost inevitably tend to become the basis for an evolutionary process.

But do we have to go to distant worlds to find other kinds of replicator and other, consequent, kinds of evolution? I think that a new kind of replicator has recently emerged on this very planet. It is staring us in the face. It is still in its infancy, still drifting clumsily about in its primeval soup, but already it is achieving evolutionary change at a rate which leaves the old gene panting far behind.

The new soup is the soup of human culture. We need a name for the

new replicator, a noun which conveys the idea of a unit of cultural transmission, or a unit of *imitation*. "Mimeme" comes from a suitable Greek root, but I want a monosyllable that sounds a bit like "gene." I hope my classicist friends will forgive me if I abbreviate mimeme to *meme*. If it is any consolation, it could alternatively be thought of as being related to "memory," or to the French word *même*. It should be pronounced to rhyme with "cream."

Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation. If a scientist hears, or reads about, a good idea, he passes it on to his colleagues and students. He mentions it in his articles and his lectures. If the idea catches on, it can be said to propagate itself, spreading from brain to brain. As my colleague N. K. Humphrey neatly summed up an earlier draft of this chapter: ". . . memes should be regarded as living structures, not just metaphorically but technically. When you plant a fertile meme in my mind, you literally parasitize my brain, turning it into a vehicle for the meme's propagation in just the way that a virus may parasitize the genetic mechanism of a host cell. And this isn't just a way of talking—the meme for, say, 'belief in life after death' is actually realized physically, millions of times over, as a structure in the nervous systems of individual men the world over."

* * *

I conjecture that co-adapted meme-complexes evolve in the same kind of way as co-adapted gene-complexes. Selection favours memes which exploit their cultural environment to their own advantage. This cultural environment consists of other memes which are also being selected. The meme pool therefore comes to have the attributes of an evolutionarily stable set, which new memes find it hard to invade.

I have been a bit negative about memes, but they have their cheerful side as well. When we die there are two things we can leave behind us: genes and memes. We were built as gene machines, created to pass on our genes. But that aspect of us will be forgotten in three generations. Your child, even your grandchild, may bear a resemblance to you, perhaps in facial features, in a talent for music, in the colour of her hair. But as each generation passes, the contribution of your genes is halved. It does not take long to reach negligible proportions. Our genes may be immortal but the *collection* of genes which is any one of us is bound to

crumble away. Elizabeth II is a direct descendant of William the Conqueror. Yet it is quite probable that she bears not a single one of the old king's genes. We should not seek immortality in reproduction.

But if you contribute to the world's culture, if you have a good idea, compose a tune, invent a spark plug, write a poem, it may live on, intact, long after your genes have dissolved in the common pool. Socrates may or may not have a gene or two alive in the world today, as G. C. Williams has remarked, but who cares? The meme-complexes of Socrates, Leonardo, Copernicus, and Marconi are still going strong.

Reflections

Dawkins is a master at expounding the reductionist thesis that says life and mind come out of a seething molecular tumult, when small units, accidentally formed, are subjected over and over to the merciless filter of fierce competition for resources with which to replicate. Reductionism sees all of the world as reducible to the laws of physics, with no room for so-called "emergent" properties or, to use an evocative though old-fashioned word, "entelechies"—higher-level structures that presumably cannot be explained by recourse to the laws that govern their parts.

Imagine this scenario: You send your nonfunctioning typewriter (or washing machine or photocopy machine) back to the factory for repair, and a month later they send it back reassembled correctly (as it had been when you sent it in), along with a note saying that they're sorry—all the parts check out fine, but the whole simply doesn't work. This would be considered outrageous. How can every part be perfect if the machine still doesn't work right? Something has to be wrong somewhere! So common sense tells us, in the macroscopic domain of everyday life.

Does this principle continue to hold, however, as you go from a whole to its parts, then from those parts to their parts, and so on, level after level? Common sense would again say yes—and yet many people continue to believe such things as "You can't derive the properties of water from the properties of hydrogen and oxygen atoms" or "A living being is greater than the sum of its parts." Somehow people often envision atoms as simple billiard balls, perhaps with chemical valences but without much more detail. As it turns out, nothing could be further from the truth. When you get down to that very small size scale, the mathemat-

ics of "matter" becomes more intractable than ever. Consider this passage from Richard Mattuck's text on interacting particles:

> A reasonable starting point for a discussion of the many-body problem might be the question of how many bodies are required before we have a problem. Prof. G. E. Brown has pointed out that, for those interested in exact solutions, this can be answered by a look at history. In eighteenth-century Newtonian mechanics, the three-body problem was insoluble. With the birth of general relativity around 1910, and quantum electrodynamics around 1930, the two- and one-body problems became insoluble. And within modern quantum field theory, the problem of zero bodies (vacuum) is insoluble. So, if we are out after exact solutions, no bodies at all is already too many.

The quantum mechanics of an atom like oxygen, with its eight electrons, is far beyond our capability to completely solve analytically. A hydrogen or oxygen atom's properties, not to mention those of a water molecule, are indescribably subtle, and are precisely the sources of water's many elusive qualities. Many of those properties can be studied by computer simulations of many interacting molecules, using simplified models of the atoms. The better the model of the atom, the more realistic the simulation, naturally. In fact, computer models have become one of the most prevalent ways of discovering new properties of collections of many identical components, given knowledge only of the properties of an individual component. Computer simulations have yielded new insights into how galaxies form spiral arms, based on modeling a single star as a mobile gravitating point. Computer simulations have shown how solids, liquids, and gases vibrate, flow, and change state, based on modeling a single molecule as a simple electromagnetically interacting structure.

It is a fact that people habitually underestimate the intricacy and complexity that can result from a huge number of interacting units obeying formal rules at very high speeds, relative to our time scale.

Dawkins concludes his book by presenting his own meme about memes—software replicators that dwell in minds. He precedes his presentation of the notion by entertaining the idea of alternate life-support media. One that he fails to mention is the surface of a neutron star, where nuclear particles can band together and disband thousands of times faster than atoms do. In theory, a "chemistry" of nuclear particles could permit extremely tiny self-replicating structures whose high-speed lives would zoom by in an eyeblink, equally complex as their slow earthbound counterparts. Whether such life actually exists—or whether we could ever find out, assuming it did—is unclear, but it gives rise to the amazing idea of an entire civilization's rise and fall in the period of a few earth days—a

super-Lilliput! The selections by Stanislaw Lem in this book all share this quality; see especially selection 18, "The Seventh Sally."

We bring this weird idea up to remind the reader to keep an open mind about the variability of media that can support complex lifelike or thoughtlike activity. This notion is explored slightly less wildly in the following dialogue, in which consciousness emerges from the interacting levels of an ant colony.

D. R. H.