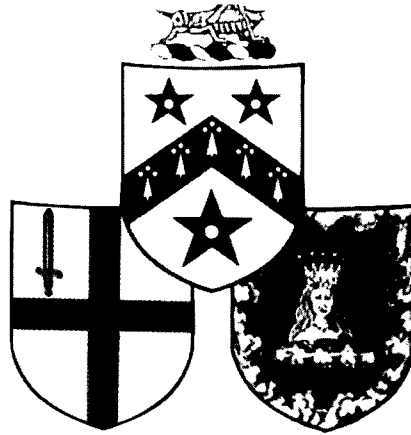


G R E S H A M
COLLEGE



THE GEOMETRY OF EVOLUTION

A Lecture by

PROFESSOR IAN STEWART MA PhD FIMA CMath
Gresham Professor of Geometry

20 January 1998

Gresham Lecture

The Geometry of Evolution

Ian Stewart 20 January 1998

The pointed conical egg of the guillemot is generally supposed to be an adaptation, advantageous to the species in the circumstances under which the egg is laid; the pointed egg is less apt than a spherical one to roll off the narrow ledge of rock on which this bird is said to lay its solitary egg, and the more pointed the egg, so much the fitter and likelier is it to survive.

D'Arcy Thompson, *On Growth and Form*

What makes evolution tick? What patterns of behaviour should we expect from an evolutionary system? What do such systems do easily, what are the genuine surprises? The application of mathematics to evolution is changing our answers to such questions, and by so doing, it is changing the way we look at the development of life on our own planet. The mathematical techniques range from the geometry of 'catastrophes' to 'artificial life'.

Earlier ages saw the rise of the 'higher' organisms as an essential feature of the evolution of today's world. The highest organism of all was humanity, and the *purpose* of the whole game was us. Biologists learned, with difficulty, to avoid imputing any kind of purpose or predetermined goal to evolution. On the molecular level, it is the result of random changes to DNA. Those changes, realised in the resulting organisms — if there are any, for many mutations fail to lead to viable organisms at all — are then subject to natural selection, and organisms that happen to survive, whether by luck or 'good design', get to propagate their genes to succeeding generations.

In this view, there is no purpose and no sense of direction to evolution — it just does whatever it does.

Mathematical modelling of evolution suggests that this exclusion of any kind of overall pattern is an over-reaction. Evolution may not have goals or purposes, but it can have a well-defined 'direction', a degree of predictability, a dynamic of its own. You can program an artificial life system *knowing* that its mutations are random, that its selection process has no inbuilt goals, no predefined notion of what is 'best' — and despite this, it will follow a distinctive series of changes, organising itself into more and more complicated 'organisms', falling into universal patterns. The very first example of artificial life, Tom Ray's 'Tierra', produced — from the simplest beginnings — things like parasites, social behaviour, even a rudimentary form of sex. None of it was programmed in explicitly — it just 'happened'.

Many other things that have long puzzled evolutionary theorists are turning out to be completely standard properties of any system that is remotely similar to evolution. One striking feature of the fossil record is 'mass extinctions', in which huge numbers of species die off simultaneously. The best known instance of a mass extinction is the death of the dinosaurs, 65 million years ago, but about twenty possible instances of mass extinction are suspected altogether, and three or four of them stand out very clearly in the fossil record. At any rate, 65 million years ago not only the dinosaurs, but innumerable other species, all died off within what in geological terms is a very short space of time. Why? Most probably this particular mass extinction was triggered by the so-called K/T meteorite, which crashed to Earth just off the coast of present-day Yucatan, in Mexico. But other mass extinctions may not have an obvious outside cause. Computer simulations of artificial life have shown that occasional mass extinctions can be the norm rather than the exception in many different kinds of 'evolutionary' system, for reasons that involve only the system's own internal dynamics. More surprisingly, they have also shown that a tendency for systems to organise themselves into more complex forms may well arise 'for free', for purely mathematical reasons. If these speculations are even close to the truth, two of the big traditional puzzles of evolution are going to turn out to be based on a

complete misunderstanding of how evolution should be expected to behave in the first place.

Darwin and his Finches

Everybody thinks they understand evolution: the idea is a simple one. However, the more closely you look at evolution, the more subtle it becomes. For this reason it will pay us to re-examine some of the usual ground before we return to the exciting, astonishing, but also highly controversial discoveries of the artificial life brigade.

According to the fossil record, life began with relatively simple organisms and slowly got more complicated. It did so in fits and starts, with occasional bursts of diversity punctuated by long periods of stasis: the reasons both for the bursts and the stasis are hotly debated, with some scientists maintaining that they are what you would expect from a complex system like life, some appealing to meteor impacts and other 'catastrophic' events, and a few disputing the evidence of the fossil record entirely and denying that either bursts or stasis have occurred. But all biologists are agreed on one overriding thing: the reason why living organisms can change, and pass on those changes to their offspring. The process involved was the brainchild of Charles Darwin, although it was arrived at independently by Alfred Wallace. Darwin called the process 'natural selection'. The *phenomenon* of evolution was already recognised, but not the mechanism behind it, which is what Darwin supplied. Nowadays we employ the term 'evolution' as a catch-all, describing both the phenomenon and Darwin's theoretical mechanism.

Evolution tells us that over long periods of time, species of organisms change. They are not created once and fixed forever: they are mutable. It also tells us why. Darwin came to his conclusions after many decades of study of living creatures. One of his most celebrated examples is 'Darwin's finches' on the Galápagos Islands, which lie on the equator, 1500 kilometres to the west of the coasts of Ecuador.

The time is 570,000 years ago. The Galápagos Islands are extremely isolated, with no large land-mass anywhere nearby. Their bird population consists only of seabirds, most of them visitors. The land, with its plants, cacti, hills, and swamps, is inhabited by reptiles — lizards, turtles — but no mammals. It would be a paradise for land-birds — except that there aren't any.

Then, by pure chance, a few bedraggled, tired finches arrive, probably blown in a hurricane. They are all of the same species, and their species evolved elsewhere to exploit a very precise environmental niche. Perhaps they were ground finches, birds that spent most of the time on the ground, eating grain. Let's suppose, for definiteness, that they were.

When finches find themselves in land-bird paradise, what do they do? They breed. There is an abundant supply of food, few competitors, no predators. The finch population must have exploded. Soon there were so many finches that the supply of grain began to run low. There were other potential sources of food: insects, cactus, berries... but *these* finches were grain-eaters.

However, the finches weren't identical. They all had ground-finch genes, but some had different genes from others. Some birds with slightly different genes from the main flock, driven to desperation as the grain ran out, found that they could eat small berries instead of grain seeds. Others evolved the ability to eat cactus. As evolution began to work upon the now diversifying range of finch abilities, the form of the finches became more specialised. The insect-eaters developed longer, thinner beaks, suitable for catching an insect in flight. The berry-eaters grew thick, short beaks. Within perhaps a hundred thousand years — maybe less — the Galápagos Islands boasted not just ground finches, but tree finches and birds that bore a closer resemblance to warblers than to finches. And this was just the beginning. Today that one species of finch has split into fourteen distinct species, each with its own lifestyle. Even today, Darwin's finches are still evolving, their genetic makeup and their form drifting gradually as their environment changes.

Darwin realised that something like this scenario must have occurred in the Galápagos Islands, because it seemed highly unlikely that fourteen separate species of bird could *each* have arrived on a hurricane. The idea that a single species might split into

several fitted many other observations that he knew about, so to Darwin, the finches of the Galápagos were the clincher. He worked out the basic mechanism of evolution without having any idea that it was based on random errors in DNA chemistry: that came a lot later. And he realised that evolution involves two very different factors. Firstly, there must be some kind of heredity — parents must be able to pass certain kinds of change on to their descendants. Secondly, the mechanism of heredity must be slightly imperfect, making the occasional 'mistake'. Given these two properties, everything else follows from the fact that on a finite Earth, all resources are limited. Therefore organisms will have to compete for resources, and the penalty for losing the competition will be a failure to pass characters on to the next generation. ('Character' is the geneticist's term for any definable feature of form, pattern, or behaviour: in ordinary language the word 'characteristic' carries the same connotations.) Either you compete effectively, or you die. You don't have to win all the time, you don't have to be the absolute best that could possibly exist: you just have to be good enough to stay around. This process is what Darwin called natural selection, often abbreviated to plain 'selection'. Darwin's great insight is that imperfect heredity plus natural selection inevitably make organisms evolve. They *change*, becoming better players at the game of survival. Usually they do this by becoming more complex, but that's a secondary observation, not an explicit feature of the theory; and sometimes they do it by becoming simpler. And since all the other players are changing too, none of them actually need become better at survival *as such*.

Mendelian Heredity

Modern biology has filled in one of the big gaps in Darwin's theory, the physical (in fact, chemical) basis of heredity. Organisms pass on characters through their DNA, and errors occur when the DNA copying process makes a mistake. This discovery has come only recently, but a thriving theory of genetics, founded upon clear mathematical principles, has been around for much longer. The theory is called Mendelism, after its discoverer, the monk Gregor Mendel. Mendel was a mathematics student at the University of Vienna. Ironically, he failed his subsidiary botany course and was therefore denied his teaching diploma. In order to pursue his studies he became a monk, succeeded *too* well, and was promoted to abbot. He was then overwhelmed with administrative duties and had to give up his scientific research. But in between, he made one of the key discoveries of his century. Mendel kept lots of pea plants, and he cross-bred them, pollinating plants with the pollen from other plants. He found that plant heredity displayed simple numerical patterns — for example when he cross-bred green-seeded pea plants with yellow-seeded ones, he got three times as many yellow-seeded offspring as green-seeded. From such result he deduced that the factors that determine the characters of the plant must be inherited from *both* parents. Nowadays these factors are called *alleles*: these are related to, but distinct from, genes, a word that has made its way into everyday language. Alleles are the different forms that a gene can take. For instance, the gene 'seed colour' has at least two alleles, yellow and green.

The tidy numerical ratios were vital clues to life's genetic mechanisms. As an example, let me explain that 3:1 ratio in peas. Mendel's idea is that the parents each have *two* alleles and the offspring inherits one from each of them — chosen at random. Call the alleles for seed colour Y and G, so that the possible pairs are YY, YG, GY, and GG. If a pea plant has alleles YY or GG, then it is clear what colour its seeds should be; but what about YG? Mendel's answer is that in such cases one particular allele always wins. That allele is said to be *dominant*, the other *recessive*. In peas, Y is dominant and G recessive, so each of the pairs YY, YG, and GY leads to offspring with yellow seeds and only GG to green. Note the numbers: three pairs yellow, one green — the magic 3:1 ratio.

A huge amount of mathematical machinery for handling this kind of calculation was developed in the early twentieth century by the statistician Ronald Aylmer Fisher. Its virtues are simplicity and pencil-and-paper accessibility: its defects are that it employs a simplifying assumption: large numbers of different individuals are homogenized into a common 'gene pool' in which only the frequencies of occurrence of alleles matters, not

who has which and in what combination. Genetics sees evolution in completely different terms, compared to Darwin: instead of organisms and characters, it focusses on genes and alleles. Organisms are a kind of secondary by-product of genes: it is only genes that really matter. This point of view, which dates from the 1930s, is called neo-Darwinism. In the modern era, that focus has been sharpened still further, and the source of all important action is seen as the molecule DNA.

Moreover, the study of evolution in terms of DNA has become highly mathematical. There are regularities and patterns even in the random mutations of DNA bases, and we can use them to trace evolutionary histories. Admittedly, the patterns are mostly statistical, and the whole area is rather controversial, for good reasons. As always, the mathematical models are only as good as their assumptions, and it now looks as if the earliest work was a bit naive in that regard. I don't consider this a major criticism: every new idea has to start somewhere, and pioneering work is always naive in retrospect. Naive or not, it's a fascinating story, and it made radical changes to our view of evolution — including our own.

Tracing Family Trees

The central idea is to employ precise mathematical techniques to trace evolutionary histories — the jargon is 'phylogenies'. Before the subject went mathematical, phylogenies were constructed on the basis of expert opinion — the gut feeling that a particular species of beetle, let us say, was evolutionarily close to another beetle species, but more distant from centipedes or wasps. The problem with such methods is that experts can disagree, and there is then no way to resolve the dispute rationally. Mathematical methods held out the hope of being more objective, because one of the great advantages of mathematics is its precision. Unfortunately this can also be one of its great disadvantages, because precision is not the same as accuracy. The accuracy of a mathematical answer is no better than the assumptions upon which it rests — but it is easy to be so impressed by the precision that you don't question the assumptions.

There are two basic types of method for tracing the 'family trees' of organisms. One is to deduce relationships by looking for common characters — for instance, all birds must be related since they have wings and feathers; bats have wings too but no feathers, so they are more distantly related to birds. The other is to ask: given two organisms, how far back in evolutionary history did they diverge from a common ancestor? The first approach, generally called 'numerical taxonomy' and pioneered by P.H.A. Sneath and R.R. Sokal, involves making a list of characters — shape of bones, pattern of veins, banding pattern of chromosomes, whatever. Then these characters are assigned numerical values. For instance, suppose we wish to distinguish between a hippopotamus, a fly, and an ant. We might draw up a table of such values, like this:

CHARACTER	HIPPO	FLY	ANT
body length (cm)	375	2	1
number of wings	0	2	0 (mostly)
number of legs	4	6	6
lives in water?	1 (yes)	0 (no)	0 (no)

and so on. The problem now is to extract from this list some quantitative measure of the overall difference between the different creatures. One way is to represent each animal's list of character values as a point in a multidimensional space, and see how the points cluster. To keep the idea simple, focus only on the first two characters in the list, length and number of wings. We can represent those characters graphically using two perpendicular axes, and locate the three animals in the resulting two-dimensional space. Visually it then becomes clear that the ant and fly are closer together in the picture than either is to the hippopotamus. A great deal of mathematical technique has been developed to make sense of this kind of clustering information in spaces with large numbers of variables, and to make the process as objective as possible.

However, it is not as objective as its proponents claim. The *calculations* are indeed objective, but the *assumptions* behind them involve human judgement as to which

characters really matter — and indeed just how to define the 'distance' between representative points. Again, people confused precision (undeniable) with accuracy (contentious).

Critics of this approach also pointed out that it had no evolutionary content. Their alternative 'phylogenetic' methods try to work out how recently two given species diverged from a presumed common ancestor. This approach can change the resulting picture dramatically. For example when applied to the three organisms lobster, barnacle, limpet, numerical taxonomy concludes that barnacle and limpet are the most closely related of the three, whereas the phylogenetic method places lobster closest to barnacle. There are two main phylogenetic schools: evolutionary taxonomy, promoted by Ernst Mayr, George Gaylord Simpson, and Theodosius Dobzhansky, and cladism, introduced by W. Hennig. Both of them see the lobster as a closer relative of the barnacle than the limpet is. In its original form, cladism was also based on measurements of characters, but it sought to deduce the *real* evolutionary tree, or 'lineage' of the organism — its list of ancestral organisms or species. It therefore looked at characters shared by groups of organisms that were presumed to be evolutionarily related, and focussed only on those characters that were unique to some group. Elaborate mathematical techniques were devised to locate such characters and deduce the family tree.

Problems with Cladism

One problem with cladism — unless you are a cladist, in which case you see it as an advantage — is that some traditional groups get split up. For instance a cow is a quadruped, and all quadrupeds evolved (long, long ago) from lobe-finned fish like the lungfish. Other kinds of fish, such as the salmon, a ray-finned fish, diverged from the lungfish lineage well before the cow did. By the rules of cladism, this prevents the salmon and lungfish from being placed in the same group — unless the cow goes in too. So either a cow is a fish, or the lungfish and salmon are not both fishes. The same problem crops up in classifying reptiles, where the crocodile is found to be a closer relative of the bird than of the lizard or snake. The cladists' viewpoint is that the family tree is just like that: too bad. The numerical taxonomists disagree, which is where the third school, that of evolutionary taxonomists, comes in — with a compromise approach that pleases neither of the other two schools.

Nowadays, however, we have an (allegedly) less contentious way to trace evolutionary lineages. Instead of tracking characters, we track DNA codes. If one organism has a sequence somewhere that goes CCGGGTTTCC and another has one in the corresponding place that goes CAGGGTTTCC, with only one mutation, then they must be more closely related than one with the sequence CGTGACTTCC, which differs from them both in many more positions. There are still some big surprises, but the evidence is distinctly less subject to personal bias in choosing characters. DNA isn't the only molecular method for tracking lineages: for example amino acid sequences in proteins can also be used. The biggest problem here is a rather interesting mathematical one: how do you define a sensible 'distance' between DNA sequences? The obvious one is what communications engineers call the Hamming distance: the number of places where the sequences differ. So the Hamming distance between

CCGGGTTTCC
CAGGGTTTCC

is 1, since there is only one different base (boldface), whereas that between

CCGGGTTTCC
CGTGACTTCC

is 4. Unfortunately, DNA does not always mutate by just changing one base. Bases may be inserted or deleted; so may whole sequences of bases. Sequences can be copied several times in a row. Moreover, sequences can also be inverted. So only one step away from CGTGACTTCC we find sequences such as

(insert a base) CGATGACTTCC
(delete a base) CGGACTTCC
(insert a sequence) CGTATTAGGACTTCC
(delete a sequence) CGTTTCC

(double up a sequence) CGTGACTTGACTTCC
 (invert a sequence) CGTTCAGTCC

The Hamming distance places all of these a long way away from the original sequence. You might decide to use something like 'the smallest number of such operations that can transform sequence 1 into sequence 2'. Apart from being very hard to calculate, this measure also has a major defect. In a literary analogy, it would place *Winnie-the-Pooh* very close to *Hamlet*. Two steps alone separate them:

- Step 1: insert the whole of *Hamlet* at the end of *Pooh*
- Step 2: delete *Pooh*.

So that won't work either. What is really needed is a way to characterise *sensible* insertions or deletions — so that when presented with "TO BE OR NOT TO BE, THAT IS THE POOH TRAP FOR HEFFALUMPS QUESTION. WHETHER 'TIS NOBLER IN THE MIND..." the mathematics would spot the inserted sequence as readily as fans of William Shakespeare and Alan Alexander Milne do.

An interesting feature of using molecular methods to trace branches of family trees is that you can deduce a certain amount of information about when particular species diverged from others. The idea is that particular regions of the genome mutate at different rates, and that mutation rates can be estimated from modern experimental data. So, in effect, the DNA mutations provide a molecular clock. There is a certain amount of disagreement about just how *regularly* the clock ticks, but on a qualitative level the idea is sensible enough. A great triumph of this approach occurred in anthropology, the study of humanity's prehistoric ancestors. Until the 1960s the fossil species *Ramapithecus* was widely considered to be a hominid — a very close relative of mankind, much closer than the great apes such as the gorilla and chimpanzee. But in 1967 V. Sarich and A.C. Wilson measured the 'immunological distance' between humans, gorillas, and chimpanzees, by seeing how strongly antisera from one of these species bound to the protein albumin in the others. Their results, interpreted via the molecular clock, indicated that humans diverged from the great apes only 5 million years ago. Other evidence showed that *Ramapithecus* and humanity diverged more than 9 million years ago. So *Ramapithecus* wasn't a hominid after all.

That figure of 5 million years has now come into dispute — though not the conclusion that *Ramapithecus* wasn't a hominid. In March 1997 Simon Easteal and Genevieve Herbert took another look at the ticking of the molecular clock. The figure of 5 million years arises from an estimate of 1.5×10^{-9} for the probability that a given base will mutate in a given year. (This means that any given DNA base will, on average, mutate once every six hundred million years. Mutations in any given base are *very* infrequent — but there are an awful lot of bases.) Easteal and Herbert argued that the mutation rate ought to be pretty much the same in all mammals; but that assumption placed the divergence of the marsupials (such as kangaroos) from the mammals at about 330 million years ago. However, fossil evidence shows conclusively that the divergence occurred no more than 125 million years ago. The two scientists concluded that the molecular clock ticks about 50% faster than had previously been assumed. This led them to revise the date at which humans and chimps diverged — it was probably 3.6 to 4 million years ago, not 5 million. This revision of history is important because it makes it possible for a known hominid, *Australopithecus afarensis*, to be the common ancestor of both chimps and humans. Another similar hominid, *A. africanus*, could then be the ancestor of gorillas. I mention all this mainly to show that science continues to refine its understanding of early human evolution, and that mathematics is proving an indispensable tool in anthropology.

Ingredients for Modelling Evolution

Evolution raises some novel problems for mathematics, because when viewed as a process it has unusual features that do not fit neatly into existing mathematical theories. Evolution has at least four ingredients:

- mutation
- selection
- development

- environment

and these interact to produce organisms that are adapted to their environment. Genes affect organisms by controlling — or at least changing — the organisms's development. Genes themselves change by random mutation. These are processes that involve the content of the organism — its own internal structure. Organisms affect the range of genes available in the next generation — the gene pool — by reproduction. Selection affects organisms by favouring those that are well adapted. These processes involve the organism's context — the environment, including other organisms, climate, terrain, and availability of mates (in sexual species). Evolution occurs when many organisms pursue these interactions and more or less systematic changes emerge. A fully realistic model of evolution must take all of these interactions into account — a daunting task.

Neo-Darwinism

Geneticists, especially those known as neo-Darwinists, try to sidestep the messiness of organisms by collapsing the evolutionary system down to something simpler, looking only at the effects 'experienced' by genes. The complex processes of selection within a changing environment are collapsed down to a single 'fitness factor' for a given allele; and phenotype — the form and behaviour of the organism — is assumed to be a direct consequence of genotype. Instead of organisms competing for the right to reproduce, neo-Darwinists see genes competing for their place in the gene-pool. Moreover, in classical genetics of the kind introduced by Fisher, an ecosystem — such as a rainforest full of different plants, insects, small animals, predators — is modelled as a homogeneously stirred pool of genes. As the organisms reproduce, those genes get mixed together in new combinations; as natural selection weeds out unfit alleles, the alleles that allow organisms to survive better tend to proliferate. Random genetic mutations keep the gene pool simmering. The mathematics focuses solely on the proportions of particular alleles in the population, and models how those proportions change in response to selection. Physicists call this kind of approach a 'mean field' theory, and resort to it only when desperate. In mean field models, a collection of distinct individuals is replaced by a homogeneous mass of identical 'average' individuals. It's like assuming that every family really does have 2.3 children — fine for some purposes, like deciding how many schools to build, but misleading for others, such as deciding how many big or small houses are likely to be needed in the next decade.

For example, a hypothetical population of slugs might have genes for green or red skins, and other genes for a tendency to live in bushes or in bright red flowers. Typical genomes include green/bush and red/flower — four possible combinations altogether. Some combinations, however, have greater survival value: for example red/bush slugs would be easily seen by birds against the green background of the bushes they inhabit, whereas red/flower slugs would be less visible. To model this system in the spirit of Fisher, we assign numerical weights, called selection coefficients, to the possible genomes. Thus red/bush might have a selection coefficient of 0.1, compared to 0.7 for red/flower. Essentially these choices indicate that a red slug living on bushes has only a 10% chance of surviving to reproduce, whereas a red one living in flowers has a 70% chance. We also assume some initial distribution for the proportions of the total slug population that correspond to each of the four pairs of alleles — say that 20% are red/bush, 15% red/flower, and so on. Fisher's mathematical scheme then lets us calculate the proportion of each allele in each succeeding generation. If some proportion becomes zero, then that particular allele dies out.

All this, of course, is no better than the assumptions that go into it — which by today's standards are unsophisticated. As well as being mean field, Fisher's genetic models are also 'linear' — they assume that the effect of an allele is proportional to the frequency with which it occurs, and that the effects of different alleles simply add up. Linear mathematics held sway in classical times because the calculations were simple enough to be done with pencil and paper. Today, most areas of science are adopting 'nonlinear' models with more complex, but far more realistic, dynamics. The same is true of frontier genetics and evolutionary theory.

Fitness Landscapes

We can capture some of the flavour of nonlinear modelling by using a geometric analogy. Imagine a plant on the side of a hill, producing seeds and scattering them randomly around it. Suppose for the sake of argument that seeds that land higher up the hill are retained, but those that land lower down are removed. Then over a period of time what you will find is a patch of plants working its way higher and higher up the hill. This image of a hill is a simple example of what Sewall Wright called a 'fitness landscape'. This is a surface, a graph that represents how the fitness of an organism depends its characters. Fitness is represented by height, and the characters determine the organism's position on the landscape. Nonlinearity implies that a typical landscape is bumpy, rather than being flat or just sloping at a constant angle. The bumps are the places where organisms are fittest, and so dominate the evolutionary behaviour; the valleys are also important, however, because they separate the bumps from each other.

Fitness is a relative concept, not an absolute one, but it's clear what such a model predicts: organisms will evolve in the uphill direction, heading towards the local peaks of fitness. There are dozens of ways in which this model of evolution is too simple to capture the reality in detail, but it emphasises the basic point: that even if the errors in heredity are random, natural selection will impart a definite directionality — towards fitter organisms. Similar models can cope with more realistic assumptions, and they lead to much the same conclusion.

Wright's image was one of the earliest biological instances of a way of thinking that has now become all-pervasive, the mathematical concept of a *phase space*. This idea, introduced by Henri Poincaré a century ago, represents dynamics as geometry. A phase space is a multidimensional mathematical space whose points represent possible states of some dynamical system — a system whose state can change over time. In Wright's model, the phase space is 'organism-space' — its coordinates form a list of the numerical values of all the relevant characters. For example, suppose that we are modelling a population of finches. Then the system is 'all possible finch phenotypes' and its states are *particular* finch phenotypes. The phase space is two-dimensional, a plane, and its two coordinates correspond to the two variables 'height' and 'wingspan'.

If we looked at twenty such variables, we would in effect be working in a twenty-dimensional phase space — a somewhat mindboggling concept that pervades modern mathematics. The word 'dimension' is introduced as an analogy: each entry in a list of twenty numbers can be varied independently of the others, so that entry behaves like an independent dimension. The geometric language proves useful because it sets up helpful analogies with spaces of two or three dimensions; precision is supplied by remembering that the actual objects under discussion are long lists of numbers.

What of dynamics? Dynamics is represented by a flow-pattern in phase space. As a representative point 'goes with the flow' its coordinates, the list of numbers that represents the state of the system, change over time. Phase spaces offer technical advantages for mathematicians, but their most important role is metaphorical: they formalise the notion of context by embedding what actually happens in a structured realm of all the things that might have happened instead. In a phase space model you can ask 'why *this* behaviour rather than *that*?' and hope to get a sensible answer.

Wright's phase space approach immediately solves one worrisome puzzle: if phenotypic variables are continuous — capable of assuming any numerical value within a given range and thus changing gradually — why do we see well-defined species with values that cluster around particular numbers, and empty gaps at other numbers? The reason is that continuously varying landscapes still have *isolated* peaks. Its main defects are

- Not all characters can be captured by continuous numerical variables.
- Fitness is not just a matter of evaluating a single number. (For example a cat is fitter than a goldfish at climbing trees, but less fit at staying underwater for hours at a time. Who wins depends on what game they are playing.)

Nevertheless, the 'fitness landscape' is a useful and insightful metaphor for certain aspects of evolution.

Catastrophe Theory

By developing some ideas of René Thom in the later 1960s, Christopher Zeeman set up a mathematical model of the fitness landscape which addresses the slippery issue of whether evolution can jump. Can a radically new organ, such as an eye, or a radically new species, *suddenly* come into being? Darwin's view on this question is generally presented as unequivocal: 'Natura non facit saltum,' *nature does not make jumps*. But he also said: 'Many species once formed never undergo any further change... and the periods during which species have undergone modification, though long as measured by years, have probably been short in comparison with the periods during which they retained the same form.' That is, the jumpiness of evolution depends on the timescale over which you look.

It is true that every so often the fossil record seems to show some very sudden changes. Are they genuine, or do they just reflect gaps in the fossil record? Certainly the fossil record is very incomplete, and apparent gaps are still being filled at a relatively rapid rate as palaeontologists dig up new specimens. The gradualist view of speciation is that over long periods of time the phenotype of a species slowly drifts, until eventually the change becomes so great that the species seems to have changed. For example, the trilobite is a creature that lived on the ocean floor and became extinct 250 million years ago. Trilobites evolved for hundreds of millions of years: new species arose, earlier ones died out. But all of the known changes are arguably gradual in character.

In 1972 Niles Eldredge and Stephen Jay Gould caused a storm with their theory of 'punctuated equilibrium'. This maintained that speciation nearly always occurs by the splitting of lineages, not by the slow drift of a single lineage, and that this splitting happens much more rapidly than the usual drift rate. They tied this theory (unnecessarily, it seems to me) to the prevailing idea that the splitting of species occurs by 'allopatric speciation', in which a small subpopulation on the edge of the geographical range moves away and becomes disconnected from the main body of the species. Once isolated, this group evolves in new ways because it is in a different environment. If the resulting new species reinvades the original territory, the fossil record *at that place* will appear to show a jump.

The gradualists disagreed completely: they maintained that almost *all* speciation takes place gradually. They accepted the possibility of splitting as well as drift — after all, the number of species alive today is a lot greater than it used to be, and the new ones must have come from *somewhere* — but saw splitting itself as a gradual separation, not a sudden jump. Eldredge and Gould, in contrast, thought that almost all species arise through rapid changes at splittings, and hardly any changes in species occur by gradual drift.

It's a complicated debate, not helped by differences of opinion about what constitutes a species. Cladists, for instance, define species in a manner that precludes drift as a mechanism for species change, because to them anything that changes gradually represents the *same* species. What mathematics does here is to suggest that the whole debate is misconceived. Anyone brought up on modern dynamical systems — the best *general* theory of how systems can change over time that we have ever had — knows that the same system can change suddenly or gradually. The sudden changes are called *bifurcations*, a term that represents a *conceptual* splitting of possible behaviours rather than an actual splitting of species — and is not confined to splitting into *two* pieces, despite its usage in everyday language. Imagine a dynamical system whose behaviour depends on external parameters — environment, perhaps. Assume that those parameters vary gradually. What does the system do? The 'obvious' answer is that the system changes gradually too: that continuous changes produce continuous effects.

Obvious — but wrong. *Most of the time* the effects will change gradually, but every so often the parameters can hit a 'critical value' at which the change becomes rapid and dramatic. When this happens, we have a bifurcation. For an example, imagine a stick being gradually bent by an external force (your hands). To begin with, it bends, and the change is just as gradual as the change in the applied force. But then, without anything terribly obvious changing, the stick suddenly snaps. After that, you can continue moving your hands gradually and the stick again moves with them.

There is nothing unusual about this dual behaviour: usually smooth, sometimes

sudden. It is what nearly all dynamical systems do.

Bifurcations occur when the state of the system changes from being stable to being unstable: the system then seeks a new stable state, which may mean a big change. The gradual behaviour occurs when stable states remain stable. Symmetry-breaking is a particular type of bifurcation behaviour, found in symmetric systems. But even asymmetric systems can, and often do, bifurcate.

In the 1960s, Thom introduced some new ideas from pure mathematics into the classification of bifurcations, and Zeeman gave them the name 'catastrophe theory' to emphasise the sudden changes involved. Catastrophe theory did not have a big impact on biological modelling — partly as a result of exaggerated criticism in its early days — but it completely revolutionised bifurcation theory. It was a bloodless revolution, accomplished under an assumed name (singularity theory), and it took place largely within mathematics, so hardly anybody noticed. That change in dynamical systems can be sudden was not new, but the possibility of classifying such changes in terms of a sequence of increasingly complex geometric forms *was* new.

The central question, from this point of view, is this: suppose that a fitness landscape changes gradually, as a result of changes in external parameters. *What should we expect the fitness peaks to do?* You might expect them to move gradually as well — this seems to be the unspoken assumption behind gradualist thinking about evolution. That expectation is correct if there is only one peak and it does not run into anything else — such as a slope. It is also correct in the especially simple kind of linear mathematics that was common a century ago. But it is *not* true — not even close to truth — for more realistic nonlinear mathematics. The reason is that in nonlinear systems peaks can be born, can be absorbed, can collide, can split.

So why not species too?

Admittedly, a fitness landscape is too simple to capture *all* of the rich reality of biology — but if anything, that reinforces its message. The model shows that both rapid *and* gradual change are natural in any system that occupies the peaks of slowly changing landscapes. There is absolutely no need to choose one or the other as being exclusive, and every reason not to. If the *simplest* nonlinear model of transitions in gradually changing fitness landscapes possesses such a rich range of dynamic behaviour, surely more complex and biologically more accurate ones should be capable of at least the *same* richness?

The gradualist/punctuationalist controversy is pointless. Most likely, the two schools of thought are both right, some of the time, and both wrong, some of the time. It's time they put their ideas together: neither will 'win' on its own.

Artificial Life

So far, the evolutionary mathematics that I've told you about has been fairly conventional, at least to a mathematician. But some aspects of evolution pose completely new problems for mathematics, and their solution demands the creation of new mathematics. It's not arrived yet, but with imagination we can see it on the way.

The need for biologically more realistic models of evolution has stimulated a very different approach to evolutionary modelling, known as the theory of complex adaptive systems, or Complexity Theory for short. Artificial life is a development within Complexity Theory. Complexity theorists try to model complicated systems of individuals *as* complicated systems of individuals. They don't take short cuts with 'average' behaviour, they don't assume everything is uniformly mixed: they accept the unique nature of the individual, and delight in it. To model evolution, they set up computer models with lots of 'virtual organisms' that obey simple rules of interaction, and see what happens. Remember the slugs and their selection coefficients? For a complexity-theoretic approach to the same problem, we set up a simulation based on a square grid, say 100 squares by 100. We decide which square corresponds to a piece of bush, or flower, or whatever. Then we populate a randomly chosen selection of these squares with 'virtual slugs', by assigning a slug 'genome' (combination of the alleles under consideration) to each such cell. For example the square 28 units long row 49 of the grid might be assigned genome red/bush, and so on. Other squares might be 'virtual predators'. Next, we give the computer rules for how these virtual organisms move around the grid and interact with each

other. For example, we might decide that at each time-step a slug moves at random to a neighbouring square, or stays put, whereas a predator 'sees' the nearest slug and moves five squares towards it, 'eating' it if it reaches the slug's own square — meaning that that particular virtual slug is removed from the computer's memory. We set up the rules so that green slugs are less likely to be 'seen' if they are on bushes rather than flowers, and so on. Then we play this mathematical computer game — the technical term is 'cellular automaton' — for ten thousand time-steps, and read off the proportions of various surviving slug alleles. In all likelihood we would run the simulation several hundred times, to ensure that any apparent mathematical patterns are independent of the particular sequence of random events that occurs in a single run. An advantage of the complexity model is that it explicitly incorporates organisms as individuals, rather than by proxy as allele proportions, and it implements natural selection by pitting predators against prey in an environment, rather than simply assigning numerical weights to the probable outcome of such a contest.

Complexity theorists have invented innumerable models in the same spirit: building in simple rules for interactions between many individuals, and then simulating them on a computer to see what happens. The provocative but apt term 'artificial life' was coined to describe such activities. A celebrated example is *Tierra*, invented by Tom Ray. In *Tierra*, short segments of computer code compete with each other inside the computer's memory, reproducing and mutating. The source of all *Tierran* 'life' is an ancestral organism, a self-replicating segment of computer code occupying 80 bits of memory. In January 1990 Ray released this organism into a primal ocean of random bits in a computer's memory and left the system to its own devices. Copies of the replicating ancestor quickly took over large regions of memory, but then occasional mutations — computer errors — began to cause changes. New replicating 'species' appeared, some smaller than the ancestor, some bigger. As time passed, the diversity of the ecosystem fluctuated: sometimes there were very few species, sometimes a lot. It was all rather confusing. Then 45-bit parasites emerged. Lacking their own copying instructions, they 'borrowed' them from nearby organisms. In some runs of the program the ancestral organism then mutated, becoming 79 bits long and resistant to the parasites, so the parasites died out. In other runs, 'hyperparasites' appeared, which subverted the parasites' method of replication and used it to replicate themselves. Some of the hyperparasites evolved into social organisms with 61 bits, which replicated only by mutual cooperation. Their existence paved the way for 27-bit cheats, which hijacked the entire program by stealing control from the social organisms.

Tierra may only be a random sea of bits in a computer's memory, but 'all life is there'. It strongly reinforces the view, central to this book, that the patterns exploited by biology arise 'for free' from mathematics. Ray did not *instruct* his bit-strings to become parasites, or hyperparasites, or to cooperate.

They did it anyway.

However, he did instruct them to reproduce, by including an explicit 'copy' command in his computer language. You've got to start somewhere, but the whole story would be a lot more convincing if the ability to replicate itself arose 'for free'. Because then you'd be modelling the origin of life, not just what happens to it once it has arisen.

Ambitious? Improbable? Of course. But science won't advance at all if we pursue only lines of attack whose success can be predicted in advance.

In 1996 Andrew Pargellis unveiled his own artificial life program: 'Amoeba'. Ray had 'played God' by seeding the computer's memory with a specially designed replicator, but Pargellis started with just a random block of computer code. Every 100,000 computational steps the program wiped out 7% of the memory slots and replaced them with randomly chosen commands. He found that about every fifty million steps a self-replicating segment of code appeared. Replication didn't have to be built into the rules — it just happened.

Systems like *Tierra* and *Amoeba*, without being given any *explicit* instruction to do so, display 'high level' patterns very similar to those found in real terrestrial evolution. These include the spontaneous appearance of replicators, spontaneous increases in complexity, rudimentary forms of symbiosis and parasitism, lengthy periods of stasis

punctuated by rapid changes — even a kind of sexual reproduction. The message is that all of these puzzling phenomena are entirely natural: they are typical properties of complex adaptive systems. Instead of being surprised when we see them in the evolutionary record, we should be surprised if we did not.

Implications

These are striking discoveries, but what is their significance? Does artificial life really tell us anything useful about real life?

I think it does. I can best describe why by invoking the concept of phase space — a geometric image in which every event that does happen is surrounded by a ghostly halo of 'nearby' events that didn't — but could have. When you set up any mathematical system, be it a classical dynamical system or the kind of thing used by devotees of artificial life, you also implicitly set up a phase space. Phase spaces are *big* — they contain all possibilities, not just a selection. If the rule system is sufficiently 'rich' — which basically means not horribly boring and obvious — then all sorts of possibilities lurk within its phase space. Now we begin to see the significance of mutations in evolution. They don't just make evolution possible: they enable the system to explore its phase space. The states that it is occupying today may change tomorrow. We also see the role of selection more clearly: it makes the exploration efficient. If all that happened were random mutations, the system would wander around in its phase space like a drunkard, tottering one step forward, two steps back. Indeed the mathematics of 'random walks' shows that such systems spend an awful lot of time revisiting old haunts. But with selection, bits of phase space that 'don't work' are eliminated. Selection helps the system to home in on the interesting regions of phase space, the places where useful things happen, the central features of the evolutionary landscape.

The phase space for real terrestrial evolution is far more complicated than that for *Tierra* or *Amoeba*. But it plays the same role. Its rules are those of the physical universe. Mutations allow life on Earth to explore the evolutionary phase space, selection cuts down the possibilities so that evolution doesn't spend all of its time wandering up dead ends or revisiting places that don't contain anything interesting. Their combined effect creates a geography of phase space, making it more like a landscape than a featureless plain; evolution homes in on the more significant features, behaving *as if* it has goals when actually it is being driven by the geography of its phase space. Evolution does not know where it is heading — but if we could see its phase space, we'd get a pretty good idea.

Stuart Kauffman, a highly original scientist who has thought deeply about such matters, sees this kind of structure as being characteristic not just of evolution, but of *any* process that has the ability to complicate and organise itself. Instead of 'phase space' he talks of the 'space of the adjacent possible'. And instead of merely pointing out that phase spaces have a geography, hence a dynamic, he believes that we may soon be able to state precise mathematical laws that govern how a system explores the space of the adjacent possible. And he is convinced that those laws will be a lot closer to 'as fast as possible without falling to bits' than they are to 'with no purpose and no sense of direction'.

FURTHER READING

- Jack Cohen and Ian Stewart, *The Collapse of Chaos*, Viking, New York 1994.
- Charles Darwin, *The Origin of Species*, Penguin Books, Harmondsworth 1985.
- Richard Dawkins, *The Blind Watchmaker*, Longman, London 1986.
- Richard Dawkins, *The Selfish Gene*, Oxford University Press, Oxford 1989.
- Richard Dawkins, *The Extended Phenotype*, Oxford University Press, Oxford 1982.
- Richard Dawkins, *River out of Eden*, Weidenfeld and Nicolson, London 1995.
- Richard Dawkins, *Climbing Mount Improbable*, Viking, London 1996.
- N. Eldredge and S.J.Gould, Punctuated equilibria: an alternative to phyletic gradualism, in T.J.M.Schopf (editor), *Models in Palaeobiology*, Freeman, Cooper & Co., San Francisco 1972.
- Paul Guinnessy, Life crawls out of the digital soup, *New Scientist* (13 April 1996) p.16.
- Stuart A. Kauffman, *The Origins of Order*, Oxford University Press, Oxford 1993.
- Stuart A. Kauffman, *At Home in the Universe*, Viking, New York 1995.
- D.Lack, *Darwin's Finches*, Cambridge University Press, Cambridge 1947.
- Roger Lewin, *Complexity*, Macmillan, New York 1992.
- Klaus Mainzer, *Thinking in Complexity*, Springer-Verlag, Berlin 1994.
- Tim Poston and Ian Stewart, *Catastrophe Theory and Its Applications*, Pitman, London 1978.
- Mark Ridley, *Evolution*, Blackwell Science, Oxford 1996.
- V. Sarich and A.C.Wilson, Immunological timescale for human evolution, *Science* **158** (1967) 1200-1203.
- C.G.Sibley and J.E.Ahlquist, DNA hybridization evidence of hominoid phylogeny: results from an expanded data set, *Journal of Molecular Evolution* **26** (1987) 99-121.
- Karl Sigmund, *Games of Life*, Oxford University Press, Oxford 1993.
- John Maynard Smith, *Evolution and the Theory of Games*, Cambridge University Press, Cambridge 1978.
- Ian Stewart and Jack Cohen, *Figments of Reality*, Cambridge University Press, Cambridge 1997.
- Ian Stewart, *Life's Other Secret*, Penguin, to appear May 1998.
- René Thom, *Structural Stability and Morphogenesis* (translated by David H. Fowler), Benjamin-Addison-Wesley, New York 1975.
- Mitchell Waldrop, *Complexity*, Simon and Schuster, New York 1992.
- Sewall Wright, *Evolution: Selected Papers* (editor W.B.Provine), University of Chicago Press, Chicago 1986.
- E.C.Zeeman, *Catastrophe Theory: Selected Papers 1972-77*, Addison-Wesely, Reading MA 1977.
- E.C.Zeeman, Evolution and elementary catastrophes, preprint, Math. Institute, University of Warwick 1988.

GRESHAM COLLEGE

Policy & Objectives

An independently funded educational institution, Gresham College exists

- to continue the free public lectures which have been given for 400 years, and to reinterpret the 'new learning' of Sir Thomas Gresham's day in contemporary terms;
- to engage in study, teaching and research, particularly in those disciplines represented by the Gresham Professors;
- to foster academic consideration of contemporary problems;
- to challenge those who live or work in the City of London to engage in intellectual debate on those subjects in which the City has a proper concern; and to provide a window on the City for learned societies, both national and international.

Gresham College, Barnard's Inn Hall, Holborn, London EC1N 2HH
Tel: 020 7831 0575 Fax: 020 7831 5208
e-mail: enquiries@gresham.ac.uk