

The Evolution and Creation of Diversity

February 21, 2002

1 Evolution versus Creation Revisited

In this lecture, we extend our earlier discussion of the evolution and the creation of diversity. In an evolutionary model, the development of new characteristics that produce greater relative fitness occurs through mutation and sexual recombination. There is no intentionality to this process. Genes that increase survival and reproductive rates are more likely to be present in a population than those that do not.¹. Even though our primary interest is the evolution and creation of diversity, we will spend much of our discussion chatting about “fitness” and “efficiency” because these are the keys to understanding evolution and creation. Having said that, for the most part “fitness” nor “efficiency” lack logically sound definitions, as each is defined contextually and therefore somewhat circularly: “fitness is the capacity to reproduce sustainable offspring.” (It’s not quite as tight a definition as one could give for say an apple.)

In our earlier handout, we identified differences between evolution and creation. Genetic evolution is severely constrained for obvious reasons. Mutation rates must be moderate in order for a sufficient number of offspring to be viable, but this limits the amount that an offspring can differ from its parents. In a creative process, leaps in space can be huge. We will suggest two ways the evolutionary systems can act as if they have larger mutation rates: sexual differentiation and neutral mutations. The evolution of sexes allows for greater mixing of characteristics than might otherwise occur. Neutral mutation exploits the complexity of our DNA sequences to expand search.

Evolution is further constrained by the fact that each step along the path to an improvement must be viable and capable of reproduction. This contrasts with creative processes in which interim viability need not be a concern. For example, at the 2002 North American Auto Show in Detroit, Chrysler exhibited a vehicle prototype with a glass roof and they claimed that drop down video monitors were an option. This is not to say that the engineers didn’t see the problem. They admitted it. They just have to figure out away around it.

¹see Dawkins *The Selfish Gene among others*

We have to be careful to not lose site of the goal: to show how evolution and creation can lead to diversification. We will discuss three natural ways in which genetic evolution can lead to diversification: *branching*, *geographic variation* and *isolation of small populations*. We will then talk about how combining these with the idea of co-evolution causes a massive diversification. We then quickly turn to one unnatural way in which we get diversity: *selective breeding*, which we will argue is a hybrid between evolution and creation. We will then briefly turn to creation and diversity.

Before we get started, we have to reach an agreement between diversity and noise. If a population converges to a single type, there will always be some diversity in the form of minor mutations around that type. So when we speak of diversity, we mean differences greater than that generated by this mutational drift.

2 A Not So Complete Description of Biological Evolution

Not that we really care that much about this, but here are the basics of genetic evolution. The finest grain we know of for sure contains quarks and bosons and things like that. Below that may be higher dimensional strings. Who knows. Fortunately, we can combine quarks and atoms until we get up to *nucleotides*. These, not genes, form the fundamental units for evolution. They have three parts: a nitrogen base (A,C,G,T), a sugar molecule, and a phosphate molecule. Don't ask me what those do. A *gene* is a sequence of nucleotides. These sequences have specific orders. Again, I don't have any idea why this matters, but suffice it to say that some orders work and others do not. The value that a *gene* assumes among these many possibilities is called an *allele*. *Chromosomes* are sequence of genes which in turn are sequences of nucleotides. Chromosomes are where much of the action takes place. They are the self replicating part. We have twenty three pairs of chromosomes, one of which is a sex chromosome. You get 22 pairs of autosomal chromosomes (non sex chromosomes): one from each parent. These form the double stranded DNA. You only get one sex chromosome.

Your genes do not directly determine your behavior, nor is that true for other species. Genes create a probability distribution over phenotypes. Phenotypic fitness determines reproduction, but genes (probabilistically) create the phenotypes, so genes are usually the appropriate unit of analysis.

3 Rugged and Dancing Landscapes

Throughout our analysis we will rely on two metaphors: *rugged landscapes* and *dancing landscapes*. Think of a rugged landscape as a relief map of a mountainous region such as the Alps, the Rockies, or the Andes. Points of high altitude correspond to points of high fitness. In this metaphor, the geographic placement of a member of

a population of some species corresponds to its genotype. Mutations and crossover create new genotypes and new locations on the landscape. Individuals with higher altitude reproduce faster. Over time, the population of species improves in average fitness.

Notice though that the population need not settle on the highest peak. Imagine a landscape with just two peaks. If the population begins in a basin of the first peak it is not likely to make the leap to the other one. In this case, we would say that the population gets stuck on a suboptimal local peak. A two peaked landscape would typically not count as *rugged*. That's why you should have in mind a mountain range with lots of valleys, peaks, and potential goat paths. It is along those goat paths that evolution creeps toward higher peaks. If we take a snapshot at any point in time, we would not expect a species to be at the global optimum. All we can expect is that it probable is not sitting in the bottom of a valley.

The rugged landscape metaphor has great intuitive appeal. Too bad it's wrong. In evolutionary systems, fitness is context dependent. The fitness of a worm depends on the characteristics and behaviors of birds (and anglers). For any fixed set of characteristics and behaviors of competitors and symbiont's, we do have a rugged landscape, but once a species related to our species changes, it shifts our species' landscape. Metaphorically, we can capture this as *dancing landscapes*.

Suppose that we have only two species: lions and gazelles. Lions eat gazelles, and gazelles eat grass. Gazelles might evolve coats that blend into the savanna and short legs, making them harder to spot. This might place them on a local hill on their landscape. But if lions then evolve ocular or olfactory capacity to identify the gazelles, those short legs place the gazelles at the bottom of a valley on their landscape. In response, they will evolve long legs. We have to be careful here. The gazelles do not choose do evolve long legs. It's just that those with long legs happen to get eaten less frequently so they reproduce.

In textbooks, the more technical term *co-evolution* is used instead of dancing landscapes. But the dancing landscapes metaphor captures co evolutionary phenomena beautifully, so we will use it.

4 A Simple Model of Evolution

To understand how evolution creates diversity, we will construct a a simple version of an evolutionary process known as a *genetic algorithm*. Genetic algorithms were developed by John Holland of the University of Michigan in the mid 1970's. The idea is to think of a solution to a problem as a sequence of variables akin to human DNA and to evolve solutions.

In this case, our set of objects Ω will be sequences of variables. To keep things simple, we will have these be binary strings.

Def'n: The set of **objects** $\Omega = \{1, 2, \dots, \omega\}$ consists of binary strings. If $a \in \Omega$ then $a = a_1 a_2 a_3 \dots a_N$ with $a_i \in \{0, 1\}$.

We will refer to a_i as the i th attribute. We're going to shift back and forth between thinking of these attributes as genes and as physical attributes. The advantage of doing the latter is that it makes the arguments easier to follow. The former has two advantages: it's accurate and it allows us to make more subtle arguments such as how neutral mutation works. To frame our discussion, we will assume that the object we are evolving is dogs. So an attribute might be whether a dog licks or not.

Think back to our lecture on the diversity of perspectives. Evolution uses our genes as an encoding of our physical characteristics and tendencies. This perspective differs substantially than the interpretations that doctors or artists use to describe the human body. DNA is a terrible way to communicate information for most practical purposes. If you wanted to say that I wear size twelve shoes, you would probably have to send a complicated description of my DNA to get that point across. That is not to say that some phenotypes, such as Downs' Syndrome cannot be explained easily using chromosomal information, but the key point here is that when I looked at my older son's chromosomal mapping with a genetic counselor, we didn't both immediately say anything like "well, he'll play basketball, he's left handed, he won't like chicken, and we'll have to do something about that big mole on his left earlobe." All we said was "it's a boy!"

4.1 Reproduction and Fitness

The likelihood that a dog reproduces could depend on many factors: its genes, its geographic environment, the attributes and size of the relevant population of dogs, the attributes and size of other species with which it interacts. For the moment, we assume that only genes matter. There are good genes and there are bad genes. We can then define a *fitness function* $F(\cdot)$ which assigns a fitness value between 0 and 1 to any object, in our case to any dog. This fitness function creates a rugged landscape. So, when we say only genes matter for fitness, this implies that the landscape does not dance.

5 More than Mutation

Mutation in this model is just the random flipping of a gene that should have been a one to a zero. As we discussed previously, this does not allow for much search. Evolution would seem rather constrained and it seems unlikely that we would have evolved frogs let alone us. Evolution has more tools than just mutation, we'll discuss two: sexual reproduction and neutral mutation. To give evolution full credit, it evolved (probably²) both of them.

²Alternatively, the nonlinearities in the neutral mutation story could be an artifact of our chemistry that it was lucky enough to be able to exploit.

5.1 Sexual Reproduction

Adding sexual reproduction has several implications. Instead of having haploid (single strand) chromosomes, you get diploid (double stranded) chromosomes. This increases robustness and genetic memory among other things. Our interest is in how sexual reproduction can speed the rate of search. The example below shows how sexual reproduction combines good attributes.

Sexual Reproduction and Search: Suppose that there is one female out of ten with a beneficial mutation (an extra attribute with value zero) and one male out of ten with a different beneficial mutation. Form twenty random pairs of dogs to create offspring. Let the probability that the female (or male) with the extra attribute with value zero equal one eighth. This implies that the probability that one of the other dogs is chosen equals $\frac{7}{8}$. The probability of a pair having both the highest fitness female and the highest fitness male equals $\frac{1}{64}$. The probability of having one of these two in a pair equals $\frac{14}{64}$. The offspring in these pairs only have a 50% chance of keeping the beneficial mutation. Therefore, ignoring mutation effects, the new population will have one eighth as opposed to one tenth of its members with at least one beneficial mutation. This is the same as would have been true with asexual reproduction, but one eighth of those (or $\frac{1}{64}$ of the total) will have *both* beneficial mutations.

John Holland refers to these good attributes as building blocks. In technological evolution, we often combine ideas. Cup holders appeared on golf carts, then lawn chairs, then cars, and now you can even buy leather arm chairs with cup holders and little coolers.

5.2 neutral mutation

The metaphor we want to keep in mind is that these dogs are collections of attributes and that evolution through mutation and crossover creates new collections of attributes. Suppose that a population has converged (or that reproduction is asexual) so that all attributional novelty comes from mutation. Let the number of attributes be ten. The object 00000000 has ten one mutant neighbors out of over one thousand possible neighbors. If there are thirty attributes, an object has only thirty neighbors out of over a billion possible objects. From a search standpoint this is lousy. It would be easy to get stuck on a local optimum.

How then to get around this problem? Think of the attributes as chromosomes. Suppose that individual chromosomes do not matter, but that pairs of chromosomes do matter. For example, the chromosomes 011111 might create the same dog as the chromosomes 111111. This is true of our own DNA, much of it is thought to be “junk” or the residue of our evolutionary path. Advocate of neutral mutation theory argue that it is not junk, that it allows for greater search. To see how this works, suppose that our objects have thirty attributes and that they can be partitioned into

six clusters of five attributes

00000 00000 00000 00000 00000 00000

Suppose that in each cluster of attributes, that the switching the first attribute has no effect, so 00000 and 10000 are no different. However, suppose that switching the other attributes has an effect. And (here's the important part) suppose that if the first attribute in a cluster changes along with another attribute that the joint effect differs from when just the second attribute is changed. For example 00100 differs from 00000 which is the same as 10000, but 10100 is different from both 00100 and 00000. This matters because it is possible to drift from 00000 to 10100 because the mutation to 10000 has no effect. Thus neutral mutation, expands the neighborhood size.

In this example, it is easy to calculate how much the neighborhood size increases. Each cluster instead of having five neighbors, now has eight. This makes for a total of forty eight neighbors instead of thirty. This is over a fifty percent increase in the size of the search space.

Suppose that the first two attributes are now neutral mutants. Flipping either one alone or the two together has no effect. This means that those two attributes can be in any of four configurations 00, 01, 10, or 11. For each of these four configurations there are three possible one attribute mutants on the other attributes in the cluster for a total of twelve neighbors:

00000 00001 00010 00100
01000 01001 01010 01100
10000 10001 10010 10100
11000 11001 11010 11100

Now we're getting somewhere. There are seventy-two neighboring objects, a more than doubling of the neighborhood size.

6 Diversity on Rugged Landscapes

I've always loved the phrase "getting our feet wet" because it's the inverse of evolution. Even though no one says "getting our fins dry" it seems more apropos of our endeavor. Anyway, we begin by looking at the causes of diversity on rugged landscapes. After our toes and insteps are sufficiently damp, we'll let those landscapes dance.

6.0.1 A Sexless Life On A Rugged Landscape

Assume first that reproduction dogs happens asexually. A dog's offspring differ from the dog only through mutation. If this were the case, we might still see diversity arise through what I will call functional branching. Functional branching corresponds to the fitness landscape having multiple peaks.

Diversity Through Functional Branching: Assume there are N attributes. Let $F(a)$ equal the larger of two numbers: $A(1)$ which is the number of a_i that equal one or $A(0)$ which is the number of a_i that equal zero. This fitness function has two optima: $a^0 = 00000 \dots 0$ and $a^1 = 11111 \dots 1$, both have equal fitness.

Let's construct a play with a specific example. Suppose that there are five attributes. Suppose that in it's life time, each dog has exactly two offspring and that the probability of an offspring surviving equals $\frac{3F(a)}{4N}$. So, if a dog had the genotype 00110, its fitness would be three, and the probability that an offspring of this dog survives equals $\frac{9}{20}$ or less than one half. Alternatively, the offspring of a dog with genes 11111, has a survival probability of $\frac{3}{4}$, so this dog expects to have one and a half offspring (the technical term is puppies) survive.

Constructing a computer program to experiment with this model takes about 45 minutes. Here are three typical runs. For each period, I show the population size (which grows), the average reproduction rate, and the percentage of dogs whose genes are all 0's or all 1's.

Period	Population	Repro Rate	% All 0's	% All 1's
1	20	0.510	0.05	0.05
10	48	0.603	0.31	0.03
20	415	0.590	0.25	0.00
30	2660	0.608	0.30	0.01
40	19656	0.612	0.30	0.02

Period	Population	Repro Rate	% All 0's	% All 1's
1	20	0.517	0.07	0.11
10	122	0.610	0.07	0.27
20	1008	0.617	0.09	0.24
30	7289	0.613	0.08	0.25
40	55352	0.612	0.09	0.24

Period	Population	Re pro Rate	% All 0's	% All 1's
1	20	0.510	0.06	0.12
10	42	0.632	0.28	0.28
20	406	0.635	0.16	0.24
30	3718	0.620	0.13	0.23
40	29150	0.613	0.11	0.22

There are many things to discuss about this model, but we want to focus on three. First, we see diversity. In each run, we get both dogs with all zero genes and dogs with all one genes. Evolution found both peaks. Even though this is not really a rugged landscape for explanatory purposes is what rugged enough to show how diversity can result from the segments of the population climbing different hills.

Second, the population grows over time. Populations need not stay stable, but they cannot continue to increase forever, at some point resources become scarce. When the population is growing, we have something much less like survival of the fittest than death of the unfit. You don't have to be that fit to survive.

Third, the average reproduction rate (recall that this is linear in fitness) levels off to a value of around 0.613, which an average of about three and one quarter bits agreeing. It would be tedious but not hard to prove that this is the expected limiting average reproduction rate in this model.

If you are puzzled why reproduction rate does not converge to 0.75 recall two features of our model : (i) the mutation rate implies that offspring of maximally fit dogs need not be maximally fit, and (ii) those dogs with four matching genes also grow in the population. Their growth rate isn't as fast as the growth rate for those dogs whose genes all match. If we vary the mutation rate and measure the average reproduction rate in the population after fifty periods, we see that the mutation rate is the main culprit. We could get rid of mutation but then we would have no exploration.

Mutation Rate	Fitness Average
10%	0.550
5%	0.612
4%	0.633
3%	0.656
2%	0.683
1%	0.714

This example seems a bit contrived because the two solutions have the same value. We'll can get rid of that assumption. All that matters is that each of the

peaks corresponds to a sustainable reproduction rate. This example takes on special relevance when we think about applying evolutionary thinking to social phenomena. We have already talked about the enormous diversity in people. Part of the reason is that our evolutionary environment is quite forgiving. We could call it either “survival of even the partially alert” or “death of only the really stupid.”³ If we date the wrong person, pay too much for a car, or choose the wrong job, we are not likely to be eradicated.

Admittedly, firms, organizations, and governments face greater selective pressures. This is thought to be especially true for markets; you will hear people say that firms that function inefficiently do not survive long in the marketplace. While market forces to militate against inefficiencies, firms are full of them. It may be that the same skills that allow an organization has chosen turn a profit (say for example placing trust in employees) cause the organization to do so inefficiently. Now, someone might argue that if they have to trust their employees to turn a profit, this is optimal. That’s not what I am saying. It may be that they can turn a larger profit with another mechanism. They just haven’t thought of it yet.

My friend John Miller tells the story of two men in the woods who see an angry bear approaching. One kicks off his boots and slips on his running shoes. The other says “you idiot, you cannot outrun a bear.” To which the first replies “I’m not trying to outrun the bear.” Moral: competition among firms is about being relatively better, of having a higher point on the landscape, not about finding the highest point.

Further, the same survival of the fittest idea applies to other non market organizations as well from charity groups to theatrical societies and even to governments. A point we will note and then abandon temporarily.⁴

6.1 Geographic Separation

So far, we have just been nibbling around the edges of the speciation story. The primary driver in speciation is geographic separation. Ernst Mayr and others built upon Darwin’s model of evolution to show how geographic segregation creates new species. Imagine that there are two identical packs of dogs living in a lovely river valley. Suppose that we take one pack and move it to a mountain location. Assume that these dogs are viable in this new ecosystem. Over many reproductive cycles, the dogs will evolve attributes that are best suited for their location.

Speciation can happen very fast. Suppose that ones are preferred to zeroes in the valley and zeroes are preferred to ones on the mountain. Suppose that there are twenty dogs in the pack, fifty attributes and a mutation rate of only 2%. First, assume asexual reproduction. This means that each offspring will have on average one attribute that gets switched. For the dogs on the mountain, initially, every attribute switch will be beneficial. After k attributes have been switched the probability of

³See the “Darwin Awards” web page for examples supporting this line of argument.

⁴Time permitting we will talk about the creative process in markets versus government. The latter does not allow as much room for experimentation for many reasons.

switching a one to a zero falls to $\frac{N-k}{N}$. To see how fast this can happen, we can run our previous model with an initial population of dogs with almost all zero genes and let the reproductive rate equal the number of ones plus one half (so that the dogs with no ones have some chance of having their offspring survive) divided by five.

Period	Population	Repro Rate
1	1000	0.054
2	109	0.262
3	58	0.331
4	43	0.409
5	33	0.552
6	40	0.615
7	50	0.692
8	66	0.764
9	97	0.823
10	158	0.842
11	259	0.873
12	446	0.882
13	783	0.883
14	1375	0.882
15	2428	0.881

Here we see that initially the population falls off to almost nothing but then it quickly bounces back up. Notice also that the reproduction rate gets very high. Metaphorically, what has happened here is that by putting the dogs on the mountain, we placed them in a fitness valley, which through genetic evolution, they escaped.

One quick caveat. Geographic segregation can happen in the same physical location. Suppose that there is a single attribute that determines whether you are awake during the day (diurnal) or at night (nocturnal) so two dogs must agree on this gene in order to reproduce. It's as though the animals work different shifts at a factory.

6.2 Small Populations

Species that get segregated can evolve into distinct species even if the ecosystems they live in do not differ in any meaningful way. Smaller population are more likely to have unlikely features. This is a consequence of the law of large numbers. If you flip a fair coin a million times, you are likely to get about half heads and half tails. If you flip it ten times you get eight heads about 9% of the time. Similarly, if you grab a million birds that are various shades of gray distributed evenly from 5 to 6 on a scale, you are likely to see a mean greyness of about 5.5. In contrast, if

you choose twenty birds, their average may be 5.8. Given that this is true for every attribute, the two populations are likely to be very different. This matters because the genetic composition determines the starting point for “search” for better genes. If you believe, as I do, that evolution creates forces for improvement, but few species are “optimal” then having different starting points matters.

7 Dancing Landscapes

The real action occurs when we let the landscapes dance. We are going to be brief about this even though it may be the most important part of the story. We can get away with this because even though it is important, it’s obvious. If we have two geographically segregated sets of co-evolving populations of two species, they are not likely to look that similar after awhile.

Why? Think it through. To do this, construct a model with two islands, and the same two species (A and B) on each island in the same proportions. Suppose that on the first island, one of the species (A) evolves something that makes it faster, but on the other island the same species evolves something to make it more powerful. Now the two landscapes for the other species (B) do not look very much alike as they are solving different problems. Therefore, they are not likely to make the same improvements, and now the landscapes for the A species will be different. Once the landscapes differ, dancing or not, the likely improvements differ and we get divergent paths.

The story is a little more complicated then that, but not much.

8 Creation

This is topic of your homework: *how creation creates diversity*. To get your mind headed in the right direction, think about the selective breeding. First dogs. This has gone on for centuries. Even though all dogs from my Great Pyrenees to my grandparents Shiatsu have identical bone structures (just of different sizes), they have been able to bend and shape them to create an amazing (ridiculous?) diversity of breeds, including some that cannot breathe well like pugs.

If you’re tired of talking about dogs, think about flowers, fruits, and vegetables. Broccoli is a mix of kale and cauliflower. The man who developed it produced all of the early James Bond films with his money. The list continues to grow from navel oranges to tangelos, pluots, and more types of apples then I dare count.

8.1 Homework

I want you to take either recipes or cars and compare the options available in 1980 versus now. And then discern which of the new solutions are due to technology (cell phones and cd players would be a yes as would some types of salt) or due to people

just thinking about the problem differently (mini-vans or fish tacos). This should be easy because you can probably use the web. If you are doing cars compare the 1980 X to the 2000 Y (or if you take the mustang you can use the same car each time). If you are doing food, find recipes for say chicken in 1980 versus those now.