

Differential Fitness Transmission: Detecting Darwinian Evolution with Genealogic Records

Thomas Miconi (corresponding author)
t.miconi@cs.bham.ac.uk
(+44) 121 414 8201

Alastair Channon
a.d.channon@cs.bham.ac.uk
(+44) 121 414 4770

School of Computer Science
University of Birmingham
Edgbaston
Birmingham B152TT, UK

November 15, 2006

Abstract

We introduce differential fitness transmission as a signature of adaptive, Darwinian evolution, which can be detected using genealogical records of a reproducing population. This method is motivated by the observation that Darwinian evolution fundamentally consists in the differential transmission of heritable, fitness-affecting traits that result in differential transmission of fitness itself: fitter parents should tend to produce fitter offspring. Based on this idea, we propose several statistics which allow us to detect the presence of differential fitness transmission under various conditions. As an experimental illustration, we apply our statistics to simple evolutionary algorithms using different selection and replacement regimes. We demonstrate that differential fitness transmission can be used to detect the presence of a force which consistently favours certain lineages at the expense of others, over more than one generation. We conclude that differential fitness transmission is a useful, practical signature of Darwinian evolution for situations in which a genealogical record of the population can be obtained.

Keywords: Detecting evolution, evolutionary statistics, evolutionary activity, evolutionary dynamics, adaptation, fitness.

1 Introduction

1.1 The problem: Are we evolving yet?

This paper is concerned with the following question. Consider a population of reproducing agents, which interact and reproduce according to the rules of their particular environment. How can we determine whether these agents are undergoing Darwinian evolution?

This question arises from the fact that when a population of reproducing agents is observed, it is not always immediately clear whether the dynamics of the population result from Darwinian evolution, or merely from random variations and stochastic effects such as genetic drift. The particular system at hand may also introduce its own effects, which may bias or alter the dynamics of the population in unpredictable ways. When this system is sufficiently complex, determining whether a population is evolving in a Darwinian sense may not be a trivial task.

Besides its conceptual implications, the question is of practical interest. It is often desirable to determine whether evolutionary adaptation is occurring within a given system, especially in the fields of evolutionary computation and artificial life. Indeed in some situations, the onset of significant adaptive evolutionary activity is by itself a major objective of the system: for example, artificial environments such as Echo [8] and Geb [6] were explicitly designed with the aim of exhibiting meaningful evolutionary activity. Being able to detect the presence of genuinely adaptive evolution is a fundamental pre-requisite for the validation of such systems.

1.2 What is evolution?

A problem which we must confront is that the term ‘evolution’ is sometimes used with two different meanings in mind:

1. Evolution as change in the makeup of populations. This is the original meaning of the word, and the one most frequently encountered in succinct definitions: for example, Ridley defines evolution as “the change in a lineage of populations between generations” in the glossary of his textbook [9]. This meaning is clearly implicit in such expressions as, for example, ‘rate of evolution’. We may label this notion as ‘evolution-the-change’.
2. Evolution as the Darwinian mechanism that explains this change, or rather the strikingly adaptive nature of this change: natural selection operating on heritable fitness-affecting characteristics, which are then differentially transmitted and thereby ‘steer’ the population in locally adaptive directions. This meaning is apparent in such expressions as ‘evolutionary theory’, ‘evolutionary explanation’ (which usually involves identifying the selective aspects of the particular trait to be explained), ‘evolutionary design’, ‘evolutionary algorithms’, etc. We can call this notion ‘evolution-the-mechanism’.

Of course this Darwinian mechanism, which explains evolutionary change, may also lead to local, temporary stability, for example if the population is currently located around a convenient local optimum: hence the concepts of ‘stabilising selection’ and ‘evolutionary stasis’ (see chapter 4 of [9]).

In this paper, we are interested in the latter of these two meanings, namely evolution-the-mechanism. Our aim is to provide a measure to detect the active presence of the Darwinian mechanism of evolution by natural selection. We want to detect the presence (or absence) of *Darwinian forces* which actively shape a population, as opposed to the mere accumulation of incidental, non-adaptive fluctuations - to catch the ‘invisible hand’ of Darwinian evolution in the act, as it were.

We are specifically *not* interested in measuring ‘progress’ or complexification, two notions sometimes associated with the idea of “measuring evolution” [4]. Neither are we trying to determine whether this evolution is somehow innovative, or successful in terms of optimisation. Our objective is simply to find out whether the trajectory of a given population of reproducing agents is the result of Darwinian adaptation through natural selection - no more, no less. A consequence is that our measure will give high marks for systems that are being actively molded by natural selection, even if this molding consists of actively maintaining the population around a given optimum. In other words, it will give high marks to systems undergoing stabilising selection. Interestingly our measure will differentiate such a system from ‘frozen’ systems in which no genetic variation occurs at all, which will be given a zero score. This is precisely what is required.

The design of such a measure requires thorough consideration of the mechanics of evolutionary processes, in order to make the notoriously subtle concept of evolution more explicit.

1.3 Differential Fitness Transmission: A signature of Darwinian evolution

We posit that the defining characteristic of Darwinian evolution is the differential propagation of heritable, fitness-affecting characteristics, which steer the population in locally adaptive directions. From this rather uncontroversial basis, we deduce that it should be possible to detect the presence or absence of Darwinian evolution from the observed genealogical dynamics of evolutionary processes. Roughly, our reasoning is that if any transmission of heritable fitness-affecting traits is occurring, then on the whole, individuals sharing a common lineage, which tend to inherit similar traits, should also tend to have some degree of similarity in their fitnesses. For example, we might expect that fitter (resp. less fit) parents should tend to have fitter (resp. less fit) children. Thus if heritable, fitness-affecting characteristics are being transmitted, a tenuous, but persistent *correlation* should exist between the fitness of individuals sharing a common lineage: the differential transmission of fitness-affecting characteristics should result in a *differential transmission of fitness*. This differential transmission of fitness is the genealogical signature which we believe characterises

the presence of active, Darwinian adaptation through natural selection based on heritable characteristics.

2 Related Work

2.1 The Bedau-Packard measure of evolutionary activity

Bedau and Packard [1, 2] have developed a set of concepts and methods to “discern whether or not evolution is taking place in an observed system”. It should be noted that Bedau and Packard’s aim in that work is somewhat different from ours here. Here we are interested in finding out whether the trajectory of a population in configuration space is being actively shaped by Darwinian evolution. Bedau and Packard, on the other hand, are specifically interested in the *innovations* produced by evolution, and in the capacities of various systems to keep on producing adaptive innovations over time. This requires a method to determine whether an apparent innovation is indeed adaptive or merely the result of random fluctuations, which clearly relates to our own concerns. To this end, Bedau and Packard introduce a set of methods to compute the “*evolutionary activity*” of components and, by extension, of systems. These methods form the basis of a classification of systems into various classes, each class being associated with a certain type of evolutionary dynamics.

The Bedau-Packard measures of evolutionary activity are based on *persistence* of adaptive innovations: they identify components that persist over time at a level that exceeds what would be expected under purely random conditions. In a short introduction to the framework [3], Bullock and Bedau provide the following rationale:

A paradigm of the idea of evolutionary activity is the notion of measuring the continued persistence of elements in an adaptive system as a proxy for their adaptive significance – if a particular element persists in the system for a long time, this is likely to be because it is being maintained by selection.

If we are to use persistence “for a long time” as a criterion for detecting evolution, we need a method to determine what “a long time” is. When do we decide that a given element has persisted long enough to be regarded as ‘adapted’? It is clearly impractical to provide a general answer to this question *a priori*, which could apply to all systems and in all situations. To tackle this problem, Bedau and Packard introduced the idea of using a neutral “shadow” of the system under study: a replication of the original system, in which birth, reproduction and death of individuals occur in synchronisation with the real system, but are applied to randomly chosen individuals. More precisely, every time a new individual is being created in the real system under study, a new individual is also created in the shadow; but with the difference that, in the shadow, the parents of the new individual are chosen randomly. Thus the neutral shadow is expected to show the behaviour that would be seen in the

system, in the absence of any selective pressure. By comparing the persistence data obtained in this “shadow” to that obtained in the real system, Bedau and Packard argue, it should be possible to detect whether selection and adaptive evolution are present.

Building upon the concept of enduring persistence as a measure of evolutionary activity, Packard and Bedau have developed a series of evolutionary statistics based on persistence information. These statistics include:

- Diversity D : The number of different components present at a given time in the population.
- Activity $a_i(t)$: The age of component i at time t , indicating how long it has persisted so far.
- Cumulated activity $A_{\text{cum}}(t)$: The sum of the ages of all individuals present at time t .
- New activity $A_{\text{new}}(t)$: Perhaps the most important statistic, which sums the activity of all ‘new’ (Bedau and Packard use the term ‘*adolescent*’) components present in the system at time t . Entities are ‘new’ if their age a_i falls within a certain interval $[a_0, a_1]$. The idea is that a_0 is the minimum age at which a new component can be regarded as adaptively significant: if a component reaches age a_0 , then we assume that it has persisted more than could be accounted for by mere random fluctuations, and that this persistence is a sign of its adaptive value. A suitable value for a_0 is to be determined empirically through comparison with a neutral ‘shadow’ system. $A_{\text{new}}(t)$ is therefore a measure of significant *innovation* taking place in the system at a given time.

Bedau and Packard have used these statistics as a basis for classification of evolutionary dynamics into three classes: Class I (“no adaptive evolutionary activity”), Class II (“bounded adaptive evolutionary activity”) and Class III (“unbounded adaptive evolutionary activity”). Class III, in particular, is expected to encompass truly ‘lifelike’ evolution, similar in creativity and endurance to that found in nature. It requires that the total evolutionary activity A_{cum} (the sum of the ages of components present at time t in the population) be unbounded, that is, grow at least linearly with t , and that new activity $A_{\text{new}}(t)$ be positive.

2.2 ‘Evolution-the-change’ and ‘evolution-the-mechanism’

Bedau and Packard’s measures are arguably the most widely known of their kind. They have been applied to several systems, including artificial ecologies such as Echo, and natural components such as the genera within the fossil record [2]. Other researchers have applied them to various systems [10, 5]. Clearly the Bedau-Packard measures of evolution have been found successful. Why, then, do we feel the need to introduce another method for detecting Darwinian evolution?

Our basic motivation is that while the Bedau-Packard method is a powerful tool for detecting evolution in terms of enduring novelty, the question of how this novelty is brought about is left open. In other words, we believe that the Bedau-Packard measures assess ‘evolution-the-change’, rather than ‘evolution-the-mechanism’. Because our objective is to detect the presence of Darwinian evolution based on natural selection of heritable characteristics (that is, evolution-the-mechanism), we feel that introducing another tool for this purpose is justified.

Let us give an example of a system which illustrates this distinction. Consider a population of individuals in which reproduction and removal occur completely at random, except for the following caveat: every now and then, a certain randomly chosen individual is allowed to reproduce more than the others. Clearly this process is not Darwinian; crucially, we should not expect the children of this ‘lucky’ individual to show any exceptional abilities. Whichever notion of ‘adaptation’ we use, no novel adaptive ability would be durably introduced into the system.

However, in this system, the activity $a(t)$ of the genetic components associated with the favoured individuals would actually be positive. The reason is that the genes of the individuals which get this ‘boost’ will become more numerous, and thus will be expected to persist for a longer time. But in the shadow system, the children would be evenly distributed among randomly chosen parents, so no particular group of genes would be favoured. Thus a difference would be created between the shadow and the real system, and a positive signal would be detected by persistence statistics, despite the absence of Darwinian evolution.

To take an illustration, let us picture a population of the proverbial gazelles, in which every now and then certain individuals receive some kind of ‘boost’; giving free rein to fancy, we could imagine that benevolent safarists equipped with hypodermic guns would randomly select a few gazelles from the herd and inject them with some performance-enhancing substance that allows these particular gazelles to be much better at whatever gazelles do, such as grazing, attracting suitable mates, and of course escaping the equally proverbial jaguars. Owing to their enhanced capacities, one might expect these ‘doped’ gazelles to obtain remarkable reproductive success: their artificially enhanced characteristics would be, in this limited sense, ‘selected’ for. But because these characteristics are not heritable, there is no reason why their children should exhibit the same success. Even if the higher reproductive success of these particular gazelles had been so remarkable as to significantly affect the gene pool of the population (which might happen if the population is very small, or the drug very potent!), the resulting change in genetic distribution could hardly be called Darwinian: the newer individuals would not obtain a much higher reproductive success than their ancestors, no matter how many genetic components they inherit from their successful parents.

However, the genes of the ‘boosted’ gazelles would persist longer than others in the real system, simply because there are more of them. But in the shadow system, because parents for every new child are chosen randomly, no such effect would occur. This would create a difference between the observed activity of

these genes between the real system and the shadow, and therefore a positive activity score. This illustrates the fact that a positive activity can be obtained in the absence of Darwinian evolution.

In summary, we believe that the Bedau-Packard measure targets a specific notion of evolution, namely ‘evolution-the-change’, rather than ‘evolution-the-mechanism’. The statistics, and the classification that relies on them, offer a comprehensive measure of persistent novelty generated within a reproductive system. The measure in this paper, by contrast, takes pretty much the opposite approach: here we focus on the question of whether or not a system is undergoing active Darwinian evolution, rather than the question of actual genetic change. In other words, we target ‘evolution-the-mechanism’ rather than ‘evolution-the-change’.

The following section provides a description and justification of our measure, based on an analysis of the Darwinian mechanism of evolution by natural selection.

3 Differential Fitness Transmission: An Indicator for Evolutionary Activity

3.1 Darwinian evolution: randomness, selection and heredity

In order to devise a reliable detector for the presence of Darwinian adaptation, it is of course necessary to make it clear what exactly we define as such. Evolution, taken in its broadest sense, simply evokes the idea that lineages may change over time, rather than being fixed for all eternity (as was commonly held before Lamarck and Darwin). Adaptive evolution occurs when this change (or, in some cases, the absence of change) is somehow adaptive with regard to the environment in which the evolution occurs, in that the newer members of these lineages find themselves better at surviving and multiplying in this environment than their immediate ancestors. Darwin’s fundamental insight is to have proposed natural selection, that is, differential reproductive success, caused by *heritable* characteristics, as the source of this phenomenon.

In short, Darwinian evolution automatically results from the existence of fitness-impacting, heritable variations. Variations that improve fitness will be propagated quickly, initiating thriving lineages; while those that reduce fitness will hinder their own propagation, creating feeble (or even quickly extinct) lineages. Thus lineages constantly branch out into variants, and the uneven distribution of these branches, being dramatically skewed towards those which result from fitness-enhancing variations, will result in the overall effect that the newer descendants of the original lineage will tend to be those better adapted to their current, local environment: heritable fitness-affecting variation will have “steered” the original lineage towards adaptive directions among all those encountered by mutational variations.

Note that although this process will usually result in a modification of the species over sufficiently long periods of time, it will also often result in temporary stasis. If a species happens to be located at a convenient local optimum in the fitness landscape, then many variations which depart from the optimum will tend to reduce the fitness of the individual. In this case the differential transmission of characteristics enforced by natural selection will actively maintain the population around the optimum: the population will be constantly steered back towards its current position. This also occurs when being close to the population average is by itself advantageous. This phenomenon, known as ‘stabilising selection’, is actually regarded as more common than directional selection (see [9], Chap. 4.4). Lungfish and horseshoe crabs, which have remained very similar (though not exactly identical) over the last 200 million years, are well-known examples of evolutionary stasis. Indeed, distinguishing the continuing action of Darwinian forces when the population remains globally stable (as opposed to mere absence of variation or reproduction, which is a passive, non-evolutionary phenomenon) is one of the challenges faced by any detector of Darwinian evolution.

3.2 The requirement of heredity

Why have we stressed the word ‘heritable’ in our discussion? It is because the requirement of heredity may be overlooked if too much attention is given to the selective aspects of evolution. Yet this requirement is absolutely central to the Darwinian mechanism of evolution. Selection is based on the idea that, owing to their differences, some individuals will have better reproductive success than others. But for this to result to translate into actual adaptive evolution, it is necessary that the differences being selected for (or against) be heritable.

It is of course perfectly possible to have selection (even adaptive selection) without evolution, if this selection is not based on heritable characteristics. Our example of the ‘doped gazelles’ is one illustration of such a situation. Let us take another example: an evolutionary process in which the genotype of each individual is determined randomly at birth, independently of its parent’s characteristics. There is still some selection, in that the individuals having obtained (by pure chance) a highly fit genotype will have more children. Furthermore, because the selection is based on the characteristics of the individual, it can be called adaptive rather than random. However, because the genotype of each individual is random, none of these changes will be transmitted to the children, thus the lineage will not be impacted. Therefore evolution cannot really be said to have taken place.

Thus the requirement of heritability can be seen as the fundamental difference between mere selection and genuine Darwinian evolution: while certain individuals may have better reproductive success than others, it will only translate into Darwinian evolution if this success is somehow related to characteristics which can be transmitted to the next generation.

3.3 “Survival of the fittest”

One example of how crucial the requirement of heredity is, relates to the expression “Survival of the fittest” and its common misusing. At the suggestion of several correspondents, including Wallace, Darwin attempted to make his argument more intuitive by adopting Herbert Spencer’s vivid term of “Survival of the fittest”. Unfortunately this has led to even more confusion. A common, misguided criticism of the concept of Darwinian evolution is that the expression “survival of the fittest” is merely a tautology: if fitness is defined after chances of survival and reproduction (as it should be), then it can be rewritten as “survivors survive”, or “reproducers reproduce”.

Of course this misunderstanding results from the fact that the expression “survival of the fittest”, taken at face value, is actually an incomplete account of the process of evolution through natural selection. While it does capture rather efficiently the concept of natural selection in the strictest sense, it overlooks the crucial aspect which allows this selection to result in adaptive evolution - that is, the requirement that this selection should bear on heritable characteristics. When the full picture is considered, no tautology exists: the complete mechanism leading from heritable fitness-impacting differences, through differential reproductive success (“survival of the fittest”), to actual adaptive evolution (change in the makeup of lineages toward better adaptation) is a valid reasoning, hinging on the testable hypothesis that such fitness-impacting heritable variations actually exist. Thus while “survival of the fittest”, taken out of context, may well be a tautology, this cannot be used as a criticism of Darwin’s theory.

3.4 Differential Fitness Transmission: A genealogic signature of Darwinian forces in a reproductive system

We have stated that Darwinian evolution is based on differential reproductive success caused by heritable variations. When some individuals possess heritable traits which impact their fitness, these traits will be transmitted to their children, skew the distribution of the branching lineages, and thereby ‘steer’ the original lineage towards traits which improve reproductive success within the current environment - that is, better adaptation¹. In fact, we are of the opinion that Darwinian evolution is synonymous with the existence of such heritable, fitness-affecting characteristics and their (necessarily) differential transmission.

But if such a situation occurs; if fitness-impacting, heritable traits are actually being transmitted and propagated, then this should have an impact on the genealogical record: individuals sharing a common lineage, being more likely to inherit common characteristics, should therefore tend to have slightly more similar fitnesses. In other words, if some fitness-affecting traits are being transmitted, then there should be some degree of *correlation* between the fitnesses (that is, the reproductive success) of individuals from a common lineage: the

¹In short: As creatures replicate, traits propagate, adaptations proliferate, and species originate.

transmission of heritable, fitness-affecting traits should result in some degree of *differential transmission of fitness*.

Differential fitness transmission is our proposed signature for Darwinian evolution. It is, quite simply, the concept that individuals from a common lineage should tend to exhibit a tenuous, but persistent correlation in fitnesses.

This correlation would certainly be quite weak. Partly this is because the impact of any single trait on a particular individual's fitness is usually small. Reproductive success is determined by very many factors (chance being often an important one). In addition to this, children will usually not inherit all of their parent's traits, so the fitness-impacting traits may not be transmitted to all children. The result is that in general, the fitness of a particular child cannot be precisely estimated from the fitness of its parents alone. Heritable fitness-impacting traits do not predictably alter evolution at the level of a particular individual; but when many individuals are considered, then over a large enough sample a tenuous correlation should emerge between the reproductive successes of individuals from a common lineage, very much in the same way that thermodynamical properties such as pressure and temperature, while not discernible (or indeed definable) at the level of single molecules (which only exhibit widely varying kinetic energies), emerge as consistent and important properties when a sufficient number of molecules are considered. More than its actual value, which should be quite low, the *enduring presence* of this correlation should be the real indicator of Darwinian evolution.

3.5 Practical considerations

As usual when calculating statistical correlations, care should be taken in only comparing what is comparable: conflating data from widely different distributions may result in artificial, spurious results. Selective conditions may change widely over the course of an evolutionary run, even with a fixed fitness function. This may wreak havoc on undiscerning evaluations of statistical correlation. For example, if strict ranking is used, surviving and reproducing entails dislodging a previous survivor; but as evolution proceeds towards an optimum, and new champions are increasingly well-adapted, it becomes increasingly difficult (and thus rare) for new individuals to dislodge previous champions. This means that the children's fitness will tend to go down (because more of them disappear without a descent) and the parent's fitness will tend to go up (because they remain in the population longer) over time. This alone would be sufficient to create a negative correlation between the fitnesses of parents and children: earlier parents would have a moderate number of children, each with a good chance to reproduce; while later parents would accumulate enormous numbers of children, which would have very low reproductive success.

To avoid this, we must ensure that we only consider quantities obtained under similar conditions. Collating data from individuals born at widely different times would violate the requirement for a 'fair comparison'. This implies that differential fitness transmission, as described above, must be calculated as a *local* quantity in time. This does not mean that the data itself should be strictly

local. For example, reproductive success for a given individual may be collated over its entire history. However the periods over which reproductive successes are measured should start at the same point in time, so that we can ensure that they are obtained over equivalent conditions.

4 Statistics for the Calculation of Differential Fitness Transmission

This section is concerned with the description of statistical indicators which may be used to detect differential fitness transmission in a genealogical record. They are clearly not exhaustive, and the very many different configurations encountered in artificial evolution may allow for different statistics to be used. However we believe that the statistics presented in this section are sufficiently general to be usable in a vast diversity of cases.

4.1 Parent-Children Fitness Correlation: A statistic for measuring differential fitness transmission

Our first proposed statistic for detecting the presence of differential fitness transmission, is simply the statistical correlation between the number of children generated by a parent and a child (Parent-Child Fitness Correlation, or PCFC). The basic procedure is as follows: using a genealogical record of the (supposedly) evolutionary process at hand, we extract a sample consisting of many parent-children pairs. Taking into account the caveats discussed in the previous section, the sample we extract consists of all parent-children pairs such that the children were born during a certain period of time (for example, at a given generation). For each parent-child pair within this sample, we compute the number of offspring generated by either member of the pair, from the birth of the child member onwards. We then apply a standard statistical formula to this sample, which allows us to estimate the correlation between the reproductive successes of parents and children.

More formally, PCFC over a certain time period $[t1, t2]$ of an evolutionary process can be calculated in the following way:

1. Scanning the genealogic record, find and store every pair of individuals (X, Y) such that Y was born during time period $[t1, t2]$, and X is a parent of Y .
2. For every stored parent-child pair (X, Y) , retrieve their respective number of children C_x, C_y , from the time of birth of Y onwards.
3. Calculate the statistical correlation between the C_x and the C_y variables, using the standard Pearson formula:

$$\text{Corr}(X, Y) = \frac{\sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y})}{(N - 1)\sigma_X\sigma_Y}$$

This measure is a straightforward formalisation of the notion of differential fitness transmission. It simply begs the question: “how much is it true that fitter (resp. less fit) parents tend to have fitter (resp. less fit) offspring?” We believe that this simple intuitive interpretation is a strong point of the statistic.

4.2 A limitation of direct PCFC

The formula which we just described for evaluating differential fitness transmission has a limitation. It attempts at measuring how much fitter (resp. less fit) parents will tend to have fitter (resp. less fit) children, with fitness being equivalent to reproductive success. However, what happens when *no* parent has a higher reproductive success than any other? Clearly in such a situation the question becomes meaningless, and the calculation itself is undefined.

Thus PCFC will not be applicable under a particular condition, namely when no selective gradient exists among the reproductive successes of individuals which do reproduce. This occurs when all these individuals which do reproduce, have either equal or randomly varying reproductive success. Evolution is very much possible under those conditions, if selection bears on the choice of *which* individuals are allowed to reproduce, with all other factors being equal or contingent. In this case the selective gradient does not lie in the particular number of children begotten by an individual, but solely in whether this individual is able to reproduce or not.

A simple example of such a system would be a genetic algorithm with non-overlapping generations, in which at every generation a certain proportion of individuals are selected (according to some evaluation function) as parents for the new individuals of the next generation. New individuals are then generated by randomly (or evenly) sampling parents from this set of selected individuals². Such an algorithm may certainly give rise to adaptive evolution, based on the differential transmission of heritable fitness-impacting traits. Yet this system would not produce any observable correlation between the fitnesses of parents and children.

This is because the selective gradient would only exist between those which reproduce, and those which do not; but if we calculate correlations among parent-children pairs, obviously individuals which do not reproduce cannot be included in the calculation as “parents”! If there is no selective gradient among the parents, then no correlation can be detected. The notion of correlation depends on the presence of a gradient in two quantities, such that higher (lower) values in one quantity are associated with higher (lower) values of the other quantity; absent or random gradient in one of these quantities implies the absence of any correlation.

²If generations are allowed to overlap, then fitter individuals may be expected to live longer, and to have more children; this would suffice to create a gradient among reproducing individuals, and thus make it possible to calculate a direct correlation between the fitnesses of parents and children.

4.3 PCFC with indirect fitness evaluation

Fortunately, even in such a situation, PCFC can still be measured, if we use an indirect measurement of fitness. Even if all individuals which do reproduce have equal or random number of children, the same does not hold for their children, which may or may not have children themselves. If natural selection does operate, there should be differences between the reproductive successes of children generated by different parents. If fitness-affecting, heritable characteristics are being transmitted, then those children which inherit “good traits” should have a higher chance of reproducing than those which inherited mediocre ones. Thus, even though the direct reproductive success of reproducing individuals carries no information about their relative genetic endowments, we can still roughly re-create this information by comparing the proportion of their children which will actually reproduce. By using the proportion of reproducing children as an indirect indicator for fitness, we can apply PCFC even in situations where the direct reproductive success of parents is fixed or random³.

This quantity can be calculated through the following procedure:

1. Scanning the genealogic record, find and store every pair of individuals (X, Y) such that Y was born during time period $[t1, t2]$, and X is a parent of Y .
2. For every stored parent-child pair (X, Y) , assign scores S_x, S_y to each member of the pair in the following way:
 - For children Y , the score is simply the total proportion of their own children which do reproduce. If Y has no children, the score is 0.
 - For parents X , the score is the total proportion of their children which do reproduce - excluding Y itself, of course!⁴ If X has no other child than Y , the score is 0.
3. Calculate the statistical correlation between the S_x and the S_y score values, using the same formula as for simple PCFC.

Notice that scores are not based on the number of children, but on whether or not reproduction occurs; similarly, we use the proportion of reproducing children, rather than their raw number, for parents. This should diminish the possible blurring effects of randomly varying reproductive successes. On the other hand, because this statistic uses an indirect evaluator for genetic endowment, the effects of random fluctuations may still be stronger than in the previous case. The consequence is that we should expect lower overall values for the calculated correlation.

³We are aware that with this method the meaning of the term ‘fitness’ is slightly modified, from a mere shorthand for ‘reproductive success’, to the notion more generally used in artificial evolution (general characteristics affecting reproductive success). Since detecting the transmission of such characteristics is precisely our objective, this does not affect our reasoning.

⁴If Y were not excluded, a spurious correlation would appear.

4.4 Siblings Fitness Correlation

In some circumstances, genealogical data may be only partially available. For example we might encounter a system which, at any time, only provides information about the current generation. In this case, the incomplete genealogical information would prevent the calculation of PCFC.

Fortunately differential fitness transmission can still be detected, in most cases, with such partial information. Differential fitness transmission implies that, if Darwinian evolution is occurring, then individuals sharing a common lineage should exhibit some small correlation in their fitnesses. But individuals sharing a common parent, from this very fact, share a common lineage. Therefore differential transmission also applies to them. As we mentioned in the previous section, if fitness-affecting, heritable characteristics are being transmitted, then children sharing common parents should exhibit slightly similar reproductive successes: those children which inherited “good traits” would have a higher chance of reproducing than those which inherited mediocre ones. These differences in the reproductive successes of children from different parents should be detectable in the genealogical record.

The problem of detecting these differences can be summarised in the following question: “how much is my own reproductive success correlated with the reproductive success of my parent’s other children?” If fitness-affecting characteristics are being transmitted, then we should expect that individuals sharing a common parent would show (on the whole and over a sufficiently large sample) a certain correlation in their fitnesses, stemming from the inheritance of common fitness-affecting characteristics. On the other hand, if no such characteristics are being transmitted and no adaptive evolution occurs, then the reproductive success of an individual should be utterly independent from that of its siblings.

This new statistic effectively measures the correlation between the reproductive successes of siblings, that is, individuals sharing a common parent. We therefore call this quantity Siblings Fitness Correlation (SFC). It can be calculated by applying the following procedure:

1. Scanning the genealogical record, find and store every pair of individuals (X, Y) such that both X and Y were born during time period $[t1, t2]$, and share a common parent.
2. For every stored siblings pair (X, Y) , assign a binary score S_x, S_y to each member of the pair, which is 1 if the individual does have children over its lifetime, and 0 if it doesn’t.
3. Calculate the statistical correlation between the S_x and the S_y score values, using the same Pearson formula as for PCFC.

This statistic has clear limitations: because it only includes data from one single generation, it is not able to confirm consistent transmission of fitness from one generation to the next. For example, let us imagine a rather contrarian system in which the reproductive success of individual would somehow be *inversely*

proportional to that of their parents: successful parents would always generate poor children, and vice versa. Such a system would hardly be expected to arise ‘by chance’, but it would certainly be easy to create on purpose; at any rate it could hardly be called ‘evolving’ in a Darwinian sense. Nevertheless, it would still produce a high SFC value, because children from a common parent would have similar reproductive successes. Thus we cannot say that positive SFC is a definite indicator for the presence Darwinian evolution: it does indicate that something potentially worth investigating is going on, but in some situations this ‘something’ may not be Darwinian evolution.

However, the opposite assertion (that a zero SFC implies the absence of Darwinian evolution) seems very solid. If SFC is zero, then there is no correlation between the fitness of an individual and that of its siblings: in other words, having a common parent does not provide any information regarding reproductive success. This seems to contradict the idea of fitness-affecting characteristics being transmitted from generation to generation. Therefore, if SFC can indeed be calculated and can be conclusively shown to be zero, then we may be confident that such fitness-affecting characteristics are not being differentially transmitted.

5 Experiments

5.1 Experimental settings

Our purpose in this section is to set up a couple of experiments in order to determine whether fitness transmission is indeed a reliable indicator of Darwinian evolution. To do this, we will use simple evolutionary systems with predictable dynamics, in which the presence or absence of evolution can be easily controlled. We will apply our calculation method to these systems and determine whether the presence or absence of Darwinian evolution was successfully detected.

To perform our experiments, we used genetic algorithms involving a population of 1000 individuals, over 100 generations. We considered two optimisation problems: the Rosenbrock function $100(x^2 - y^2)^2 + (1 - x)^2$ (using genomes of 2×12 bits) and a very simple OneMax problem over 20 bits. The Rosenbrock function is a commonly used test function in the field of optimisation. The purpose of the simple OneMax problem is to examine the behaviour of different algorithms on very easy problems, when the the global optimum is discovered quickly.

New individuals were created either by applying bitwise mutation to a parent, or (with 60 % probability) by applying one-point crossover between two suitably chosen parents, and then applying bitwise mutation to the resulting offspring. The probability of mutating (flipping) each bit is the inverse of the total number of bits in the genome, rounded to the closest higher percent; thus, on average, each genome should undergo about one mutation. As explained below, we tested different methods of selection and replacement.

In a normal genetic algorithm, Darwinian evolution can be expected to oc-

cur. As a point of comparison, we need to compare a “neutral” version of the genetic algorithm, which preserves as many features of the algorithm as possible, while effectively removing Darwinian evolution. We could simply use a normal genetic algorithm, with random selection, but in this case the absence of fitness transmission is rather obvious. Can we find a more challenging test case?

We chose to use genetic algorithms with uncorrelated landscapes, in which every new individual was attributed a random genotype (and therefore a random phenotype) at birth, regardless of the genetic make-up of its parents. This is different from purely random selection in that selection still occurs, and is still based on fitness; however the randomness of the reproductive process turns the problem into a random search on a random landscape, thus preventing any meaningful evolution: no innovation can propagate throughout the population. Fitness-affecting traits are still present, but not heritable. A satisfactory measure of evolutionary activity should be able to detect the absence of real evolution and return a zero value for this situation.

5.2 A simple genetic algorithm

We first describe the calculation of fitness transmission in a standard genetic algorithm, quite similar to the original model proposed by Holland [7], except that we use tournament selection rather than roulette wheel selection. In this new algorithm, an entire new population is created at every generation. One or two parents for each new individual were selected from the previous population, using tournament selection. A individual was then generated as previously described. The process was iterated until the new population was filled. In this system, as in most genetic algorithms with non-overlapping generations, all individuals have the same lifetime of exactly one generation.

For every generation G , fitness transmission is calculated as follows:

1. Find and store every pair of individuals (X, Y) such that Y was born at generation G , and X is a parent of Y .
2. For every stored parent-child pair (X, Y) , retrieve their respective number of children over the history of the run, from generation G onwards.
3. Calculate the statistical correlation between the C_x and the C_y variables.

By only considering individuals born at the same generation, and comparing their reproductive success with that of their parents over the same period, we avoid the potential errors and biases mentioned earlier: because the values were obtained over the same time period, they can be fairly compared and correlations can be safely calculated.

Figure 1 shows the results of these calculations, applied to the “fossil record” generated by our simple genetic algorithm. This figures shows the results for the Rosenbrock function optimisation problem with 20 bits, both with normal reproduction and with reproduction based on random phenotypes. The top graph shows the results of 5 different run for each of these reproduction methods,

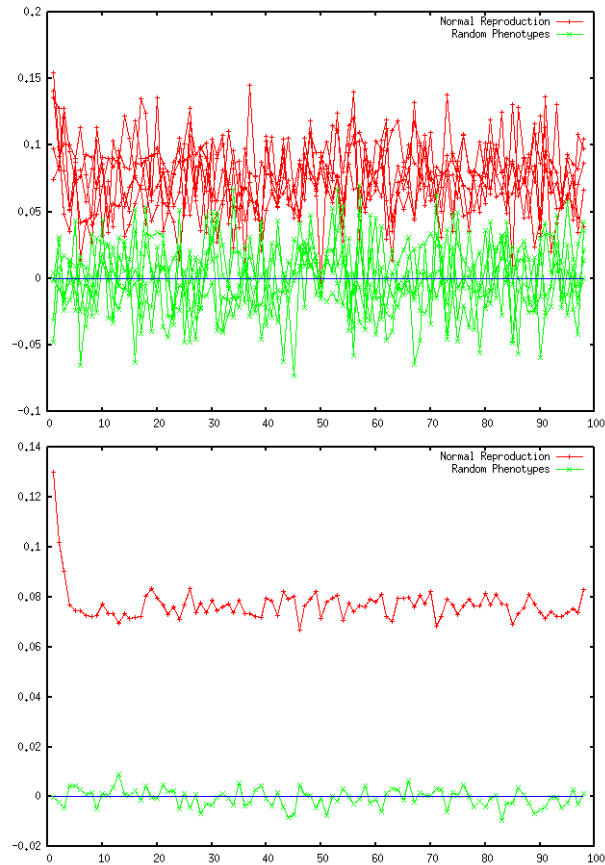


Figure 1: PCFC: Rosenbrock function, non-overlapping generations, 5 different runs (top) and average of 50 different runs (bottom).

while the bottom graph shows average curves over 50 runs. Figure 2 shows the same data for the OneMax problem. In the normal selection case, the correlation between the number of children of parents and children is distinctly positive (especially at the very beginning at the run) and stabilises to a positive plateau. The enduring positive value indicates that the population is constantly, actively maintained in the vicinity of the global optimum (which is reached quite early in the OneMax problem) through active evolutionary forces. Even though the optimum has been reached, mutation constantly disperses the population, and Darwinian evolution constantly drives it back. Stabilising selection results in a positive value for differential fitness transmission.

In the case of random genotypes, as expected, no meaningful fitness transmission occurs.

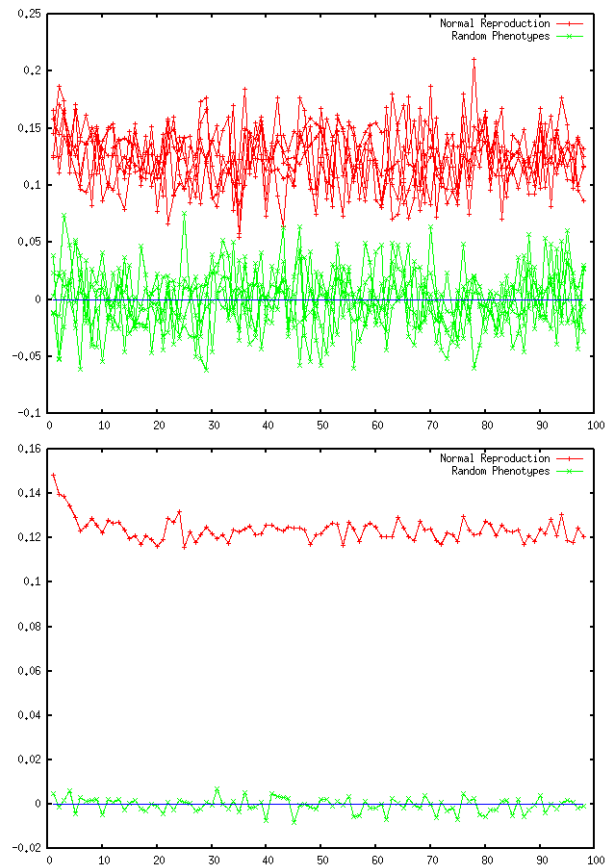


Figure 2: PCFC: OneMax function, non-overlapping generations, 5 runs (top) and average of 50 runs (bottom).

5.3 Removing mutation

That the enduring presence of fitness transmission in this case is caused by mutation can be seen quite readily. If we set the mutation rate to zero, then the population converges totally: all individuals end up sharing the exact same genome, and diversity disappears. From this point on, all individuals having exactly the same genotype, reproduction patterns become effectively random. The result is that evolutionary activity, as indicated by fitness transmission, quickly goes to zero (with noise oscillations) after an initial phase of high activity: Natural selection exerts no active force upon the system, because the population has basically frozen and no genetic variation exists to drive evolution. This illustrates the capacity of differential fitness transmission to distinguish between active stabilising selection on the one hand, and passive stillness caused by ab-

sence of genetic variation on the other.

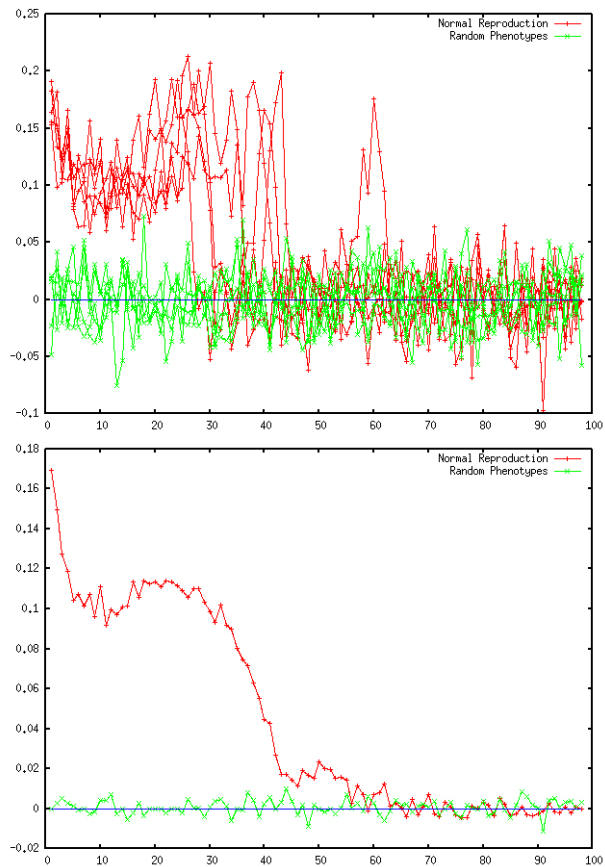


Figure 3: PCFC: Rosenbrock function, non-overlapping generations, without mutation, 5 runs (top) and averages of 50 runs (bottom). The absence of mutation leads the population to ‘hard’ convergence: all individuals share the same genome and no variation occurs. From then on, no meaningful evolution occurs.

Unsurprisingly, this process occurs much faster with the OneMax problem, since the global optimum is reach very quickly.

5.4 Removing selective gradient among parents

What happens when no selective gradient exists among parents? To investigate this question, we modified our algorithm as follow: at every generation, a set of survivors is selected from the population through strict ranking selection, and the parents for the next generation are then *randomly* selected from among this set of survivors. Thus, while there is some selection and potentially some

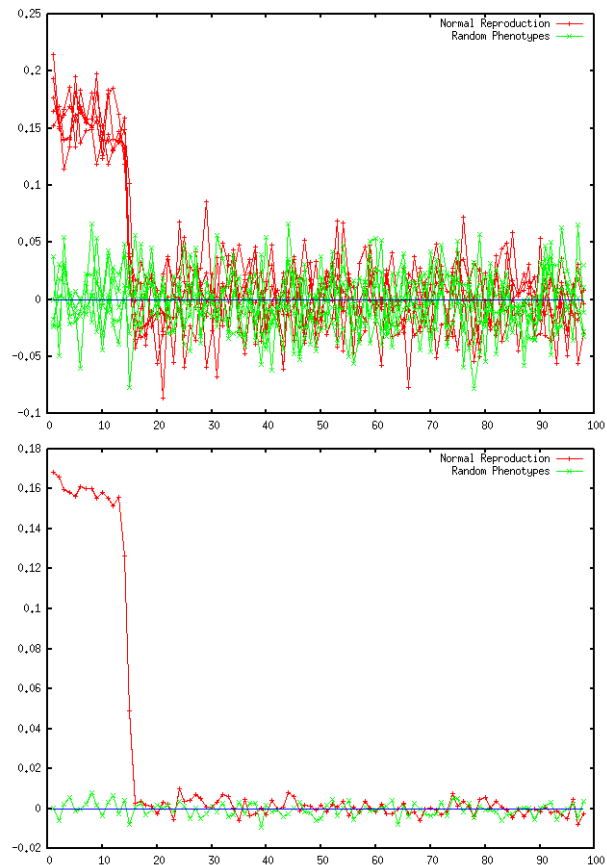


Figure 4: PCFC: OneMax function, non-overlapping generations, without mutation, 5 runs (top) and averages of 50 runs (bottom).

perfectly valid Darwinian evolution, no selective gradient exists between parents, because the selective gradient lies solely between those which become parents and those which do not.

As expected, when such an algorithm is used, calculating direct PCFC to the genealogical record returns a null signal. This can be seen in Figure 5 which show the result of this calculation for the Rosenbrock function. Similar results are obtained with the OneMax function.

However, when using PCFC with indirect fitness evaluation, the situation is different: a positive signal does emerge. Indirect fitness evaluation allows us to retrieve the fitness gradient among parents which cannot be obtained by assessing direct individual selection. This signal is actually quite strong in the case of the OneMax function, as can be seen in Figure 6. However, the signal is much weaker in the case of the Rosenbrock function, except at the beginning of

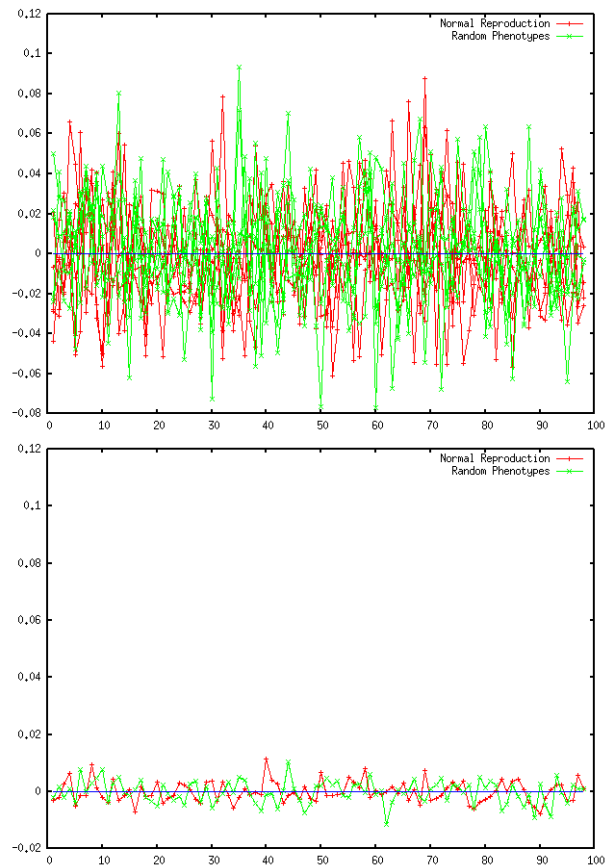


Figure 5: PCFC: Rosenbrock function, non-overlapping generations with ranking-based survival and random parent selection, 5 runs (top) and averages over 50 runs (bottom). The elimination of a selective gradient among parents removes any correlation.

the run (Figures 7). In this case, an initially strong signal quickly decreases to a very low value, which is difficult to notice in individual runs (although a weak, but distinctly non-zero signal can be observed when averaging over many runs).

6 Limitations of Genealogic Analysis

We have described how genealogical records could be used to detect differential fitness transmission, and Darwinian evolution, through statistical methods. These methods are applicable in a wide range of circumstances. However there is one extreme situation in which differential fitness transmission can apparently

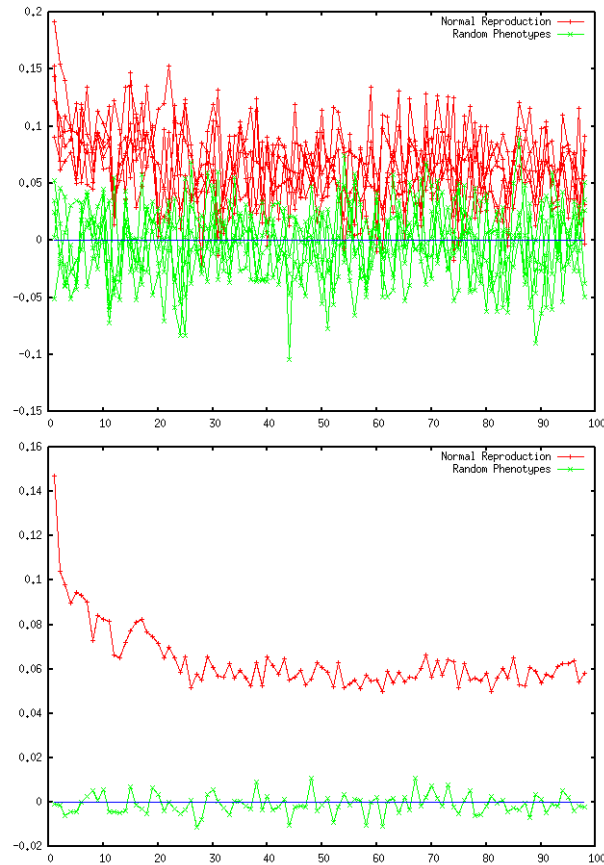


Figure 6: PCFC with indirect fitness estimation: OneMax function, non-overlapping generations with ranking-based survival and random parent selection, 5 runs (top) and averages over 50 runs (bottom). The signal is recovered, and Darwinian evolution can be demonstrated.

not be detected by genealogic analysis. This situation occurs when all individuals present at any given time share the exact same genealogic tree - in other words, when there is never more than one lineage in the population.

Genealogic analysis basically consists in comparing the relative growth of various lineages. This implies that there should actually be several lineages within the population, otherwise no comparison is possible. However, this is not the case when all individuals present at any time within the population share the same entire genealogic tree. In this case, there is only ever one single lineage present at any time within the population. This can occur in a non-overlapping generational system, such that all individuals from every new generations have exactly the same parents from the previous generation: at

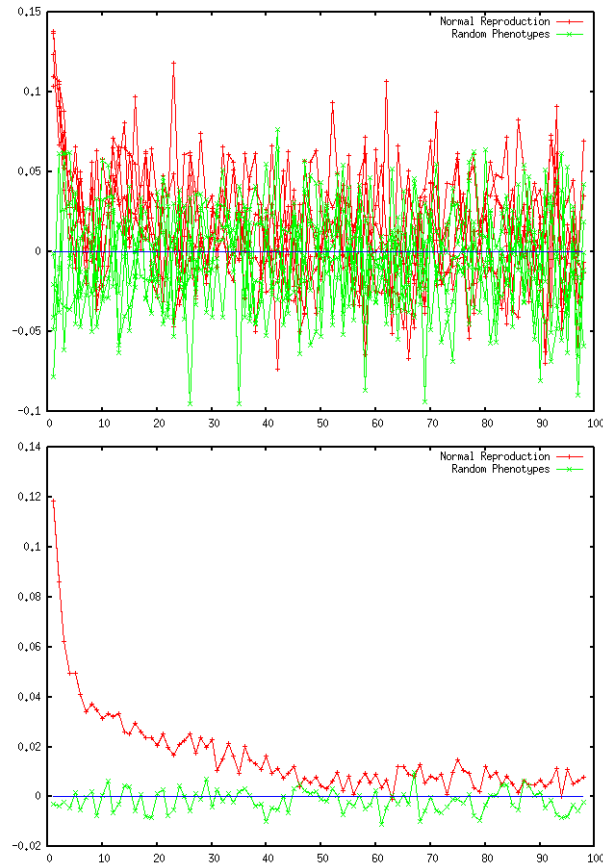


Figure 7: PCFC with indirect fitness estimation: Rosenbrock function, non-overlapping generations with ranking-based survival and random parent selection, 5 runs (top) and averages over 50 runs (bottom). The initially high signal goes to a very low, almost (but not quite) nil value.

every generation, some individuals are chosen to serve as parents for the next generation, and *all* the individuals from the new generation inherit material from *all* of those selected parents. In this case, Darwinian evolution may occur (heritable characteristics are certainly being differentially transmitted), but genealogic analysis will not be able to detect whether the selection of parents is adaptive or random.

We believe that this situation is sufficiently exotic to preserve the usefulness of genealogical analysis. At any rate such situations can be easily detected in any system for which a genealogical record exists.

7 What Differential Fitness Transmission Measures, And What It Doesn't

Having proposed differential fitness transmission as a signature for Darwinian evolution, it is important to make it clear what exactly this signatures reveal and what it doesn't, in order to avoid confusion.

Differential fitness transmission is not a measure of performance or optimisation. This is readily seen from the fact that even random search can produce some amount of optimisation, even though no evolutionary activity or fitness transmission is present. Similarly, we saw that for some algorithms, fitness transmission can remain positive even after the global optimum has been reached, as long as Darwinian forces keep acting on the population.

Differential fitness transmission is not a measure of inventiveness, creativity or novelty. Whether or not the newly created individuals are novel or not is not considered: only their fitness (and the correlation between their fitness and their parents' and children's fitness) is taken into account. Again, this is seen from the fact that positive fitness transmission is compatible with the fact that the population is being maintained around a single optimum, where no adaptively significant innovation is possible.

Differential fitness transmission is simply the answer to this question: "all other things being equal, is it true that individuals from a common lineage will tend (ever so slightly) to have similar fitness?" Since nothing is assumed about the nature of fitness, except for its basic observational property (reproductive success), we can attempt the following, general statement: fitness transmission simply measures *the presence or absence of some force which consistently and actively favours some lineages at the expense of others* (over at least two generations). We believe that in general, this constitutes a good signature of Darwinian evolution by the operation of natural selection on heritable characteristics.

8 Possible Continuations: A Direct Measure of Evolution

Differential fitness transmission is not a direct equivalent of Darwinian evolution; rather, it is a convenient signature which is expected to be present when actual evolution is. This signature allows us to detect evolution with genealogical records. However, if we have access to a complete *genetic* record, keeping track of the genetic material present in each individual, then we can attempt a direct measurement of evolution itself, rather than a particular signature of it.

Let us go back to our original argument. We have stated that adaptive evolution was the result of the differential transmission of heritable, fitness-affecting characteristics. Genes are defined as the support of heredity: they are the informational message which is transmitted from one individual to another and allows the latter to exhibit some characteristics of the former, these characteristics being encoded by the transmitted genes. Let us assume that we have a

complete record of the genetic material of all individuals over the history of a population⁵.

Genetic components are, by definition, heritable. In fact, under the (strong) assumption that our genetic record is indeed complete, they cover the total field of heredity within the system. The question which must be answered, therefore, is whether or not some of the genetic components present in a population do indeed affect fitness: if possessing certain genes reliably entails higher (or lower) reproductive success, then Darwinian evolution is occurring - otherwise, it is not. The questions of whether Darwinian evolution occurs in a system, and whether some genetic components causally affect fitness, can be regarded as basically equivalent. Detecting evolution with genetic records can thus be reduced to the following question: "Are there some genes which are significantly and reliably associated with above- or under-average reproductive success?"

This is not an easy question to answer. One reason is that in this question the term 'gene' can really cover any substring of the genetic message, including disjoint ones: a gene here is really a *schema*, in the sense of Holland [7]. Because any novel schema can be the source of some evolutionary adaptation, we should keep track of all possible schemas in the system. But a string of N elements contains 2^N different schemas, so the potential number of schemas can grow very large as soon as genomes and populations reach a significant size. Keeping track of all schemas present at any time in the population is clearly not a trivial task. It becomes extremely difficult when the genomes vary in length, requiring some sort of mechanism to identify homologous genes between different genomes.

This problem is greatly reduced when the genomes consists of discrete, complex elements (such as rules within a rule-set), for in this case we may track these elements directly (although combination information, which may be crucial, will not be available): see [5] for an example, in the context of Bedau-Packard statistics. In the general case of bitstring genomes, however, some kind of finesse would be needed to keep track of all possible fitness-affecting schemas.

Another problem is to determine when exactly a suspected association between the presence of a gene and a fitness impact becomes 'significant'. What test should we use? How many instances do we need? What margin of error should we allow for?

On the other hand, this measure would have several advantages. First, it has the conceptual benefit of actually tracking the process of evolution itself, rather than a signature or a proxy for it. Accordingly, it would work in a wide range of situations, including those special cases where genealogic methods fail (as described in section 6). An additional advantage is that this measure would make it easy to track truly evolutionary innovation: an innovation is simply a new fitness-affecting gene that was not encountered previously. Thus this measure would identify both 'evolution-the-change' and 'evolution-the-mechanism' in the same process.

⁵This implies that there is heredity within the system, that this heredity can be reduced to an informational support, and that we have complete access to this support.

9 Conclusion

We have shown that differential fitness transmission is a useful signature of Darwinian evolution, which can be detected in genealogical record by using simple statistics. We believe that this signature may be more suitable for this purpose than previously suggested methods for detecting evolution. We have introduced two such statistics, Parent-Children Fitness Correlation and Siblings Fitness Correlation. We have applied these statistics to the genealogical records generated by real evolutionary algorithms, demonstrating their capacity to detect the presence or absence of adaptive evolution. We hope that these methods will prove useful to experimenters and researchers in the fields of artificial evolution and artificial life.

10 Acknowledgements

This research is being conducted at (and funded by) the School of Computer Science at the University of Birmingham, under the supportive supervision of Dr Alastair Channon. Part of this research was originally conducted at (and funded by) the University of Portsmouth, in Dr David Brown's Intelligent Systems Research Group. This research benefited greatly from challenging discussions with Dr John Bullinaria and Dr Richard Dearden.

References

- [1] Bedau, M. A., & Packard, N. H. (1992). Measurement of evolutionary activity, teleology, and life. *Artificial Life II: Proceedings of the Workshop on Artificial Life* (pp. 431–461). Redwood City, CA: Addison-Wesley.
- [2] Bedau, M. A., Snyder, E., & Packard, N. H. (1998). A classification of long-term evolutionary dynamics. *Artificial Life VI: Proceedings of the Sixth International Workshop on the Synthesis and Simulation of Living Systems* (pp. 228–237). Cambridge, MA: MIT Press.
- [3] Bullock, S., & Bedau, M. (2006). Exploring the dynamics of adaptation with evolutionary activity plots. *Artificial Life*, 12, 1–5.
- [4] Burtsev, M. (2003). Measuring the dynamics of artificial evolution. *Proceedings of the Seventh European Conference on Advances in Artificial Life (ECAL 2003)* (pp. 580–587). Berlin: Springer.
- [5] Channon, A. D. (2006). Unbounded evolutionary dynamics in a system of agents that actively process and transform their environment. *Genetic Programming and Evolvable Machines*, 7(3), 253–281.
- [6] Channon, A. D., & Damper, R. I. (1998). Perpetuating evolutionary emergence. *From Animals To Animats 5: Procs 5th Intl Conf on the Simula-*

tion of Adaptive Behavior (SAB '98) (pp. 534–539). Cambridge, MA: MIT Press.

- [7] Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press.
- [8] Hrabar, P. T., Jones, T., & Forrest, S. (1997). The ecology of Echo. *Artificial Life*, 3(3), 165–190.
- [9] Ridley, M. (1993). *Evolution*. Blackwell Scientific.
- [10] Standish, R. K. (2002). Diversity evolution. *Artificial Life VIII: Proceedings of the Eighth International Workshop on the Synthesis and Simulation of Living Systems* (pp. 131–137). Cambridge, MA: MIT Press.