The Puzzle of the Persistent Question Marks: A Case Study of Genetic Drift

Inman Harvey

CSRP 278, April 1993

Cognitive Science Research Paper

Serial No. CSRP 278

The University of Sussex
School of Cognitive and Computing Sciences
Falmer
BRIGHTON BN1 9QH
England, U.K.

This paper will appear in:

S. Forrest, (ed.),
Genetic Algorithms: Proceedings of the Fifth International Conference (GA93),
Morgan Kaufmann, San Mateo, CA.

The Puzzle of the Persistent Question Marks: A Case Study of Genetic Drift

Inman Harvey*
School of Cognitive and Computing Sciences
University of Sussex
Brighton BN1 9QH, England

Abstract

In an important paper, Hinton and Nowlan (1987) demonstrate the Baldwin effect in a simple Genetic Algorithm. The ability of the phenotype to adapt, coupled with the evolutionary process, allows behavioural goals to become over time genetically specified; this seems Lamarckian but is not. In that paper, as a subsidiary point, the slowness of fixation of the last few goals is commented on, and a later paper by Belew (1989) attempts an analysis. In this paper I show that genetic drift is the explanation for this slowness phenomenon. Using a diffusion equation approach, I give an analysis of genetic drift for genetic algorithms, where it is too often ignored. Critical relationships between mutation rate, population size, and forces of selection are given which decide whether genetic drift will be of significance or not.

1 Introduction

In an important and elegant paper, Hinton and Nowlan (1987) demonstrate with a deliberately simple example the Baldwin effect, wherein the ability of a phenotype to adapt in its lifetime (ability to 'learn') alters the fitness landscape of the corresponding genotype. This has the consequence that selection within a population moves the genotypes towards the region where the adaptations, that were originally made in the lifetime of the phenotypes, are genetically fixed. This has the appearance of Lamarckism, but is not so, as there has been no direct flow of information from the adapted phenotype to the genotype.

The model chosen as an example uses genotypes with a number of genes that can be specified as incorrect, correct, or open to adaptation during the lifetime of the phenotype. The evaluation function only favours those phenotypes that, within a finite lifetime, find a perfect solution through a combination of 'correct' genes, and 'adaptive' genes which successfully adapt. It is demonstrated that with the application of a standard genetic algorithm (GA) to the population as specified, the number of incorrect alleles on the genotype rapidly decreases to zero; the number of correct alleles increases at first rapidly and then slows down; the number of undecided (adaptive) alleles decreases slowly. If the same experiment is tried out only with correct and incorrect genes, and no adaptive ones, then the 'needle in a haystack' nature of the single perfect solution means that only random search works, and takes an unreasonably long time.

The main thrust of Hinton and Nowlan's paper is endorsed here, but a subsidiary matter that is mentioned as an aside there is taken up as the main point for investigation here in this paper:

> One interesting feature of [the figure] is that there is very little selective pressure in favor of genetically specifying the last few potential connections, because a few learning trials is almost always sufficient to learn the correct settings of just a few switches.

The figure in question indicates that there could be an asymptote at a relative frequency of about 0.45 below which the number of undecided alleles will not fall.

My own re-implementation of the model usually shows an asymptote at between 0.05 and 0.2. A typical run is shown in figure 1, showing the dramatic changes in the first 50 generations, and the longer term behaviour over 500 generations. The variations between runs is indicated in table 1, showing the values at the end of 20 runs of 500 generations each. The re-implementation by Belew (1989) shows 'an almost steady-state' at about 0.3. He asserts that the curve is 'in fact asymptotically approaching ... 0.0'. This I will demonstrate to be false, in the general case; the analysis of what is really happening shows that the combination of genetic drift and the hitch-hiking effect so completely swamps the selective pressures that some of the genes are com-

*Email: inmanh@cogs.susx.ac.uk

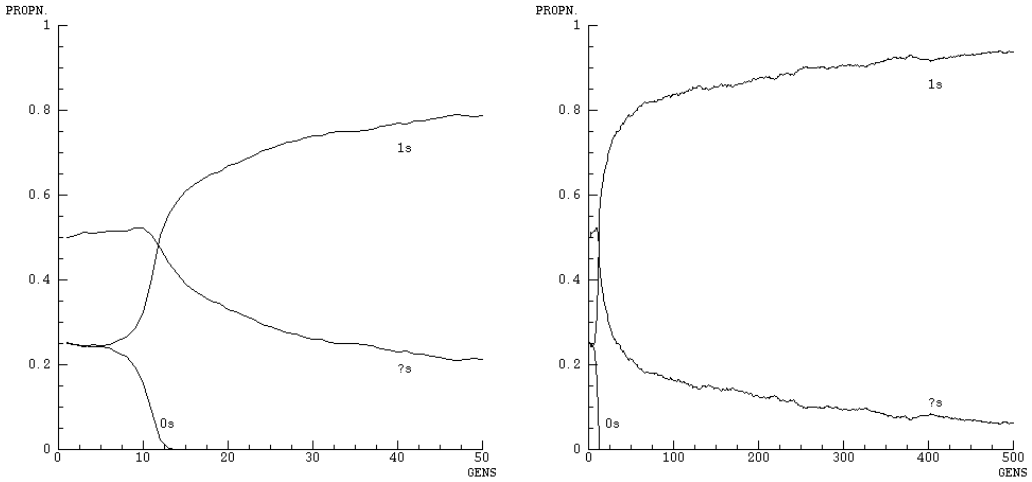

Figure 1: The proportions of incorrect, correct, and undecided (adaptive) alleles (**0**s, **1**s, **?**s) in the whole population, against generations. On the left, the first 50 generations of a run, and on the right the same continued for 500 generations.

pletely converged to the undecided value, rather than the 'correct' one.

Strictly speaking, if there is even the smallest amount of mutation in the system, applied independently at each locus on each genotype, then if you are willing to wait long enough you will see *any* population state; even a population entirely composed of incorrect alleles, or entirely composed of undecideds. This would be a transitory phenomenon, and the necessary timespans are way beyond those being considered here. The diffusion equation analysis of genetic drift given below gives a picture of those circumstances under which genetic drift can be expected to be a significant force.

Genetic drift — the consequences of random fluctuations in relatively small populations — is a matter of fundamental concern to population geneticists, but appears to be almost completely ignored by people using GAs. Since 1000 is often 'relatively small' in this context, and GAs frequently use population sizes of 100 or less, to ignore genetic drift is commonly, as in this example, to ignore one of the fundamental processes underlying the phenomena.

I include a number of graphs (figures 3 and 4) indicating the relative influences of selection, mutation and genetic drift for different parameter values. I also give a reworking of standard genetic drift analysis taken from population genetics theory, and adapted to the haploid models common in GAs.

2 The model

For fuller details of the model used in the demonstration I refer you to the original paper, and to a subsequent analysis by Belew (1989). As a brief summary, the model has a population of 1000, each with genotypes with 20 genes having possible values **0**, **1** and **?**. In the initial population these are randomly selected with probabilities 0.25, 0.25 and 0.5. The derived phenotypes are taken to be a set of 20 switches, which undergo a series of up to 1000 trials each. The allele **0** at a particular gene specifies that the corresponding switch is set incorrectly, **1** specifies that the switch is set correctly, and a **?** indicates that the corresponding switch is flipped randomly at each trial. The series of trials on a phenotype is stopped when all switches happen to be set correctly, on trial number i, or alternatively at $i = 1000$, the final trial, if there is no success. Of course, if any of the alleles in the genotype are **0**, i.e. some switch is genetically fixed at the incorrect position, inevitably the trials will run the full course until $i = 1000$.

The fitness F is then calculated from i by the formula $F = 1 + 19(1000 - i)/1000$. This gives an all-perfect phenotype, which needs no trials to reach success, a fitness of 20; while one which never succeeds ($i = 1000$) either through being born without a chance (one or more alleles of **0**) or through failing despite having a chance, has a fitness of 1. The necessary equations to calculate the expected fitness are given in the appendix. If q is the number of undecided alleles in a genotype which otherwise is correct, then for $q > 14$ the expected fitness is near to 1; for $q < 5$ the expected fitness is near to 20. The sharp transition is shown in figure 2.

2.1 Early stages ...

At each generation the relative attained fitnesses of each member of the population determine the proba-

Table 1: The final proportions of undecided alleles after 20 runs each of 500 generations, with no. of loci converging or converged on **?**. 4 runs† have in fact completely converged at all 20 loci, only one run‡ does not yet have a locus with **?** fixed.

Propn of ?s at 500 gens.	Loci having >50% ?s	Loci having 100% ?s	Propn of ?s at 500 gens.	Loci having >50% ?s	Loci having 100% ?s
0.063	1	1	0.108	2	2
0.109	2	2	0.093	2	1
0.082	1	1	0.150	3	3†
0.123	2	2	0.150	3	3†
0.118	2	2	0.112	2	2
0.074	1	1	0.100	2	2†
0.107	2	2	0.093	2	1
0.200	4	4†	0.121	2	1
0.134	3	2	0.115	2	2
0.092	2	0‡	0.115	1	1

bility of that member contributing to the reproductive pool for the next generation. In the early stages, virtually all the members will have the same minimum fitness. Something similar will happen also at the later stages, after the incorrect (**0**) alleles have been eliminated; virtually all members will have small q-values, and hence, because of the flatness of the curve for $F(q)$ at small q, nearly identical fitnesses. At both these stages there is very little selective pressure.

However, as Figure 1 indicates, typically around generations 5 to 15 successful members emerge with a fitness nearly 20 times as great as that of the original random members. This enormous selective differential operates near-exponentially for a few generations, giving the sharp swings indicated in the figure. If the fitness function is adjusted to give a spread of fitnesses from 1 to 2, rather than 1 to 20, this transition is typically delayed until perhaps generation 50, but due to its fundamentally exponential nature it is then a similarly sharp transition.

2.2 ... Hitch-hiking ...

During this transition the genetic material of the first high-scorers dominates the reproductive pool. By marking the genetic material of the first 'winner', and then tracing the marked genes in later generations as they are selected and recombined with others, it can be seen that typically within 10 generations of appearance 50% or more of the whole genetic pool is derived from that first winner. Hence the accidental pattern of **1**s and **?**s in that first winner has a strong chance of dominating future generations after selection has ceased to be a major force — the 'hitch-hiking' effect.

2.3 ... then Genetic Drift

Once that has happened, genetic drift will allow the proportion of **?**s at any one locus vary until it has reached either 0% or 100%, when in the absence of mutation change will cease; and even in the presence of low mutation a stable state can be expected. In the complete absence of selection, then since expected changes from generation to generation do not alter the expected mean, from an initial position of x% **?** alleles one can expect x% of the time convergence to all **?**s, and $(100-x)$% of the time convergence to all **1**s.

The run shown in figure 1, which is also the first example in table 1, has at 500 generations one locus 100% converged to **?**s. Hence the appearance in the graph of a long-term trend towards no **?**s is deceptive, as the asymptote will be at 5%. Table 1 gives an idea of the variations in these figures over 20 separate runs.

In (Hinton and Nowlan 1987), no mention is made of any mutation, and I am unable to account for the much higher asymptote indicated at 45%. It may be an artefact of some idiosyncracies in the programming of the algorithms, or the fact that the graph is hand-drawn may show that it is meant to be loosely indicative only.

In Belew's paper (Belew 1989) the asymptote appears to be at about 30%. This is significantly outside the range covered in my simulations. Mutation is mentioned, without specifying a rate. In the two-bit coding for each locus there described, what is in my terminology **0**, **1**, and **?** translate into respectively **10**, **11**, and either **00** or **01**. The early selection to eliminate **0**s (in my terminology) would eliminate **1**s from the left-hand bit of each pair; the occasional mutation in these left-hand bits will be swiftly eliminated by strong selection. The right-hand bits would then distinguish between (in my terminology) **1**s and **?**s. No mutation rate here can explain the high asymptote shown.

Belew gives three versions of an explanation for this asymptote, which can be summed up as suggesting that selective pressures are so low that 'the probability of producing more than an average number of offspring is infinitesimal'. For a selection as implemented

q	F(q)	q	F(q)
0	20.000		
1	19.962	11	4.965
2	19.924	12	3.140
3	19.848	13	2.113
4	19.696	14	1.568
5	19.392	15	1.287
6	18.784	16	1.144
7	17.569	17	1.072
8	15.233	18	1.036
9	11.649	19	1.018
10	7.868	20	1.009

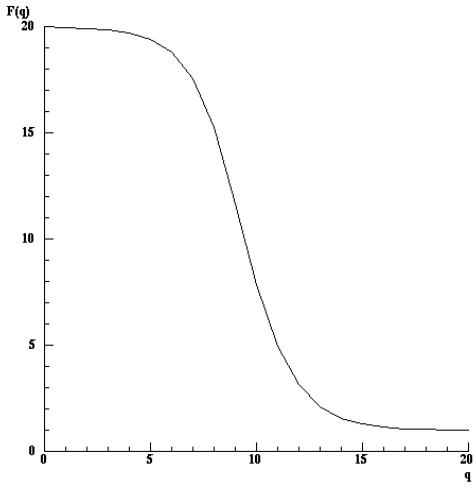

Figure 2: The expected fitness $F(q)$ of a gene with q undecided alleles and $(20-q)$ correct ones. For $q=0$ it has been assumed that success was on the 'zero-th' trial, to give a fitness of 20.

in a standard GA, this would not only be the wrong answer, but it would be almost diametrically opposite to the truth. At a single locus which has exactly 300 out of the 1000 population set to the one value, the probability of the next generation having exactly 300 set again, in the absence of selection, is

$$\binom{1000}{300}(0.3)^{700}(0.7)^{300} = 0.0421$$

I.e. only some 4% of the time will exactly 300 of the next generation have this same allele at this locus, and of the balance some 48% of the time more than 300 will (and some 48% of the time less than 300 will). It is this variation, the *improbability* of remaining at the same proportion from generation to generation, which constitutes genetic drift.

An additional factor to increase genetic drift in this particular model is the fact that the fitness evaluation is non-deterministic, and has a very large variance. Where the fitness is deterministic this would obviously have no variance. In most non-deterministic evaluations, by taking a reasonably large sample size the variance is normally reduced to insignificance, as the variance of a sample of size N is $1/N$ times the variance of the population it is sampled from.

However, Belew mentions that he uses the GENESIS (Grefenstette 1983) GA simulation facility. Into this (or at least the later versions) is built the ingenious selection algorithm, due to Baker (1987), which "guarantees that the number of offspring of any structure is bounded by the floor and the ceiling of the (real-valued) expected number of offspring". In other words, although the expected number of offspring is maintained at the correct value, the variation about this value — and it is this variation which is associated with genetic drift — is reduced to a minimum. Nevertheless, this variation is still significant, particularly at the massive early transition stage.

3 Genetic Drift

If a coin is tossed 100 times, then on average it will be heads 50 times, but it is unlikely to be exactly 50. The same holds if 100 random selections are made with replacement from 50 heads and 50 tails, without turning any over. In the selection case, repetition of the process will on each occasion on average give you the same result as on the previous occasion, but the variance allows significant change in this average over time. If at any stage the selection resulted in all heads (or tails), then future change would be impossible.

The consequence is that, in the absence of mutation, and even without any selection, a population will eventually converge to all one value or the all the other. This also holds true for low values of mutation and/or selective bias, and the critical values which permit genetic drift to be significant can be calculated. The theory of genetic drift is analysed in the field of population genetics, but usually for a diploid population with fitnesses affected by dominant and recessive alleles. In general it seems to be ignored in GAs.

4 A Diffusion equation approach to Genetic Drift

Goldberg and Segrest (1987) give a finite Markov chain analysis of genetic drift. The alternative approach is using a diffusion approximation. The following starts from Roughgarden (1979), where an analysis is given for diploid systems, and makes the necessary alterations for a standard haploid genetic algorithm.

The underlying basis for this approach is that of considering one experiment with a particular set of parameters such as population size, selective bias, and mutation rate; considering one particular locus, and

Table 2: The selection s for a gene with q undecided alleles and $(20-q)$ correct ones is calculated from $s(q) = (F(q) - F(q-1))/F(q)$. The population size N is 1000.

q	F(q)	s(q)	2sN
0	20.000		
1	19.962	0.00190	3.80
2	19.924	0.00191	3.82
3	19.848	0.00383	7.66
4	19.696	0.00772	15.44
5	19.392	0.01568	31.36
6	18.784	0.03237	64.74
7	17.569	0.06916	138.31
8	15.233	0.15335	306.70

taking a census across the population to find the distribution of the possible alleles at this locus, *after* sufficient generations have passed for any initial transitional phenomena to have died away.

For instance, in a particular experiment the proportion of **0**s at this locus will be x in the range $[0.0, 1.0]$. But because of the stochasticity, a series of experiments will give different values of x. Hence a whole ensemble of such experiments are considered — using the same parameter settings for the whole ensemble. The census results on different members of the ensemble will vary, but the probabilities of different census results can be calculated analytically.

The size of the population N is assumed large enough for it to be valid to make a continuous approximation to the discrete steps actually taken — the proportions of any allele can in fact only change in steps of size $1/N$. It is assumed that the proportions in a population of each allele at one locus can be analysed independently of what is happening at other loci, which are taken to be either fixed or with no interdependence on this locus. An ensemble of populations is considered, all acting under the same forces of selection, mutation and drift. It is assumed that from any starting position, this ensemble will spread out under these forces until some equilibrium is reached. This equilibrium will be shown to have strikingly different features depending on the values of the parameters. The figures later on characterise the features of such an ensemble, and hence give a perspective on what might plausibly happen in any one individual population.

The diffusion can be analysed with the same equations as are used for physical processes. A 'diffusion equation' is introduced which approximates the Markov chain. In the analysis of a physical system of diffusion we let $\rho(x,t)$ denote the density of particles at location x at time t. (The translation to our ensemble of populations is: let $\rho(x,t)$ denote the proportion of populations in the ensemble that at time t give a census return of x for the proportion of 1s at the relevant locus.) The flow across a surface at x is $J(x,t)$. The change in density at a location is equal to the spatial derivative of the flow.

$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{\partial}{\partial x}J(x,t) \tag{1}$$

In the current context the expression for $J(x,t)$ contains a term for external forces — mutation and selection — and another term for diffusion.

$$J(x,t) = M(x)\rho(x,t) - \frac{1}{2}\frac{\partial}{\partial x}V(x)\rho(x,t) \tag{2}$$

In a short time interval Δt, $M(x)\Delta t$ is the average distance travelled from a point x under force of mutation and selection. $V(x)\Delta t$ is the variance of the distances travelled.

We are interested in the equilibrium distribution $\hat{\rho}(x)$, where it exists. At equilibrium $J(x)$ in (2) will be constant, and in this context zero. Hence

$$\frac{1}{2}\frac{d}{dx}V(x)\hat{\rho}(x) = M(x)\hat{\rho}(x) \tag{3}$$

Introduce $g(x) \equiv V(x)\hat{\rho}(x)$.

$$\begin{aligned}
\frac{1}{2}\frac{d}{dx}g(x) &= \frac{M(x)}{V(x)}g(x) \\
\frac{1}{g(x)}dg(x) &= 2\frac{M(x)}{V(x)}dx \\
d\ln[g(x)] &= 2\frac{M(x)}{V(x)}dx \\
\ln[g(x)] &= 2\int\frac{M(x)}{V(x)}dx + constant \\
g(x) &= c\exp\left(2\int\frac{M(x)}{V(x)}dx\right) \\
\hat{\rho}(x) &= \frac{c}{V(x)}\exp\left(2\int\frac{M(x)}{V(x)}dx\right)
\end{aligned} \tag{4}$$

where c is an appropriate normalizing constant to make the area under $\hat{\rho}(x)$ equal to 1 in the range of x from 0 to 1.

So far the calculations have followed Roughgarden exactly, but now we make adjustments appropriate for a haploid GA. With a mutation rate of m then Δx due to mutation in one generation is

$$\Delta x_{mut} = x(-m) + (1-x)m = m(1-2x) \tag{5}$$

To calculate the Δx in one generation due to selection, we shall assume that the schema fitness of the allele **0** is f_0 and of allele **1** is f_1. The average fitness $\overline{f}$ depends on the proportion x of **1**s in the current population, $\overline{f} = (1-x)f_0 + xf_1$. We shall define the selective force s in favour of allele **1** as

$$s \equiv \frac{f_1 - f_0}{f_0} \tag{6}$$

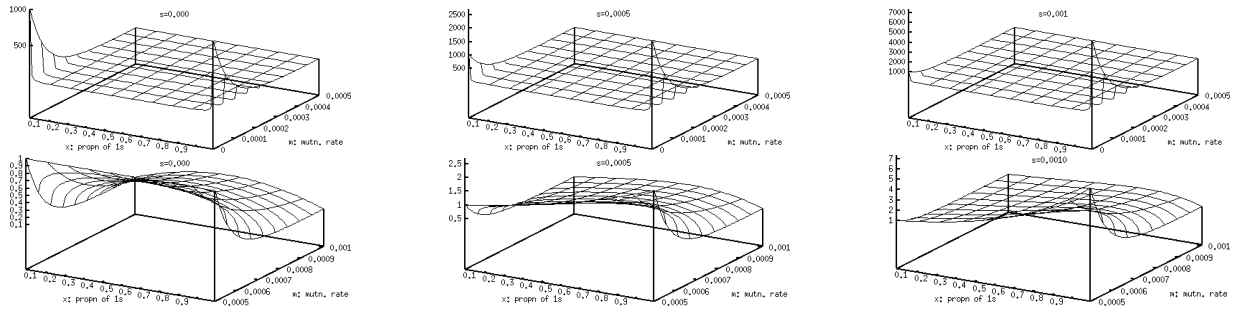

Figure 3: Equilibrium distributions, varying m for particular values of s. The horizontal scale is the proportion x of the allele being selected for, in the range $x = 0.001$ to 0.999. the vertical scale varies from graph to graph, as the constant c in eqn. 11 has here been set to 1; whereas it should normalise the graph so that the area underneath is unity. Hence for the U-shaped curves, only the general shape is indicative.

$$\Delta x_{sel} = x\left(\frac{f_1}{\overline{f}} - 1\right) = \frac{sx(1-x)}{1+sx} \tag{7}$$

Using the population size N, we now convert Δx to a new time scale where N generations equals one unit of time.

$$\begin{aligned} M(x) &= N(\Delta x_{mut} + \Delta x_{sel}) \\ &= mN(1-2x) + \frac{sNx(1-x)}{1+sx} \end{aligned} \tag{8}$$

On calculating $V(x)$ we use the fact that the variance of Δx over one generation is $x(1-x)/N$. Converting to the same time scale as above we have

$$V(x) = N\frac{x(1-x)}{N} = x(1-x) \tag{9}$$

Substituting (8) and (9) into (4) we have the following (the constants on integration can be assimilated into the normalizing constant c):

$$\hat{\rho}(x) = \frac{c}{x(1-x)}\exp\left(W(x)\right)$$

where we define $W(x)$ to be

$$\begin{aligned} &\equiv 2N\int\left(\frac{m(1-2x)+\frac{sx(1-x)}{1+sx}}{x(1-x)}\right)dx \\ &= 2mN\int\frac{dx}{x(1-x)} - 4mN\int\frac{dx}{1-x} + 2sN\int\frac{dx}{1+sx} \\ &= -2mN\ln\left(\frac{1-x}{x}\right) + 4mN\ln(1-x) + 2N\ln(1+sx) \end{aligned} \tag{10}$$

So we have $\hat{\rho}(x)$

$$\begin{aligned} &= \frac{c}{x(1-x)}\left(\frac{1-x}{x}\right)^{-2mN}(1-x)^{4mN}(1+sx)^{2N} \\ &= cx^{2mN-1}(1-x)^{2mN-1}(1+sx)^{2N} \\ &= \frac{c(1+sx)^{2N}}{[x(1-x)]^{1-2mN}} \end{aligned} \tag{11}$$

Here it is clear that the term on the top relates to the forces of selection; whereas the denominator, symmetrical in x and $(1-x)$, shows different characteristics depending on whether the exponent is positive or negative — which depends on the relationship between mutation rate and population size.

4.1 Varying the mutation rate

In the case of zero selective force, $s = 0$, this becomes

$$\hat{\rho}(x) = cx^{2mN-1}(1-x)^{2mN-1} \tag{12}$$

The behaviour of eqn. 12 varies dramatically according as to whether $2mN < 1$ or $2mN > 1$. In the former case of low or non-existent mutation the curve is the U-shaped $\frac{1}{x(1-x)}$, demonstrating that the population will converge completely on one allele or the other. In the latter case of mutation significantly high in relation to the population size, then the reverse will happen and the distribution will be centred on $x = 0.5$. If selection is positive rather than zero, then either the U-shaped curve or the humped curve, as appropriate, will be skewed towards the side favoured by selection.

The graphs in figures 3 and 4 are indicative only of the general shape. In particular, for the U-shaped curves demonstrating genetic drift, the constant c in (12) would need to be zero for the area under the curve to be unity. This gives a vertical bar at $x = 0$ and at $x = 1$ (emphasized in the figures), with zero elsewhere. For this reason the graphs are only shown for x =0.001 to 0.999.

4.2 Varying selection

When m=0, then (11) becomes

$$\hat{\rho}(x) = \frac{c(1+sx)^{2N}}{x(1-x)} \tag{13}$$

which is the limit of a U-shaped curve. The denominator is symmetrical in x and $1-x$. For an indication

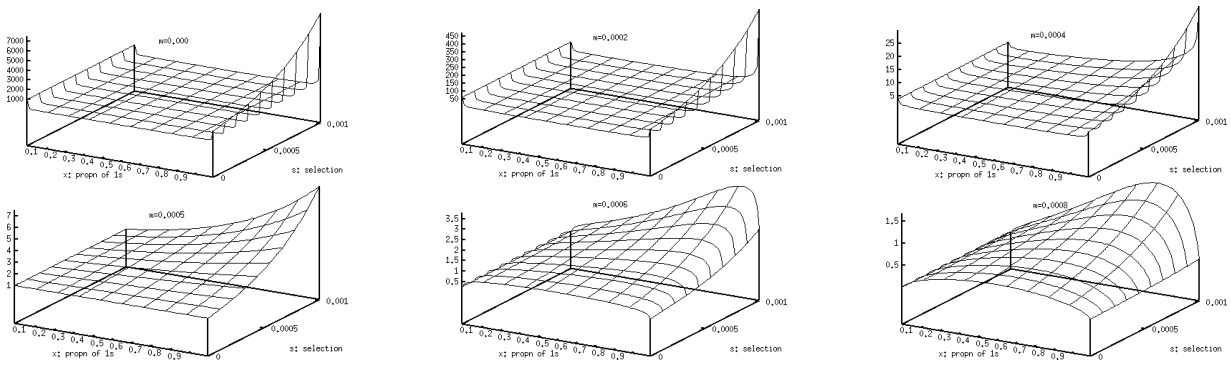

Figure 4: Equilibrium distributions, varying s for particular values of m. See caption to figure 3

of the relative proportions of the converged population that settle at $x = 0$ or $x = 1$, the numerator should be considered for these values; although not much reliance should be placed on this, as it is exactly here that the diffusion approximation breaks down with a finite size population.

Nevertheless, for $s = 0$ the numerator is constant, and as s increases the numerator is more at $x = 1$ than it is at $x = 0$. $(1 + s)^{2N}$ becomes $\mathcal{O}(e)$ when $2sN = 1$, and increases exponentially as s increases above this value. Hence when $2sN \ll 1$ we can expect the two arms of the U-shaped curve to be nearly equal in size (i.e. selection is insignificant), and when $2sN \gg 1$ the arm that selection favours will predominate.

Returning to the Hinton and Nowlan example, the value of s can be calculated when there are q undecided alleles and $20-q$ correct ones. The selective force s at a locus which could change one undecided allele to a correct one can be calculated from the schema fitnesss in figure 2. For any q, it is $(F(q-1)-F(q))/F(q)$, which is calculated in table 2. It can be seen that, whereas convergence on the 'wrong' value can be expected for $2sN \ll 1$, in this case convergence on the wrong value occurred for q as great as 4, and hence $2sN$ as big as 15. This can be explained as due to the hitch-hiking effect, where before genetic drift could take over as the population settled down to equilibrium, the swamping of the population by the very high selection in favour of the first successful genotype has resulted in near-convergence on the 'wrong' value.

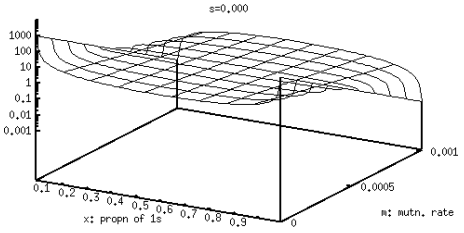

Figure 5: Eqn. 12 for $c = 1$, and m ranging from 0 to 0.001. Here the vertical axis is log-scaled, and selection is zero. The transition between U-shaped and $\bigcap$-shaped curves at $2mN = 1$ can be clearly seen.

5 Conclusion

A puzzling anomaly in the Hinton and Nowlan paper has been explained as the result of genetic drift, due to low selective forces, following a period when very high selective forces and the 'hitch-hiking effect' have distorted the proportions of alleles at temporarily irrelevant loci. By analysing this case in detail, the significance of genetic drift has been brought out, and using the diffusion equation approach some more general results demonstrated.

These can be summarised as the following, where m is the mutation rate, N is the population size, and s the *selective force* in favour of a particular allele 1, at a binary locus, is defined as $(f_1 - f_0)/f_0$, f_1 and f_0 being the schema fitness of alleles 1 and 0:

> If and only if $2mN < 1$ then the population will converge at this locus on one value or another; and if $2sN \ll 1$ it will be almost equally likely to converge on the 'wrong' value as the 'right' one.
> Where 'hitch-hiking' takes place, even for $2sN$ somewhat larger than 1, convergence on the wrong value can still happen.

"The case has been an interesting one," remarked Holmes when our visitors had left us, "because it serves to show very clearly how simple the explanation may be of an affair which at first sight seems to be almost inexplicable." Sir A. Conan Doyle, The Adventure of the Noble Bachelor, 1892.

Acknowledgments

This work was supported by a grant from SERC, the UK Science and Engineering Research Council. It arose from discussion at an Alergic meeting at Sussex.

A Appendix: The Hinton & Nowlan model

A.1 Expected fitness of potential winner

To calculate the expected fitness of a genotype composed of q question-marks and $(20-q)$ **1**s.

$$\begin{array}{lrcll} \text{Define} & p & = & 1/2^q & \text{prob. success on one trial.} \\ & r & = & 1-p & \text{prob. failure on one trial.} \\ & R & = & r^{1000} & \text{prob. failing all 1000 trials.} \end{array}$$

If success comes on the ith trial, for $i \leq 1000$, then the actual fitness is then given by $1 + 19(1000-i)/1000$. The chance of first succeeding on the ith trial, which necessitates failing the preceding $(i-1)$ trials, is given by $r^{i-1}p$.

Hence the expected fitness $F(q)$, bearing in mind the chance R of failing all 1000 trials with a resulting fitness of 1, is given by:

$$\begin{aligned} F(q) &= R + \sum_{i=1}^{1000} r^{i-1} p \left(1 + \frac{19(1000-i)}{1000}\right) \\ &= R + 20p \sum_{i=1}^{1000} r^{i-1} - \frac{19p}{1000} \sum_{i=1}^{1000} i r^{i-1} \end{aligned}$$

But we can use:

$$\sum_{i=1}^{1000} r^{i-1} = \frac{1-r^{1000}}{1-r}$$

and by multiplying each side by r and then differentiating w.r.t. r:

$$\begin{aligned} \sum_{i=1}^{1000} i r^{i-1} &= \frac{d}{dr}\left(\frac{r(1-r^{1000})}{1-r}\right) \\ &= \frac{1 - 1001 r^{1000} + 1000 r^{1001}}{(1-r)^2} \end{aligned}$$

Substituting we get:

$$F(q) =$$

$$R + 20p\frac{(1-r^{1000})}{1-r} - \frac{19p(1-1001r^{1000}+1000r^{1001})}{1000(1-r)^2}$$

$$= 20 - \frac{19(1-r^{1000})}{1000(1-r)} + 19r^{1000}$$

This is used to calculate the figures in table 1, although care must be taken with the precision in computing as very small numbers are involved in the intermediate calculations.

A.2 Expected number of winners at start

In a member of the initial random population, the probability of having no **0**s, i.e. of being a potential winner, is $(3/4)^{20}$. Such a member will be all **?**s and **1**s, with **?**s being twice as likely as **1**s at any locus.. The chance of having exactly q **?**s in a potential winner, given by the binomial expansion of $(2/3+1/3)^{20}$, is

$$\binom{20}{q}\left(\frac{2}{3}\right)^q\left(\frac{1}{3}\right)^{20-q}$$

and then the probability of actual success is $1-(1-0.5^q)^{1000}$.

So the probability of this initial random member being a winner is

$$\left(\frac{3}{4}\right)^{20}\sum_{q=0}^{20}\binom{20}{q}\left(\frac{2}{3}\right)^q\left(\frac{1}{3}\right)^{20-q}\left(1-(1-0.5^q)^{1000}\right)$$

$$\simeq 0.000558$$

This figure differs from the value 0.028 given in (Belew 1989). The probability of there being no winner in an initial random population of 1000 is $(1-0.000558)^{1000} \simeq 0.572$.

References

J.E. Baker (1987). Reducing bias and inefficiency in the selection algorithm. In J. J. Grefenstette, editor, *Genetic Algorithms and their Applications: Proceedings of the Second International Conference on Genetic Algorithms*, pp. 14–21. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Richard K. Belew (1989). When both individuals and populations search. In J. D. Schaffer, editor, *Proceedings of the Third International Conference on Genetic Algorithms*, San Mateo, California: Morgan Kaufmann.

David E. Goldberg and Philip Segrest (1987). Finite markov chain analysis of genetic algorithms. In J. J. Grefenstette, editor, *Genetic Algorithms and their Applications: Proceedings of the Second International Conference on Genetic Algorithms*, Hillsdale, N.J.: Lawrence Erlbaum Associates.

J.J. Grefenstette (1983). A User's Guide to GENESIS. Technical Report CS-83-11, Computer Science Department, Vanderbilt University.

Geoffrey E. Hinton and Steven J. Nowlan (1987). How learning can guide evolution. *Complex Systems*, 1:495–502.

Jonathan Roughgarden (1987). *Theory of Population Genetics and Evolutionary Ecology: an Introduction.* New York: Macmillan.