

Extrema Selection: Accelerated Evolution on Neutral Networks

Terry Stewart

Carleton University

Institute of Interdisciplinary Studies (Cognitive Science)

Ottawa, Canada

tcstewar@chat.carleton.ca

Abstract- A new modification to the genetic algorithm is presented which is specifically designed to increase the rate of evolution on fitness functions with high degrees of neutrality (mutations that do not change the individual's fitness). Instead of allowing random genetic drift to occur when most of the population has reached the same fitness, the "reproduction fitness" of individuals is set to their distance from the population centroid. This has the theoretical effect of spreading the population quickly across the neutral network, and thus finding regions of higher fitness more quickly than it would otherwise. A series of experiments are described which show a significant improvement using this method on the NKp family of fitness functions, and that this improvement is correlated to the degree of neutrality.

1 Introduction

The concept of the "fitness landscape" seems fundamental to any theory of artificial (or, indeed, natural) evolution. However, some recent developments, drawing inspiration from natural evolution, have called into question one particular hidden assumption in the common interpretation of the "fitness landscape".

This assumption is that fitness landscapes are made up of peaks and valleys, and in order to get from one peak to another peak, one must travel through a valley. In other words, populations are often stuck at "local maxima", and in order to improve any farther, they must first decrease their fitness. This is often referred to as the problem of premature convergence. We envisage the population clustered around a peak, and in order to reach the next peak we require either a highly unlikely multiple mutation, or the chance event of a low-fitness individual reproducing and mutating in the correct direction. This can even be seen as an explanation for the "punctuated equilibrium" model of evolution, where a population stays at one particular fitness level for a long time, and then suddenly jumps to a new, higher, fitness value.

Certain researchers have questioned this point of view. In particular, evidence from the RNA to protein mapping problem has shown that these complex fitness functions have a high degree of what is termed "neutrality" (Huynen

et al., 1996), (Forst et al., 1995). This means that there are a great number of mutations, which do not affect the fitness of a particular individual. By following a long sequence of these neutral mutations (i.e. by moving along a "neutral network"), one can arrive at a new, higher fitness value without ever requiring a super-mutation. Punctuated equilibrium, then, is not a population sitting stagnant on the top of a fitness hill. Rather, the individuals in the population are spreading along this neutral network, randomly drifting until a new, higher fitness value is found.

Interestingly, this point of view also diverges from standard evolutionary theory in that evolution can continue to be significant even after the population has converged. Furthermore, it argues that having redundancy in the genome is a good thing. We should not worry about having a highly efficient encoding, since we do not want every single mutation to immediately affect the fitness. This leads to a new view as to the use of "junk" DNA, which would lead to the possibility of future mutations which may improve fitness. These are the neutral pathways that evolution can follow to find new, better solutions without decreasing in fitness. For more information on these possibilities, the reader is directed to (Harvey and Thompson, 1996).

Another way of seeing the difference between this new view and the standard "landscape" view is to recognise that the standard view is inherently two-dimensional, while actual fitness landscapes are extremely high-dimensional. The practical import of this observation is that there are many more direct paths between two points in a high-dimensional space than there are in a low-dimensional space. The higher the dimensionality, the more paths, and the more likely it is for there to be a path from one point in the space to another of higher fitness which does not involve decreasing fitness at any point.

However, this paper shall not endeavour to further argue for neutral network theory. It is sufficient to recognise that the neutrality of a fitness function may be a significant issue when evolving solutions. With this in mind, the remainder of this paper describes a novel modification to the standard genetic algorithm which is specifically designed to take advantage of neutral networks. This new approach has been examined using a landscape known to exhibit neutrality,

and shows a significant improvement over a standard GA. Furthermore, the amount of improvement is correlated to how neutral the landscape actually is.

2 Extrema Selection

The observation that leads to this new approach is that, if neutral network theory is correct, then much of the time that an evolutionary algorithm is running, it is exploring a neutral network. This corresponds to the long periods where there is no obvious improvement in the solutions found by the GA. In this situation, most of the individuals in the population have the same fitness values, and the ones that do not generally have a very low chance of reproducing. This means that all of the reasonably fit individuals in the population have the same chance of reproducing. In other words, there is a random drift within the population, without the presence of any selection pressure. This results in the following question: Is there a better way to explore a neutral network?

As an alternative, I propose the following methodology: individuals are selected for breeding based on their distance from the population centroid. In other words, extreme individuals, far from the mathematical mean of the population, that still have the same fitness as everyone else, would be more likely to produce offspring. This would have the effect of biasing the system towards exploring new areas of the search space, rather than randomly walking over much of the same space. The name “Extrema Selection” is given to this approach.

In order for extrema selection to work, the population must be constrained to be on the current neutral network, so mutations that result in lower fitness values must be ignored. Thus, extrema selection is meant to be paired with a “thresholding” function that arbitrarily stops individuals with a low fitness from reproducing. This thresholding can either be strict (any individual with a fitness less than the maximum fitness of the population is not allowed to breed), or more lenient (any individual with a fitness less than 0.9 times the maximum is not allowed to breed). Note that this is the only way in which the actual fitness of the individuals enters into the picture.

It should be noted that extrema selection can be thought of as creating a new fitness value that replaces the old one for the purposes of selection. First, the standard fitness is calculated for each member of the population. Then, all individuals with a low fitness have their fitness values set to zero. Finally, all other individuals have their fitnesses set to the distance (squared Euclidean) from the individual to the population mean. By looking at extrema selection in this manner, we can see that it is compatible with such approaches as tournament selection or roulette selection. Indeed, any particular style of genetic algorithm could be modified to incorporate extrema selection.

Some confusion does arise here, in that we are using two different definitions of the word “fitness”. Generally, fitness is both how good a solution the particular individual is to the problem at hand, as well as a measure of its relative likelihood of producing genetically related offspring. With extrema selection, we are separating these definitions. It is important to remember that while discussing extrema selection, fitness refers only to how good a solution the individual is (the value of the fitness function for that genome). Two individuals in the population may have the exact same fitness, but very different chances of producing offspring.

This stretches the definition of “fitness” considerably. However, it still seems to be a good term to use, since fitness is still the major deciding factor in breeding. The “extremeness” of an individual merely decides who, among equals, should breed for the best chance of producing future, better individuals.

This is not to say that extrema selection sits in the background and waits for a population to all be on a neutral network, and then kicks in. Extrema selection is always active. This means that the typical evolutionary process looks very different than it does in the standard model. Generally, there will be a large number of individuals in the population with the same fitness. There will also be a large number with lower fitness, all of which have no chance of breeding. Eventually, a new individual will arise with a higher fitness than any current individual. This new genome will suddenly be the only one on the new, more fit neutral network, and so will be the only one able to breed. The next generation will arise pre-converged on that new location, and then proceed to explore that new neutral network.

3 Accelerated Evolution

To determine if extrema selection actually results in an improved evolutionary algorithm, a series of experiments was performed. In these experiments, a large number of genetic parameters were varied, and for each combination of parameters, a number of evolutionary runs were made. By analysing the results of these runs, we can see that extrema selection does, on the particular fitness function in question, provide a significant improvement.

For all of these experiments, the NKp family of fitness functions was used. This family, introduced in (Barnett, 1998), is a generalisation of the standard NK fitness function, and was chosen for these experiments because it is known to be highly neutral. Most importantly, its neutrality can be modified independently of its ruggedness (as measured by the auto-correlation function). Since ruggedness is the traditional measure of how “difficult” a particular fitness function is, this allows us to examine the effect of neutrality on its own.

3.1 Experiment 1: The Genetic Algorithm

The first experiment was meant to determine if extrema selection can perform better than standard selection over a range of different evolutionary algorithms. The following five parameters were varied:

- Extrema selection vs. normal selection
- Thresholding vs. no thresholding
- Roulette vs. tournament
- Crossover vs. no crossover
- Generational vs. steady-state

This leads to 32 different settings for these parameters. For each possible combination, 100 evolutionary runs were performed. In each case, the population size was 100, mutation rate 0.01 (these values were varied in the next experiment).

Each run was performed on the same NKp fitness function, generated with $N=60$, $K=12$, $p=0.99$. The measured value for each run was the number of individuals evaluated before finding a fitness greater than 0.1. If a particular run went to 20,000 individuals before finding a sufficiently good result, the run was terminated and a value of 20,000 was recorded.

The following table gives the mean number of individuals examined for the different parameters. It can be clearly seen that extrema selection outperforms the standard selection methods by a very significant margin.

	Extrema Selection		Standard Selection	
		Thresh.		Thresh.
RG	20000	4580	13763	9451
RGC	20000	5859	13063	13217
RS	20000	3095	11322	9642
RSC	20000	3042	8360	10496
TG	20000	18036 (5070)	7876	12336
TGC	20000	20000 (5364)	9940	13080
TS	20000	4483	6884	8563
TSC	20000	5261	6673	8996

Table 1: Number of individuals required to find one of fitness 0.1 or more. (R=Roulette; T=Tournament; S=Steady-State; G=Generational; C=Crossover)

The two values in parentheses are the results from modifying the standard generational tournament selection to make sure that each member of the population participates in at least one tournament each generation. Without this modification, advantageous mutations are often lost. This is because when a very good new individual is found, all other individuals have their fitnesses set to zero. Under standard tournament selection, it is common for that one good individual to not be chosen for breeding at all. However, it should also be noted that tournament selection is generally

used with steady-state situations, not generational, so this is not a particularly important issue.

An analysis of these results (via ANOVA) does show that extrema selection's improvement over standard selection is statistically significant with $p < 0.01$. The first column of numbers (extrema selection without thresholding) is unsurprising, since without thresholding the actual fitness of an individual is irrelevant, and so evolution has nothing to work with. The final column (standard selection with thresholding) serves to show that thresholding by itself does not generally improve the evolution, and in fact most of the time makes it worse.

3.2 Experiment 2: Mutation and Population Size

A further experiment was performed to determine the effect of varying the population size and the mutation rate. Populations of 10, 100, and 1000 individuals were tried, along with mutation rates of 0.001, 0.003, 0.01, and 0.03. For each combination, 100 runs were made using extrema selection, and 100 runs of standard selection. Both roulette and tournament selection were also examined, and in all situations there was a steady-state population with crossover.

The results for this experiment are in the following table. In each cell, the top number is with standard selection, and the bottom number is extrema selection. Again, it is clear that extrema selection improves over standard selection methods.

	m=0.001	M=0.003	m=0.01	m=0.03
N=10	18820	16967	11802	5685
	10715	6054	5053	2812
N=100	17474	14912	8910	5115
	10179	5964	3466	2748
N=1000	19126	19189	19146	19945
	13079	10096	7076	6231

Table 2: Roulette steady-state with crossover, varying mutation rate and population size for normal selection (top) and extrema selection (bottom).

	m=0.001	M=0.003	m=0.01	m=0.03
N=10	16522	10474	5447	4259
	10764	6213	4216	3553
N=100	15290	12756	5742	5575
	8759	6597	5411	12471
N=1000	19233	18445	18831	19987
	20000	20000	20000	20000

Table 3: Tournament steady-state with crossover, varying mutation rate and population size for normal selection (top) and extrema selection (bottom).

For the roulette settings, extrema selection gives a statistically significant improvement in all cases. For tournament selection, it gives a general improvement, exception in situations of large populations or very high

mutation rates (note that with a population size of 1000, the run is only going to 20 generations before halting).

From this experiment and the previous one, we can make the fairly strong claim that for finding solutions on an NKp landscape with $N=60$, $K=12$, and $p=0.99$, extrema selection will generally give a significant improvement to any evolutionary paradigm.

3.3 Experiment 3: Neutrality

A final experiment was run to compare the improvement of extrema selection to the degree of neutrality exhibited by the fitness function. In this case, the parameters of evolution were fixed at roulette selection on a steady-state population of 100 individuals, with a mutation rate of 0.01 and crossover. With this configuration, both extrema selection and normal selection were investigated on NKp fitness function of varying values of p .

As proved in (Reidys and Stadler, 1998), the auto-correlation function (“ruggedness”) of NKp fitness landscapes does not vary for different values of p . The auto-correlation function is a measure of the similarity between neighbouring points on the fitness landscape. It is generally used as a measure of the difficulty of evolving a solution to a particular problem, since the theory of the GA search algorithm is based on the assumption that similar genomes will have similar fitness.

However, (Barnett, 1997) proves that NKp fitness functions have a degree of neutrality based on the value of p . Thus, with NKp landscapes, we can vary the neutrality without varying the ruggedness. Indeed, this is the reason that this family of landscapes was chosen for these experiments.

In order to see if the improvement that extrema selection gives is related to the neutrality of the fitness function, we can compare the amount of improvement shown by extrema selection to the neutrality of the landscape. The neutrality of the landscape is defined as the probability that a random mutation will be neutral.

In order to determine the improvement on the different landscapes, each method of evolution (extrema selection and normal selection) was run for 10000 individuals, and the improvement was taken to be the difference in the fitness between the best individuals found. The results are shown in Figure 1. The thick line represents the average improvement, while the thin line gives the neutrality of the NKp landscape for that particular value of p .

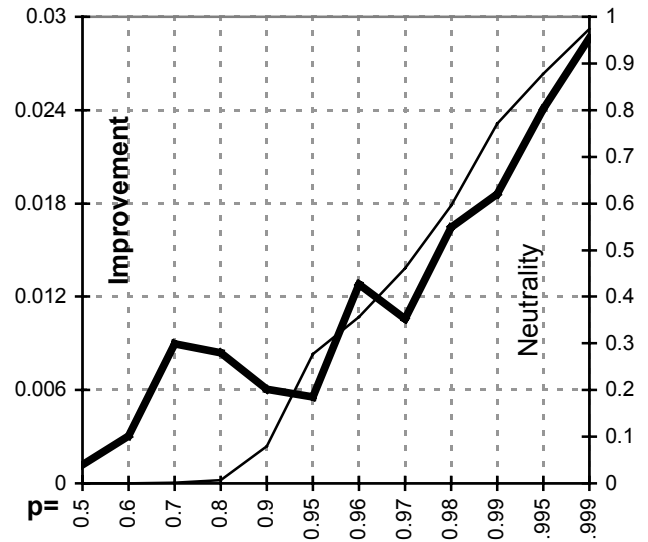


Figure 1: Comparison of the improvement and the degree of neutrality for NKp landscapes of different p-values.

As can be seen, the amount of improvement provided by extrema selection correlates well with the degree of neutrality of the fitness function. It should also be noted that for p values below 0.95, the difference between the results with extrema selection and the results with normal selection is not large enough to be considered statistically significant ($p>0.05$). For p values greater than 0.95, they are strongly significant. Note that the reason for this difference is that the absolute value of the fitness reached for NKp landscapes of low p is higher than that reached by the more neutral fitness functions.

This experiment gives strong evidence that the degree of neutrality is an important factor in extrema selection’s success. On NKp fitness functions of equal ruggedness, extrema selection will only give a significant improvement if the degree of neutrality is sufficiently large. Furthermore, it does exhibit a dose effect, in that the greater the neutrality, the greater the benefit given by extrema selection.

At this point, it is unknown whether or not this benefit extends to non-NKp landscapes.

4 Summary of Results

From the three experiments presented in this paper, it is clear that extrema selection can provide significant improvements in the efficacy of a genetic algorithm. The key factor that led to the development of extrema selection was neutral networks: large connected areas of the fitness function with equal fitness. Extrema selection is specifically designed to search those flat areas quickly, rather than relying on genetic drift. There is strong

evidence of an improvement given by extrema selection on NKp fitness functions, and this improvement is related to how neutral the network is. Source code for all experiments is available upon request.

5 Implications for Artificial Evolution

It is certainly clear from this paper that if someone is trying to evolve high-fitness values on an NKp fitness landscape with a large p-value, then they should certainly use extrema selection. However, I hope a stronger point has been made here.

Since extrema selection was designed to deal with neutrality in general, and not just NKp fitness functions, it is likely that it would be useful in other situations. Indeed, I think there is a reasonable theoretical argument that any fitness function which exhibits a high degree of neutrality would benefit from extrema selection. Certainly, more evidence is needed before we can conclude that it is of general advantage in all such situations. The next step is to conduct experiments with real, complex fitness functions, rather than ones using specifically designed artificial landscapes.

6 Implications for Natural Evolution

After seeing the results of extrema selection on artificial evolution, one is forced to wonder whether or not something like this could be at work in natural evolution. Granted, it is unlikely that the exact same process could be used, since there seems to be no natural way to relate reproduction chances to the genomic difference between an individual and the population's average genome. However, a living creature's phenotype is a lot more complicated than a single fitness value. Indeed, it might be argued that a process such as selection based on sexual characteristics (such as a peacock's feathers) could be considered to be a form of extrema selection. If an individual is extreme in one respect in their phenotype, and yet they are still able to meet the basic functioning standards of the species, then perhaps mating with them could spread the population quickly along the current neutral network?

I offer no evidence or further argument along these lines. However, it does seem to be an interesting new question to ask, and I hope that further work will address it.

7 Further Work

Other than the more theoretical question asked in the previous section, there is much more that needs to be done to examine the usefulness of extrema selection.

First of all, other fitness functions need to be examined. This will lead to a broader claim that extrema selection will be useful in non-NKp landscapes. While doing this,

evidence can be gathered as to the correlation between the improvements found (if any) and the degree or neutrality of the particular landscape.

Solid evidence for extrema selection's utility will come from actually making use of extrema selection on a real problem. The standard tasks of controlling a robot or evolving hardware would be suited well to extrema selection, since both generally involve a fair degree of redundancy, and thus neutrality.

To tidy up a few loose ends, I also plan to look at how extrema selection interacts with elitism and rank-based selection, neither of which were considered in this paper.

Acknowledgements

The work presented in this paper would not have occurred were it not for a talk by Dr. Inman Harvey about neutral networks. Nor would it have occurred without my discussions with Dr. Harvey and Lionel Barnett, and I owe my understanding of neutral networks to them. This work was also performed under a scholarship from the Canadian National Science and Engineering Research Council.

Bibliography

Barnett, L. (1997). Tangled Webs: Evolutionary Dynamics on Fitness Landscapes with Neutrality. MSc in Evolutionary and Adaptive Systems, COGS, University of Sussex, UK.

Barnett, L. (1998). Neutrality and Ruggedness: The NKp Family of Fitness Landscapes. In *Proc. of the Sixth Annual Conference on Artificial Life (ALIFE98)*. 18-27. MIT Press, Cambridge, USA.

Forst, C.V., Reidys, C. and Weber, J. (1995). Evolutionary Dynamics and Optimization: Neutral Networks as Model-Landscapes for RNA Secondary-Structure Folding-Landscapes. Lecture Notes in Artificial Intelligence, vol. 929: Advances in Artificial Life (Moran, F., Moreno, A., Merelo, J.J., and Chacon, eds.), Springer-Verlag, Berlin.

Harvey, I. and Thompson, A. (1996). Through the Labyrinth Evolution Finds a Way: A Silicon Ridge. In *Proc. of the First International Conference on Evolvable Systems (ICES96)*. 406-422. Springer-Verlag, Berlin.

Huynen, M.A., Stadler, P.F. and Fontana, W. (1996). Smoothness within Ruggedness: The Role of Neutrality in Adaption. In *Proc. National Academy of Sciences USA*, 93: 397-401.

Reidys, C.M. and Stadler, P.F. (1998). Neutrality in Fitness Landscapes. Santa Fe Institute Working Paper 98-10-089.