

# **Not Measuring Evolvability: Initial Investigation of an Evolutionary Robotics Search Space**

**Tom Smith<sup>\*1</sup>, Phil Husbands<sup>2</sup> and Michael O'Shea<sup>1</sup>**

Centre for Computational Neuroscience and Robotics (CCNR)

<sup>1</sup>School of Biological Sciences, <sup>2</sup>School of Cognitive and Computing Sciences

University of Sussex, Brighton, UK

\* toms@cogs.susx.ac.uk

## **Abstract-**

**This paper investigates the underlying search space of a difficult robotics problem. Previous work on the development of neural networks incorporating a model of gaseous neuromodulation (the *GasNet*) suggested that such networks are well suited to evolutionary design for some problems [9, 27]. Networks allowed to use the gaseous signalling mechanism evolved significantly faster than networks with the mechanism disabled, implying a significant difference between the two search spaces. In this paper, we investigate this difference using a series of standard techniques for predicting the “difficulty” of search in fitness landscapes. We show that in this instance, measures based on random sampling do not discriminate between the two search spaces, due to the highly skewed nature of the fitness distributions, similar to those found in other difficult optimisation problems. It may be that such metrics are not useful as measures of difficulty for a class of complex problems.**

## **1 Introduction**

In previous work, we have investigated the incorporation of mechanisms based on current research in gaseous signalling in the brain [6, 23], into a more standard neural network [9, 27]. The resulting *GasNets* were used for controlling a robot engaged in a complex task, and proved faster to evolve than corresponding networks not allowed to use the gaseous signalling mechanism. This speed difference must be due to differences in the two fitness landscapes underlying the evolutionary process. In this paper we use a series of standard techniques for predicting the difficulty of search in such spaces in an attempt to identify the differences producing the faster evolutionary search.

The evolutionary robotics method applies the techniques of stochastic optimisation to the production of controllers for robots operating autonomously in complex environments. Two chief concerns are the development of new behaviours and the development of new controller types. However, most evolutionary robotics practitioners pay little attention to the search space underlying the behaviours and controllers they are trying to develop. Studies comparing, for example, two types of neural network usually report simply the number of generations required before success. It is understandable that researchers are loathe to devote large amounts of time on in-

vestigation of the search space, but as evaluation of genotypes becomes ever more expensive, it becomes ever more important to have available a good comparative method. For complex problems such as evolutionary robotics [3] and evolutionary hardware [29], identification of easily evolvable systems is crucial. One possible method is analysis of the actual performance of the systems over their lifetime, identifying useful operational features. A second approach, used here, is analysis of the search space underlying the representation of the problem and solution.

This paper presents a preliminary exploration into the different search spaces underlying a comparative study of two particular types of neural network [9], using statistical measures of the search spaces in order to predict the known differences in evolvability. The results presented show that a number of standard measures of search space difficulty simply fail to discriminate between the two spaces. A breakdown of the evolutionary process identifies the reasons for this failure; the two spaces only differ at high levels of fitness not reached with the random sampling techniques many standard measures rely on. Further work is currently looking at biased sampling of the space through collection of both weighted sample sets and online evolution sample sets.

The paper is laid out as follows: The first section summarises the notion of a search space, and section 3 briefly introduces the *GasNet* approach. Section 4 outlines theoretical measures for approximating the difficulty of searching for good solutions in such a space, while section 5 applies these metrics to the spaces underlying the *GasNet* neural networks. The paper concludes with discussion.

## **2 Search Spaces**

The notion of evolvability in evolutionary computation has strong connections with the idea of the optimising process as searching some large space. The classic metaphor of the genetic *search landscape* [36] views the space in which search takes place as a high-dimensional landscape, with one dimension per genotype bit and an extra dimension - *height* - representing the fitness of the genotype at that point. The search space defined by a two bit representation can thus be viewed as a three-dimensional landscape, across which genotypes move during the search process. Mutation operators typically produce small movements in this landscape, while crossover of two different genotypes generally has some re-

sult intermediate between the genotypes.

This view of the search space leads naturally to the identification of the major problems with which any search process will have to cope; ruggedness and modality [14]. Highly *epistatic* problems where fitness is dependent on multiple inter-gene interactions will produce a rugged landscape. Similarly, a high degree of *modality*, i.e. large numbers of local optima, will be seen as large numbers of hill-tops in the landscape with no neighbours of higher fitness. The feature of *neutrality* is also seen, as plateaus in the landscape where many genotypes have the same fitness/height [16, 11].

An alternative view, explicitly based on defining the search space through the operators used, argues that the fitness landscape metaphor is misleading in the identification of local optima [12, 17]. Mutation operators acting on more than one bit, and other complex operators such as crossover, may not ‘see’ these hill-tops as local optima at all. A more accurate view of the search space would be as a connected graph, with vertices connected through the action of mutation and other operators.

However, it is unclear what explanatory or interpretive power this picture provides especially when search processes can provide very high connectivity between vertices. For instance, representations using real-valued genotype bits and Gaussian mutation operators acting on one or more bits simultaneously could lead to fully connected hypergraphs, useless as a visualisation aid. For these reasons, the authors prefer the original “fitness landscape” analogy. The rest of the paper thus follows the fitness landscape picture, rather than the connected hypergraph analogy.

## 2.1 Search Space Difficulty

Work on characterising problems in terms of their difficulty (see section 4) usually focuses on boiling the fitness search space down to a single parameter, typically a measure of the ruggedness of the space. However, it is clear that such attempts cannot work in general; epistasis, modality, deceptivity, and neutrality may all affect the problem difficulty. A further problem is the possible anisotropy and non-homogeneity of the space, implying that calculated measures may not be equal across the space, or at different fitness levels.

These caveats aside, there is no doubt that knowledge of the problem difficulty, however approximate, is invaluable when tackling the complex problems faced by practitioners of artificial evolution and other stochastic search processes. The next section introduces previous results which this work attempts to explain in terms of search space properties.

## 3 The Problem: *GasNets*

The *GasNet*, introduced by Husbands *et al.* [9, 27], incorporates a mechanism based on the neuron-modulating properties of a diffusing signalling gas into a more standard sigmoid-unit neural network. In previous work the networks have been used in a variety of evolutionary robotics tasks, comparing

the speeds of evolution for networks with and without the gas signalling mechanism active.

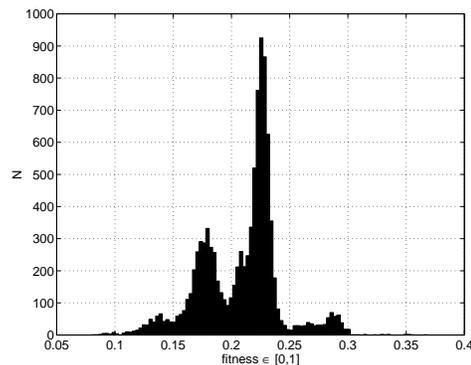


Figure 1: The fitness distribution histogram of a single genotype evaluated 10,000 times, ranging from [0.07, 0.37]. Possible fitness  $\in [0, 1]$ . The bimodal (or even 4-modal) nature of the distribution is due to the controller successfully completing the task on a varying number of the trials in each evaluation.

In the work presented here, we are using the two types of network for control of a robot engaged in a visual shape discrimination task; from random starting orientation and position in a black arena the robot must navigate to a white triangle while ignoring a white square. However, we can consider the two scenarios simply as a two distinct mappings from genotype to phenotype. As in many problems requiring controllers to provide sensor-to-motor mappings over time, fitnesses are extremely time consuming to evaluate (in the work presented here, generating a sample of  $10^6$  fitnesses takes around 24 hours on a Pentium II 700MHz machine) and inherently extremely noisy (figure 1). Success in the task was taken as an evaluated fitness of 1.0 over thirty successive generations of the genetic algorithm. For further information on the actual networks and tasks see [9, 27].

## 3.1 The Genetic Algorithm

A distributed asynchronous updating genetic algorithm was used, with a population of 100 arranged on a  $10 \times 10$  grid. Fitness was awarded on the fraction of the distance moved towards the triangle over a series of 8 runs with different initial conditions. Parents were chosen through rank-based roulette-wheel selection on the mating pool consisting of the 8 nearest neighbours to a randomly chosen grid-point. The child solution was a mutated copy of the parent (the mutation operator applied a 3% mutation probability per bit, and a small probability per genome of adding or deleting a network node. No crossover was used.) and placed back in the mating pool using inverse rank-based roulette-wheel selection.

## 3.2 Results

Over a series of twenty evolutionary runs both with and without the gas signalling process turned on, networks allowed to

use the gaseous signalling mechanism reached success significantly faster (table 1). This speed difference was seen in several different evolutionary robotics scenarios [9].

Gas mean rank (median generations)	16.88 (675)
No gas mean rank (median generations)	24.13 (1228)
Mann-Whitney U	127.5
2-tailed P	0.049

Table 1: Mean ranking and median generations required for success, for conditions with/without the gas signalling enabled. The results for the Mann-Whitney test for similarity between the ranked distributions are also shown, with the probability for identical distributions,  $P < 0.05$ .

## 4 Methods: Search Space Metrics

A large number of metrics have been developed as proxies for search space “difficulty”<sup>1</sup>. Development of such measures tends to highlight the inadequacies of existing methods when applied to some theoretical landscape, then proposing a new metric which does not suffer the same problem. However, it should be stressed that the existence of a global measure applicable in all situations is not possible in principle through the “No Free Lunch” theorem [35]. The following section describes some standard methods, focusing on the ruggedness, or epistasis, of the search space.

### 4.1 Correlation Lengths

Weinberger [33, 34] proposed the autocorrelation function  $\rho(d)$  as a mathematical definition of landscape ruggedness:

$$\rho(d) = \frac{\langle f(x)f(y) \rangle - \langle f \rangle_c^2}{\text{var}_c[f]} \quad (1)$$

the correlation between the fitnesses  $f$  of two sets of points  $x, y$  separated by distance  $d$ , typically the Hamming distance. The measure is defined over the entire space  $c$ , but Weinberger introduced the *random walk* method as an approximation for  $\rho(d)$ . A random point, or genotype, is created and the walk constructed through successive application of a mutation operator. For a walk of length  $l$ :

$$\rho(d) \approx \sum_{s=1}^{d=l-s} \frac{\langle f_s f_{s+d} \rangle - \langle f \rangle^2}{\text{var}[f]} \quad (2)$$

the correlation between the fitnesses  $f$  of two sets of points  $s, s + d$  separated by  $d$  applications of the mutation operator. One advantage of the random walk method for approximating the autocorrelation function is the explicit use of the mutation

<sup>1</sup>The definition of difficulty here is slightly ambiguous, but is usually taken to mean “time required to find good solutions”, alternatively “the goodness of solution found after a certain time”. The definition is usually used only comparatively; Naudts and Kallel [22] among others bemoan the lack of standard test-problems and difficulty measures.

operator to generate the genotype sets. Non-standard operators can be used, even if Hamming distance cannot be calculated, as with variable length representations where mutation operators may increase or decrease genotype length.

Weinberger defines the correlation length as the measure of ruggedness in the system. The longer the correlation length the smoother the space, or the further one can move before the fitnesses at the start points are not correlated with the fitnesses at the end points:

$$\tau = -\frac{1}{\ln(\rho(1))} \quad (3)$$

Others define the length to be the distance at which the correlation function falls below an arbitrary value, usually 0.5. Weinberger showed that this correlation length correctly predicted increasing ruggedness for increasing epistatic interaction  $K$  in Kauffman’s  $NK$  landscapes [33]. Other measures have been developed for correlation analysis of fitness landscapes, notably Lipsitch [18] and Hordijk [7].

### 4.2 Information Content

An alternative method for approximating landscape ruggedness through random walk is given by Vassilev [32, 31], based on the idea of the *information content*  $H(\epsilon)$  of a system as a measure of how difficult it is to describe that system:

$$H(\epsilon) = -\sum_{p \neq q} P_{[pq]} \log_6(P_{[pq]}) \quad (4)$$

All points along the random walk can be characterised by the fitness of the nodes on either side. A local optima, either minima or maxima, has fitness increasing then decreasing (or vice versa); in this case the block is referred to as  $pq$ . Nodes where fitness consistently increases or consistently decreases can be characterised by  $pp$ .

The parameter  $\epsilon$  defines neutral fitness; if the fitness difference between neighbouring nodes is less than this value, the nodes are taken to be equal in fitness. Thus as  $\epsilon$  increases from 0 to the maximum fitness difference between nodes along the walk, the amount of fitness change, or entropy, decreases to zero.  $H(\epsilon)$  is thus a measure of the entropy of the random walk, as defined by the relative probabilities  $P$  of encountering blocks  $pq$ . The information content of the system is defined as the value of  $\epsilon$  at which  $H(\epsilon) = 0$ , or the maximum fitness difference between neighbouring points on the walk.

For systems with high epistasis the information content will be high, and low for systems with low epistasis. Vassilev shows the measure correctly predicts ruggedness for the  $NK$  landscapes [31].

### 4.3 Other Measures

The metrics of landscape ruggedness outlined above use samples of the fitness space based on random walks. Many other measures have been developed: Fitness-distance correlation

[13] requires knowledge on the distribution/position of fitness optima, and uses the correlation of genotype fitness with distance to known optima as a metric. Epistasis variance [4] calculates the level of variance in the system with respect to a set of first order approximations to the system. Site-wise optimisation [21], density of states [24] and other techniques investigate the distribution of genotypes in the space at different fitness levels. However, few of these approaches are applicable to the evolutionary search space investigated in this paper. The positions of the global optima are unknown, the fitness evaluation of robot performance over time in a real environment is not easy to project onto a first order approximation, and the variable length permutable representation used means calculating the distance between two arbitrary solutions in the space is not a simple task.

One more promising area of interest lies in the idea of biased sampling. It is likely that the search spaces of difficult problems contain large proportions of solutions with low/zero fitness; only biased sampling procedures will allow analysis of the search space at higher levels of fitness. Weinberger [34] introduced the notion of a *random adaptive walk*, where the next step on the walk is accepted only if the fitness is higher than that of the current genotype (in contrast to the random walk where the next step is accepted regardless). In effect the walk is a simple hill-climb, and will halt at a local optimum. The lengths and heights of the walks are thus a measure of the modality of the space; spaces containing large numbers of local optima will produce short adaptive walks, while spaces without local peaks will allow long walks to the global optima. This issue is discussed further in section 6.

#### 4.4 Landscape Neutrality

This section has so far concentrated on landscape ruggedness as a proxy for difficulty in finding good solutions. However, the importance of non-adaptive *neutral* evolution has also been argued for [16, 10, 5, 2]. In high-dimensional spaces, local optima and other epistatic effects may have less impact on the search process. Instead, evolution can proceed through randomly searching a sequence of neutral changes, a *neutral network*. In an extension to Kauffman’s *NK* landscapes [15], Barnett has developed the tunably neutral set of *NKp* landscapes and shown that altering the level of neutrality *p* can affect the evolutionary dynamics for systems with fixed epistatic interaction *K* [1]. It may be possible to design systems in such a way as to encourage this neutral evolution, e.g. through redundancy in the genotype-phenotype mapping [25].

Little research has been done on calculating the degree of neutrality for a given problem; most work has focused on the effects of adding some known degree of neutrality. It is also unclear to what extent neutrality is useful as a potential escape from local optima, and to what extent it is harmful in the sense of producing flat plateaus that must be explored through random drift. Theoretical work on search space difficulty should certainly not ignore the issue of neutrality. Neutral networks

in the search spaces investigated here are discussed further in a companion paper [26].

## 5 Results: *GasNet* Search Spaces

The significant difference shown between the speeds of evolution for the neural networks with and without the gas signalling mechanism active must be measurable through differences in the two search spaces. The following section looks at the metrics introduced in section 4, and investigates whether they predict differences between the two search spaces.

### 5.1 Random Sampling

Straightforward random sampling of  $10^6$  genotype fitnesses in each of the spaces showed no discernible differences (table 2 and figure 2). The vast majority of fitnesses found through such sampling are extremely low, with the maximum only at 53% of the maximum obtainable fitness. Note, an evolutionary run hitting success in 500 generations will have sampled  $5 \times 10^4$  genotypes, only 5% of the random sample used here.

Fitness	Gas N	No-gas N
0.0-0.099	992,927	993,434
0.1-0.199	6,521	6,042
0.2-0.299	508	48
0.3-0.399	35	35
0.4-0.499	8	9
0.5-0.599	1	0
Max. Fitness	0.53	0.50
Mean Fitness	0.022	0.022
Fitness $\sigma$	0.024	0.024

Table 2: Random sample ( $10^6$ ) of network fitnesses with/without gas. Note: Fitness lies in range [0.0,1.0]; no fitnesses above 0.53 were obtained through random sampling.

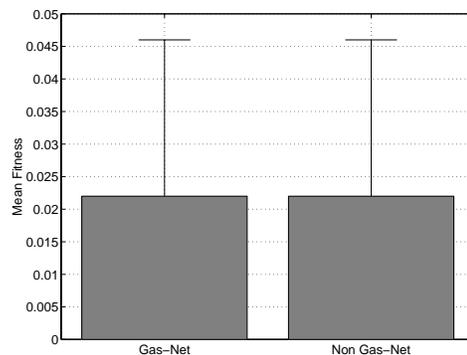


Figure 2: Mean fitnesses of random sample ( $10^6$ ) with/without gas signalling.

An extension of the random sampling procedure was to take a subset of the sample (100 genotypes), and apply  $10^4$  random mutations to each genotype, see figure 3. Again, no

significant differences were seen, and mean and best mutations were still extremely low.

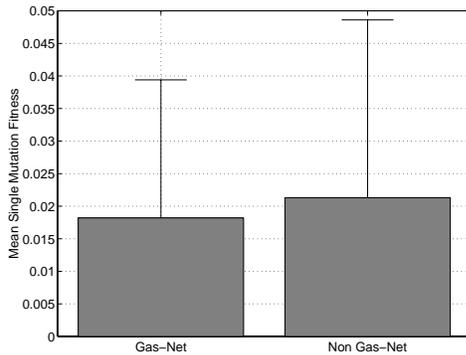


Figure 3: Mean fitnesses of  $10^4$  single mutations from each of a random sample of 100 genotypes. Data shown for with and without gas signalling conditions. Error bars show standard deviations.

## 5.2 Random Walks

One hundred random walks of  $10^4$  steps were performed (another sample of  $10^6$  genotypes), and the correlation lengths and information stability measures described in section 4 calculated, see figures 4 and 5. No significant differences were observed between the conditions.

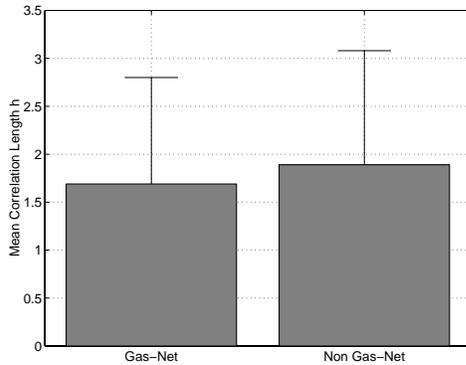


Figure 4: Mean correlation lengths for 100 random walks of  $10^4$  steps. Data shown for with and without gas signalling conditions. Error bars show standard deviations.

One point to note here is the high standard deviations in the correlation lengths, and to a lesser extent in the information stabilities. These are much larger than were seen with such measures calculated on Kauffman's  $NK$  landscapes [31], and indicate a large degree of anisotropy in the system. However, calculation of Stadler and Grüner's coefficient of anisotropy [28] does not show the same level of anisotropy. This is discussed further in section 6.

## 5.3 Modality

One thousand random adaptive walks [34] were performed from random starting points, stopping when a local optimum

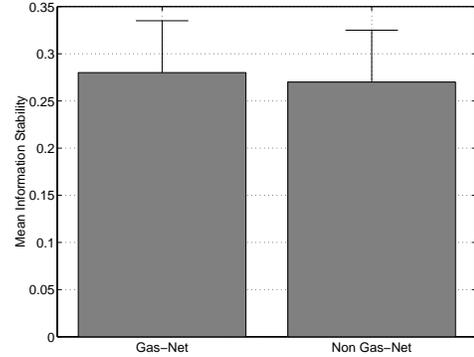


Figure 5: Mean information stability for 100 random walks of  $10^4$  steps. Data shown for with and without gas signalling conditions. Error bars show standard deviations.

was reached (figures 6 and 7). At each step, fifty genotypes were generated, and the next step chosen randomly from among the genotypes with higher fitness than at the current step. Due to the mutation operators used, it is impossible to state definitively if a genotype is a local optimum - the walk was terminated if none of the fifty generated genotypes had higher fitness. Again no difference was seen between the conditions.

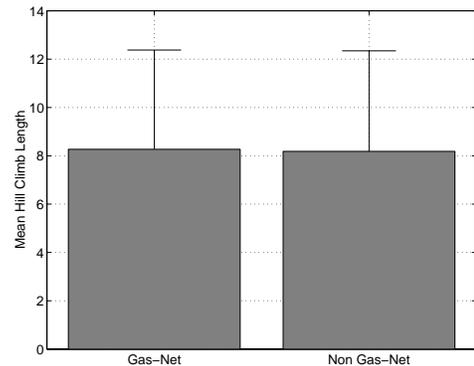


Figure 6: Mean random adaptive walk length for  $10^3$  walks. Data shown for with and without gas signalling conditions. Error bars show standard deviations.

Table 3 shows the distribution of the fitnesses of the final step on the adaptive walks. Note that we have obtained a significantly fitter sample here than the random sample shown in section 5.1, with maximum fitness now at 0.86. However, the vast majority of walks still terminate at extremely low fitnesses.

## 5.4 GA Sampling

The analysis above shows no significant differences between the spaces. However even with the random adaptive walk sampling method, fitnesses found were on average extremely low. Table 2 shows that no fitnesses above 0.53 were found through random sampling, even with large sample sizes. The

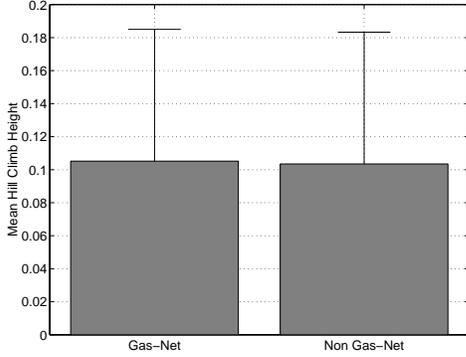


Figure 7: Mean random adaptive walk height for  $10^3$  walks. Data shown for with and without gas signalling conditions. Error bars show standard deviations.

Fitness	Gas N	No-gas N
0.0-0.099	605	627
0.1-0.199	271	273
0.2-0.299	97	63
0.3-0.399	23	27
0.4-0.499	2	8
0.5-0.599	1	2
0.6-0.699	0	0
0.7-0.799	0	0
0.8-0.899	1	0
Max. Fitness	0.86	0.57
Mean Fitness	0.11	0.10
Fitness $\sigma$	0.08	0.08

Table 3: Final genotype fitnesses on  $10^3$  random adaptive walks. Data shown for with and without gas signalling conditions. Error bars show standard deviations.

results from the GA runs show that there must be some difference between the spaces; one hypothesis is that the spaces are only significantly different in certain areas, e.g. above a certain fitness. To test this hypothesis, the performance of the genetic algorithm over the 40 runs was broken down into the numbers of generations required to reach certain fitnesses (figure 8 and table 4).

We see that only above a fitness score of 0.5 is any real difference seen between the performance of the genetic algorithm on the two conditions, with statistically significant differences appearing only above fitnesses of 0.7. Thus all random sampling methods used will simply not be sampling the parts of the space which show differences. Increasing the sample size will not solve the problem; calculated statistics will still be swamped by the huge fraction of the space at near zero fitness. Section 6 continues the discussion introduced in section 4.3 on biasing the sample towards areas of the space that show differences.

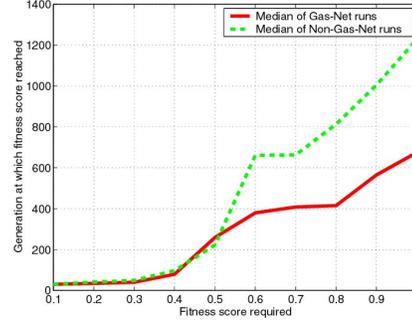


Figure 8: Median generations before given fitness level reached, for both with/without gas conditions.

Fitness level	Gas	No-gas
0.1	30	30
0.2	34.5	41
0.3	40	49
0.4	79	97
0.5	259	222
0.6	379.5	661
0.7	408	663
0.8*	415	813
0.9*	565	940
1.0*	675	1228

Table 4: Median generations before given fitness level reached, for both with/without gas conditions (significant difference \* $p < 0.05$ ).

## 6 Discussion

In this paper we have presented an analysis of the search spaces underlying two different styles of neural network used as robot controllers in a difficult problem. An earlier comparative study found significant differences in the number of generations required to evolve successful networks for the two setups [9]. However, in this paper we have demonstrated that the application of several standard problem difficulty metrics shows no differences between the spaces.

The primary reason for the failure of the metrics used is the highly skewed nature of the distribution of fitnesses over the space. Fewer than 1% of the genotypes obtained through random sampling had fitness above 0.1, and none had fitness higher than 0.53 out of a maximum possible fitness of 1.0. Thus statistics derived from random walks and other random sampling procedures are massively biased towards the lower fitness areas of the space. From a closer examination of the performance of the genetic algorithm, we saw that the differences between the two search spaces only emerged at fitnesses above 0.5. Even with the large samples used here -  $10^6$  genotypes corresponds to 10,000 generations of a genetic algorithm with a population of 100 - random sampling methods will produce very small numbers of genotypes with fitness above 0.5, and so will be unable to distinguish be-

tween the two spaces. Similar distributions have been found in other difficult problems, e.g. evolutionary hardware experiments [30], job-shop scheduling [8] and evolutionary design optimisation [20]. It is likely that in many complex problems the distribution of fitnesses will follow this pattern; a large number of solutions will have extremely low fitness.

A second problem was the large variance in statistics derived from the random sampling; the space is highly anisotropic (section 5.2). Although calculation of Stadler and Grüner's coefficient of anisotropy over the random walks [28] fails to confirm this result, the reliance of the measure on the variance of the mean fitnesses over a number of walks must be viewed with suspicion in the light of the fitness distributions presented here. As most of the proposed difficulty metrics make the explicit assumption that the space is isotropic, results from anisotropic spaces such as seen here should be treated with caution.

It is not possible to accurately describe the shape of a fitness landscape in two dimensions, but a rough sketch may be useful in explaining the results discussed in the previous two paragraphs. The landscape mostly consists of a rugged plateau with fitness in the range  $[0.0, 0.1]$  (the sea?); the height of the plateau varies due to the inherent noise in fitness evaluation. Out of this plateau rise unevenly distributed small peaks (islands?), with varying heights. Random walks (swims?) and other random sampling procedures will collect the majority of their sample from the plateau of low fitness, only occasionally sampling the peaks. Thus statistics based on such samples will simply tell us about the nature of the plateau, and not about the peaks.

The failure of random sampling methods in spaces with such highly skewed fitness distributions leads naturally to the idea of biased sampling, as discussed by Naudts and Kallel in the context of theoretical problem landscapes [22]. The random adaptive walks in section 4.3 are a simple way of producing such samples. Although the samples collected were still heavily weighted towards low fitnesses, higher fitness genotypes were obtained. Such adaptive walks are simple hill climbers; increasing the number of genotypes tested at each step, and allowing several steps of failure before terminating the walk would turn the sampling procedure into a real hill-climb optimisation process. At some critical point, such procedures must start to show the differences that exist in the spaces investigated here. However, there is a real concern that the time taken to collect such a sample may well approach the time taken to solve the problem. For instance if the required biased sample involves collecting samples at or near the optimum, we will have effectively solved the problem merely in the act of description. A useful analogy could be drawn here with Marr's type II systems [19]; the system may not be reducible to a simpler level of description than the system itself. It may well be that we can not characterise complex problem spaces in advance, and will have to employ the "suck-it-and-see" optimisation approach that prevails today.

## Acknowledgements

The authors would like to thank Andy Philippides, Anil Seth and all the members of the CCNR (<http://www.cogs.susx.ac.uk/ccnr/>) for constructive discussion. We would also like to thank the Sussex High Performance Computing Initiative (<http://www.hpc.sussex.ac.uk/>) for computing support. TS is funded by a British Telecom sponsored BBSRC Case award.

## Bibliography

- [1] L. Barnett. Tangled webs - evolutionary dynamics on fitness landscapes with neutrality. Master's thesis, School of Cognitive and Computing Sciences, University of Sussex, 1997.
- [2] L. Barnett. Ruggedness and neutrality: The NKp family of fitness landscapes. In C. Adami, R.K. Belew, H. Kitano, and C.E. Taylor, editors, *Artificial Life VI: Proceedings of the Sixth International Conference on Artificial Life*. MIT Press / Bradford Books, 1998.
- [3] D. T. Cliff, I. Harvey, and P. Husbands. Explorations in evolutionary robotics. *Adaptive Behaviour*, 2(1):71–104, 1993.
- [4] Y. Davidor. Epistasis variance: A viewpoint on GA-hardness. In G.J.E. Rawlins, editor, *Foundations of Genetic Algorithms (FOGA'1)*, pages 22–35. Morgan Kaufmann, 1991.
- [5] I. Harvey and A. Thompson. Through the labyrinth evolution finds a way: A silicon ridge. In T. Higuchi, editor, *Proceedings of the First International Conference on Evolvable Systems: From Biology to Hardware (ICES'96)*. Springer-Verlag, 1996.
- [6] C. Hölscher. Nitric Oxide, the enigmatic neuronal messenger: Its role in synaptic plasticity. *Trends in Neurosciences*, 20:298–303, 1997.
- [7] W. Hordijk. A measure of landscapes. *Evolutionary Computation*, 4(4):335–360, 1996.
- [8] P. Husbands, M. McIlhagga, and R. Ives. Experiments with an ecosystems model of integrated production planning. In T. Bäck, D. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*. Oxford University Press/ Institute of Physics, 1997.
- [9] P. Husbands, T.M.C. Smith, N. Jakobi, and M. O'Shea. Better living through chemistry: Evolving GasNets for robot control. *Connection Science*, 10(3-4):185–210, December 1998.
- [10] M. Huynen. Exploring phenotype space through neutral evolution. *Journal of Molecular Evolution*, 43:165–169, 1996.
- [11] M.A. Huynen, P.F. Stadler, and W. Fontana. Smoothness within ruggedness: The role of neutrality in adaptation. *Proceedings of the National Academy of Sciences, USA*, 93:394–401, 1996.
- [12] T. Jones. One operator, one landscape. Santa Fe Institute Technical Report 95-02-025, Santa Fe Institute, 1995.
- [13] T. Jones and S. Forrest. Fitness distance correlation as a measure of problem difficulty for genetic algorithms. In L.J. Eshelmann, editor, *Proceedings of the Sixth International Conference on Genetic Algorithms (ICGA95)*, pages 184–192. Morgan Kaufmann, CA, 1995.
- [14] L. Kallel, B. Naudts, and C.R. Reeves. Properties of fitness functions and search landscapes. 2000.
- [15] S.A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, 1993.
- [16] M. Kimura. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, 1983.
- [17] J. Kingdon and L. Dekker. The shape of space. In *Proceedings of the First IEE/IEEE International Conference on Genetic Algorithms in Engineering Systems (GALESIA95)*, pages 543–548. IEE, London, 1995.

- [18] M. Lipsitch. Adaptation on rugged landscapes generated by iterated local neighboring genes. In R.K. Belew and L.B. Booker, editors, *Proceedings of the Fourth International Conference on Genetic Algorithms (ICGA91)*, pages 128–135. Morgan Kaufmann, 1991.
- [19] D. Marr. Artificial intelligence - a personal view. Technical Report AIM 355, MIT AI Lab, 1976.
- [20] M. McIlhagga, P. Husbands, and R. Ives. A comparison of search techniques on a wing-box optimisation problem. In H.-M. Voigt, W. Ebeling, I. Rechenberg, and H.-P. Schwefel, editors, *Proceedings of the Fourth Conference on Parallel Problem Solving from Nature: PPSN IV*, volume 1141 of *Lecture Notes in Computer Science*, pages 614–623. Springer-Verlag, 1996.
- [21] B. Naudts. *Measuring GA-Hardness*. PhD thesis, University of Antwerp, 1999.
- [22] B. Naudts and L. Kallel. A comparison of predictive measures of problem difficulty in evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 4(1):1–15, April 2000.
- [23] A.O. Philippides, P. Husbands, and M. O’Shea. Four-dimensional neuronal signaling by nitric oxide: A computational analysis. *Journal of Neuroscience*, 20(3):1199–1207, 2000.
- [24] H. Rosé, W. Ebeling, and T. Asselmeyer. The density of states - a measure of the difficulty of optimization problems. In H.-M. Voigt, W. Ebeling, I. Rechenberg, and H.-P. Schwefel, editors, *Proceedings of the Fourth Conference on Parallel Problem Solving from Nature: PPSN IV*, volume 1141 of *Lecture Notes in Computer Science*, pages 208–217. Springer-Verlag, 1996.
- [25] M. Shackleton, R. Shipman, and M. Ebner. An investigation of redundant genotype-phenotype mappings and their role in evolutionary search. In *Proceedings of the 2000 Congress on Evolutionary Computation: CEC2000*, pages 493–500. IEEE, San Diego, USA, July 2000.
- [26] T.M.C. Smith, P. Husbands, and M. O’Shea. Neutral networks in an evolutionary robotics search space. In *Proceedings of the 2001 Congress on Evolutionary Computation: CEC2001*. IEEE, Korea, 2001. To appear.
- [27] T.M.C. Smith and A. Philippides. Nitric oxide signalling in real and artificial neural networks. *BT Technology Journal*, 18(4):140–149, October 2000.
- [28] P.F. Stadler and W. Grüner. Anisotropy in fitness landscapes. *Journal of Theoretical Biology*, 165(3):373–388, 1993.
- [29] A. Thompson. *Hardware Evolution: Automatic Design of Electronic Circuits in Reconfigurable Hardware by Artificial Evolution*. Distinguished Dissertation Series. Springer-Verlag, 1998.
- [30] A. Thompson and P. Layzell. Evolution of robustness in an electronics design. In J. Miller, A. Thompson, P. Thomson, and Fogarty, T., editors, *Proceedings of the Third International Conference on Evolvable Systems: From Biology to Hardware (ICES’2000)*, volume 1801 of *Lecture Notes in Computer Science*, pages 218–228. Springer-Verlag, 2000.
- [31] V. Vassilev, T. Fogarty, and J. Miller. Information characteristics and the structure of landscapes. *Evolutionary Computation*, 8(1):31–60, 2000.
- [32] V.K. Vassilev. Information analysis of fitness landscapes. In P. Husbands and I. Harvey, editors, *Proceedings of the Fourth European Conference on Artificial Life: ECAL97*. MIT Press / Bradford Books, 1997.
- [33] E. Weinberger. Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biological Cybernetics*, 63:325–336, 1990.
- [34] E. Weinberger. Local properties of Kauffman’s N-K model: A tunably rugged energy landscape. *Physical Review A*, 44(10):6399–6413, 1991.
- [35] D. Wolpert and W. MacReady. No free lunch theorems for search. Technical report, The Santa Fe Institute, 1995.
- [36] S. Wright. The role of mutation, inbreeding, crossbreeding and selection in evolution. In D.F. Jones, editor, *Proceedings of the Sixth International Congress on Genetics*, volume 1, pages 356–366. 1932. Reprinted in Ridley, M. (1997). *Evolution*. pp32-40. Oxford University Press.