Analysis on Topologies of Fitness Landscapes with Both Neutrality and Ruggedness Based on Neutral Networks

Yoshiaki Katada
Setsunan University
Faculty of Engineering
Neyagawa, Osaka 572-8508, JAPAN
katada@ele.setsunan.ac.jp

Kazuhiro Ohkura
Hiroshima University
Graduate School of Engineering
Higashi-Hiroshima, Hiroshima 739-8527, JAPAN
kohkura@hiroshima-u.ac.jp

ABSTRACT

Fitness landscapes which include neutrality have been conceptualized as containing *neutral networks*. Since the introduction of this concept, EC researchers have expected that a population can move along neutral networks without getting trapped on local optima. On the other hand, it has been demonstrated in tunably neutral NK landscapes that neutrality does not affect the ruggedness, although it does reduce the number of local optima. These show that the effects of neutrality are still contentious issues. This paper investigates the effects of neutrality and ruggedness on topologies of fitness landscapes. A neutral network of a fitness landscape is described in a mathematical form based on Harvey's original definition with minor modifications. Our results demonstrate that landscapes with a higher degree of neutrality have the larger sizes of neutral networks. For landscapes with the lowest degree of ruggedness, all networks lead to the networks of the highest fitness via any networks. For landscapes with a higher degree of ruggedness, there are few contact points between the networks of high fitness and the ones of the highest fitness, which seem to be *isolated*, *deceptive* or *rugged*.

Categories and Subject Descriptors: I.2.8 [ARTIFICIAL INTELLIGENCE]: Problem Solving, Control Methods, and Search–*Heuristic methods*

General Terms: Theory

Keywords: fitness landscape, ruggedness, neutrality, analysis

1. INTRODUCTION

Selective neutrality is caused by highly redundant mappings from genotype to phenotype or from phenotype to fitness. Fitness landscapes which include neutrality have been conceptualized as containing *neutral networks* [2]. Harvey [2] first introduced the concept of neutral networks into the GA community. His definition is as follows: "A neutral network of a fitness landscape is defined as a set of connected points of equivalent fitness, each representing a separate genotype: here connected means that there exists a path of single (neutral) mutations which can traverse the network between any two points on it without affecting fitness." This concept is central to the majority of research in this field.

GECCO'09, July 8–12, 2009, Montréal Québec, Canada.
ACM 978-1-60558-325-9/09/07.

EC researchers have expected that a population on a fitness landscape with neutrality can move along neutral networks without getting trapped on local optima even though the landscape also includes ruggedness. One question arises on this: does neutrality diminish local optima? It has been demonstrated in a tunably neutral NK landscape[1, 3] that increasing neutrality does not affect the ruggedness, although it does reduce the number of local optima [1, 3, 4]. This means that the effects of both ruggedness and neutrality on landscapes are still contentious issues and these must be taken into account when genetic operators are designed.

This paper investigates the effects of neutrality and ruggedness on topologies of fitness landscapes in test problems. The paper is organized as follows. The next section describes a neutral network of a fitness landscape in a mathematical form based on Harvey's original definition with minor modifications. Section 3 analyzes topologies in and between neutral networks for a tunably neutral NK landscape and discusses the relationship between topologies of fitness landscapes and GA difficulties. Conclusions are given in the last section.

2. A FORMAL DEFINITION OF A NEUTRAL NETWORK

The expression, "single (neutral) mutations", in Harvey's original definition of a neutral network mentioned in the previous section is ambiguous. Because the amount of a single mutation is not defined, the genetic distance between a parent and the (neutral) mutant has various values depending on the mutation operator applied to the parent. Therefore, a neutral network is not uniquely determined by the original definition.

Thus, we consider the minimum genetic distance between them. When we use binary representations, the minimum genetic distance is described by $\min \mathrm{H}(x^g, y^g)$, where $\mathrm{H}(\cdot,\cdot)$ is the Hamming distance between a parent and the offspring, x^g and $y^g \in \Phi_g$, Φ_g is the set of genotypes. We assume one point mutation applied to parents then $\min \mathrm{H}(x^g, y^g) = 1$.

We describe a neutral network caused by redundant mappings from genotype to phenotype in a mathematical form based on the above consideration.
At first, two individuals, x^g and z^g, are connected, $x^g \sim z^g$, if there exists $\{x_i^g\}_{i=0}^n \subset \Phi_g$, *s.t.* $x^g = x_0^g$, $z^g = x_n^g$, $f_g(x_i^g) = f_g(x^g)$, $\mathrm{H}(x_i^g, x_{i+1}^g) = 1$, where f_g is the mapping from genotype to phenotype and assumed to be surjective and not injective. Thus, a neutral network of a genotype z^g is $\Phi_g^{'}(z^g) = \{x^g \in \Phi_g | x^g \sim z^g\}$.

We extend this definition to redundant mappings from phenotype to fitness. Two individuals, x^g and z^g, are connected, $x^g \sim z^g$, if there exists $\{x_i^g\}_{i=0}^n \subset \Phi_g$, $s.t.$ $x^g = x_0^g$, $z^g = x_n^g$, $(f_p \circ f_g)(x_i^g) = (f_p \circ f_g)(x^g)$, $\mathrm{H}(x_i^g, x_{i+1}^g) = 1$, where f_p is the mapping from phenotype to fitness and assumed to be surjective and not injective. Addition to this assumption, there are two cases on f_g, which is either bijective, or surjective and not injective. In both cases, however, $f_p \circ f_g$ is surjective and not injective only if f_p is surjective and not injective. Thus, a neutral network of a genotype z^g is described in the both cases as follows: $\Phi_g^*(z^g) = \{x^g \in \Phi_g | x^g \sim z^g\}$.

3. THE ANALYSIS OF TOPOLOGIES IN A TUNABLY NEUTRAL NK LANDSCAPE

3.1 Terraced NK Landscape

A terraced NK landscape (TNK) is the tunably neutral NK landscape proposed by Newman and Engelhardt [3]. A terraced NK landscape has three parameters: N, the length of the genotype; K, the number of epistatic linkages between genes; and F, a constant integer for tuning neutrality. The neutrality of the landscape can be tuned by changing the value of F. The neutrality of the landscape is maximized when $F = 2$, and is effectively non-existent as $F \to \infty$. The ruggedness of the landscape increases with the increase of K and is maximized when $K = N - 1$. In this function, a fitness value is assigned to a genotype directly so no phenotype is defined. Thus, it is also assumed that f_g is bijective, and f_p is investigated.

3.2 Simulation Conditions

Computer simulations were conducted by exhaustive enumeration of Φ_g. This enumeration and their clustering into neutral networks are so time consuming that we conducted simulations with the small genotype space. We conducted 10 independent runs for each problem under the landscape parameters, $N = 10$, $K = \{0, 9\}$ and $F = \{2, 8\}$. Here, $|\Phi_g| = 2^{10} = 1024$.

3.3 Simulation Results

Table 1-2 show the sizes of the neutral networks and the number of them for different values of K and F. For $K = 0$, all genotypes belong to an arbitrary neutral network. The sizes of all networks are equal to 16 for $F = 2$ and 2 for $F = 8$. This result suggests that for $K = 0$ the sizes of the neutral networks become larger with the decrease of F, that is, with a higher degree of neutrality. For $K = 9$, the sizes of the neutral networks are not constant, some of which are larger for $F = 2$ than for $F = 8$. That is, there are a small number of large networks for $F = 2$ and a large number of small ones for $F = 8$. Many genotypes do not belong to any neutral networks and stand alone (20% for $F = 2$ and 66% for $F = 8$).

The total number of neutral networks is so large that the whole networks do not appear in this manuscript due to space restrictions. Thus, topologies among these neutral networks are only described. For $K = 0$, all networks have some portals to the networks of higher fitness. This means that individuals always reach the networks of the highest fitness whichever networks they pass through if we assume that they move along neutral networks by one point mutation and then reach a portal to a neutral network of higher fitness. For $K = 9$, the number of portals between networks is very small. For $F = 8$, there is no portal among the networks of the highest several fitness values. Moreover, there are long distances between the networks where there is no contact point. These can be considered that the networks are isolated or "deceptive". It would be difficult for individuals to reach the networks of the highest fitness.

Table 1: The sizes of the neutral networks and the number of them in the TNK for $K = 0$

$F = 2$		$F = 8$	
size	number	size	number
16	64	2	512

Table 2: The sizes of the neutral networks and the number of them in the TNK for $K = 9$

$F = 2$				$F = 8$	
size	number	size	number	size	number
1	210	11	1	1	680
2	39	14	1	2	92
3	10	15	1	3	19
4	11	16	1	4	15
5	6	22	1	5	4
6	2	72	1	6	2
7	1	212	1	11	1
8	2	226	1	-	-
9	1	-	-	-	-

4. CONCLUSIONS

Our results can be summarized as follows: (1) The sizes of the neutral networks become larger for landscapes with a higher degree of neutrality. (2) All networks have some portals to the networks of higher fitness. Thus, individuals reach the networks of the highest fitness through them for landscapes with the lowest degree of ruggedness. (3) Most of networks of the high fitness do not have many portals to networks of the highest fitness for landscapes with a higher degree of ruggedness. This would be considered that the networks are isolated. Future work will investigate whether these results are observed in real-world problems which are expected to have landscapes with both neutrality and ruggedness and to be with the small genotype space

5. REFERENCES

[1] L. Barnett. Tangled webs: Evolutionary dynamics on fitness landscapes with neutrality. Master's thesis, School of Cognitive and Computing Sciences, Sussex University, Brighton, UK, 1997.

[2] I. Harvey and A. Thompson. Through the labyrinth evolution finds a way: A silicon ridge. In *Proceedings of the first International Conference on Evolvable Systems: From Biology to Hardware*, pages 406–422, 1996.

[3] M. Newman and R. Engelhardt. Effect of neutral selection on the evolution of molecular species. In *Proc. R. Soc. London B.*, pages 256:1333–1338, 1998.

[4] T. Smith, P. Husbands, P. Layzell, and M. O'Shea. Fitness landscapes and evolvability. *Evolutionary Computation*, 10(1):1–34, 2002.