Smoothness, Ruggedness and Neutrality of Fitness Landscapes: from Theory to Application

Vesselin K. Vassilev[1], Terence C. Fogarty[1], and Julian F. Miller[2]

[1] School of Computing
Napier University
Edinburgh, EH14 1DJ, UK
E-mail:{v.vassilev, t.fogarty}@dcs.napier.ac.uk

[2] School of Computer Science
University of Birmingham
Birmingham, B15 2TT, UK
E-mail: j.miller@cs.bham.ac.uk

Summary. The theory of fitness landscapes has been developed to provide a suitable mathematical framework for studying the evolvability of a variety of complex systems. In evolutionary computation the notion of evolvability refers to the efficiency of evolutionary search. It has been shown that the structure of a fitness landscape affects the ability of evolutionary algorithms to search. Three characteristics specify the structure of landscapes. These are the landscape smoothness, ruggedness and neutrality. The interplay of these characteristics plays a vital role in evolutionary search. This has motivated the appearance of a variety of techniques for studying the structure of fitness landscapes. An important feature of these techniques is that they characterize the landscapes by their smoothness and ruggedness, ignoring the existence of neutrality. Perhaps, the reason for this is that the role of neutrality in evolutionary search is still poorly understood.

In this chapter some recent results on the spectral properties of the algebraic structures of fitness landscapes are summarized to provide a basis for studying the landscape structure. This approach is further employed to introduce an information analysis that characterizes the structure of fitness landscapes in terms of their smoothness, ruggedness and neutrality. The findings are finally applied in a study of the fitness landscapes generated by evolving digital circuits using an idealized model of a field-programmable gate array. The landscapes of this engineering problem are quite different from many recently studied landscapes that tend to be defined over simplified combinatorial and optimization problems. The difference originates from the genotype representation that is a configuration defined over two completely different alphabets. This makes the study of the corresponding landscapes much more involved. It is shown that the circuit evolution landscapes are products of subspaces with different characteristics. They are landscapes with vast neutrality and sharply differentiated plateau.

A. Ghosh et al. (eds.), *Advances in Evolutionary Computing*

1 Introduction

The notion of a fitness landscape being a collection of genotypes arranged in an abstract metric space was introduced by Wright [1]. A fitness landscape has been defined as a space in which each point is a genotype assigned with a value, and two genotypes are adjacent if one can result from the other after a single mutation. Wright's original intent was to visualize biological evolution as a population flow on a surface in which the altitude level of a point qualifies how well the corresponding organism is adapted to its environment. The fact that evolutionary adaptation can be considered as a walk on a landscape implies the importance of developing a suitable mathematical framework for studying the features of landscapes referred to as landscape theory. In evolutionary computation the theory of landscapes has its own history. Related work can be traced back to studies on the genotype epistasis [2,3], schemata and deception [4–9], landscape modality [10,11], landscape ruggedness [12–15]. Various authors have considered that the structure of landscapes affects the ability of evolutionary algorithms to search [12,16–18]. Further, Jones [19] proposed a model of landscapes general enough to represent a variety of mutation and recombination landscapes. A fitness landscape has been defined on a graph for which each vertex consists of one or more configurations, and two vertices are connected by an edge if the configuration(s) from one of these vertices are obtained by applying the move operator to the configuration(s) from the other. Consequently the number of configurations in a vertex is specified by the nature of the move operator. For instance, the vertices of mutation spaces are single configurations while those of "two parent" recombination spaces are all possible pairs of configurations. A major problem associated with the model above is the difficulty in studying the similarities of mutation and recombination spaces [20]. Work on the mutation recombination isomorphism approach can also be found in [21].

A model of landscapes was given in the algebraic (graph) theory of landscapes outlined in [22,20]. In this model the landscape underlying structure is a hypergraph in which each vertex is a configuration, and each hyperedge is the set of all configurations that can be obtained by applying the move operator to a configuration. To construct a recombination space the model employs P-structures [23,24] that are mappings of pairs of configurations to the hyperedges of the hypergraph associated with this space. The fitness landscape results from the combination of the following objects:

1 A set of configurations that are often referred to as *genotypes.*
2 A cost function that evaluates the configurations, known in evolutionary computation as a *fitness function.*
3 A topological structure that allows relations within the set of configurations.

The set of configurations consists of the encoded elements of the "search" space. The relationship between the configurations is defined by a move op-