Theoretical Computer Science 545 (2014) 59–75

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

# The component model for elementary landscapes and partial neighborhoods

Darrell Whitley [a], Andrew M. Sutton [b], Gabriela Ochoa [c], Francisco Chicano [d]

[a] Computer Science, Colorado State University, Fort Collins, CO, USA
[b] Institut für Informatik, Friedrich-Schiller-Universität Jena, Jena, Germany
[c] Computing Science and Mathematics, University of Stirling, Stirling, Scotland, UK
[d] Lenguajes y Ciencias de la Computación, University of Málaga, Málaga, Spain

ARTICLE INFO

*Article history:*
Received 30 October 2012
Received in revised form 28 March 2014
Accepted 28 April 2014

*Keywords:*
Fitness landscape analysis
Stochastic local search
Elementary landscapes

ABSTRACT

Local search algorithms exploit moves on an adjacency graph of the search space. An "elementary landscape" exists if the objective function $f$ is an eigenfunction of the Laplacian of the graph induced by the neighborhood operator; this allows various statistics about the neighborhood to be computed in closed form. A new component based model makes it relatively simple to prove that certain types of landscapes are elementary. The traveling salesperson problem, weighted graph (vertex) coloring and the minimum graph bisection problem yield elementary landscapes under commonly used local search operators. The component model is then used to efficiently compute the mean objective function value over partial neighborhoods for these same problems. For a traveling salesperson problem over $n$ cities, the 2-opt neighborhood can be decomposed into $\lfloor n/2 - 1 \rfloor$ partial neighborhoods. For graph coloring and the minimum graph bisection problem, partial neighborhoods can be used to focus search on those moves that are capable of producing a solution with a strictly improving objective function value.



## 1. Introduction

A *fitness landscape* for a combinatorial problem instance is defined by a triple $(X, N, f)$. In this definition, $X$ is a set of *candidate solutions*, and $f$ is an *objective function* $f : X \to \mathbb{R}$ that maps each candidate solution to a real value. The objective is to either minimize or maximize $f$. Fitness landscapes are typically associated with local search methods that use a neighborhood move operator to define adjacency between points in the search space. We define a *neighborhood operator* as a function $N$ that maps candidate solutions in $X$ to subsets of $X$. Given a candidate solution $x \in X$, $N(x)$ is the set of points in $X$ (i.e., the neighbors of $x$) that are adjacent to $x$. We say that a neighborhood is regular if the number of neighbors of each solution $x \in X$ is same and we denote with $d = |N(x)|$ this number. We say that a neighborhood is symmetric if the neighborhood relationship is symmetric, that is, $x \in N(y)$ if and only if $y \in N(x)$.

Grover [6] originally showed that for certain NP-hard problems there exist landscapes where it is possible to compute the mean objective function value over the set of neighbors $N(x)$ without explicitly evaluating any of the neighbors of $x$. He showed there exist neighborhoods for the traveling salesperson problem, graph coloring, minimum graph bisection, weight partitioning, as well as not-all-equal satisfiability where this calculation is possible. Stadler [13] named this class of

*E-mail addresses:* whitley@cs.colostate.edu (D. Whitley), andrew.sutton@uni-jena.de (A.M. Sutton), gabriela.ochoa@cs.stir.ac.uk (G. Ochoa), chicano@lcc.uma.es (F. Chicano).

http://dx.doi.org/10.1016/j.tcs.2014.04.036
0304-3975/

problems "elementary landscapes" and Stadler showed that for these problems the objective function $f$ is an eigenfunction of the Laplacian of the graph induced by the neighborhood operator. It can also be shown that a landscape with a symmetric neighborhood operator is elementary if and only if the time series generated by a random walk on the landscape is an AR(1) process [14,5].

Other problems, such as maximum $k$-satisfiability, NK-landscapes, subset sum, and the quadratic assignment problem [16,2] can be shown to be expressible as a superposition of a small number of elementary landscapes. Maximum 3-satisfiability, taken with the traditional Hamming neighborhood, can be expressed as a superposition of three elementary landscapes; it is not only possible to compute the mean, but also to compute other statistical moments in polynomial time, including variance, skew and kurtosis. These statistics can be computed in polynomial time even over generalized neighborhoods that are exponentially large [16].

Some elementary landscapes correspond to problems whose objective functions are linear combinations of components drawn from some finite set $C$. In these cases, the objective function can be characterized as a *discrete linear subset problem* introduced by Papadimitriou and Steiglitz [9]. Each candidate solution to such a problem is defined by some subset of $x \subseteq C$, and the objective function is a weighted sum over the components in $x$. A landscape is *elementary* when the set of candidate solutions and the objective function are coupled with a neighborhood operator that moves components in and out of solutions with uniform frequency. We give a rigorous definition of this *component model* in Section 2.1.

In this paper, we also answer the question of whether or not, using the component model, a neighborhood can be partitioned in such a way to explicitly calculate how components are sampled in the partitions of the neighborhood. We show how the component model can be used to derive conditions under which there will exist partial neighborhoods that retain some of the properties that characterize the full neighborhood. We give a formal definition of partial neighborhood in Section 2.1.

From a theoretical point of view, the existence of elementary landscapes and the ability to compute statistical information about neighborhoods and partial neighborhoods of elementary landscapes is inherently interesting. It is too soon to show that this information can be leveraged to build new and improved local search algorithms. But there have already been some breakthroughs. We can now find the improving moves in a unit distance Hamming neighborhood in $O(1)$ time without explicitly generating any of the neighbors; this result holds for NK-landscapes, maximum $k$-satisfiability, and all pseudo-Boolean functions where the objective function taken together with the Hamming neighborhood operator is a superposition of elementary landscapes [17]. We can also use the average of the neighborhood two moves ahead as a surrogate for the objective function. Thus, instead of optimizing the objective function $f$ directly, we can instead optimize $\operatorname{avg} N(x)$, which corresponds to the expected value of the next move, one move ahead. In some cases, this can yield better results than optimizing $f$ directly [18].

From a practical point of view, the results we present on partial neighborhoods are useful because they provide statistical information about neighborhood operators that make limited changes to a current solution. In some cases, it may be desirable to limit the changes of a local search move operator so that it is not too disruptive. One example is in the planar traveling salesperson problem (TSP) where a partial neighborhood corresponds to the set of all tour inversions of length $\ell$ (see Section 4.1). An instance of planar TSP is given by a set of points $P$ in the Euclidean plane, and a candidate solution is a Hamiltonian circuit through $P$. Let $H \subseteq P$ denote the set of points that lie on the convex hull of $P$. If the points in $H$ appear in a candidate solution in the same order they do on the convex hull, then their relative order is already correct. Indeed, it is possible to design randomized search heuristics that exploit this property [8]. Thus any inversion move that destroys this order is likely to be too disruptive for a local search operator. From such a solution, one may be interested in performing only short inversions of up to length $\ell = |P| - |H|$. This corresponds to a set of partial 2-opt neighborhoods that we introduce in this paper.

We are only just beginning to understand how new and more detailed information about the local search landscape graph can be exploited by search algorithms. The current paper provides not only a foundation for better understanding elementary landscapes, but also for extracting statistical information about partial neighborhoods.

In the next section we briefly review basic mathematical properties of elementary landscapes. We then formally introduce the "component model" for elementary landscapes, providing a more rigorous and solid formulation than the one presented in the previous work [19,21]. Moreover, we approach the characterization from a new perspective that links the component model with the concept of discrete linear subset problems. In Section 3 we provide proofs that the traveling salesperson problem, the weighted graph (vertex) coloring, and the minimum graph bisection are elementary landscapes. These results are known and are included to make the paper self-contained. However, the proof in Section 3.2 for the weighted graph coloring is novel and more rigorous than the previous proof [21]. In Section 4 we analyze partial neighborhoods of the three problems. For each case we extend and generalize the previous results and provide alternative approaches that simplify the mathematical developments. We conclude the paper in Section 5.

## 2. Elementary landscapes

For a fitness landscape $(X, N, f)$, the neighborhood operator can be represented by a $|X| \times |X|$ *adjacency matrix*

$$\mathbf{A}_{xy} = \begin{cases} 1 & \text{if } y \in N(x); \\ 0 & \text{otherwise.} \end{cases}$$