# Non-parametric Estimation of Properties of Combinatorial Landscapes

Anton Eremeev[1] and Colin R. Reeves[2]

[1] Sobolev Institute of Mathematics (Omsk Branch)
Discrete Optimization Laboratory
Omsk, Russia
[2] School of Mathematical and Information Sciences
Coventry University
Coventry, UK

**Abstract.** Earlier papers [1,2] introduced some statistical estimation methods for measuring certain properties of landscapes induced by heuristic search methods: in particular, the number of optima. In this paper we extend this approach to non-parametric methods which allow us to relax a critical assumption of the earlier approach.
Two techniques are described—the jackknife and the bootstrap—based on statistical ideas of resampling, and the results of some empirical studies are presented and analysed.

## 1 Introduction

Many heuristic search methods for combinatorial optimization problems (COPs) are based on neighbourhood search or its modern developments (simulated annealing, tabu search etc.). Such methods quickly converge to a local optimum, and the search has to begin afresh—either by starting from a new point, or by using some 'meta'-technique to circumvent the current impasse.

The idea of neighbourhood search (NS) is important, because the idea of a local optimum, and associated properties such as its basin of attraction, can be given a precise meaning. One of the properties that make a problem instance difficult is the number of optima induced by the combination of the fitness function and the neighbourhood operator. However, it is possible to envisage an instance with many local optima whose basins of attraction are minimal in size, while the global optima have very large basins. So the distribution of basin sizes is clearly important too.

In earlier papers [1,2] we showed how, on the assumption of equally-sized basins—an *isotropic* landscape—it was possible to estimate the number of optima using data obtained from random restarts of a NS procedure. This paper will extend the earlier work to deal with the case of non-isotropic landscapes.

### 1.1 Terminology

In what follows, we are concerned with some finite *search space*, denoted by $\mathcal{X}$, with the objective of optimizing some real-valued function $f : \mathcal{X} \mapsto \mathbb{R}$. Our NS

S. Cagnoni et al. (Eds.): EvoWorkshops 2002, LNCS 2279, pp. 31–40, 2002.

procedure uses a neighbourhood function

$$N : \mathcal{X} \mapsto 2^{\mathcal{X}}$$

to generate a sequence of points $x_0, x_1, \ldots, x_n$, terminating at a local optimum. The strategy followed is assumed to be a 'best improving' one[1], i.e. the search algorithm finds $x_{i+1}$ such that

$$x_{i+1} = \arg \max_{y \in N(x_i)} f(y) \quad \text{and} \quad f(x_{i+1}) > f(x_i),$$

stopping when the second condition cannot be met. We assume that the procedure is used $r$ times, each commencing from a (different) random initial solution. We shall refer to the number of optima in a landscape as $\nu$, and denote by $k$ the number of *distinct* optima seen. The search can be thought of as a function

$$\mu : \mathcal{X} \mapsto \mathcal{X}$$

where if $x$ is the initial point, $\mu(x)$ is the local optimum that it reaches. Each optimum $x_1^*, \ldots, x_\nu^*$ then has a basin of attraction whose normalized size is

$$p_i = \frac{|\{x : \mu(x) = x_i^*\}|}{|\mathcal{X}|}$$

(where the expression $|\cdot|$ means, as usual, the cardinality of a set). As is standard in statistical literature, estimators of a parameter will be denoted by a ˆ symbol; maximum likelihood estimators will be denoted by the superscript $^{ML}$.

## 2 Parametric Estimation

As explained in [1], the distribution of the random variable $K$ for the isotropic case is

$$P[K = k] = \frac{\nu!}{(\nu - k)!} \frac{S(r,k)}{\nu^r}, \quad 1 \leq k \leq \min(r, \nu),$$

where $S(r,k)$ is the Stirling number of the second kind [3]. This distribution is commonly used in population statistics in ecology, where it is associated with a procedure known as the *Schnabel census*. From this, given observations $(k, r)$, it is possible to use the statistical principle of maximum likelihood [4] to obtain an estimate $\hat{\nu}^{ML}(k, r)$. (Computationally, several standard numerical methods can be used—see [1] for details.) Experiments reported in [1,2] showed that the estimates based on this approach were good for isotropic landscapes, but were quite sharply biased (negatively) when basin sizes are not identical.

Improvements could be made by assuming a basin size distribution for $\{p_i\}$ and using the *frequencies* of observing local optima to fit the parameters of the assumed distribution. However, in [2], tractable distributions (e.g. gamma) appear not to be good models for observed data, while more realistic distributions (e.g. lognormal) are very difficult to use. (On the other hand, Kallel and Garnier [5] have recently reported greater success using the gamma distribution.)

[1] Using (say) a 'first improving' strategy does not change the number of optima, although it may affect the effective basin sizes.