On Confidence Intervals for the Number of Local Optima

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

[1] Omsk Branch of Sobolev Institute of Mathematics,
13 Pevtsov str. 644099, Omsk, Russia.
eremeev@iitam.omsk.net.ru

[2] Coventry University, School of Mathematical & Information Sciences,
Priory Street. Coventry CV1 5FB, UK
C.Reeves@coventry.ac.uk

Abstract. The number of local optima is an important indicator of optimization problem difficulty for local search algorithms. Here we will discuss some methods of finding the confidence intervals for this parameter in problems where the large cardinality of the search space does not allow exhaustive investigation of solutions. First results are reported that were obtained by using these methods for NK landscapes, and for the low autocorrelation binary sequence and vertex cover problems.

1 Introduction

Many heuristic search methods for combinatorial optimization problems (COPs) are based on local search. Typically, search is focused on a *neighbourhood* of the current point in the search space, and various strategies can be used for deciding on whether or not to move to one of these neighbours, and if so, to which. In practice, such methods quickly converge to a local optimum, and the search must either begin again by starting from a new point, or use some 'metaheuristic' (e.g., tabu search) to guide the search into new areas. While such local search methods can be arbitrarily bad in the worst case [1], experimental evidence shows that they tend to perform rather well across a large range of COPs. However, experience suggests that the type and size of the neighbourhood can make a big difference to practical performance of the local search methods. Some recent theoretical results also suggest that the local optima are tightly connected with the attractors of the simple genetic algorithm [2].

The neighbourhood structure can be described by a graph, or by its related adjacency matrix, which enables us to give a precise meaning to the idea of a *landscape* (which results from the interaction of this structure with a particular problem instance), and to its characteristics, such as local (and global) optima, and to related properties of an optimum such as its basin of attraction. Formally, we have an optimization problem: given a function (the objective or goal function) of a set of decision variables, denoted by the vector x,

$$g : \mathcal{X} \mapsto I\!R,$$

find

S. Cagnoni et al. (Eds.): EvoWorkshops 2003, LNCS 2611, pp. 224–235, 2003.

$$\arg\max_{x \in \mathcal{X}} g.$$

We define a neighbourhood function

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

whence the induced landscape $\mathcal{L}$ is the triple $\mathcal{L} = (\mathcal{X}, g, d)$ where d denotes a distance measure $d : \mathcal{X} \times \mathcal{X} \to I\!R^+ \cup \{\infty\}$ for which is required that

$$\left.\begin{array}{l} d(x,y) \geq 0 \\ d(x,y) = 0 \Leftrightarrow x = y \\ d(x,u) \leq d(x,y) + d(y,u) \end{array}\right\} \quad \forall\, x, y, u \in \mathcal{X}.$$

Usually, the distance measure d is defined by the neighbourhood structure:

$$y \in N(x) \Leftrightarrow d(x,y) = 1.$$

A *locally optimal* vector $x^o \in \mathcal{X}$ is then one such that

$$g(x^o) \geq g(y) \quad \forall\, y \in N(x^o).$$

We shall denote the set of such optima as $\mathcal{X}^o$. (For a fuller discussion on landscapes, see [3] or [4]; for some illustrations of the way landscapes are formed, and how they depend on neighbourhood structure, see [5].)

Many things make an instance difficult, but one of them is the *number* of optima induced in the landscape $\nu = |\mathcal{X}^o|$. Besides this an estimate of ν might be useful in defining stopping criteria for local search based metaheuristics, and in experimental testing of the 'big valley' conjecture [6] (for such an estimate may show how representative is the sample of the local optima found).

A number of approaches have been proposed for estimation of ν: one may apply an assumption that the distribution of basin sizes is isotropic [7], or fit certain type of parametric distribution of the basin sizes (exponential, gamma, lognormal etc.) [8,9]. Nonparametric estimates, such as the bootstrap or the jackknife, can also be employed [9,10]. Assuming a particular type of distribution of basin sizes one can obtain the maximal likelihood estimate for ν, or a confidence interval for it. It is interesting that some of these approaches are actually based on the same methods as used by ecologists to estimate the number of animals in a population (see e.g. [11]).

It would be interesting to see how ν depends on the problem size: theoretical results show that many optimization problems have an exponential number of local optima in the worst-case (see [1] and references therein), but experimental results have usually been limited to low-dimensional problems where complete enumeration of solutions is possible. At present, information on experimental evaluations of the parameter ν is scarce even for widely used benchmarks.

In this paper we suggest several methods of statistical evaluation of this parameter in problems where the large cardinality of the search space does not allow exhaustive investigation of solutions. The focus is on the theoretical basis and experimental testing of the methods on some well-known benchmarks.