Chapter 10
Stochastic Search Methods

Christian Jacob
University of Calgary, Canada

10.1. Introduction: Exploring and Searching

Sophisticated search techniques form the backbone of modern machine learning and data analysis. Computer systems that are able to extract information from huge data sets (data mining), to recognize patterns, to do classification, or to suggest diagnoses, in short, systems that are adaptive and — to some extent — able to learn, fundamentally rely on effective and efficient search techniques. The ability of organisms to learn and adapt to signals from their environment is one of the core features of life. Technically, any adaptive system needs some kind of search operator in order to explore a feature space which describes all possible configurations of the system. Usually, one is interested in "optimal" or at least close to "optimal" configurations defined with respect to a specific application domain: the weight settings of a neural network for correct classification of some data, parameters that describe the body shape of an airplane with minimum drag, a sequence of jobs assigned to a flexible production line in a factory resulting in minimum idle time for the machine park, the configuration for a stable bridge with minimum weight or minimum cost to build and maintain, or a set of computer programs that implement a robot control task with a minimum number of commands.

These few examples show that learning means exploration of high-dimensional and multi-modal search spaces. The many dimensions make visualization as well as data analysis extremely difficult, hence, designing an appropriate search technique is a complex job. Multi-modality means that there is no single global maximum (optimum). There are many local maxima and several "interesting"

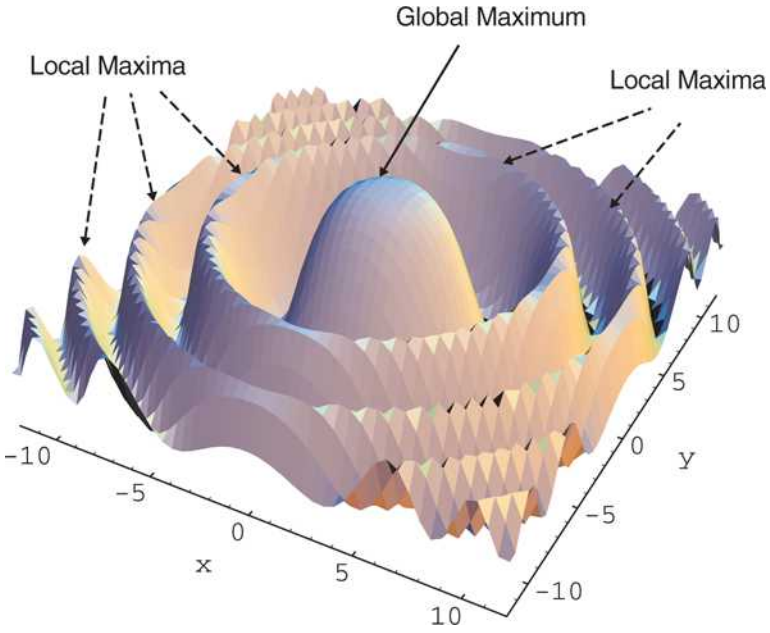

Fig. 10.1. Example of a multi-modal function.

(global) maxima, of which at least one has to be discovered by the search algorithm. Figure 10.1 shows a simple example of a multi-modal function for two dimensions.

In general, finding the global maximum (or minimum[1]) of an objective function that has many degrees of freedom and is subject to conflicting constraints is an NP-complete problem, since the objective function will tend to have many local extrema [144, 65]. Therefore, any procedure for solving hard optimization problems — these tend to be the practically relevant ones — should sample the multi-modal search space in such a way that there is a high probability of discovering near-optimal solutions. This technique should also lend itself to efficient implementation.

10.1.1 Categories of Search Techniques

Many search techniques that are more or less tailored to solve specific optimization problems have been developped: from calculus-based methods such as gradient strategies [443] in the versions of Gauss-Seidel, extrapolating gradient or simplex [409], [468, pp. 58] to enumerative techniques such as exhaustive search and dynamic programming [245], as well as stochastic methods such as

[1] Please note that searching for the maximum of a function $f(x)$ is the same as searching for the minimum of $-f(x)$.