Editorial Introduction
Special Issue on Estimation of Distribution Algorithms

Estimation of Distribution Algorithms (EDAs) are a set of Evolutionary Algorithms characterized by (i) the use of explicit probability models to recover the information of the selected individuals and to sample new solutions, and (ii) the possibility of naturally incorporating prior knowledge about the optimization problem to be solved. The EDA term was first coined by Mühlenbein and Paaß (1996); seminal papers about EDAs were written three years later (Etxeberria and Larrañaga (1999); Mühlenbein and Mahnig (1999); Pelikan et al. (1999)). Since then, there has been a growing interest in EDAs, which now constitute an established discipline in the field of Evolutionary Computation (EC). Evidence of its establishment is the great number of papers on EDAs published in the main EC conferences and in EC-related journals, as well as the tutorials given in the PPSN, CEC and GECCO conferences, and the edited book by Larrañaga and Lozano (2002).

The birth of EDAs has helped to put forward the relationship between EC and related disciplines: probabilistic graphical models, estimation of distributions, statistical physics, information theory, etc. The cross-fertilization between EC and these related communities will bring together new and exciting research results in the future.

This special issue brings together a set of papers that are representative of current research in EDAs. The six papers accepted could be divided into groups of two. The first two deal with the development of new algorithms and the relationship between EDAs and other algorithms coming from different disciplines. The next two papers are about the use of new probability models in EDAs to efficiently solve particular problems. Finally, the last two papers are theoretical analyses of EDAs.

Mühlenbein and Höns relate EDAs to algorithms previously developed in statistics, artificial intelligence, and statistical physics. The authors view all of these algorithms as minimizers of the Kullback-Leibler divergence between an unknown distribution $p(\mathbf{x})$ and a class of distributions. After that, the authors concentrated on the problems that must be faced in order to learn appropriate distributions for three particular algorithms.

Gallagher and Frean deal with continuous EDAs. The authors develop a new algorithm based on a stochastic gradient descent on the Kullback-Leibler divergence between a probability density and a model of the objective function. They associate the designed algorithm with a continuous version of the population-based incremental learning algorithm and the generalized mean shift clustering framework.

The article by Peña et al. suggests that EDAs be used in globally multimodal problems. These problems characteristically present several global optima. As these kinds of problems are difficult for most common EDAs, the authors propose a new probabilistic graphical model which consists of 'multinets'. This kind of Bayesian network efficiently and elegantly solves the multimodal problem.

Santana describes a new probability model to be used in EDAs. After analyzing the limitations of the most common probability models used in EDAs, the author proposes a new probability model based on what in statistical physics is known as the Kikuchi approximation. The author studies problems where this kind of model is superior to the ones previously used.

Shapiro analyzes the drift phenomenon in EDAs. The author deals with this effect when the estimation of the parameters of the probability model in the EDA is not carried out by means of a soft estimator. Some ideas to avoid this phenomenon are related to the scaling of the parameters of the algorithm with the system size in a strongly problem-dependent way. The author exemplifies the result using the size of the population as a parameter for several common algorithms and objective functions.

Gao and Culberson present a paper which investigates the complexity issues related to the representation of sampling distributions in EDAs. Particularly, they pay attention to additive decomposable functions and the graphs produced by these functions. Since the probability models of some EDAs are based on these graphs, they analyze the complexity of these kinds of graphs. Their results provide insight into how to design efficient EDAs and what the limitations of the algorithms are.

P. Larrañaga and J. A. Lozano, Guest Editors
December 2. 2004