Probabilistic Algorithms*

Jason Brownlee
jasonb@CleverAlgorithms.com
The Clever Algorithms Project
http://www.CleverAlgorithms.com

Abstract

The Clever Algorithms project aims to describe a large number of Artificial Intelligence algorithms in a complete, consistent, and centralized manner, to improve their general accessibility. The project makes use of a standardized algorithm description template that uses well-defined topics that motivate the collection of specific and useful information about each algorithm described. A collection of algorithms for the project have been described, all of which are classified as Probabilistic Algorithms under the adopted algorithm taxonomy. This report provides a point of reflection on the preparation of these algorithms.

Keywords: `Clever, Algorithms, Project, Probabilistic, Findings`

1 Introduction

The Clever Algorithms project aims to describe a large number of algorithms from the fields of Computational Intelligence, Biologically Inspired Computation, and Metaheuristics in a complete, consistent and centralized manner [2]. The project requires all algorithms to be described using a standardized template that includes a fixed number of sections, each of which is motivated by the presentation of specific information about the technique [7]. This report provides an overview of the Probabilistic Algorithms in the Clever Algorithms project. Section 2 provides background information and reviews common themes for the general class of algorithm and summarizes those probabilistic algorithms that have been described for the Clever Algorithms Project.

2 Probabilistic Algorithms

2.1 Background

The algorithms to be described in the Clever Algorithms project are drawn from a diverse set of subfields of Artificial Intelligence, such as Computational Intelligence, Biologically Inspired Computation, and Metaheuristics. The majority of the algorithms selected for description in the project are optimization algorithms [5]. The recently completed algorithms that have been described for the Clever Algorithms project are referred to as Probabilistic Algorithms. They

---

are differentiated from the remainder of the algorithms described in the project that have been designated a taxonomy including swarm, stochastic, immune, physical, neural, and evolutionary algorithms [5].

2.1.1 Probabilistic Models

Probabilistic Algorithms are those algorithms that model a problem or search a problem space using an probabilistic model of candidate solutions. Many Metaheuristics and Computational Intelligence algorithms may be considered probabilistic, although the difference with these so-called probabilistic algorithms is the explicit (rather than implicit) use of the tools of probability in problem solving. The vast majority of such algorithms are referred to as Estimation of Distribution Algorithms.

2.1.2 Estimation of Distribution Algorithms

Estimation of Distribution Algorithms (EDA) also called Probabilistic Model-Building Genetic Algorithms (PMBGA) are an extension of the field of Evolutionary Computation that model a population of candidate solutions as a probabilistic model. The generally involve iterations that alternate between creating candidate solutions in the problem space from a probabilistic model, and reducing a collection of generated candidate solutions into a probabilistic model.

The model at the heart of an EDA typically provides the probabilistic expectation of a component or component configuration comprising part of an optimal solution. This estimation is typically based on the observed frequency of use of the component in better than average candidate solutions. The probabilistic model is used to generate candidate solutions in the problem space, typically in a component-wise or step-wise manner using a domain specific construction method to ensure validity.

Pelikan, et al. provide a comprehensive summary of the field of probabilistic optimization algorithms, summarizing the core approaches and their differences [18]. The edited volume by Pelikan, Sastry, and Cantu-Paz provides a collection of studies on the popular Estimation of Distribution algorithms as well as methodology for designing algorithms and application demonstration studies [21]. An edited volume on studies of EDAs by Larraaga and Lozano [12] and the follow-up volume by Lozano et al. [13] provide an applied foundation for the field.

2.2 Described Algorithms

This section lists the Probabilistic Algorithms currently described for inclusion in the Clever Algorithms project. It is proposed that these algorithms will collectively comprise a chapter on 'Probabilistic Algorithms' in the Clever Algorithms book.

1. **Population-Based Incremental Learning**: [6]

2. **Univariate Marginal Distribution Algorithm**: [8]

3. **Compact Genetic Algorithm**: [3]

4. **Bayesian Optimization Algorithm**: [1]

5. **Cross-Entropy Method**: [4]

3 Extensions

There are other algorithms and classes of algorithm that were not described from the field of Probabilistic Intelligence. Some areas that may be considered for algorithm description in follow up works include:

- **Extensions to UMDA**: Extensions to the Univariate Marginal Distribution Algorithm such as the Bivariate Marginal Distribution Algorithm (BMDA) [19, 20] and the Factorized Distribution Algorithm (FDA) [14].

- **Extensions to cGA**: Extensions to the Compact Genetic Algorithm such as the Extended Compact Genetic Algorithm (ECGA) [10, 11].

- **Extensions to BOA**: Extensions to the Bayesian Optimization Algorithm such as the Hierarchal Bayesian Optimization Algorithm (hBOA) [17, 16] and the Incremental Bayesian Optimization Algorithm (iBOA) [22].

- **Bayesian Network Algorithms**: Other Bayesian network algorithms such as The Estimation of Bayesian Network Algorithm [9], and the Learning Factorized Distribution Algorithm (LFDA) [15].

- **PIPE**: The Probabilistic Incremental Program Evolution that uses EDA methods for constructing programs [24].

- **SHCLVND**: The Stochastic Hill-Climbing with Learning by Vectors of Normal Distributions algorithm [23].

4   Conclusions

This report provided a point of reflection for the batch of Probabilistic Algorithm descriptions prepared for the Clever Algorithms project. All described algorithms were assigned to the 'Probabilistic Algorithms' kingdom in the chosen algorithm taxonomy. This report highlighted the commonality for all described Probabilistic Algorithms and provided a definition suitable for use in the proposed book and website.

References

[1] Jason Brownlee. Bayesian optimization algorithm. Technical Report CA-TR-20101123-1, The Clever Algorithms Project http://www.CleverAlgorithms.com, November 2010.

[2] Jason Brownlee. The clever algorithms project: Overview. Technical Report CA-TR-20100105-1, The Clever Algorithms Project http://www.CleverAlgorithms.com, January 2010.

[3] Jason Brownlee. Compact genetic algorithm. Technical Report CA-TR-20101118c-1, The Clever Algorithms Project http://www.CleverAlgorithms.com, November 2010.

[4] Jason Brownlee. Cross-entropy method. Technical Report CA-TR-20101122a-1, The Clever Algorithms Project http://www.CleverAlgorithms.com, November 2010.

[5] Jason Brownlee. Data-driven optimization algorithm selection. Technical Report CA-TR-20100112-1, The Clever Algorithms Project http://www.CleverAlgorithms.com, January 2010.

[6] Jason Brownlee. Population-based incremental learning. Technical Report CA-TR-20101118a-1, The Clever Algorithms Project http://www.CleverAlgorithms.com, November 2010.

[7] Jason Brownlee. A template for standardized algorithm descriptions. Technical Report CA-TR-20100107-1, The Clever Algorithms Project http://www.CleverAlgorithms.com, January 2010.

[8] Jason Brownlee. Univariate marginal distribution algorithms. Technical Report CA-TR-20101118b-1, The Clever Algorithms Project http://www.CleverAlgorithms.com, November 2010.

[9] R. Etxeberria and P. Larraaga. Global optimization using bayesian networks. In *Proceedings of the Second Symposium on Artificial Intelligence (CIMAF-99)*, pages 151–173, 1999.

[10] Georges Harik. Linkage learning via probabilistic modeling in the extended compact genetic algorithm (ecga). Technical Report 99010, Illinois Genetic Algorithms Laboratory, Department of General Engineering, University of Illinois, 1999.

[11] Georges R. Harik, Fernando G. Lobo, and Kumara Sastry. *Scalable Optimization via Probabilistic Modeling*, chapter Linkage Learning via Probabilistic Modeling in the Extended Compact Genetic Algorithm (ECGA), pages 39–61. Springer, 2006.

[12] Pedro Larraaga and Jos A. Lozano, editors. *Estimation of distribution algorithms: A new tool for evolutionary computation.* Springer, 2002.

[13] Jos A. Lozano, Pedro Larraaga, Iaki Inza, and Endika Bengoetxea, editors. *Towards a new evolutionary computation. Advances in estimation of distribution algorithms.* Springer, 2006.

[14] Heinz Mühlenbein, Thilo Mahnig, and Alberto Ochoa Rodriguez. Schemata, distributions and graphical models in evolutionary optimization. *Journal of Heuristics*, 5(2):215–247, 1999.

[15] H. Mhlenbein and Th. Mahnig. Fda–a scalable evolutionary algorithm for the optimization of additively decomposed discrete functions. *Evolutionary Compilation*, 7(4):353–376, 1999.

[16] M. Pelikan and D. E. & Goldberg. Escaping hierarchical traps with competent genetic algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, number 511–518, 2001.

[17] Martin Pelikan and David E. Goldberg. Hierarchical problem solving and the bayesian optimization algorithms. In *Genetic and Evolutionary Computation Conference 2000 (GECCO-2000)*, pages 267–274, 2000.

[18] Martin Pelikan, David E. Goldberg, and Fernando G. Lobo. A survey of optimization by building and using probabilistic models. *Computational Optimization and Applications*, 21:5–20, 2002.

[19] Martin Pelikan and Heinz Muehlenbein. Marginal distributions in evolutionary algorithms. In *Proceedings of the International Conference on Genetic Algorithms Mendel*, 1998.

[20] Martin Pelikan and Hans Mühlenbein. *Advances in Soft Computing: Engineering Design and Manufacturing*, chapter The Bivariate Marginal Distribution Algorithm, pages 521–535. Springer, 1999.

[21] Martin Pelikan, Kumara Sastry, and Erick Cantu-Paz, editors. *Scalable Optimization via Probabilistic Modeling: From Algorithms to Applications.* Springer, 2006.

[22] Martin Pelikan, Kumara Sastry, and David E. Goldberg. iboa: The incremental bayesian optimization algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2008)*, pages 455–462, 2008.

[23] S. Rudlof and M. Koppen. Stochastic hill climbing with learning by vectors of normal distributions. In *First On-line Workshop on Soft Computing*, Nagoya, Japan, 1996.

[24] Rafa Saustowicz and Jrgen Schmidhuber. Probabilistic incremental program evolution: Stochastic search through program space. In *Proceedings of the 9th European Conference on Machine Learning Prague*, pages 213–220, 1997.