Theory of Evolutionary Algorithms: A Birds Eye View

A. E. Eiben [a] and G. Rudolph [b]

[a] *Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands*

[b] *Department of Computer Science, Dortmund University, Germany*

In this paper we consider the most important questions, research topics and technical tools used in various branches of evolutionary algorithms. The road map we give is to facilitate the readers' orientation in evolutionary computation theory. In the meanwhile, this survey provides key references for further study and evidence that the field of evolutionary computation is maturing rapidly, having many important results and even more interesting challenges.

Key words: Evolutionary algorithms, convergence, running time, dynamical behavior, schemata, operator analysis

1 Introduction

The term *evolutionary algorithm* (EA) stands for a family of stochastic problem solvers based on principles that can be found in biological evolution. Within this paradigm, achieving a solution to a given problem is seen as a survival task: possible solutions compete with each other for survival (and the right to reproduce), and this competition is the driving force behind the progress that supposedly leads to a(n optimal) solution. This idea has appeared several times independently over the last four decades, but the early attempts from the fifties and sixties did not receive much follow-up [1]. Development in the seventies and eighties was more coherent, but it took place along three rather independent lines of research. This led to three streams that are traditionally called Genetic Algorithms (GAs), Evolution Strategies (ES), and Evolutionary Programming (EP) [2–9]. The term evolutionary algorithm was proposed in 1990, meant as a superclass, containing all aforementioned variants and also all other techniques based on the evolutionary perception on problem solving. Since the early nineties several EAs have been proposed, Genetic Programming (GP) being probably the most influential new stream [10,11]. The borders between the different streams are loosening up in the last years, while each style having a number of particular features [12]. The emerging field of studying and applying evolutionary algorithms is called *evolutionary computation* (EC) having its specific research goals and aims; the most comprehensive collection of knowledge on the subject is the *Handbook of Evolutionary Computation* [13].

It might be clear from this brief summary that the theory of EAs can be given a double interpretation. On the one hand, it concerns the theory of the general evolutionary mechanism, underlying all representatives of the EA family. Several convergence results based on Markov chains concerning general search procedures with a population of candidate solutions, undergoing reproduction and selection belong to this category (even if they are published under the name "GAs"). On the other hand, it refers to theoretical studies of particular issues that arise within a specific style of EAs. A well-known example is the schema theory of genetic algorithms.

There exist several good overviews of EC related theory, containing extensive bibliographies [14,15]. The main goal of this paper is not the merged reproduction of such overviews, rather we are to provide a road map of different areas of interest in EC, where theoretical activities are taking place or are likely to emerge.

2 Theoretical Questions

What are the typical theoretical questions in EC? Like in any problem solving paradigm, the main issue is whether the algorithm reaches a(n optimal) solution. Obviously, no unconditional 'yes' can be expected, so the question is mostly reformulated as "under which assumptions can it be guaranteed that the algorithm reaches a(n optimal) solution". Immediately related to this question is the issue of the type of guarantee. In particular, the stochastic nature of EAs prevents crisp guarantees, turning the question into a probabilistic one. Technically speaking, almost sure convergence, convergence in probability, or convergence in mean are some options.

From a purely theoretical point of view, guaranteeing, for instance, convergence with probability 1 is satisfactory. Practically, however, the speed of convergence is just as important. The number of expected search steps needed to reach a(n optimal) solution is an important implementation independent measure of algorithm efficiency. In EAs, each new candidate solution is generated by mutating and/or recombining old solutions, and each newborn solution is immediately evaluated, i.e., its fitness value is calculated. Therefore, the number of fitness evaluations is the most commonly used measure to assess efficiency.

Note that in the foregoing we tacitly assumed that an EA is applied to an optimization task. Many other types of tasks in, for instance, machine learning, search, and constraint satisfaction, can be seen as, or transformed to, an optimization task. Nevertheless, envisioning EAs as optimizers is too narrow of a view [16]. From a broader perspective, EAs are adaptive systems having a 'basic instinct' to increase the average and maximum fitness of a population, thus optimizing, but are not optimizers in the strict sense. From this perspective, population dynamics is a typical issue for theoretical investigations. For instance, the development of the populations gene distribution over time is an important issue. In GAs, where typically bit-strings of fixed length are evolved, it frequently occurs that the population converges [1] to a relatively good bit-string before having approached the actual solution sufficiently. Analysis of such premature convergence is essential in the genetic algorithm field.

Additionally to the question whether, and if yes, how fast and by what kind of population dynamics a solution can be reached, there are important issues regarding the means to desired goals. An answer to this question in ES is self-adaptation, meaning that the algorithm adjusts itself (its own technical parameters, called strategy parameters) to the problem while running on the same problem. Practice indicates that self-adaptation is indeed a powerful tool, but only little theoretical work has been devoted to analyzing this phenomenon. In GAs the emphasis traditionally lies on the search operators: mutation and recombination (crossover). The notion of schemata, later generalized to formae, and the effects of search operators in preserving, respectively destroying schemata is one of the key issues in GA theory.

There are of course further theoretical questions (especially in future) but it is certainly feasible to say that the limit behavior, running time, and dynamical behavior of evolutionary algorithms are the key topics of evolutionary computation theory in its current stage.

3 Tools and Methods

The above overview of theoretical questions showed a variety of issues. Accordingly, the technical/theoretical tools that are used, or can be used, for answering the arising questions are also diverse. Without claiming to be complete, the following methods are relevant.

[1] The term 'convergence' in GAs mostly denotes the phenomenon that the population approaches a state where it consists of multiple copies of the same bit-string. This differs from the traditional use of the word, standing for the approximation of a solution.

3.1 Schema Theory

The so-called schema theory represents an early attempt to explain the behavior of a specific evolutionary algorithm named the simple genetic algorithm [3]. First published in 1975, this theory was considered fundamental to the understanding of GAs until the early 1990s. The reasons for this change of opinion were as follows: First, schema theory cannot explain the dynamical or limit behavior of EAs. Second, it is implicitly assumed that the problem is separable to some extent. The ignorance of this assumption has led to the "building block hypothesis" which allegedly explains the working mechanism of GAs. Alternative explanations do exist [17]. Third, the advent of Markov chain theory in the field of evolutionary computation.

3.2 Markov Chains Theory

Since the population of an EA only depends on the state of the previous population in a probabilistic manner, it is clear that Markov chains are appropriate to model and analyze evolutionary algorithms. First theoretical results, basing on qualitative models, concerning the limit behavior of EAs were available in 1991 [18]. About the same time there appeared the first papers presenting the exact transition matrices of the Markov chains associated with certain evolutionary algorithms [19,20]. Although the entire information about the evolutionary process is contained in these transition matrices, the degree of aggregation is too high to allow a simple derivation of detailed answers to particular questions (like the expected time of visiting the optimum for the first time). As a consequence, only simple versions of evolutionary algorithms have been successfully examined in this manner by now (see [15] for a summary of the results).

3.3 Dimensional Analysis

The observation that the exact Markov model is isomorphic to the associated EA but offers only little chances to extract important aspects has led to the idea of approaching EAs via dimensional analysis [21,22]. This methodology is borrowed from engineering sciences [23]. Dimensional analysis tries to identify the important dimensions or key features of a complex system and establishes a functional relationship between them. When applied to evolutionary algorithms, isolated measures for iterated selection, crossover, and mutation operators (like takeover time, mixing time and others) are put into some functional relationship which choice is validated (or not) by simulations. Needless to say, these functional relationships are a result of "good guessing." But these descriptive models may give some clues for a more detailed theoretical study—an avenue that has apparently not been entered yet.

3.4 Order Statistics

The theory of order statistics [24,25] has proved useful in determining the convergence rates of ESs for convex fitness functions [26–30]. Moreover, if the population size is infinitely large there is a close theoretical relationship to the theory of quantitative genetics [30].

3.5 Quantitative Genetics

At a first glance, it seems obvious that an analysis of biologically inspired dynamical systems should exploit the results developed in theoretical biology. The problem, however, is that the theoretical questions raised in evolutionary computation usually differ from those raised in theoretical genetics. An exception was detected by Mühlenbein & Schlierkamp-Voosen [31], who presented a specific evolutionary algorithm that can be analyzed via a theory originally developed for quantitative genetics [32,33]. Although this approach is limited to additively separable fitness functions and infinitely large populations, it contributes a piece to the mosaic of evolutionary computation theory that is under constant development.

3.6 Orthogonal Functions Analysis

Orthogonal functions like Fourier, Walsh, and Haar functions [34] have been used as a tool for constructing fitness functions that are either hard or easy for a specific evolutionary algorithm [35]. Occasionally, Walsh transforms played an important role in the analysis of evolutionary algorithms that were modeled by quadratical dynamical systems.

3.7 Quadratical Dynamical Systems

The quadratical dynamical systems (QDS) model has been classically used to model various natural phenomena in physics and biology [36]. As shown in [37] and subsequent papers, the simple genetic algorithm can be cast into a QDS, provided the population size is infinitely large. Since the simulation of a QDS is PSPACE-complete [36], this approach does not lead to an efficient method of analysis. As a consequence, most work in this field is devoted to the determination of the systems' eigenvalues and their stability. Moreover, it can be shown [36] that, in general, the predictions of the QDS approach are only sufficiently accurate for extremely large populations.

3.8 Statistical Physics

Physicists have developed various tools to cope with stochastic systems they encounter in statistical physics. Not surprisingly, there is some work of casting biological and EA models into their theoretical framework (see e.g. [38,39]).

4 Concluding Remarks

This paper has provided a road map to evolutionary computation theory. Rather than producing an immense list of references, we have outlined the most important questions, research topics and technical tools used in various branches of EAs. Hereby we hope to facilitate the readers' orientation in a field that for a long time had the reputation of "childish games" among theoretical computer scientists. We are convinced that this survey and those works in the corresponding bibliography sufficiently demonstrate that evolutionary computation does have a theory with many important results and even more interesting challenges.

References

[1] D. B. Fogel. *Evolutionary Computation: The Fossil Record.* IEEE Press, New York, 1998.

[2] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison Wesley, Reading (MA), 1989.

[3] J. H. Holland. *Adaptation in Natural and Artificial Systems.* MIT Press, Cambridge (MA), 2nd edition, 1992.

[4] I. Rechenberg. *Evolutionsstrategie '94.* Frommann-Holzboog Verlag, Stuttgart, 1994.

[5] D. B. Fogel. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence.* IEEE Press, New York, 1995.

[6] H.-P. Schwefel. *Evolution and Optimum Seeking.* Wiley, New York, 1995.

[7] T. Bäck. *Evolutionary Algorithms in Theory and Practice.* Oxford University Press, New York, 1996.

[8] Z. Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs.* Springer, Berlin and Heidelberg, 3rd edition, 1996.

[9] M. Mitchell. *An Introduction to Genetic Algorithms.* MIT Press, Cambridge (MA), 1996.

[10] J. R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection.* MIT Press, Cambridge (MA), 1992.

[11] W. Banzhaf, P. Nordin, R. E. Keller, and F. D. Francone. *Genetic Programming: An Introduction.* Morgan Kaufmann, San Francisco (CA), 1998.

[12] T. Bäck and H.-P. Schwefel. An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1(1):1–23, 1993.

[13] T. Bäck, D. Fogel, and Z. Michalewicz, editors. *Handbook of Evolutionary Computation.* Institute of Physics Publishing Ltd and Oxford University Press, Bristol and New York, 1997.

[14] T. Bäck, J.N. Kok, J.M. de Graaf, and W.A. Kosters. Theory of genetic algorithms. *Bulletin of the EATCS*, 63:161–192, October 1997.

[15] G. Rudolph. Finite markov chain results in evolutionary computation: A tour d'horizon. *Fundamenta Informaticae*, 35, 1998. to appear.

[16] K. A. De Jong. Are genetic algorithms function optimizers? In R. Männer and B. Manderick, editors, *Parallel Problem Solving from Nature, 2*, pages 3–13. North Holland, Amsterdam, 1992.

[17] H.-G. Beyer. An alternative explanation for the manner in which genetic algorithms operate. *BioSystems*, 41(1):1–15, 1997.

[18] A. E. Eiben, E. H. L. Aarts, and K. M. Van Hee. Global convergence of genetic algorithms: A markov chain analysis. In H.-P. Schwefel and R. Männer, editors, *Parallel Problem Solving from Nature*, pages 4–12. Springer, Berlin and Heidelberg, 1991.

[19] T. E. Davis. *Toward an extrapolation of the simulated annealing convergence theory onto the simple genetic algorithm.* PhD thesis, University of Florida, Gainesville, 1991.

[20] A. E. Nix and M. D. Vose. Modeling genetic algorithms with Markov chains. *Annals of Mathematics and Artificial Intelligence*, 5:79–88, 1992.

[21] D. E. Goldberg, K. Deb, and D. Thierens. Toward a better understanding of mixing in genetic algorithms. *Journal of the Society for Instrumentation and Control Engineers (SICE)*, 32(1):10–16, 1993.

[22] D. Thierens. Dimensional analysis of allele-wise mixing revisited. In H.-M. Voigt, W. Ebeling, I. Rechenberg, and H.-P. Schwefel, editors, *Parallel Problem Solving From Nature — PPSN IV*, pages 255–265. Springer, Berlin, 1996.

[23] D. C. Ipsen. *Units, Dimensions, and Dimensionless Numbers.* McGraw-Hill, New York, 1960.

[24] H. A. David. *Order Statistics.* Wiley, New York, 1970.

[25] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *A First Course in Order Statistics.* Wiley, New York, 1992.

[26] H.-P. Schwefel. *Evolutionsstrategie und numerische Optimierung.* Doctoral dissertation, Technische Universität Berlin, 1975.

[27] H.-G. Beyer. Toward a theory of evolution strategies: Some asymptotical results from the $(1 \overset{+}{,} \lambda)$–theory. *Evolutionary Computation*, 1(2):165–188, 1993.

[28] H.-G. Beyer. Toward a theory of evolution strategies: The (μ, λ)–theory. *Evolutionary Computation*, 2(4):381–407, 1994.

[29] H.-G. Beyer. Toward a theory of evolution strategies: On the benefits of sex — the $(\mu/\mu, \lambda)$ theory. *Evolutionary Computation*, 3(1):81–111, 1995.

[30] G. Rudolph. *Convergence Properties of Evolutionary Algorithms.* Kovač, Hamburg, 1997.

[31] H. Mühlenbein and D. Schlierkamp-Voosen. Predictive models for the breeder genetic algorithm I: Continuous parameter optimization. *Evolutionary Computation*, 1(1):25–49, 1993.

[32] M. G. Bulmer. *The Mathematical Theory of Quantitative Genetics.* Clarendon Press, Oxford, 1980.

[33] D. S. Falconer. *Introduction to Quantitative Genetics.* Longman Scientific & Technical, Harlow (UK), 3rd edition, 1989.

[34] K. G. Beauchamp. *Applications of Walsh and Related Functions.* Academic Press, London, 1984.

[35] A. D. Bethke. *Genetic algorithms as function optimizers.* PhD thesis, University of Michigan, 1980.

[36] S. Arora, Y. Rabani, and U. Vazirani. Simulating quadratic dynamical systems is PSPACE-complete. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing (STOC'94)*, pages 459–467. ACM Press, New York, 1994.

[37] M. D. Vose and G. E. Liepins. Punctuated equilibria in genetic search. *Complex Systems*, 5(1):31–44, 1991.

[38] A. Prügel-Bennett and J. L. Shapiro. Analysis of genetic algorithms using statistical mechanics. *Physical Review Letters*, 72(9):1305–1309, 1994.

[39] T. Aßelmeyer and W. Ebeling. Unified description of evolutionary strategies over continuous parameter spaces. *BioSystems*, 41(3):167–178, 1997.