
Preface

The last decade of the 20th century has witnessed a surge of interest in numerical, computation-intensive approaches to information processing. The lines that draw the boundaries among statistics, optimization, artificial intelligence and information processing are disappearing, and it is not uncommon to find well-founded and sophisticated mathematical approaches in application domains traditionally associated with ad-hoc programming. Heuristics has become a branch of optimization and statistics. Clustering is applied to analyze soft data and to provide fast indexing in the World Wide Web. Non-trivial matrix algebra is at the heart of the last advances in computer vision.

The breakthrough impulse was, apparently, due to the rise of the interest in artificial neural networks, after its rediscovery in the late 1980s. Disguised as ANN, numerical and statistical methods made an appearance in the information processing scene, and others followed. A key component in many intelligent computational processing is the search for an optimal value of some function. Sometimes, this function is not evident and it must be made explicit in order to formulate the problem as an optimization problem. The search often takes place in high-dimensional spaces that can be either discrete, or continuous or mixed. The shape of the high-dimensional surface that corresponds to the optimized function is usually very complex. Evolutionary algorithms are increasingly being applied to information processing applications that require any kind of optimization. They provide a systematic and intuitive framework to state the optimization problems, and an already well-established body of theory that endorses their good mathematical properties. Evolutionary algorithms have reached the status of problem-solving tools in the backpack of the engineer. However, there are still exciting new developments taking place in the academic community. The driving idea in the organization of this compilation is the emphasis in the contrast between already accepted engineering practice and ongoing explorations in the academic community.

After the seminal works of Holland, Goldberg and Schwefel, the field of evolutionary algorithms has experienced an explosion of both researchers and publications in both the application-oriented and the fundamental issues. It is

obviously difficult to present in a single book a complete and detailed picture of the field. Therefore, the point of view of this compilation is more modest. Its aim has been to provide a glimpse of the large variety of problems tackled with evolutionary approximations and of the diversity of evolutionary algorithms themselves based on some of the papers presented at the Frontiers on Evolutionary Algorithms Conference within JCIS 2002 and complemented with some papers by well-known authors in the field on topics that were not fully covered in the sessions. Following the general trend in the field, most of the papers are application-oriented. However, we have made an effort to include some that refer to fundamental issues as well as some that provide a review of the state of the art in some subfield.

As the subtitle “From industrial applications to academic speculation-” suggests, the organization of the compilation follows an axis of nearness to practical applications. We travel from industrial day-to-day problems and practice to the more speculative works. The starting collection of papers is devoted to immediate applications of clear economical value at present.

- The chapter by T. Bäck is an example of successful consulting with a toolbox of computational methods that include evolutionary algorithms addressing nontrivial industrial problems. Although the emphasis of the chapter is on Evolutionary Strategies, Bäck’s work is a bright example of a host of evolutionary solutions to everyday problems being developed at both universities and the industry RD labs.
- The general approach of Deschaine and Francone is to reverse engineer a system with Linear Genetic Programming at the machine code level. This approach provides very fast and accurate models of the process that will be subject to optimization. The optimization process itself is performed using an Evolutionary Strategy with completely deterministic parameter self-adaptation. The authors have tested this approach in a variety of academic problems. They target industrial problems, characterized by low formalization and high complexity. As a final illustration they deal with the design of an incinerator and the problem of subsurface unexploded ordnance detection.
- Nowadays there is a big industry of 3D computer modeling based on several 3D scanning methods. The rendering of these 3D structures from a cloud of scanned points requires a triangulation defined on them which may be very costly, depending on the number of scanned points, and subject to noise. Smooth and efficient approximations are therefore desired. In the chapter by Weinert et al. we find the application of Evolution Strategies to the problem of finding optimal triangulation coverings of a 3D object described by a cloud of points. The authors introduce a special encoding of the triangulation on a real-valued vector, a prerequisite for the application of Evolution Strategies. This encoding consists of the modeling of the triangulation as grid of springs and masses of varying coefficients. These coefficients and the Z coordinate of the mass points result in a problem

encoding that is closed under conventional Evolution Strategy genetic operators.

- Another practical application domain of current interest is the exploitation of hyperspectral images, especially those that arise in remote sensing. The recent advances in hyperspectral sensors and the space programs that include them in modern and future satellites imply that a large amount of data will be available in the near future. Fast, unsupervised analysis methods will be needed to provide adequate preprocessing of these data. Graña, Hernandez and d’Anjou propose an evolutionary algorithm to obtain an optimal unsupervised analysis of hyperspectral images given by a set of endmembers identified in the image. The identification is based on the notion of morphological independence, and Morphological Associative Memories serve as detectors of this condition.
- The processing of digital images is already an exploding application domain of computational methods. One of the issues of current interest, especially in the medical image domain and Magnetic Resonance Imaging (MRI), is the correction of illumination inhomogeneity (bias). Algorithms for illumination correction may be parametric or nonparametric. The latter are more computationally demanding. The formers require an appropriate modeling framework. Fernandez et al. present a gradient-driven evolution strategy for the estimation of the parameters of an illumination model given by a linear combination of Legendre polynomials. The gradient information is used in the mutation operator and seems to improve the convergence of the search, when compared with similar approaches.
- Job shop scheduling is a classical operations research problem and a recurrent problem in many industrial settings, ranging from the planning of a small workshop to the allocation of computing resources. Varela et al. propose an encoding that allows the modular decomposition of the problem. This modular decomposition is of use for the definition of new genetic operators that always produce feasible solutions. In addition, the new genetic operators benefit from the local/global structural tradeoffs of the problem, producing an implicit search of local solutions, akin to the local search in memetic approaches, but carried out in a parallel fashion.

The next batch of chapters includes works that present interesting and innovative applications of evolutionary approaches. The emphasis is on the departure of the application from the conventional optimization problems. Topics range from archeology to mobile robotics control design.

- The starting work is a fascinating application of evolution to the creation of a computational model that explains the emergence of an archaic state, the Zapotec state. The model is composed of the ontogenies evolved by specific agents dealing with the data about the sites in the Oaxaca valley, embodied in a GIS developed in the project. One of the basic results is the search for sites that may have been subject to warfare. A GA-driven

Rough Set data mining procedure was realized and its results compared with a decision tree approach.

- Phylogenetic is the search of evolutionary pathways between species based on biological data. The phylogenetic relations take the form of trees. Species are characterized by several attributes, and the construction of phylogenetic trees is somehow reminiscent of decision tree construction. Attributes usually consist of phenotypic data, although recent approaches also use genetic data. Measures of the quality of phylogenetic trees are based on the parsimony evolutive relation representation. These parsimonious objective functions have been used to guide heuristic search algorithms applied to phylogenetic tree construction. C.B. Congdon proposes an evolutionary approach to their construction. A GA is applied because only binary valued attributes are considered. A canonical tree is introduced to compare phylogenies, and the genetic mutation and crossover operators are defined accordingly. Besides the comparison of the evolutionary approach with standard algorithms, the effect of the genetic operators is studied.
- An active area in evolutionary robotics is the field of evolutionary development of robotic controllers. The need to test these controllers on the real robot to evaluate the fitness function imposes stringent constraints on the number of fitness evaluations allowable. Therefore, the convergence problems of conventional evolutionary approaches are worsened because of the poor sampling of the fitness landscape. Becerra et al. introduce Macroevolutionary Algorithms for the design of robot controllers in the domain of mobile robotics. Robot controllers take the form of artificial neural networks and the intended task is robust wall following in food or poison rewarding environment. The Macroevolutionary Algorithms partition the population into races that may evolve independently and, sometimes, become extinct. The chapter studies the setting of the colonization parameters that produce different exploitation/exploration balances.
- Parsing based on grammars is a common tool for natural language understanding. The case of sublanguages associated with a specific activity, like patent claiming, is that many features of the general language do not appear, so that simplified grammars could be designed for them. Learning of grammars from a corpus of the sublanguage is possible. Statistical learning techniques tend to produce rather complex grammars. Cyre applies evolution algorithms to the task of finding optimal natural language context-free statistical grammars. Because of the obvious difficulties in coding entire grammars as individuals, Cyre's approach is a GA whose individuals are grammar rules, endowed with bucket-brigade rewarding mechanisms as in the classical Holland classifier systems. The evolutionary algorithm uses only mutation in the form of random insertion of wildcards in selected rules. The discovery of new rules is performed by instantiating the wildcard and evaluating the resulting rules. The fitness of a rule is the number of parsed sentences it has contributed to parse. Rules with small fitness

are deleted in a culling step. Positive results are reported in this chapter with some large corpora.

- Discovering the spatial structure of proteins from their spectral images is a result that may be influential to pharmacological and biological studies. Gamalielsson and Olsson present the evaluation of protein structure models with off-lattice evolutionary algorithms. The type of evolutionary algorithms applied are evolution strategies with and without fitness sharing for premature convergence avoidance. The encoding of the protein structure is carried out by means of the angles between the residues. The main experimental argument of the paper is to study the effect of the fitness function definition. The best results are obtained with a fitness function that assumes knowledge of the actual spatial structure. This is equivalent to a supervised training problem. Unsupervised structure discovery is realized by fitness functions defined on characteristics of the composing amino acids.
- Classification is the most basic intelligent process. Among the diversity of approaches, the decision trees and related rule-based systems have enjoyed a great deal of attention, with some big success. Riquelme presents the generation of hierarchical decision rules by evolutionary approaches, comparing it to classical C4.5 decision trees over a well-known benchmark collection of problems. The hierarchical decision rules possess some nice intrinsic features, such as the parsimonious number of tests performed to classify a data pattern. They are in fact a decision list, which is constructed incrementally with the evolutionary algorithm serving as the rule selector for each addition to the list. Individuals correspond to candidate rules. Continuous-valued attributes are dealt with by the definition of intervals that quantize the attribute value range. Crossover and mutation are accordingly defined to deal with interval specifications. The fitness function computation involves the correctly classified examples, the erroneously classified examples and the coverage of the problem space by the rule.
- Evolvable hardware is an active field of research that aims at the unsupervised generation of hardware fulfilling some specifications. A fruitful area in this field is that of evolving designs of gate circuits implementing Boolean functions specified by truth tables, with great potential for application to circuit design. Hernandez Aguirre reviews the evolutionary approaches developed to handle this problem, which include classical binary GA and modifications, Ant Colony Systems and variations of the GP. Future lines of research include the design of appropriate platforms, because most present work is performed in an extrinsic mode, while the desired goal would be to perform the evolutionary search embedded in the hardware being optimized, that is, in an intrinsic way.
- System identification is the estimation of the parameters and structure of a system processing an input signal, on the basis of the observed input/output pairs. It is used in the context of designing control for processes

whose models are unknown or highly uncertain. Montiel et al. present an approach to system identification using breeder genetic algorithms, an evolutionary algorithm with features of Genetic Algorithms and Evolutionary Strategies. They present a learning strategy and experimental results on the identification of an IIR filter as the unknown system that show great promise.

We have clustered under the label of Issues in Evolution Algorithm Foundations a collection of papers that deal with some fundamental aspects of evolutionary algorithms. Fundamental properties are usually related with convergence properties and domain of application.

- The starting work is the state of the art review on multiobjective optimization by C.A. Coello et al. This tutorial paper provides a comprehensive introduction to the history and present state of the field, giving a clear picture of the avenues for future research. A special emphasis is made on the approaches followed in the literature to introduce elitism in multiobjective evolutionary algorithms (MOEA). Elitism poses specific problems in MOEA, because of the need to preserve nondominated solutions, and the subtleties that appear when trying to combine them with the new generations of solutions. In addition, a case is made for the use of constrained single objective optimization problems as benchmarks for MOEAs, and, conversely, of the power of MOEAs as constraint satisfaction optimization algorithms.
- Premature convergence is one of the key problems in GAs, trapping them in local optima. Kubalik et al. lead us through a good review of approaches to avoid premature convergence. They propose and test a GA with Limited Convergence (GALCO) to solve this convergence problem. The GALCO imposes a restriction of the difference between the frequencies of ones and zeros of each gene across the population. The replacement strategy is designed so as to ensure the preservation of the convergence restriction. No mutation is performed. Only one crossover operator is applied to each generation. Empirical evaluations over deceptive functions are provided.
- When dealing with dynamic environments, the fitness function driving the evolutionary algorithm involves probing this environment, a process that may not result in a steady response. A time-varying fitness function appears in control-related applications, namely in mobile robotics. Two questions arise: (1) how much information do we need about the time evolution of the fitness response to ensure appropriate knowledge of it to drive the evolutionary algorithm? and (2) how to synchronize the evolutionary algorithm and the environment? That is, how frequent must the sampling of the environment be to ensure its tracking by the evolutionary algorithm. Bellas et al. deal with these questions in the setting of evolution based learning of time-dependent functions by artificial neural networks. Their results provide insights to more complex and realistic situations.

The closing collection of chapters includes the more speculative approaches that induce glimpses of the open roads for future developments. Some of the approaches are loosely related to evolutionary algorithms except for the fact that they are population-based random global optimization algorithms.

- Molecular Computing deals with the realization through molecular interaction of complex computational processes. The work of Liu and Shimohara presents a molecular computing method based on the Rho family of GTPases, that can be realized in situ (on living cells). They apply it at the simulation level to a 3SAT problem, obtaining linear dependencies of the execution time and space requirements on the number of clauses and propositions. The justification lies in the fact that the computational units are the molecular pathways that grow exponentially with the number of molecules.
- Evolutionary games play a central role in the Artificial Life paradigm. Cases and Anchorena present several developments of the theory of evolutionary games that try to bridge the conceptual chasm between Dynamical Systems and Artificial Life, two rich research areas that involve divergent modeling of dynamical systems. Among the propositions in the paper is the formalization as a grammatical model of two-person evolutionary games.
- Al-kazemi and Mohan present a discrete version of the Particle Swarm Optimization (PSO) that involves the partition of the population of particles into coherent subpopulations, the definition of repulsive and attractive phases and a greedy local search. PSO is a random, population-based search algorithm, where particle motion can be assimilated to mutations in evolutionary algorithms. The results to benchmark difficult discrete and continuous functions improve over other enhancements of PSO and GA.

As indicated above, the present compilation started with the FEA'2002 workshop, embedded in the JCIS'2002 celebrated in Research Triangle Park, NC. Most of the chapters correspond to extended versions of selected papers presented at the workshop. Some chapters have been requested of the authors with the aim of obtaining a view of some specific issue not present at the workshop. We want to express our gratitude to the members of the scientific committee that volunteered their time and insights to evaluate the papers submitted to the workshop:

Jarmo Alander, Enrique Alba, Thomas Bäck, Helio J.C. Barbosa, Hilan Bensusan, Peter Bentley, Maumita Bhattacharya, Stefano Cagnoni, Erick Cantu-Paz, Yuehui Chen, Carlos A. Coello Coello, Marie Cottrell, Kelly Crawford, Alicia d'Anjou, Dipankar Dasgupta, Kalyanmoy Deb, Marco Dorigo, Gerry V. Dozier, Richard Duro, Candida Ferreira, Alex Freitas, Max Garzon, Andreas Geyer-Schulz, Christophe Giraud-Carrier, Robert Ghanea-Hercock, David Goldberg, Manuel Graña, Darko Grundler, Francisco Herrera, Vasant Honavar, Frank Hoffmann, Spyros A. Kazarlis, Tatiana Kalganova, Sami Khuri, Hod Lipson, Evelyne Lutton, John A.W. McCall, J.J. Merelo, Jae

C. Oh, Bjorn Olsson, Ian C. Parmee, Frank Paseman, Andres Perez-Urbe, Jennifer L. Pittman, Alberto Prieto, Robert G. Reynolds, Leon Rothkrantz, Marco Ruso, Francisco Sandoval, Jose Santos, Marc Schoenauer, Shigeyoshi Tsutsui, J. Luis Verdegay, Thomas Villmann, Klaus Weinert, Man Leung Wong, Xin Yao, and Yun Seog Yeun.

Finally, we acknowledge the financial support of the Ministerio de Ciencia y Tecnología of Spain through grants TIC2000-0739-C04-02, TIC2000-0376-P4-04, MAT1999-1049-C03-03, DPI2003-06972 and VEMS2003-2088-c04. The Universidad del Pais Vasco has supported us through grant UPV/EHU 00140.226-TA-6872/1999. Manuel is grateful to Caravan, Juan Perro, Amaral, Terry Pratcher and Stanislaw Lem for adding spice to our lives.

San Sebastian, Spain
January 2004

Manuel Graña
Richard Duro
Alicia d'Anjou
Paul P. Wang

Information Processing with Evolutionary Algorithms
From Industrial Applications to Academic Speculations

Grana, M.; Duro, R.J.; d'Anjou, A.; Wang, P.P. (Eds.)

2005, XVIII, 334 p. 137 illus., Hardcover

ISBN: 978-1-85233-866-4