Special issue on evolutionary computing and complex systems

Alexandru-Adrian Tantar · Emilia Tantar ·
Pascal Bouvry · Oliver Schütze · Carlos A. Coello Coello ·
Pierre Del Moral

Published online: 3 May 2013
© Springer-Verlag Berlin Heidelberg 2013

1 Introduction

The large-scale applicability and use of evolutionary computing for complex real-life systems determined a need to ensure strong analytical and theoretical grounds. The special issue, with respect to these concerns, aims at building a bridge between probability, statistics, set oriented numerics and evolutionary computing. A strong interest for identifying new common and challenging research topics is considered, addressing both theoretical and applied aspects in highly complex systems. Uncertainty, error propagation or poor analysis and design, all may lead to catastrophic failures or losses in, for example, high-sensitivity, large scale complex systems; examples

can be found by looking at financial markets, surveillance and defense systems, etc. Therefore, algorithms capable of delivering robust solutions while subject to erroneous or abnormal inputs, resilient to failures or with performance guarantees, are of a foremost importance. To this end, a collection of 13 papers is included in this issue, carefully selected from a total of 80 submissions we received. A series of high impact directions are brought into discussion like the design of efficient algorithms for highly complex systems, new algorithms, e.g. extended compact genetic algorithm, multi-objective artificial physics algorithm or service oriented algorithms, performance measures in dynamic optimization, Bayesian networks learning or high-dimensional optimization. Real-world application examples are given towards the end, with a view of mobile large-scale networks and GPGPU-based track-before-detect systems.

2 Algorithms, novel approaches and paradigms

A study on using heuristics in chemotherapy scheduling can be found in the article of Ochoa and Villasana. An optimal control/optimization view is considered, with tumor progression regarded as a dynamical system where control can be applied via drug concentrations to deliver. The model also considers drugs that are not cytotoxic and incorporates the tumor and body's natural defense system interaction, looking at different stages of the tumor cell cycle and the immune system response. A concise introduction to the biomedical background is given, followed by a description of the model used for patients (differential equations based), initial conditions, mathematical formulation and objective function. Next, an outline of the heuristics under study is given, i.e. covariance matrix

---

A.-A. Tantar (✉) · E. Tantar · P. Bouvry
Computer Science and Communications, University
of Luxembourg, 6 Rue Richard Coudenhove-Kalergi,
1359 Luxembourg city, Luxembourg
e-mail: Alexandru.Tantar@uni.lu

E. Tantar
e-mail: Emilia.Tantar@uni.lu

P. Bouvry
e-mail: Pascal.Bouvry@uni.lu

O. Schütze · C. A. Coello Coello
Depto. de Computación, CINVESTAV-IPN, Av. Instituto
Politécnico Nacional No. 2508, Col. San Pedro Zacatenco,
07300 Mexico city, DF, México
e-mail: Schuetze@cs.cinvestav.mx

C. A. Coello Coello
e-mail: CCoello@cs.cinvestav.mx

P. Del Moral
INRIA Bordeaux, Sud-Ouest, Bordeaux Mathematics Institute,
Université Bordeaux I, 351, cours de la Libération,
33405 Talence cedex, France
e-mail: Pierre.Del-Moral@inria.fr

adaptation evolution strategies, differential evolution and a hybrid particle swarm pattern search algorithm. A series of metrics like the total application time of a drug, duration of treatment, immune system health, convergence or inter-run diversity are used for comparison; no parameter tuning is made, each algorithm is used in its default configuration. The results and subsequent analysis show that the differential evolution algorithm outperforms its counterparts in both solution quality and search efficiency. At the same time, it is shown that the particle swarm optimization, while offering the least competitive results with respect to tumor size reduction, leads to a broader range of alternative solutions, with larger treatment times, and less impact on the immune system health.

A look at community detection in complex networks, nonlinear optimization perspective, is given in the work of Li et al. An overview of the community detection field is given in the introduction, followed by a concise problem description. A set of algorithms, also described in the first part of the paper, is later used to analyze the performance of an extended compact genetic algorithm, including an n -ary alphabet encoding variant; mutual information based clustering is also introduced in order to improve performance. Last, experiments on several benchmarks are presented, e.g. networks where all vertexes have the same degree, equally or exponential distribution sized communities, undirected and unweighted graphs; real-world networks are also analyzed. Among other future work directions, the authors mention dealing with more complex cases where, for example, overlapping, hierarchical or dynamic communities need to be identified.

Wang et al. introduce an artificial physics based multi-objective optimization approach. Solutions are modeled with attributes like position, mass, velocity or momentum, relying on attractive vs. repulsive rules to explore the problem space. The inner nature of the sampling allows individuals to escape local optima as to evolve towards the global optimum; a per individual rank inferred from a mass function is used for archiving; neighborhood size is, in this model, adjusted to fit the state expressed by the current population. Analysis and experiments are carried on several classical benchmarks, e.g. ZDT and DTLZ, the results being compared with the ones obtained by NSGA, SPEA or MOPSO.

3 Machine learning

The article of María A. Franco et al. offers a systematic and detailed analysis of two genetics-based machine learning algorithms, namely BioHEL and GAssist. The two approaches share a similar structure and (mainly) differ in the learning paradigm used, i.e. an iterative (separate-and-

conquer) rule learning approach, respectively a Pittsburg (classifier defined as a standalone ruleset) learning system. The article provides a comprehensive view of the different other studies already available in the literature on genetics-based machine learning, followed by a thorough presentation of both BioHEL and GAssist. A comparison of the two algorithms is made using 35 real-world datasets, suggesting that (i) GAssist is more fit to deal with small problems, i.e. with either less classes and/or less instances, and (ii) BioHEL scores better on problems with opposite characteristics though while showing a tendency to overfit the data and to generate larger rule sets. Furthermore, the C4.5, PART, IBk, Naïve Bayes and an SVM (both polynomial and Gaussian kernels) classifier are used to run additional tests. It is thus shown that BioHEL obtains results comparable to the ones of SVM and GAssist (in terms of accuracy) while requiring less computational resources.

An Artificial Bee Colony (ABC) based algorithm is described in the work of Ji et al. looking at learning Bayesian network structure from data. A very concise discussion of Bayesian networks is given, followed by an outline of the standard and the extended (chemical communication of honeybee swarms) ABC algorithm. The approach relies on a constructive procedure where, starting from an empty Bayesian (graph structure) network, new edges are accepted if and only if an improvement is observed. The fitness of each solution is evaluated based on the $k2$ metric. Experiments are run on different datasets, comparing the approach to Ant Colony Optimization and Tabu Search.

4 Complexity and performance measures

Ren and Wu propose an evolutionary algorithm for high-dimensional function optimization. The approach relies on a cooperative coevolution Artificial Bee Colony (ABC) algorithm with orthogonal experimental design. A random grouping strategy is used to break the initial problem into subcomponents, relying internally on factor analysis to guide the search. Experiments are run on several continuous benchmark functions, e.g. shifted Ackley, Rastrigin, Griewank, etc., with a setup considering 1,000 dimensions. The algorithm is compared to several cooperative coevolution paradigms, i.e. a standard ABC approach as well as differential evolution and particle swarm variants.

A discussion on the performances of optimization algorithms used in dynamic problems is given in Ben-Romdhane et al. While performance metrics made the object of extensive studies in static optimization setups, only a limited use of these metrics was considered for

dynamic problems. A description of five algorithms used inside this study is first given, i.e. random immigrants (RI), hyper-mutation (HM), local search (LS), random search (RS) and a steady state genetic algorithm (SSGA). Next, several performance measures are presented, including fitness-based classical metrics (average best-of-generation, accuracy or offline performance), as well as behavior based measures. Among this last class, stability and fitness degradation are discussed. Last, experimental results on a dynamic knapsack problem (dynamic cyclic XOR based environment) are presented. The subsequent analysis shows that fitness-based measures are in agreement for HM, LS and RS, with a low stability for LS, and with HM as the best for tracking the optimum. Furthermore, it is concluded that relying on fitness-based measures alone when judging results may lead to biased or erroneous conclusions.

A study on synchronicity and neighborhood size in particle swarm optimization (PSO) is given in Rada-Vilela et al. The effect different PSO neighborhood structures have is analyzed by considering synchronous and (random) asynchronous setups. An in-depth discussion of PSO convergence is given from an interconnecting network topology perspective, namely with ring or star graphs used for information exchange among particles. Synchronous exchanges are defined as a perfect information update mapping, i.e. with all particles having a complete view of the best solution(s) each particle explored. At the opposite end, asynchronous communication is limited in terms of what information can be accessed (partial view of neighbors or best position readings from past updates only). The random communication model extends the asynchronous approach by updating only a subset of all particles, selected w.r.t. a uniform distribution. A particle, in this setup, can be selected more than once or not at all. An overview of related work is given, followed by experiments and results on large-scale optimization functions. A discussion and conclusions on trade-offs between synchronous vs. asynchronous and neighborhood size are given.

An automatically tuned, increasing population CMA-ES is discussed in Liao et al. offering a view of the current state-of-the-art on continuous benchmark function optimization. A collection of 10-dimensional continuous SOCO and CEC'05 functions are used to determine optimal parameter sets, later analyzed vs. CMA variants like the MA-LSch-CMA and the PS-CMA-ES. Additional experiments look at how the algorithm behaves when compared to Sep-iCMA-ES-tsc or to variants trained with, for example, uni-modal or hybrid SOCO benchmark functions alone. Among other aspects, it is shown that the found parameters tend to determine a more intensive exploration of the search space during the initial phases and after a restart. Last but not least, it is argued that the hybrid SOCO

functions alone are not sufficient for constructing a robust, high-performing parameter tuning.

5 Large instances and parallel algorithms

Xu and Lei address coverage problems in large-scale camera networks with mobile nodes by relying on a constrained particle swarm algorithm. The article namely looks at maximizing the area covered by a sensor (camera) network via optimal placement (location and orientation degrees of freedom, subject to moving distance constraints). Three particle swarm algorithm variants are investigated, looking at how constraints can be handled, penalties or specific operator design, among others. All cameras are assumed to be of a similar design, with a fan shaped field of view. A region of interest, defined as (i) a two-dimensional area, (ii) via a set of points or (iii) by using multiple curves, needs to be covered. Different experiments are carried, addressing point and route coverage, different grid resolution, size and weight configurations.

A service oriented evolutionary algorithm is proposed by García-Sánchez et al. dealing with the question of integration and scalability across heterogeneous systems. After a brief introduction and state of the art on service oriented architectures, evolutionary computing frameworks and examples (basic GA, service oriented NSGA-II, loose coupling, distributed execution), implementation details are given. The framework features language independent interoperability, hence allowing to bridge heterogeneous systems, with a direct reuse of existing code or with services being added or removed at runtime. Overhead and results are discussed, a comparison with MALLBA, ECJ and Algorithm::Evolutionary being included. The source code is also provided online by the authors for the interested reader.

Segura et al. investigate on the scalability and robustness of parallel hyper-heuristics in frequency assignment problems. A multiobjectivisation perspective is adopted, implying the mapping of an initially mono-objective problem onto a multi-objective formulation; a main aim of such an approach is to modify the fitness landscape with, as a result, a lower probability to get trapped in local optima. The paper offers a general view of nowadays wireless communication networks with a short discussion of the importance and implications frequency assignment has. It is stressed that several real-world, difficult points need to be addressed, including data transmission, interference or signal quality. Among other aspects, the paper introduces a structured view of hyper heuristics, looking at automated design by iterative parameter or algorithm refinement. A hybrid optimization algorithm is proposed, relying on

parallel evolutionary computing concepts, and referred to by the authors as a dynamic-mapped island-based model. A population structured in demes is simultaneously evolved on multiple islands, connected by all to all or undirected ring topologies. One of the nodes acts as a master and runs a (scoring strategy based) hyper heuristic that collects solutions and decides what topology mapping and low level heuristic (memetic approach) to use for each island. The authors analyze the performance of the approach for two (real) network instances while using different parameter configurations. A scalability analysis is also conducted, showing that the algorithm tends to be more sensitive to migration related settings when the number of islands is increased; at the opposite end, a more robust approach is obtained when only a few islands are used.

A code reordering application for GPGPU based (low signal-to-noise ratio) track-before-detect systems is presented in the work of Mazurek. A brief outline of what tracking stands for, algorithms and processing devices is first given, including some general technical details. A local random extraction and insertion operator is proposed, aimed at optimizing assembly code via reordering. The overall approach follows an evolutionary design where solutions are associated to code; the algorithm needs to

evolve a tracking solution where the fitness quality is given by the median execution time. While offering an improvement with potential application in real time systems, the approach is mentioned to require long optimization runs due to a need to compile and run candidate code solutions. As the authors note, this aspect needs to be considered carefully, also subject to further investigation.

Acknowledgments Connecting theory and practice is an aim of foremost importance when a solid theoretical understanding and an effective solution are simultaneously needed, e.g. in security and defence applications, large-scale problems or when dealing with real-life constraints. Achieving such a desideratum implies the combined effort of bridging notions from theory and practice, and the will to acquire a sound description of the problem and the algorithm(s) under study. With respect to these considerations, we would like to end by acknowledging the authors for their extraordinary contribution and exceptional work. It is our strong belief that the time spent in providing studies at this level of detail and in running such extensive experiments will be fully valued by numerous researchers from different fields. We would also like to express our sincere gratitude to all the reviewers for their careful and untiring efforts that made possible to ensure the quality this special issue. Last but not least, we would like to thank Dr. Giovanni Acampora and the Soft Computing staff for their extraordinary support and kind guidance during the entire editing process.