Jornada de Seguimiento de Proyectos, 2005
Programa Nacional de Tecnologías Informáticas

Hybridizing Evolutionary Computation and Ant Colony Optimization. Applications to Fuzzy Rule Learning and Bioinformatics Problems. TIC2003-00877

Oscar Cordón *
Dept. of Computer Science and Artificial Intelligence. E.T.S.I. Informática
University of Granada. 18071-Granada. Spain

Abstract

The goal of this research project is to analyze the possibilities of hybridization of different metaheuristics, namely Evolutionary Computation and Ant Colony Optimization, as well as other Soft Computing techniques such as fuzzy logic and fuzzy clustering. We aim at proposing new hybrid algorithms to solve some real-world applications, such as fuzzy rule learning and several Bioinformatics problems. More specifically, we will analyze the different components existing in the isolated metaheuristics to respectively explore and exploit the search space. This way, we will be able to develop hybrid algorithms combining complete instances of the individual metaheuristics, on the one hand, and to identify a set of exploration and exploitation components (useful for the search diversification and intensification, respectively), that can be appropriately combined in order to obtain better search and optimization algorithms, on the other hand.
Keywords: Metaheuristics, Evolutionary Computation, Ant Colony Optimization, Hybridization of Metaheuristics, Multiobjective and Multimodal Evolutionary Algorithms, Fuzzy Rule-Based Systems, Fuzzy Clustering, Bioinformatics.

1 Research project goals

The goal of the project is to study the chances for the hybridization of different metaheuristics [4], specifically, Evolutionary Computation (EC) [1] and Ant Colony Optimization (ACO) [2], proposing new hybrid algorithms with a better trade-off between search space exploration and exploitation, as well as between accuracy and efficiency, to get better solutions for both classical combinatorial optimization problems and real-world applications such as fuzzy modeling and Bioinformatics.

This global goal can be divided into five specific objectives:

1. Study of the different instances of EC and ACO, as well as other metaheuristics, and identification of their exploration and exploitation components.

*Email: ocordon@decsai.ugr.es

2. Proposal of specific hybrid algorithms from the study developed in the previous objective. Taking the theoretical and practical study of the comparison of the analyzed metaheuristics as a base, we aim at proposing second generation hybrid search and optimization algorithms allowing us to improve the results obtained by the isolated algorithms in classical combinatorial optimization problems. Namely, our aim is to work in three specific research lines:

   (a) Incorporation of new evolutionary algorithm (EA) diversification components into ACO algorithms.

   (b) Integration of parallel EA components into ACO algorithms.

   (c) Integration of multiobjective and niching techniques into ACO algorithms.

3. Design of hybrid algorithms between EAs and ACO for the inductive learning of fuzzy rule-based system (FRBS) knowledge bases in real-world applications. This objective collects two different research lines:

   (a) Automatic learning of fuzzy knowledge bases by means of the EC-ACO hybrid algorithms designed.

   (b) Automatic learning of fuzzy knowledge bases by evolutionary meta-learning algorithms for the data base definition, and ACO algorithms for the rule base derivation.

4. Design of hybrid algorithms between EAs and fuzzy clustering, as well as between EAs and ACO to integrate genetic information of different kinds, obtained from different sources, for the modeling and prediction of genetic regulatory networks. This objective involves the following stages:

   (a) Design of hybrid algorithms (multiobjective and multimodal EAs, hybridized with fuzzy clustering) for pattern discovery to detect protein regulatory sites and DNA sequence promoters.

   (b) Design of inference algorithms of causal connections among groups of co-expressed and co-regulated genes between EC-ACO hybrid algorithms devoted to the genetic network derivation from DNA sequence analysis and oligonucleotide microarrays. Synthesis and description of the inferred pathways by means of FRBSs.

5. Validation of the proposed hybrid algorithms in the mentioned application areas.

The estimated schedule for the achievement of the previous goals is shown in Table 1. Notice that, each activity is associated with the corresponding described objective. Besides, its starting and end month is also reported. We should remind that the project is to be developed in a 36 month time.

2 Successful achievements of the project

The work developed till now in the research project mainly corresponds to the second, third and fourth objectives described in the previous section. The next three subsections are devoted to describe the results obtained in each of these three areas, together with the publications generated from them.

Table 1: Estimated research project schedule

Activities/Tasks	Start	End
O1. Study of existing bibliography	m1	m4
O1. Analysis of the exploration and exploitation components of EAs and ACO algorithms	m5	m8
O2.1. Hybridization of ACO algorithms and EA specific components	m9	m12
O2.2. Design of parallel ACO algorithms based on parallel EAs	m13	m16
O2.3. Integration of multiobjective and multimodal techniques into ACO	m17	m20
O3.1. Automatic learning of fuzzy knowledge bases by means of EC-ACO hybrids	m9	m22
O3.2. Automatic learning of fuzzy knowledge bases by meta-learning EC-ACO hybrids	m23	m28
O4.1. Design of multiobjective and multimodal EAs-fuzzy clustering hybrid algorithms to detect protein regulatory sites and DNA sequence promoters	m9	m22
O4.2. Design of inference algorithms of causal connections among groups of co-expressed and co-regulated genes between EC-ACO hybrid algorithms	m23	m28
O5. Validation of the proposed hybrid algorithms in the application areas considered	m29	m36

2.1 Proposal of specific hybrid algorithms beetween EAs and ACO

Three research lines have been developed within this first area, corresponding to the three subobjectives distinguished in the previous section.

2.1.1 Incorporation of new EA diversification components into ACO algorithms

We have gone on working with our proposal of the Best-Worst Ant System (BWAS), ACO algorithm which incorporates EA components to improve its performance. The relation between the three distinguished components of the algorithm has been analyzed in deep, in order to determine that they three are actually needed in order to get a proper performance. As a result of this work, two publications have been developed, corresponding to two book chapters:

1. S. Alonso, O. Cordón, I. Fernández de Viana, F. Herrera, La Metaheurítica de Optimización Basada en Colonias de Hormigas: Modelos y Nuevos Enfoques. G. Joya, M.A. Atencia, A. Ochoa, S. Allende (Eds.), Optimizacion Inteligente: Técnicas de Inteligencia Computacional para Optimización, Servicio de Publicaciones de la Universidad de Málaga (2004) pp. 261-313.
2. S. Alonso, O. Cordón, I. Fernández de Viana, F. Herrera, Integrating Evolutionary Computation Components in Ant Colony Optimization Evolutionary Algorithms: An Experimental Study. Recent Developments in Biologically Inspired Computing, L. Nunes de Castro, F.J. Von Zuben (Eds.), Idea Group Publishing (2004) pp. 148-180.

On the other hand, the incorporation of Dr. Damas to the project team, opened new research lines within this area. Dr. Damas' area of expertise is the application of metaheuristics (mainly, EAs) to a real-world problem, the registration of 3D medical images, on which he made his PhD two years ago, advised by Dr. Cordón. Since 3D image registration can be formulated both as a combinatorial optimization problem and as a numerical optimization problem, it has become a very interesting benchmark for the application of hybrid metaheuristics. More specifically, up to now in this project we have worked with EAs establishing an appropriate trade-off between intensification and diversification, such as CHC [3], as well as with hybrid techniques considering the use of local search optimizers, namely Scatter Search

[5] and Iterated Local Search [6] paradigms. Seven accepted (two national conference contributions, four international ones, and one international journal paper) and three publications currently submitted to international journals have been originated from this work:

1. O. Cordón, S. Damas, J. Santamaría, Registrado de Imágenes 3D Mediante CHC, Actas del Tercer Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB'04), Córdoba (España) (2004) pp. 91-97.
2. O. Cordón, S. Damas, J. Santamaría, R. Martí, Una Aplicación de Scatter Search al Problema del Registrado de Imágenes Médicas, Actas del Cuarto Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB'05), Granada (España) (2005). In press.
3. O. Cordón, S. Damas, J. Santamaría, A Scatter Search Algorithm for the 3D Image Registration Problem, 8th International Conference on Parallel Problem Solving from Nature (PPSN VIII), Lecture Notes in Computer Science 3242, Birmingham (UK) (2004) pp. 471-480.
4. O. Cordón, S. Damas, J. Santamaría, Solving the 3D Image Registration Problem by Scatter Search with Path-Relinking, Proceedings of INFORMS Annual Meeting, Denver (USA) (2004) pp. 113.
5. O. Cordón, S. Damas, J. Santamaría, A Scatter Search-based Optimizer for the Registration of 3D Surfaces, Proc. IEEE Congress on Evolutionary Computation (CEC), Edinburgh (UK) (2005). In press.
6. O. Cordón, S. Damas, J. Santamaría, R. Martí, 3D Inter-Subject Medical Image Registration by Scatter Search, Proceedings of the Second International Workshop on Hybrid Metaheuristics (HM2005), Lecture Notes on Computer Science 3636, Barcelona (España) (2005) pp. 90-103.
7. O. Cordón, S. Damas, Image Registration with Iterated Local Search, Journal of Heuristics (2005). In press.
8. O., Cordón, S. Damas, J. Santamaría, R. Martí, Scatter Search for the 3D Point Matching Problem in Image Registration, Submitted to INFORMS Journal on Computing (June 2005).
9. O. Cordón, S Damas, J. Santamaría, A Fast and Accurate Approach for 3D Image Registration using the Scatter Search Evolutionary Algorithm, Submitted to Pattern Recognition Letters (November 2004).
10. O. Cordón, S. Damas, J. Santamaría, Feature-based Image Registration by means of the CHC Evolutionary Algorithm, Submitted to Image Vision and Computing (June 2004).

We should remark that, the availability of a real-world optimization problem will allow us to study the performance of the EC-ACO hybrid techniques developed in the first stages of the research project in a significantly much more realistic way than when working with the classical traveling salesman problem (TSP) and quadratic assignment problem (those initially proposed to be considered in the project). This is one of the first tasks to be done in the second part of the project.

2.1.2 Integration of parallel EA components into ACO algorithms

The second part of the second publication mentioned in the previous subsection, was devoted to the proposal of different parallel ACO algorithms based on the usual parallelization strategies followed in EC. However, the results obtained were not promising at all. As a consequence, we decided to give up with this working line to focus more in the remaining project research lines.

2.1.3 Integration of multiobjective and niching techniques into ACO algorithms

We have specially focused on the design of multiobjective ACO algorithms. A review of the existing bibliography on the topic has been developed, finding a large number of existing proposals (around fifteen) of multiobjective ACO algorithms, but applied to very different problems. From them, those based on the Pareto-optimality definition have been selected as the most representative, resulting in ten different algorithms, which have been adapted to solve

the same benchmark, the bi-criteria TSP. Besides, their performance has been compared to two state-of-the-art multiobjective EAs, SPEA2 and NSGA-II. This way, a realistic comparison between the different approaches has been established for the very first time. Our conclusion is that some multiobjective ACO algorithms outperform SPEA2 and NSGA-II when no local optimizer is considered, but the latter EAs present the best performance when every algorithm is hybridized with a local search procedure. The results of the work developed comprise an international conference contribution and a paper submitted to an international journal:

1. C. García-Martínez, O. Cordón, F. Herrera, An Empirical Analysis of Multiple Objective Ant Colony Optimization Algorithms for the Bi-criteria TSP. 4th International Workshop, ANTS2004, Brussels (Bélgica). M. Dorigo y otros (Eds.), LNCS 3172 (2004) pp. 61-72.
2. C. García-Martínez, O. Cordón, F. Herrera, A Taxonomy and an Empirical Analysis of Multiple Objective Ant Colony Optimization Algorithms for the Bi-criteria TSP. Submitted to European Journal of Operational Research (August 2004).

2.2 Design of EC-ACO hybrid algorithms for the inductive learning of FRBS knowledge bases in real-world applications

The two research lines developed in this second area are described in the next two subsections.

2.2.1 Automatic learning of fuzzy knowledge bases by EC-ACO hybrid algorithms

The project team has worked on the design of FRBSs by the learning of their rule bases using the BWAS algorithm. Very promising results have been obtained both in fuzzy modeling applications (related to electrical distribution problems) and in mobile robotics. The latter application is very interesting as the interpretability of the generated knowledge base is an important requirement. This research line has concluded in four publications: a national conference contribution, an international conference one, a paper recently published in an international journal, and another one submitted to another journal:

1. M. Mucientes, J. Casillas, Generación eficiente de un controlador difuso con alta interpretabilidad para navegación de robots móviles, Actas del XII Congreso Español sobre Tecnología y Lógica Fuzzy (ESTYLF 2004), Jaén (2004) pp. 401-406.
2. M. Mucientes, J. Casillas, Obtaining a fuzzy controller with high interpretability in mobile robots navigation, Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2004), Budapest, Hungary (2004) pp. 1637-1642.
3. J. Casillas, O. Cordón, I. Fernández de Viana, F. Herrera, Learning Cooperative Linguistic Fuzzy Rrules using the Best-worst Ant System Algorithm, Intl. Journal of Intelligent Systems 20 (2005) pp. 433-452.
4. M. Mucientes, J. Casillas, Fast Learning of Fuzzy Controllers with High Interpretability in Mobile Robotics. Submitted to IEEE Transactions on Fuzzy Systems (2005).

2.2.2 Automatic learning of fuzzy knowledge bases by means of meta-learning EC-ACO hybrid algorithms

On the one hand, we have worked both on the hybrid learning of fuzzy knowledge bases by means of the said methodology (a wrapping EA is considered to learn the FRBS data base, while the BWAS is applied to derive the fuzzy rules for each possible data base definition). On the other hand, the later adjustment of those knowledge bases using EAs has also be considered by either tuning the membership functions and selecting cooperative rule sets, or

jointly performing both tasks. Three publications have been obtained in this line: a national conference, an international conference, and an international journal paper:

1. J. Casillas, O. Cordón, F. Herrera, P. Villar, Aprendizaje Híbrido de la Base de Conocimiento de un Sistema Basado en Reglas Difusas mediante Algoritmos Genéticos y Colonias de Hormigas, XII Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF2004), Jaén (2004) pp. 357-362.
2. J. Casillas, O.Cordón, F. Herrera, P. Villar, A Hybrid Learning Process for the Knowledge Base of a Fuzzy Rule-Based System, Proc. Tenth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU2004), Perugia, Italy (2004) pp. 2189-2196.
3. J. Casillas, O. Cordón, M.J. Del Jesus, F. Herrera, Genetic Tuning of Fuzzy Rule Deep Structures Preserving Interpretability and its Interaction with Fuzzy Rule Set Reduction, IEEE Transactions on Fuzzy Systems 13:1 (2005) pp. 13-29.

2.3 Design of EAs-fuzzy clustering and EC-ACO hybrid algorithms to integrate genetic information for the modeling and prediction of genetic regulatory networks

Finally, in this third area, the following two working lines are being developed.

2.3.1 Design of hybrid algorithms for pattern discovery to detect protein regulatory sites and DNA sequence promoters

To do so, we have considered the hybridization of the knowledge representation capability of fuzzy logic, used to model the uncertainty existing in the biological patterns to be uncovered, and the search ability of multiobjective and multimodal EAs, considered to extract a non dominated solution set for the problem. As a result of this work, four publications have been developed: an international conference contribution, an international journal paper, and two papers submitted to other two international journals:

1. R. Romero-Zaliz, O. Cordón, C. Rubio, I. Zwir, A Multiobjective Evolutionary Fuzzy System for Promoter Discovery in E. coli. Proceedings of the First International Workshop on Genetic Fuzzy Systems, Granada (2005) pp. 68-75.
2. V. Cotik, R. Romero-Zaliz, I. Zwir, A Hybrid Promoter Analysis Methodology for Prokaryotic Genomes, Fuzzy Sets and Systems 152 (2005) pp. 83-102.
3. I. Zwir, R. Romero-Zaliz, H. Huang, E.A Groisman, Extracting Promoter Features from Genomic Datasets: Towards an Annotation of Genome Regulatory Regions. Submitted to Genome Research (May 2005).
4. I. Zwir, H. Huang, E.A. Groisman, Analysis of Differentially-Regulated Genes within a Regulatory Network by GPS Genome Navigation. Submitted to Bioinformatics (2005).

2.3.2 Design of EC-ACO hybrid inference algorithms of causal connections among groups of co-expressed genes for genetic network derivation

As this task is scheduled for the twenty third month of the project, the work developed on it is only preliminary. Specifically, we have worked on the selection of the most complete and biologically meaningful set of genes which become differentially expressed between control and treatment temporal microarray experiments in a real-world problem studying the host response over time to systemic inflammatory insults. Different statistical approaches have been applied to retrieve differentially expressed genes over treatment and time. These gene expression profiles are the seed from which we will be able to build genetic networks, that will

be validated by means of additional data taken from databases such as Gene Ontology, as well as with clinical data. By now, we have obtained an international conference publication in this line, as well as two papers submitted to international journals:

1. C. Rubio, O. Cordón, I. Zwir, Identifying Meaningful Temporal Gene Expression Patterns in the Inflammation and the Host Response to Injury. Proc. Affymetrix User Group Meeting, Edinburgh (2004).
2. C. Rubio-Escudero, R. Romero-Záliz, O. Cordón, J.P. Cobb, I. Zwir, Identifying Gene Profiles by Reverse Problem Solving: From Grouping Gene Expressions to Combining Microarray Analysis Methods. Submitted to Bioinformatics (July, 2005).
3. R. Romero-Záliz, O. Cordón, C. Rubio-Escudero, J.P. Cobb, I. Zwir, A Multi-objective Evolutionary Conceptual Clustering Methodology for Gene Annotation from Network Databases. Submitted to Journal of Machine Learning Research (July, 2005).

3 Indices to measure the quality of the obtained results

In summary, a total number of 28 publications has been developed in this first half of the research project, distributed as follows: 4 international journal papers, 9 papers submitted to international journals, 2 book chapters, and 13 (4 national and 9 international) conference contributions. All these publications are available in the web page of the Research Group "Soft Computing and Intelligent Information Systems" (SCI2S), to which the project team members belong (`http://sci2s.ugr.es`).

On the other hand, as regards the staff under formation, the project team members are advising three different PhD dissertations in the project topics:

- The PhD dissertation of Ms. Rocío C. Romero Zaliz, advised by Drs. Cordón and Zwir, has been finished yet, and will be presented next September, 14th, in the E.T.S.I. Informática of the University of Granada. It is entitled "Reconocimiento de Perfiles de Regulación Genética mediante Algoritmos Evolutivos Multiobjetivo" and is directly related to the objective 4.1 of the project, described in Section 1.

- Ms. Cristina Rubio, whose incorporation to the project team has been recently asked, is also being advised by Drs. Cordón and Zwir in the other project objective related to Bioinformatics, namely objective 4.2. Her work is related to biologically significant gene selection and genetic networks design for the systemic inflammatory insults response problem described in Section 2.3.2.

- Finally, Mr. Jose Santamaría is developing his PhD dissertation under the advise of Drs. Cordón and Damas on the application of hybrid metaheuristics to the 3D medical image registration problem described in Section 2.1.1 (project objective 2.1). The most of the publications mentioned in that section as regards this topic correspond to his PhD work.

We should also notice that we have got an additional grant from the Spanish Ministry related to the Bioinformatics research lines developed in the current project. It is within the "Acciones Complementarias del Programa Nacional de Cooperación Internacional en Ciencia y Tecnología del Ministerio de Educación y Ciencia" Action, in the "a7) Ayudas para el establecimiento de nuevas relaciones científicas en el ámbito de las prioridades geográficas de cooperación consignadas en el Plan Nacional" modality, and is entitled "Procesos de Regulación Genética sobre Microarrays usando Extracción de Conocimiento" (Ref. BIO2004-270-E). Its

aim is to encourage the Bioinformatics research lines with some US research groups on the topic. Two research project team members, Drs. Cordón and Zwir, with the latter being the Grant Coordinator, as well as Ms. Cristina Rubio, whose incorporation to the project team has been recently asked, participate in the grant. We have started to work with the said research groups and four publications submitted to international journals have been already developed with them (see Sections 2.3.1 and 2.3.2).

Moreover, some collaborations with other research groups have arisen from the project work. The global project scope, the study of metaheuristics and their applications, is directly related to one of the Research Networks funded by the Spanish Ministry, the Spanish Network on Metaheuristic Procedures (HEUR) (Ref. TIC2002-10886E). The members of the project team are also members of the Network, which has generated different collaborations with other Spanish research groups. Namely, as regards the current project, we have worked with the Network Coordinator, Dr. Rafael Martí, from the University of Valencia. From this joint work, the following results have been obtained:

- Organization of a special session on "Scatter Search" in the Fourth Spanish Conference on Metaheuristics and Evolutionary and Bioinspired Algorithms (MAEB2005), within the First Spanish Computer Science Conference (CEDI2005), that will be held in Granada next September 2005.
- Development of three joint publications: a contribution to the previous session (national conference), an international conference contribution, and a paper submitted to an international journal, INFORMS Journal on Computing. All of them have been already mentioned in Section 2.

References

[1] T. Bäck, D.B. Fogel, Z. Michalewicz (Eds.), Handbook of Evolutionary Computation, IOP Publishing Ltd and Oxford University Press, 1997.

[2] M. Dorigo, T. Stützle, Ant Colony Optimization, MIT Press, Cambridge, MA, 2004.

[3] L.J. Eshelman, The CHC Adaptive Search Algorithm: How to Safe Search when Engaging in Non Traditional Genetic Recombination, In Foundations of Genetic Algorithms 1, G.J.E. Rawlins (Ed.), Morgan Kaufmann Publishers, San Mateo, pp. 265-283, 1991.

[4] F. Glover, G. Kochenberger (Eds.), Handbook of Metaheuristics, Kluwer Academic Publishers, 2003.

[5] M. Laguna, R. Martí, Scatter Search: Methodology and Implementations in C. Kluwer Academic Publishers, Boston, 2003.

[6] H.R. Lourenco, O.C. Martin, and T. Stützle. Iterated Local Search. In F. Glover and G. Kochenberger, Editors, Handbook of Metaheuristics, Kluwer Academic Publishers, 2003, pp. 321-353.