Coevolution for Classification

Catalin Stoean[1], Ruxandra Stoean[1], Mike Preuss[2], D. Dumitrescu[3]

[1] *Department of Computer Science, Faculty of Mathematics and Computer Science, University of Craiova, Romania*

{catalin.stoean, ruxandra.stoean}@inf.ucv.ro

[2] *Chair of Algorithm Engineering, Department of Computer Science, University of Dortmund, Germany*

mike.preuss@cs.uni-dortmund.de

[3] *Department of Computer Science, Faculty of Mathematics and Computer Science, Babes-Bolyai University of Cluj-Napoca, Romania*

ddumitr@cs.ubbcluj.ro

1. INTRODUCTION

The evolutionary computation (EC) community has approached classification, a field with important, and sometimes even vital, practical applications, not only from the side of hybridization with non-evolutionary classification techniques, e.g. fuzzy or neural network-based, but also from the part of standalone evolutionary algorithms (EAs) constructed especially for the particular task. An evolutionary classification technique is concerned with the discovery of IF-THEN rules that model the correspondence between given samples and corresponding classes. Given an initial set of training samples, the system learns the patterns, i.e. evolves the classification rules, which are then expected to predict the class of previously unseen examples.

There are two classical approaches to evolutionary classification techniques. The first direction ([Mic96]) is represented by the Pittsburgh school that developed an evolutionary system that considers an individual to represent an entire set of rules. Rule sets are evolved using a canonical evolutionary algorithm and the best individual from all generations represents the solution of the classification problem. The opposite related approach is the work of the Michigan school ([Hol86], [Mic96]). Here, each individual encodes only one conjunctive rule in first-order logic and the entire population represents the rule set. Thus, detection and maintenance of multiple solutions (rules) in a multiple sub-populations environment is required. As a canonical EA cannot evolve non-homogeneous individuals, the Michigan approach suggested doubling the EA by a credit assignment system that would assign positive credit to rules that cooperate and negative credit to the others.

A third known approach would also be the one characterized by a genetic programming approach to rule discovery ([Frei02a], [Frei02b]). The internal nodes of the individual encode mathematical functions (e.g. AND, OR, +, -, *. <, =) while the leaf nodes encode the attributes. Given a certain individual, the output of the tree is computed and if it is greater than a given threshold, a certain class of the classification task is predicted.

The existing evolutionary classification techniques have quite intricate engines and thus their application is not always straightforward: they use complex credit assignment systems that penalize or reward good rules, as well as very complicated schemas of the entire system.

The purpose of this chapter is hence to put forward a novel evolutionary classification framework that has proven to be simpler and yet competitive. Discussion envisages two techniques based on the state-of-the-art EC field of coevolution. Coevolution between individuals assumes two opposite interactions: cooperative and competitive. Analogously, coevolution for classification assumes two possible and opposed manners of solving the task. Within both approaches, the solution of a classification problem is regarded as a set of IF-THEN conjunctive rules in first order logic (as in the Michigan approach), where the condition part is made of attributes (indicators) and the conclusion part is represented by the class.

2. OBJECTIVES

The mission of this chapter is three-fold. First and foremost, we target to present two novel applications of coevolution to classification, from both a cooperative and a competitive perspective. Subsequently, we aim to achieve an objective comparison between the two opposite proposed techniques that lie within the same framework. Finally, we outline the practical applications of presented approaches from a binary problem, i.e. breast cancer diagnosis, to a multi-class task, i.e. iris plants differentiation.

3. COOPERATIVE COEVOLUTION FOR CLASSIFICATION

Cooperative coevolution implies a decomposition of a candidate solution of the problem to be solved into a number of components and each of these parts is treated by an EA. The EAs evolve separately and interactions between populations exist only in the moment when fitness is computed; when an individual is evaluated, collaborators from all the other populations are selected in order to form a complete solution that can be measured.

Cooperative coevolution was introduced around a decade ago as an alternative evolutionary approach to function optimization [Pot94]. For this issue, one considers as many populations as the number of variables the specific function has, i.e. each variable of the function represents a component of the solution and is separately treated using an EA. Any EA can be used to handle the components of the solution. Several functions with multiple local optima and one global optimum were considered and the cooperative coevolutionary algorithm proved to be effective [Pot94, Wie03]. The technique was recently successfully applied to develop a rule-based control system for agents; two species were considered, each consisting of a population of rule sets for a class of behaviours [Pot01].

To the best of our knowledge, there has been no attempt in applying cooperative coevolution to classification based on individuals that encode simple conjunctive IF-THEN rules in first order logic. Nevertheless, other classification evolutionary models for coadapted components are Holland's classifier system (the Michigan approach earlier discussed) [Hol86] and the REGAL system [Gio00], where stimulus-response rules in conjunctive form (such as in the present approach) were evolved by EAs. In [Gio00], problem decomposition is performed by a selection operator, complete solutions are found by choosing best rules from each component, a seeding operator maintains diversity and fitness of individuals within one component depends on their consistency with the negative samples and on their simplicity [Pot00].

Within the proposed technique [Sto06a, b, c], the final rule set is imagined as to have one rule for each category of the classification problem. A natural decomposition of the problem solution is to assign rules of the same outcome to a population; thus, the number of species equals the number of classes.

When the quality of an individual is measured, collaborators (individuals) from all the other populations have to be selected with the aim of forming complete solution(s) that may be easily evaluated. A complete solution of the problem consists of a set of rules, one for each class of the considered problem. The set of rules is applied to the training data and accuracy is computed. The accuracy represents the evaluation of the considered individual (rule). When a set of rules is applied to a sample, the distances from the sample to each of the rules are computed. The class of the sample is considered to be the class of the rule that was closest. Accuracy is computed as the percent of correctly classified samples over all considered samples.

Validation has been achieved on both a binary classification task, i.e. breast cancer diagnosis, in order to get a first glance on the performance capability of the approach, as well as on a problem with multiple outcomes, i.e. iris plants recognition, with the aim to verify the capacity of the algorithm to deal with several populations that coevolve. Obtained results have been competitive to state-of-the-art classification techniques from different fields of artificial intelligence, e.g. support vector machines, classification trees, logistic regression etc.

In the form considered so far, the coevolutionary classifier obtains a number of rules equal to the number of categories of the classification problem. However, it may happen that more rules are needed for one class as there may be the case that several objects in the data set that have the same class are very different. In this sense, in the future, the canonical EA will have to be replaced with a multi-modal evolutionary technique. The goal is to acquire all optima, i.e. the best possible rules, for each category of the considered problem. This represents, without a doubt, an important means of improving the classifier based on cooperative coevolution.

4. COMPETITIVE COEVOLUTION FOR CLASSIFICATION

Within the competitive model, two species evolve complementarily through an inverse fitness interaction process, similarly to the predator-prey systems. This means that success on one side is regarded as failure on the other side; the latter species will have to react in order to maintain its chances of survival. This competitive interaction between species will eventually drive them to evolution. In terms of computational entities, one species will correspond to certain tests and the other to the proposed solutions. Competition is achieved through encounters between one individual from a population and one from the other. The two selected individuals are checked against each other and if the solution passes the test then the former is rewarded and the latter is penalized; if it fails, credits are assigned in an inverse manner. Each individual has a history of its encounters and fitness is computed on this basis.

The competitive paradigm has been applied to a wide range of problems, from the primary field of game-playing to process control [Par97a], path planning [Par97b], constraint satisfaction [Par94b] and classification. As classification is concerned, known techniques involve the evolution of neural networks [Par94a], decision trees [Sie94], cellular automata rules [Jui96], [Par97c] and the use of genetic programming for the problem of intertwined spirals [Jui98]. Again, it has to be stated that to the best of our knowledge, the competitive coevolution between simple IF-THEN rules and the training set has not been achieved yet.

The proposed evolutionary classifier considers the training data set as one fixed species and the potential rules of assessment as the other evolving one. The two species evolve together, through the inverse fitness interaction: As candidate rules fit certain samples descriptions, these records receive weaker evaluation scores and are therefore not selected for encounters any more. Consequently, other samples, more difficult to assess, will be more often selected for competitions and rules will have to evolve through adaptation to the new records that must be given a verdict. The personal history of individuals is also important as it offers lifetime evaluations of individuals and allows them to keep up with the changing rankings in the other species.

Encounters are performed only between solutions and samples of the same class. Such an encounter returns the distance between the individual (rule) and the training sample as negative for the most recent fitness of the former and positive for that of the latter. Additionally, the individuals that are selected for recombination belong to the same class. Through these restrictions, we maintain diversity within the EA, such that final rules for each class are distinct from those with different outcomes.

The proposed algorithm has been tested against the same two data sets and has proven to provide viable results, as well.

Future work envisages the inclusion of some mechanism for sample selection, because, as the second population encodes the entire training set, runtime increases with the number of records in the data set.

5. DISCUSSION

The two related algorithms have demonstrated the ability to solve classification tasks of different content, dimension and number of outcomes in a well performing manner. Nevertheless, a drawback of the cooperative algorithm is the fact that the number of populations must increase with the number of classes of the task, while a flaw of the competitive approach is the high necessary runtime to encode a large data set.

The realistic and immediate implication of the application of the proposed classification framework would be the incorporation into medical decision making systems, as a means of checking the consistency of the assessment, which takes into account a large number of variables and more than often exhibits low generalization. Additionally, an online integration with electronic records would provide a real time learning and evaluation.

REFERENCES

[Eib03] Eiben, A.E., Smith J.E., *Introduction to Evolutionary Computing*, Springer-Verlag, 2003

[Frei02a] Freitas A.A., *A Survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery*, Advances in Evolutionary Computation, 819-845, 2002
[Frei02b] Freitas A.A., *Data Mining and Knowledge Discovery with Evolutionary Algorithm*s, 2002
[Gio00] Giordana A., Saitta L., Zini F., *Learning Disjunctive Concepts by Means of Genetic Algorithms*, Proceedings of the 11th International Conference on Machine Learning, 96–104, 2000
[Hol86] Holland J.H., *Escaping Brittleness: The possibilities of General Purpose Learning Algorithms Applied to Parallel Rule-Based Systems*, Machine Learning, 2, 593–623, 1986
[Jui96] Juille H., Pollack, J.B., *Dynamics of Co-evolutionary Learning*, Proceedings of the 4th International Conference on Simulation of Adaptive Behavior, 526-534, 1996
[Jui98] Juille H., Pollack, J.B., *Coevolutionary Learning: a Case Study*, Proceedings of the 15th International Conference on Machine Learning, 1998
[Mic96] Michalewicz Z., *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, 1996
[Par94a] Paredis J., *Steps towards Co-evolutionary Classification Neural Networks*, *Artificial Life IV*, 102-108, 1994
[Par94b] Paredis J., *Coevolutionary Constraint Satisfaction*, Proceedings of the 3rd Conference on Parallel Problem Solving from Nature, LNCS, 866, 46-55, 1994
[Par97a] Paredis J., *Coevolutionary Process Control*, Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms, 1997
[Par97b] Paredis J., Westra R., *Coevolutionary Computation for Path Planning*, Proceedings EUFIT 1997, 1, 394-398, 1997
[Par97c] Paredis J., *Coevolving Cellular Automata: Be Aware of the Red Queen!* Proceedings of the 7th International Conference on Genetic Algorithms, 393-399, 1997
[Pot94] Potter M.A., De Jong K.A., *A Cooperative Coevolutionary Approach to Function Optimization*, Proceedings of the 3rd Conference on Parallel Problem Solving from Nature, 249-257, 1994
[Pot00] Potter M.A., De Jong K.A., *Cooperative Coevolution: An Architecture for Evolving Coadapted Subcomponents*, Evolutionary Computation, 8(1), 1-29, 2000
[Pot01] Potter M.A., Meeden L.A., Schultz A.C., *Heterogeneity in the Coevolved Behaviors of Mobile Robots: The Emergence of Specialists*, Proceedings of the 17th International Conference on Artificial Intelligence, 1337-1343, 2001
[Sie94] Siegel E.V., *Competitively Evolving Decision Trees Against Fixed Training Cases for Natural Language Processing*, Advances in Genetic Programming, 409-423, 1994
[Sto06a] Stoean C., Stoean R., El-Darzi E., *Breast Cancer Diagnosis by Means of Cooperative Coevolution*, 3rd ACM International Conference on Intelligent Computing and Information Systems, accepted, 2006
[Sto06b] Stoean C., Dumitrescu D., Preuss M., Stoean R., *Cooperative Coevolution for Classification*, Bio-Inspired Computing: Theory and Applications 2006, 289- 298, 2006
[Sto06c] Stoean C., Preuss M., Dumitrescu D., Stoean R., *Cooperative Evolution of Rules for Classification*, IEEE Post-proceedings Symbolic and Numeric Algorithms for Scientific Computing 2006, in press, 2006
[Wie03] Wiegand R.P., *Analysis of Cooperative Coevolutionary Algorithms*, PhD thesis, Department of Computer Science, George Mason University, 2003