

Biological Applications of Genetic and Evolutionary Computation

Marylyn D. Ritchie and Jason H. Moore

Center for Human Genetics Research, Department of Molecular Physiology and Biophysics,
Vanderbilt University, 519 Light Hall, Nashville, TN 37232
{ritchie, moore@chgr.mc.vanderbilt.edu}

1 Introduction

Due to modern technologies, the wealth of data in biological research has greatly increased over the past decade. To accommodate this explosion of information, researchers are turning to a number of analytical disciplines to find efficient ways to analyze these data. The field of Genetic and Evolutionary Computation (GEC) is a prime example of this. Many features of GEC make it very appealing for the analysis of biological data. Often the data consist of many more independent variables than samples (individuals). In addition, we are often interested in modeling complex, nonlinear relationships between those variables. Since this creates a vast search space with lots of local minima, hill climbing search techniques are usually not very successful. Thus, algorithms in the area of GEC are a more suitable approach. Since the field of GEC has benefited from borrowing many of its ideas from biological sciences, it is now “repaying the debt” by helping to solve biological problems.

This year marks the third annual workshop on Biological Applications of Genetic and Evolutionary Computation (BioGEC). This workshop is organized in conjunction with the 2004 Genetic and Evolutionary Computation Conference (GECCO) in Seattle, WA and is intended to explore and critically evaluate the application of GEC to biological problems. The primary goal of this workshop is to bring biologists and computer scientists together to foster an exchange of ideas that will move the field forward in unpredictable ways. This area of research is highly interdisciplinary in nature, and so it is important to make sure that scientists from very different fields are able to communicate with one another on topics of interest. To facilitate this conversation, we have solicited for papers discussing relevant topics in the area of BioGEC. The papers that are published in these proceedings are in the form of commentaries, essays, perspectives, and reviews that focus on ideas and concepts in the field rather than on specific data and results. This decision was made with the hope that such work would facilitate discussion at the workshop regarding topics that are of major importance in the field of BioGEC. Since this field is so interdisciplinary, researchers come from inherently different backgrounds with different theories, beliefs, and languages. One of the most important challenges in

this field is getting researchers to communicate so that they are mutually working toward the same goals. This is what we hope will occur at the workshop as a result of more philosophical papers.

We received many interesting papers for this workshop, and six were selected for publication in the workshop proceedings. Clare Congdon discusses several current issues for researchers working in the field of BioGEC. She points out some of the challenges that result in the diverse background of computer scientists and biologists that make collaboration and communication such an important challenge. Leo Caves et al. introduce a very interesting perspective on the “debt” that is often mentioned as being repaid to biology from the EC field. Has the field of EC actually borrowed enough from biology such that the debt is worth repaying? The authors give their opinion on this very question. The essay by Marylyn Ritchie points out the importance of model validation in the field of BioGEC. This paper gives a few examples of model validation approaches and applications to facilitate discussion, but is not intended to be a comprehensive review of the topic. E.C. Keedwell and A. Narayanan share their experiences with the use of genetic algorithms for DNA array analysis. This is a fast-growing area of bioinformatics in which there are many success stories, as well as many important challenges still ahead. The review by Leon Poladian et al. first reviews evolutionary algorithms as applied to the study of phylogenetics. Next, they ask whether the use of evolutionary algorithms and multi-objective optimization can provide a more powerful approach for the field. Finally, Frederik Vandecasteele et al. discuss evolutionary computation application for a slightly different field of biology, the assembly of efficient ecosystems. They review recent advances in this fairly new field and discuss the promise that evolutionary computation brings to ecological research.

2 Conclusion

The field of BioGEC has been gaining momentum in recent years due to the influx of biological data and the need for more comprehensive data analysis. With the increase in biological applications of evolutionary computation comes an increase in the challenges facing this field of research. The goal of this workshop is to bring biologists and computer scientists together to improve communication such that the field will continue to move forward. We hope that this collection of papers in addition to the three presentations at the workshop will create lively discussions and result in exciting new collaborative research.