Thoughts on Using Evolutionary Computation to Assemble Efficient Ecosystems

Frederik P.J. Vandecasteele[1,2], Thomas F. Hess[1], Ronald L. Crawford[2]

[1] Dept. of Biol. and Agric. Engineering, University of Idaho, Moscow, ID 83844-0904
[2] Environmental Biotechnology Institute, University of Idaho, Moscow, ID 83844-1052
{vand8442, tfhess, crawford}@uidaho.edu

Abstract. Evolutionary computation techniques have recently been used for the first time to construct efficient real world biological ecosystems. These experiments have opened a whole new area of application of evolutionary computation. This paper discusses various aspects of this promising new type of ecological research.

1 Merging Evolutionary Computation with Experimental Ecology

Ecology is the field within biology that studies the interactions of organisms with one another and with their physical environment. Ecosystems are made up of communities of organisms and the nonliving factors with which they interact. Such communities typically consist of populations of many different species. These organisms live in complex ecological systems, which can have high levels of non-linear interaction. The actions of every member of an ecosystem can be advantageous, disadvantageous or neutral for one or more of the other members and organisms can form intricate interacting networks with each other. In addition to this, experimental ecological data is often rather noisy in nature. These elements make the study of ecosystems particularly challenging and trying to deliberately assemble ecosystems with a specific predetermined function is seldom attempted.

The field of evolutionary computation (EC) offers a range of flexible and robust search and optimization techniques capable of dealing with noisy and nonlinear systems. They can also be used on systems without knowing the exact governing dynamics, treating systems as black boxes. Even though these characteristics make EC techniques very appealing for studying microbial ecology, they have only rarely been used in this field.

Recently, we've introduced the idea of using EC techniques to assemble efficient real world ecosystems [1,2,3]. With this approach, EC is used to search a set of individual organisms for the subset of organisms that together form an ecosystem that optimally performs a specific predetermined function.

2 Experimental Challenges and Opportunities

2.1 Possible Application Areas

Using EC to construct efficient ecosystems will prove to be a valuable technique with many areas of application. In ecology, it's generally assumed that groups of organisms working together can often perform certain functions better than each member organism could individually. Examples of such functions would be degradation or production of certain chemicals. For example, in industrial fermentation, microorganisms are used to produce chemicals, while in bioremediation they are used to degrade contaminants. In both instances, the large majority of processes are currently being performed with cultures of one single organism. Having a tool to design better performing groups of organisms holds the promise of increased process efficiency.

Similarly, EC could be used in the field of restoration ecology, where damaged ecosystems are repaired by correcting environmental conditions and reintroducing essential species. The approach could also be used to assemble self-contained ecosystems, which would be of interest for space technology.

The approach will undoubtedly lead to new fundamental ecological insights and help us to better understand how ecosystems function. Traditionally, researchers interested in a particular ecological process would seek out efficient ecosystems in nature and then try studying those. Now we have a tool to predefine a process of interest and then construct ecosystems that perform this process well from arbitrary sets of candidate organisms.

Even though changes in ecosystem composition are not considered to be evolution in the strictest sense, it can still be argued that the environment selects for the composition of ecosystems. In nature, environmental conditions represent a selective pressure that will result in the ecosystems (metagenomes) best adapted to those particular conditions. Constructing efficient ecosystems using EC can be regarded in the same way, but it also goes further. Using EC, it's actually possible to obtain ecosystems that perform functions that can not be selected for in nature. An example of this is given in [3].

2.2 Types of Ecosystem Assembly

It's possible to distinguish three main types of ecosystem assembly, each of which can be optimized using EC.

The most straightforward method of assembly is combining fixed amounts of different organisms together at one single point in time. A bit string, encoding for the presence or absence of corresponding organisms can represent an ecosystem as the subset of organisms from a set of candidate organisms.

An elaboration on this basic idea is to allow the number of organisms of each species to be added to an ecosystem to vary. Such ecosystems can also be represented as bit strings, with subsections of the strings mapping to amounts of organism, or as a string of integer or real values.

While in the previous two types of ecosystem assembly, all organisms are combined at one point in time, a third type defines the points in time at which each species is added to the ecosystem and thus also the sequence of additions. This can be combined with either fixed or variable amounts of organism to be added.

2.3 Types of Evolutionary Computation

Genetic Algorithms seem like the most obvious EC technique for these types of optimization, especially for simple bit strings encoding presence or absence of organisms. Alternatively, Evolution Strategies could be used too.

2.4 High Cost of Fitness Evaluation

An essential part of using EC to construct ecosystems as we suggest is that fitness values of solutions need to be determined experimentally. To asses how well a particular subset or ecosystem performs a predetermined function, that ecosystem needs to be assembled in a controlled environment after which fitness values are measured. This poses a severe challenge with a number of implications.

Often, by necessity, researchers using EC in ecology will deal with small population sizes. Special care will need to be taken when designing the evolutionary algorithms to account for this.

Assembling ecosystems for fitness evaluation can be labor intensive while it may be desirable to increase the population sizes. Therefore, automation by using robotic equipment will prove to be very valuable.

It may also prove to be useful to investigate the application of hybrid optimization techniques, using elements from EC combined with traditional modeling. In such cases, modeling or interpolation could help in reducing the number of fitness evaluations.

2.5 Comparison to Non-evolutionary Computation Optimization Techniques

As with any application of EC, a comparison of evolutionary optimization in this field with other, non-EC optimization techniques will be interesting. However, the nature and complexity of ecosystems (noisy and non linear) leads us to believe that EC techniques will be superior to most other optimization methods. Other techniques that come to mind are random searches, hill climbing, neural networks, particle swarm optimization and simulated annealing.

2.6 Demonstrating Successful Optimization

Experimental data of assembled ecosystems can have a high degree of variability, both within batches of experiments (generations) and between. When trying to demonstrate the successful optimization of a particular system, it may therefore not be enough to look at trends in average and maximum fitness values obtained for each generation. In fact, a more sound approach is to eliminate between generation variability by reassessing fitness values of the best solutions in each generation, in a single batch of experiments. The solution with the median fitness value in each generation may then serve as an approximation of the average fitness per generation.

Given a particular system (one set of candidate organisms and one ecological function) an interesting question is how repeatable the optimization process is for that system. It seems like this question can only be answered by repeatedly optimizing the same system.

One can wonder how universally applicable the use of EC techniques is to assemble efficient ecosystems. It's hard to see how this can be answered other than by performing many experiments on many different systems.

We propose one particular metric for assessing the success of an optimization run. Using fitness values obtained from randomly generated ecosystems, it's possible to determine the distribution of fitness values under random conditions. It's now possible to calculate the number of random experiments that need to be performed to have a 95% chance of obtaining at least the highest fitness value of the EC optimization technique. If this number is higher than the total number of EC evaluations, then the EC technique can be considered to be more efficient than a random search.

2.7 Fitness Landscape Analysis

We consider fitness landscape analysis to be a valuable tool in this type of research. Autocorrelation functions describing the correlation between similarity in ecosystem species composition and similarity in ecosystem function can be an indicator of problem difficulty. This can be especially valuable, considering that many applications will only have a small number of experimental fitness evaluations. Additionally, autocorrelation functions are of fundamental ecological interest, since they address some intuitive ecological notions of ecosystem structure versus function.

2.8 Elucidating Mechanisms of Efficient Ecosystems

Using EC techniques to construct efficient ecosystems will help us to better understand and study the process going on in such systems. Once an efficient combination of organisms has been identified, additional work needs to be done to remove the organisms with only a neutral effect. Studies of metabolism and population dynamics can then be performed on that reduced set of core organisms to identify the biological mechanisms leading to the overall efficiency. There are many possible reasons why a group of organisms can perform better than a single species. For example, one organism could perform the bulk of the function, with other organisms catering to it by producing specific nutrients or by changing the environment so that it becomes favorable for the main organism. Different organisms could also perform different parts of the overall function, without hampering each other's functioning.

3 Current Applications

We have performed a number of experiments that for the first time in the peer-reviewed literature described the use of EC to assemble efficient ecosystems. All experiments have been performed with microorganisms, which are relatively easy to manipulate in a lab environment. All experiments involved using a Genetic Algorithm and representing ecosystems as bit strings encoding for the presence or absence of the corresponding organisms. The functions we've studied include dye degradation [1], biomass production [2] and minimal growth [3].

4 Conclusion

We have described the use of EC to assemble efficient real world biological ecosystems. This constitutes a completely new area of research both in the fields of EC and ecology. This paper addressed and described various aspects of this promising new type of ecological research.

Acknowledgement

F.V. is supported by a fellowship from the Inland Northwest Research Alliance (INRA) Subsurface Science Research Institute which is funded by the Department of Energy under contract DE-FG07-02ID14277.

References

1. Vandecasteele, F.P.J.: Constructing Efficient Microbial Consortia Using a Genetic Algorithm. Biological Applications for Genetic and Evolutionary Computation (BioGEC) – 2003 Genetic and Evolutionary Computation Conference Workshop Program (2003) 69-71
2. Vandecasteele, F.P.J., Hess, T.F., Crawford, R.L.: Constructing Microbial Consortia with Optimal Biomass Production Using a Genetic Algorithm. 2003 Genetic and Evolutionary Computation Conference – Late-Breaking Papers (2003) 299-302

3. Vandecasteele, F.P.J., Hess, T.F., Crawford, R.L.: Constructing Microbial Consortia with Minimal Growth Using a Genetic Algorithm. In: G.R. Raidl et al. (Eds.): Applications of Evolutionary Computing - EvoWorkshops 2004. Lecture Notes in Computer Science Vol. 3005. Springer-Verlag, Berlin Heidelberg New York (2004) 123-129