GENETIC ALGORITHMS

Calabretta R.
Institute of Cognitive Sciences and Technologies, National Research Council, Italy

Keywords: Simulation, computational models of evolution, evolutionary computation, evolutionary biology, optimization, complex and adaptive systems, neural networks, artificial life.

Contents

Summary

Evolution by natural selection is a natural and powerful mechanism by means of which the structure of organisms is progressively (i.e. from one generation to another) modified to give a better performance in the environment where organisms live. This is the principal process responsible for the origin and evolution of the huge variety of living forms which exist on earth. Genetic algorithms (GAs) are computer simulation models of evolution that allow us to reproduce on the computer populations of interacting agents (e.g., molecules, organs, organisms), and to study the emergence and the behavior of evolving complex dynamical systems such as biological systems (e.g., cells, body apparatus, ecologies) and social systems (e.g., economic, social, and political organizations). As a result, GAs can be used as a tool not only for scientific modeling but also for industrial optimization. In the former case, GAs can be used for studying and understanding how evolution works in nature. In the latter case, it is potentially possible to use GAs for solving any optimization problems which need to be resolved (e.g., optimization of airfoil parameters). More generally, by integrating GAs with other modeling tools such as neural networks, the simulation of multi-level adaptive systems and, therefore, the interdisciplinary study of complex systems become possible.

1. Introduction

Charles Darwin spent more than 15 years before he decided to publish the book which would eventually become very famous, *On the Origin of Species*. The reason he dedicated so much time to revision was that Darwin was aware that his revolutionary

UNESCO – EOLSS SAMPLE CHAPTERS

ideas would be received negatively.

Darwin's theory was revolutionary in indicating evolution by means of natural selection as the principal responsible process for the origin and evolution of the huge variety of living forms which exist on earth. Evolution by natural selection is, in fact, a natural and powerful mechanism by means of which the structure of organisms is progressively (i.e. from one generation to another) modified to give a better performance in the environment where organisms live. The process is so powerful that, in certain appropriate conditions, it gives rise to new species.

For this process to happen, there is no need for any engineer to be involved in planning and designing the features of organisms which allow them to do what they are able to do (e.g., wings for flying, fins for swimming, legs for walking, etc.). On the contrary, features emerge *as a consequence* of the evolutionary process.

The process is, in fact, based only on the selective reproduction of the fittest organisms (i.e., those which are the most suitable to an environment), genetic recombination of their chromosomes, and on some random genetic mutations.

Imagine a population of organisms that live in an environment and reproduce by sexual reproduction. Offspring receive from parents a copy of their genes after they have been mixed with each other (genetic crossover) and have received, sometimes, some random changes (genetic mutations). The successive development of the organisms is under the control of the program written in the DNA but the outcome will also depend, in a decisive way, on the interaction with the physical environment in which an organism lives and with the other organisms that live in the same environment. Each organism, as a result of inherited mutated parents' gene products, will have more or less adapted to the environment in which it lives and will compete with the other ones for the available resources and for being sexually selected. The organisms which are the most suitable to the environment will be able to find food and, therefore, to survive, and, once selected, to mate with co-specifics.

Adaptation of an organism depends, therefore, on smooth running and coordination of many different levels of organization of the organism itself and on a dynamic interaction among these levels, the environment and the other organisms. Although this process is very complex, the solutions obtained in nature by means of natural selection exhibit features of robustness and efficiency.

To study and understand this interaction and, therefore, understand how complex systems evolve in nature it has, at least until recently, been very hard. In the case, for example, of the evolution of new species, hypotheses are based on only incomplete fossil records and on following inferences. The situation changed with the discovery of the computer and the increase in computer processing power that has furnished a new powerful tool of investigation which is simulation.

Simulation as research methodology offers, in fact, many advantages in the specific field of the study of evolutionary processes and, also, in the more general evolving complex systems, which includes social and political evolving systems. As stressed by

UNESCO – EOLSS SAMPLE CHAPTERS

Mitchell, "a computer program can simulate the evolution of populations of organisms over millions of simulated generations, and such simulations can potentially be used to test theories about the biggest open questions in evolution."

Genetic algorithms (GAs) are simulation models of evolution that can be used for studying and understanding how adaptation occurs in nature. GAs have almost always the shape of a computer software that allows us to reproduce on the computer populations of interacting agents (e.g., molecules, organs, organisms), and to study the emergence and the behavior of evolving complex dynamical systems such as biological systems (e.g., cells, body apparatus, ecologies) and social systems (e.g., economic, social, and political organization). The way of working of GAs allows us to use them in all the cases (i.e., processes) in which adaptive behavior emerges as a result of interaction among components without hand-design.

As pointed out by Forrest, "although the computational setting is highly simplified compared with the natural world, genetic algorithms are capable of evolving surprisingly complex and interesting structures."

As a result, GAs can be used not only as a tool for studying and understanding how evolution works in nature but also as a tool for industrial optimization. In the former case, it is potentially possible to use GAs (alone or with other simulation tools; see below, section 5) for scientific modeling and simulate on one's own home personal computer the evolution of, for example, the earth, biological life, homo sapiens, society, and so on! In other words, it is possible to try to review millions of years of evolution in a few hours, all of which only depends on the processing speed of a personal computer. In the latter case, it is potentially possible to use GAs for solving any optimization problems which need to be resolved (e.g., optimization of airfoil parameters).

This paper focuses mainly on the utility of using GAs for scientific modeling and, specifically, for reproducing in a computer self-organizing complex systems. In the next section, it will be pointed out that GAs are not the only algorithms that draw their own inspiration from Darwinian evolution; on the contrary, there are several evolutionary algorithms that together form the research field of evolutionary computation.

Moreover, a brief guide on which hardware and software are necessary and what to do in order to implement a GA is given together with a theoretical analysis on how GAs work. Then, the use of GAs together with other computational models of organisms, such as neural networks, is considered by pointing out advantages for such a multi-level integrated modeling. Available resources on the web are furnished and, as a conclusion, perspectives on possible future studies are drawn.

2. Historical Development of Evolutionary Computation

In the late fifties some evolutionary biologist started using the first available slow and expensive computers for running programs in order to study and understand natural evolution. Later in the sixties, some computer scientists used biological evolution as a source of inspiration and imagined an evolutionary process which took place in a computer.

UNESCO – EOLSS SAMPLE CHAPTERS

The idea underlying this approach was the following: as happens in nature where organisms' design solutions for surviving, reproducing and adapting to the environment are found across generations by means of natural selection, in the same way it is possible to find good solutions for a given problem at hand by means of software that reproduces natural selection in a computer.

Based on this general idea, different evolutionary algorithms were developed independently by different researchers and new research fields sprang up having different features and goals. They were: Evolutionary Strategies, Evolutionary Programming, Genetic Algorithms, and, more recently, Genetic Programming.

Evolutionary strategies (ES) were born in the early sixties at the Technical University of Berlin when Rechenberg and Schwefel thought of using simulated evolution for optimizing parameters of physical objects such as airfoils. Their algorithm was characterized by a population of only one parent and one offspring and by using genetic mutation as the only genetic operator. This research field is still an active one, and in the last few years has started to interact with the other evolutionary algorithm research fields.

In about the same years *Evolutionary programming* (EP) was developed by Fogel, Owens, and Walsh, "a technique in which candidate solutions to given tasks were represented as finite-state machines, which were evolved by randomly mutating their state-transition diagrams and selecting the fittest." Only mutations were used as a genetic operator for increasing variability in these algorithms.

While ES and EP were developed by researchers with industrial optimization in mind, *Genetics Algorithms* were designed as a tool for scientific modeling in the early sixties by John Holland and his students at the Department of Computer and Communication Sciences of the University of Michigan. According to Holland, the goals in reproducing natural evolution in computers were two: first, to formally study and understand how adaptation occurs in natural systems; second, to try to develop ways for transferring natural mechanisms of evolution in computer systems.

Holland's algorithm was the first to work on a population of individuals rather than on only two individuals, and to use not only the genetic operator of mutation but also those of crossover and inversion.

In the early nineties, John Koza thought of using a genetic algorithm for evolving the most suitable computer program which was able to solve a particular problem and, in this way, he created a new research field called *Genetic Programming.*

Nowadays (2002), different algorithms often no longer have the features that characterized them when they were thought of by their inventors and share some of their features. Evolution strategies, evolutionary programming, genetic programming, and genetic algorithms are referred to as evolutionary algorithms and all of them belong to the more general research field called evolutionary computation.

UNESCO – EOLSS SAMPLE CHAPTERS

3. How to Implement a GA

3.1. What (Hardware and Software) we Need in Order to Implement a GA

GAs are a simulation tool that can be as easy to use as a computer game. In fact, GAs very often have the shape of software that runs in a simple personal computer. In other words, in order to use GAs we need, as basic technology requirements, a normal personal computer and specific software that runs on it. Of course, the more powerful the computer's processor is, the quicker the simulation speed is.

It is enough only to choose the parameters of simulation (e.g., population size; see below, section 3.3) and run the program in the same way which is done for the common computer program or for computer games. A simulation can, in certain cases, finish in a few hours.

GAs software can be programmed by users or can be downloaded from the World Wide Web (see below, section 6). Different programming languages can be used to program GAs. The most common among those are C and C++ languages.

Of course, the computer software can be very complex in the sense that it can integrate within itself different simulation tools such as, for example, genetic algorithms and neural networks (see below, section 5).

In other more complex cases, GAs have the shape of software used together with hardware other than computer such as robots. In this case, the GA is software that no longer controls artificial organisms living in the computer but physical robots moving in the environment. Very often the process is the following: the "brain" of the robot (under the shape of a neural network) is allowed to evolve in the computer and then, this "brain" is downloaded in the robot and used to control the robot's movements.

In different cases, one could use the same computer – while the simulation is running- for other tasks or could see in the computer monitor what is taking place from one generation to another, for example, the organism's fitness or energy level, the size of population, the amount of food available in the environment, and so on. In both cases, it is possible to save the data obtained as files and use them whenever you want in order, for example, to display the behavior of a given organism belonging to a given generation or to compare the performance of organisms belonging to different generations.

With simulation data, very different analyses can be done. For example, it is possible to reconstruct the phylogenetic tree of organisms, that is, the sequence of parents from one generation to another of a chosen organism (phylogenies are "the history of descent of species from their common ancestors"). By using the phylogenetic tree of the best organism of the last generation, it is possible, for example, to highlight which order of steps was necessary to obtain the final result.

UNESCO – EOLSS
SAMPLE CHAPTERS

-
-
-

TO ACCESS ALL THE **16 PAGES** OF THIS CHAPTER,
Visit: http://www.eolss.net/Eolss-sampleAllChapter.aspx

Bibliography

Calabretta R., Galbiati R., Nolfi S. and Parisi D. (1996). "*Two is better than one: a diploid genotype for neural networks*". Neural Processing Letters 4, 149-155. [Examines the role of diploidy in short-term adaptation to new environments in asexual organisms - observed that diploidy is useful in producing individuals that are able to adapt to non-stationary environments, especially in the case of fast changing environments.]

Calabretta R., Nolfi S., Parisi D. and Wagner G.P. (1998). *A case study of the evolution of modularity: towards a bridge between evolutionary biology, artificial life, neuro- and cognitive science.* Proceedings of the Sixth International Conference on Artificial Life, (eds. Adami C., Belew R., Kitano H. and Taylor C.), pp. 275-284, Cambridge, MA, USA: MIT Press.[Study of efforts to create connect evolutionary biology and cognitive science with simulated artificial life.]

Darwin C. (1859). *The Origin of Species*. New York, USA: John Murray. [The landmark work on evolution in the natural world. One of the most important books in the history of scientific thought.]

Forrest S. (1993). "*Genetic algorithms: principles of natural selection applied to computation.* Science 261, 872-878. [Good overview of the evolutionary aspects of genetic algorithms.]

Futuyma D.J. (1998). *Evolutionary Biology*. Sunderland, MA, USA: Sinauer. [A classic in evolutionary biology.]

Goldberg D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA, USA: Addison Wesley. [Classic work detailing how genetic algorithms can learn to identify optimal outcomes.]

Holland J.H. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI, USA: The University of Michigan Press (Second edition: Cambridge, MA, USA: MIT Press, 1992). [A classic in the field of genetic algorithms detailing how natural evolutionary processes can be emulated in computer-based systems.]

Langton C.G. (1989). *Artificial Life*. Reading, MA, USA: Addison Wesley. [A classic in the field of artificial life detailing the basic elements required to define artificial life.]

Mitchell M. (1996). *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press. [An excellent primer on genetic algorithms.]

Rumelhart D. and McClelland J. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA, USA: MIT Press. [Interesting effort to examine linkages between computer-programs that learn and human cognition.]

Wagner G.P. and Altenberg L. (1996). *Complex Adaptations and the Evolution of Evolvability*. Evolution 50, 967-976. [Details connections between complexity, adaptation and evolution.]

UNESCO – EOLSS SAMPLE CHAPTERS

Biographical Sketch

Raffaele Calabretta received his laurea in Pharmacy in 1987 from the University of Rome "La Sapienza", Italy, and became permanent researcher of the Italian National Research Council (C.N.R.) in 1988. Since 1990 he has been a researcher in the Department of Pharmaceutical Studies at the University "La Sapienza" and, since 1993, in the Department of Neural Systems and Artificial Life of the Institute of Cognitive Sciences and Technologies of the C.N.R. in Rome. He was Professor of Educational Technologies at the University of L'Aquila and member of the Center for Computational Ecology at Yale University, where in 1997 he spent one sabbatical year as visiting scholar at the Department of Ecology and Evolutionary Biology. Currently, he is an affiliate of the New England Complex Systems Institute (Cambridge, MA). His current research interest is in simulating evolution of modularity in complex organisms by means of artificial neural networks as computational models of the brain and of genetic algorithms as computational models of evolution. He is scientific coordinator of a C.N.R. Italy-USA bilateral project "Evolutionary mechanisms for the origin of modular design in artificial neural networks. His activity also concerns science popularization: on May 2006 he published a scientific novel on emotions called Il film delle emozioni (Gaffi Editore, Rome; in Italian)."

UNESCO – EOLSS
SAMPLE CHAPTERS