

BACK TO BASICS – NATURE'S EVOLUTIONARY INTELLIGENCE MODELING IN GENETIC ALGORITHMS

INTRODUCTION

Biological evolution is the process describing how organisms evolve over time to adapt to the environment for better survival. This adaption happens in a most optimal way possible and the definition of optimality is quite dynamic decided by environment. Definitely, to develop such adaption is the result of a very long learning process evolving as a part of fundamental natural instincts. Understanding the way it happens across generations can be leveraged effectively for many AI applications in practice. Large class of such applications are contributed by selection of optimal alternative over large other options and learning such optimal alternatives through experimentations. Genetic algorithms represents such natural evolution process.

Organisms produce off springs through biological reproduction. Typically, these offspring's have similar characteristics by virtue of inheritance; however, some may have some random variations due to the biological process of mutations. Some offspring survive and produce next generations while some do not. Those organisms that are better adapted to the environment have higher chance of survival. Therefore, over time these generations become more and more adapted to the environment because the fittest of the organisms survive. This is the well-known principle of 'Survival of the fittest'. Based on this mechanism of biological evolution, John Holland from University of Michigan (1970's) developed a set of algorithms mimicking the evolution process. These algorithms are called Genetic Algorithms (GA). Genetic algorithms are the search procedures that operate on a population of solutions, which mimics natural selection to find increasingly better solutions. In this article, we will discuss the foundation of genetic algorithm, various biological operations in process of evolution, representing the problem in a genetic domain, defining various biological operators and fitness function, and applications of genetic algorithm.

EVOLUTION

The origin of life on the earth and present day plants and animals are discussed since very long by scientists and philosophers. There are three important theories of organic evolution –

- A. Theory of inheritance of acquired characters by French Biologist Lamarck [1]. This is based on two ideas –

1. Use and disuse – Individuals lose characteristics they do not require (or use) and develop characteristics that are useful.
 2. Inheritance of acquired traits – Individuals inherit the traits of their ancestors.
- Well-known example illustrating this theory is giraffes. Giraffes stretching their necks to reach leaves high in trees, strengthen and gradually lengthen their necks.
- B. Theory of Natural selection by Charles Darwin [2]. Most of the organisms produce large number of individuals having some variations in their characteristics. These individuals compete with each other and there is a struggle for existence. The individuals best suited to their environment survive, while others get eliminated (Survival of the fittest). This means nature selects only a suitable variation. Such suitable variations are transmitted from generation to generation with changes in the characters, according to the change in the environment. This results in the origin of new species after several years.
- C. Theory of Mutation by a Dutch botanist Hugo De Vries in 1901 [3]. This refers to the sudden changes in individual's (mutant) characters. These changes in mutants are inherited over generations provided such mutant survives through the competition of fittest. Thus, this theory explains how an offspring can have a character that is not present in parents. Theory of natural selection explains how such new character continues over multiple generations.

Today, the theories of Darwin and De Vries are combined to explain the method of evolution. To summarize, mutations bring sudden changes in organism's characteristics and these changed characteristics are transmitted to succeeding generations. Useful changes in characteristics keep on continuing through generations by principle of natural selection or survival of the fittest. These results in formation of new species, which are more, adapted to the environment. Figure 1 shows the journey of evolution from big bang to the present human beings, who are highly intellectual and highly adapted to the environment.

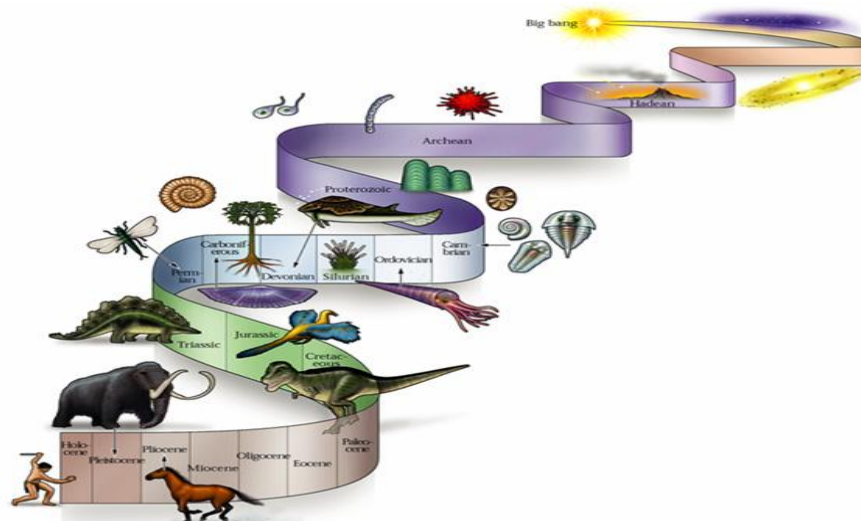


Figure 1: Biological evolution [4]

The individual's characteristics are biologically referred as genetic information. Every cell's nucleus of a living organism carries the complex chemical structure called as DNA (Deoxyribo-Nucleic Acid). This DNA carries the genetic information over generations. The basic unit which carries the signature of an individual is called as the Gene and it is a distinct portion of a cell's DNA as depicted in Figure 2. Genes are coded instructions in terms of sequence of chemical compounds, for making everything the body needs, especially proteins. These are the genes, which provide special characteristics to each individual organism and carry the inherited characteristics from one generation to next. During this evolution process, the genes keep on changing through several mechanisms so that resultant generation is more 'superior' to the initiator. Genes are packaged in bundles called chromosomes. For example, humans have 23 pairs of chromosomes (a total of 46).

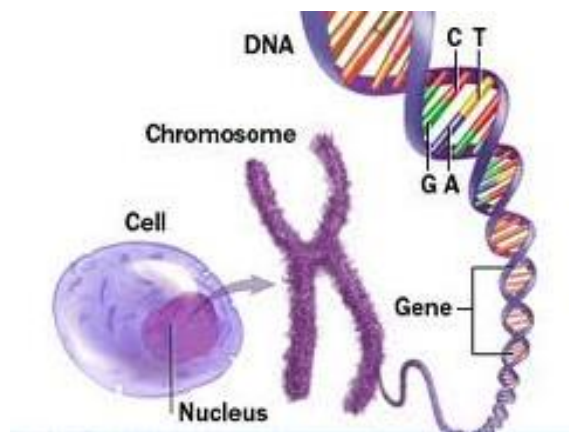


Figure 2: DNA, Gene, and Chromosome in cell nucleus

In a process of reproduction, the parent's chromosomes combine to form the offspring's chromosomes and the process is called as crossover of chromosomes. Mutation is another process as described by Hugo De Vries due to which sudden changes in genes appear providing entirely new features to the offsprings which are not present in parents. The change in a single gene of a chromosome is called as point mutation or gene mutation. Mutations involving change in the structure of chromosome are called as chromosomal mutations. This is illustrated in detail with example in further sections.

ALGORITHMIC REPRESENTATION OF EVOLUTION

As discussed earlier, biological evolution is the process of searching for the better and better characteristics of organisms such that organisms are better adapted to the environment. John Holland developed the simulation model of this evolution process in around 1970. The goal for this simulation was to understand processes in natural systems and to design artificial systems retaining the robustness and adaptation properties of natural systems. This set of algorithms, called as Genetic Algorithms, provides a generic procedure to find increasingly better solutions.

Genetic Algorithm is an iterative process, where each iteration represents a generation. The population in each generation is represented by the set of chromosomes. This forms the typical architecture of GA as shown in Figure 3 **Error! Reference source not found.**. Crossover operations take two parent chromosomes from the current generation and manipulate their genes to form a new chromosome to be used in the new generation. Parents for crossover are selected because “Parents with better fitness have better chances to produce offspring”. Mutation is then adds entirely new feature to the resultant chromosome.

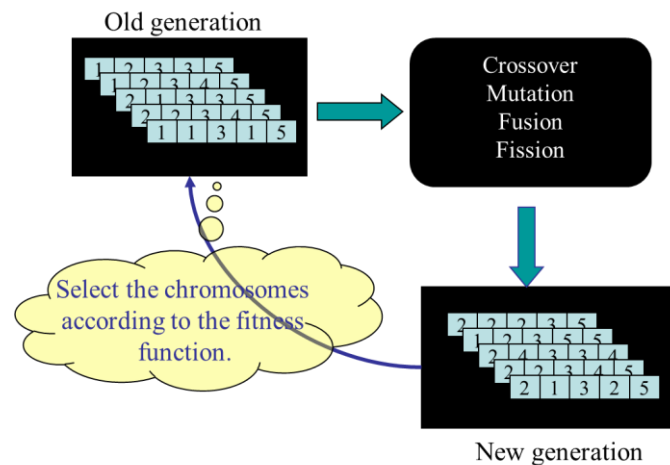
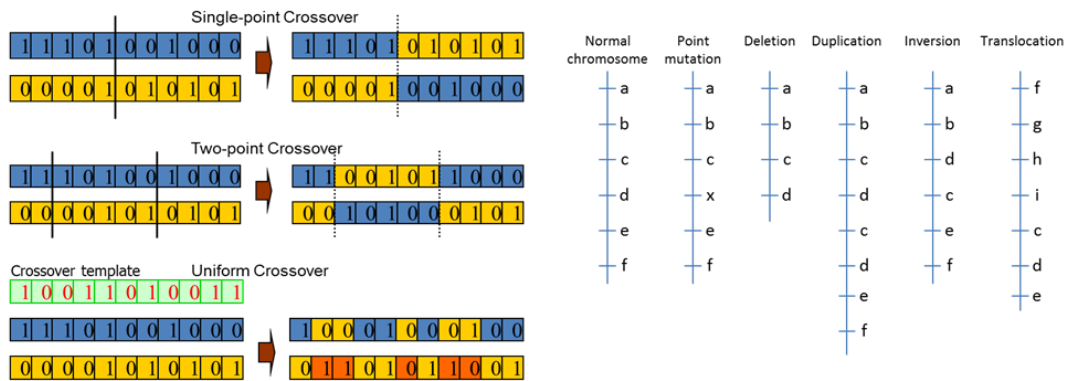


Figure 3: Genetic Algorithm architecture

Genetic algorithm (GA) is therefore a search heuristics, which can find a better solution to the problem over generations. Applying genetic algorithms to solve specific problem of finding a better solution or optimal solution requires consideration of five primary components [5].

1. Representing candidate solution of a problem as a chromosome
2. First generation will be a population of such candidate solutions. A method is required to initialize this population
3. Defining the chromosome crossover and mutation processes to change the chromosomes. This provides generating new solution state from existing solutions
4. Defining an evaluation function to quantify how well an individual solution fits the environment - “fitness” function
5. Other parameters needed to implement the above components, including population size, number of generations allowed, and stopping criteria.

Once the population of parents is formed, the further steps for evolution are chromosome crossover, mutation, and forming the next generation. There are numerous crossover operations including single point crossover, two point crossover, and uniform crossover as depicted in Figure 4(a).



(a) Crossover [6]

(b) Mutation [7]

Figure 4: Varieties of chromosome crossover & mutations

Similarly mutation operation has multiple ways like point mutation, deletion of some part of chromosome, duplicating some part of chromosome, inversion or translocation in chromosome, etc. as shown in Figure 4(b). However, the chance of such mutation to occur should be very low as it is responsible to add entirely 'new' feature to the offspring. The other parameters needed to implement the GA are population size, number of generations allowed, and stopping criteria. These parameters are entirely problem specific.

MODELING THE PROBLEM IN GA – TRAVELING SALESMAN

In its simplest form, the traveling salesman problem is to find a tour of a given set of cities so that each city is visited only once by a salesman and the total distance traveled is minimum. Such a process of finding the best combination suiting certain requirements is called as combinatorial optimization. Traveling salesman (TSM) problem is one of the best examples of such combinatorial optimization. As the number of cities increases, the possible solution number increases exponentially. For 10 cities, possible solutions are 181,000 whereas for 50 cities, it become $\approx 10^{62}$!

A chromosome for travelling sales man problem will represent the ordered list of the cities visited. Representing each city by corresponding unique number, example chromosomes might be (3 5 7 2 1 6 4 8) for 8 cities. This chromosome represents that a salesman is starting his tour from city 3, then visiting city 5, etc., finally visiting the city 8 and back to city 3. Initially many individual solutions are (usually) generated randomly to form an initial population. The fitness function is simply the total distance travelled by the sales man for the tour represented by a chromosome. For a problem size of 37 cities, a population of 200 would show how the population fitness is converging towards the

better value (minimum) over number of generations, illustrated in Figure 5, also shows the shortest trip found connecting these 37 cities.

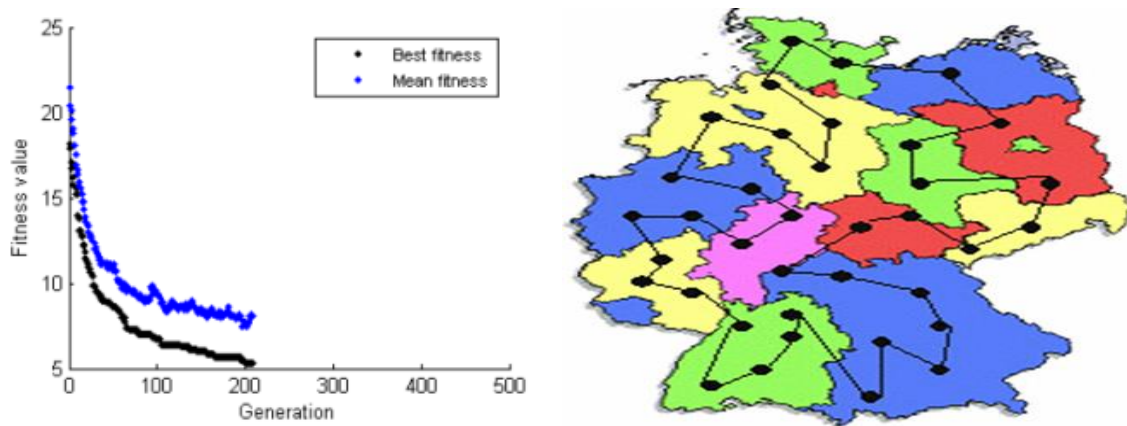


Figure 5: Population fitness convergence and shortest trip found by GA [4]

We can terminate the evaluation process if there is no change in population fitness over last few, say, 5 generations.

FEATURES AND OTHER APPLICATIONS OF GA

The simplicity of GA lies in its simple computations and potential for massive parallelism. Some features of GA are –

- Concept is easy to understand
- Highly modular structure, therefore building blocks can be used in hybrid applications
- Supports multi-objective optimization
- Good for “noisy” environment. Always results in an answer, which becomes better and better with time
- Often, GAs can rapidly locate good solutions, even for large search spaces.
- Capable of quickly finding promising regions of the search space but may take a relatively long time to reach the optimal solution.
- Good heuristics for combinatorial problems
- Usually emphasize combining information from good parents (crossover)
- Different GAs use different Representations, Mutations, Crossovers, and Selection mechanisms
- Can easily run in parallel

Considering these promising features, there are numerous applications of GA [8-11] –

- Optimization–numerical and combinatorial optimization problems, e.g. traveling salesman, routing, graph coloring, and partitioning
- Robotics - trajectory planning
- Machine learning – designing neural networks, classification and prediction

- Signal processing – filter design by optimizing the filter for desired frequency response
- Design – semiconductor layout, aircraft design, communication networks
- Automatic programming – evolve computer programs for specific tasks, design cellular automata and sorting networks
- Economics – Development of bidding strategies, emergence of economics markets

REFERENCES

- [1] Inheritance of acquired characters, Jean-Baptiste Lamarck, <http://www.britannica.com/EBchecked/topic/328430/Jean-Baptiste-Lamarck/273180/The-inheritance-of-acquired-characters>
- [2] Theory of Natural Selection - Genetics and Evolution, http://www.biology-online.org/2/10_natural_selection.htm
- [3] Mutation theory by Hugo De Vries, <http://biozoom.blogspot.in/2011/02/mutation-theory.html>
- [4] Global Optimization Genetic Algorithm, Olesya Peshko, http://www.cas.mcmaster.ca/~cs777/presentations/3_GO_Olesya_Genetic_Algorithms.pdf
- [5] Genetic Algorithms to Minimize Earliness and Tardiness of Outbound Trains in a Railyard, Robert G. Randall and Mary E. Kurz, available at www.informs-ras.org/Presentations/02_Randall.ppt
- [6] Chromosomal crossover, http://en.wikipedia.org/wiki/Chromosomal_crossover
- [7] Mutation, <http://en.wikipedia.org/wiki/Mutation>
- [8] APPLICATIONS OF GENETIC ALGORITHMSn available at http://neo.lcc.uma.es/TutorialEA/semEC/cap03/cap_3.html
- [9] Genetic Algorithms and their applications, Lawrence "David" Davis, President, Tica Associates, available at <http://www.informatics.indiana.edu/fil/CAS/PPT/Davis/>
- [10] Passive optical network planning in local access networks — an optimisation approach utilising genetic algorithms, H Paul and J Tindle, BT Technol J Vol 14 No 2 April 1996
- [11] Successful application of genetic algorithms to network design and planning, K F Poon, A Conway, G Wardrop and J Mellis, BT Technol J Vol 18 No 4 October 2000
- [12] http://en.wikipedia.org/wiki/Genetic_algorithm

AUTHOR

DR. UMESH HIVARKAR



Umesh Hivarkar's 24+ years of experience is characterized by multidisciplinary skills, having a unique combination of R&D profile, academics, and industry. He did his doctoral research in fiber optics & sensors, Instrumentation, Integrated optics, and Modeling & Simulation. His core expertise is in modeling and simulation, analytics, process & system automation and control, Maths and statistical analysis, application specific algorithm development, Machine vision, global optimization techniques, integrated optics, data acquisition/signal conditioning, radar signals and systems, speech/voice processing, data mining, PC-based instrumentation.