Volume 1 No. 3, July 2011 ISSN-2223-4985
International Journal of Information and Communication Technology Research

http://www.esjournals.org

Artificial Life Simulation in Bioinformatics

MABAYOJE M.A., MUFTAU O. W., ABIKOYE O.C.
Department of Computer Science,
Faculty of Communication and Information Sciences,
P.M.B 1515,
University of Ilorin,
Kwara State, Nigeria

ABSTRACT

Computer simulation of life is often used to study essential properties of living systems (such as evolution and adaptive behaviour).Artificial life is generated with the help of substituting blood corpuscles (blood cells). However, no truly safe and effective artificial blood product is available. Meanwhile, [2][7] has been a motivation for creating complex systems that are similar to biological system.

The aim of this work is to simulate and implement artificial life through the use of computer. Microsoft Visual Basic 6 as the background programming language was used for the implementation. Red blood cells are considered as the living systems, the cell structure is analyzed and its major constituent, DNA (deoxyribonucleic acid) is also described based on how it replicates and mutates.

Keywords: *Artificial, Life, Stimulation, System, Cell, Blood*

1. INTRODUCTION

A life is a powerful tool for studying, simulating and recreating biological system and process.

Artificial life is a blanket term used to refer to human attempts at setting up systems with life like properties all biological systems possess such as self reproduction/replication, adaptability, mutational variation, and optimization of external states and so on.

The term is commonly associated with computer simulation-based artificial life, preferred heavily to robotics, because of its case of reprogramming, inexpensive hardware and greater design space to explore.

The term artificial life often shortened to a-life or a life was originally carried by computer scientist Christopher Langton at the International Conference on the synthesis and simulation of living systems at the Los Alamos National Laboratory in New Mexico (USA) in 1987.

There are many courses related to computer science, artificial life was first studied and popularized by John Von Neumann.

In the late 1940s, he presented a lecture "The General and Logical theory of Automata", which introduced theoretical objects called "automata", state machines which underwent transformations based on well defined rules, integrating internal and external information.

Von Neumann developed such automata in high detail using nothing but pencil and graph paper – his early automata were represented as cells undergoing state changes on an infinite 2-dimensional grid. Throughout his last days, Von Neumann worked on cellular automata and his theories of self-replicating machines, developing the first cellular automata with Stamslaw Ulam during the 1950s.

Significance of Research

With the help of artificial life (in blood cells), the aim is to simulate some aspects of natural evolution (e.g. mutation, adaptation etc) in the computer.

The following are the objectives of the study:

(1) To understand evolutionary processes via computer simulation of simple (artificial) life form.
(2) To emulate the evolutionary process species undergo in nature.
(3) To emulate the survival of the fittest (natural selection), with the fittest candidate(s) at the end of the simulation being the solution.
(4) To emphasis solution on red blood related problems such as sickle-cell anemia, blood cancer etc. In other word, to proffer solution to blood related problems like anemia and cancer of the blood.

2. RELATED LITERATURE

Artificial life studies the logic of living systems in artificial environments. The goal is to study the phenomenal of living systems in order to come to an understanding of the complex information processing that defines such systems.

While artificial life is, by definition, alive, artificial life is generally referred to as being confined to a digital environment and existence.

Essentially, artificial life is a way to observe evolutionary design in action. This approach can be rigorously justified from a theoretical standpoint as Daniel

Dennett says, the evolutionary mechanism, properly understood is “substrate neural”[1][6].

This technology has recently progressed to the point that biologists can use it for their own research. Pioneering work of this sort was done by Ecologist Tom Ray, whose artificial life platform “Tierra” allowed him to observe and study the evolution of a virtual ecology [10].

Today, one of the most advanced artificial life systems is known as “avida”, originally developed by theoretical physicist Christopher Adami and computer scientist Charles Ofria.

In both Tierra and Avida, the organisms are computer programs running on their own virtual hardware. The program that defines the genome (sequence of genetic material (Chromosome)) of a digital organism is a sequence of commands in a simplified Computer Language. Neither system attempts to deal with the origin of life, both begin with a hand coded program that simply has the ability to replicate. Nor do they model any particular biological organism; rather they model key elements that are common to all forms of life.

To a first approximation, those organisms may be thought of as computer viruses, with the difference that they self-replicate, mutate and adapt by natural selection to a computational environment.

As Dennett has emphasized, “evolution will occur whenever and wherever three conditions are met: replication, variation (mutation), and differential fitness (competition) [2][6]. And evolution is exactly what one observes the original ancestor divides and mutates, creating divergent, novel lines organisms as the generations pass. The core evolutionary mechanism of variation, inheritance and natural selection is instantiated.

New variants can arise through random mutations in an organism’s genome and then be inherited in the next generation, if the organism successfully replicates.

Natural selection occurs as organisms in a population compete with each other to survive and replicate. If variations arise that give an organism some competitive advantage, it will be naturally selected and will spread in the population. In avida, they do not program what will happen, but simply set initial conditions and then observe the effects of the evolutionary processes.

In this regard, I set the initial conditions for the simulation in order to obtain the concept of evolutionary processes.

3. METHODOLOGY

We set the initial population of cells. We also set initial value of food for the newly generated population of cells to feed on.

Also set the initial value at which the food will be regenerated, if eventually there is no enough within the area.

In addition, each individual is created with certain value of energy to ease its adaptability in the area. Initial time (cycle) is set at which food will be generated in the life time of individual cell. Initial amount of energy is specified based on the food being generated at the stipulated cycle (time).

We set initial time for individual cell, starting from when it is created. Another number of cycles are specified to control or monitor the time at which each cell is multiplying. Also, initialize input value for the mutation which specifies the number of mutation of bases.

Euclidean distance (metric) is used to calculate the distance of the creatures (cells) within their area. It count the distance between the cell’s current point and the food co-ordinates. To survive within the area, individual cells are assigned a certain amount of speed. But it now depends on individual’s ability to think quickly.

The Euclidean distance is given by the following formula;

$$E_D = \sqrt{(Px - qx)^2 + (Py - qy)^2}$$

Where Px, qx, Py and qy represent co-ordinates respectively.

3.1 DNA STRUCTURE AND REPLICATION

DNA is composed of nucleotides (Francis Crick and James Watson). A nucleotide comprises a phosphate, a sugar and a base. The sugar is deoxyribose. The bases are purines-adenine (A) and guanine (G) and pyramidines-cytosine (C) and thymine (T). The hereditary of living organisms is contained in the DNA.

DNA is able to replicate, which is a process in which DNA makes exact copies of itself. However, the replication type used is semi-conservative replication which gives two daughter duplex DNAs, each of which contains one strand (blue) and another strand (red).

In addition, DNA replication is an act of duplicating a cell’s genome and this is required every time a cell divides. Replication requires specialized proteins (food) for carrying out the job.

Meanwhile, proteins tend to support evolution. They contain amino acids common to all living cells. When amino acid sequences into living cells, evolutionary relationships are formulated. The process of evolution produces DNA sequences that encode proteins with very specific functions. Two daughter DNA strands are formed from one and exactly the same as the parent, preserving the integrity of the genes as the cells divide. The DNA molecule is a double strand that contains base pairs on the inner part.

The bases are the information that contains the information part of the DNA, because they form sequences. The genetic code of the DNA is triplet which is used in the system.

4. RESULTS

Initial populations of randomly generated candidates comprise the first generation. The candidates start their lives as cells. They start roaming about in the designated area. They move around until they found source to start working on (that is food). Once this is done, they start traversing back and forth, up and down the source; until it is eaten (the food is depleted).

They continue moving around within designated area based on their energy looking for food for survival and adaptation. Once the food is depleted, they start searching for a new source (food). Once they are getting close to the source as a result of their range value (i.e. radius of the cell), their speed got increased to acquire the food before another candidates get to the source (food).

They move randomly around but less frequently preserving and expending their energy. The size of each individual is based on the source. Once a candidate exhausts its energy and could not withstand the hazard environment, it dies off.

Mutations are applied directly to a parent of a creature. Offsprings (i.e. two daughter cells) are therefore reproduced asexually (binary fission) with varying mutations. Although these mutations can be controlled, unlike nature and be forced to move more rapidly or slowly.

The mutation alters the genotype of individual candidate (cell) in a random way. These new candidates compete with old candidates for their place in the next generation (that is survival of the fittest). These mutations improve some individuals and hurt some other candidates. The less-fit individuals die off, leaving the more fit to reproduce and create even better versions of themselves.

Table 1: Variables used in the Simulation

Creat.	Lifetime	Size	Range	Speed	Parent	AncestralID	DNA Sequences
0	2417	24	39	0.5	n/a	n/a	10301221101321113122100З
1	1674	22	36	0.62	n/a	n/a	303003032210130130122102
2	2381	28	40	0.5	n/a	n/a	221233232321030012313000
3	2478	30	37	0.64	n/a	n/a	222030231311332021130322
4	2355	24	28	0.62	n/a	n/a	033102121311101321130103
5	2684	24	44	0.58	n/a	n/a	112030032310133322222003
6	2373	32	40	0.46	n/a	n/a	013202233201003222331202
7	1727	16	28	0.58	n/a	n/a	023310200100331201313201
8	1307	30	44	0.38	n/a	n/a	030100130331232311032223
9	1534	26	39	0.4	n/a	n/a	122131222100001332022323
10	1857	24	33	0.42	n/a	n/a	012000120311101333021300
11	2464	28	45	0.52	n/a	n/a	112112113212111033200321
12	2137	16	24	0.7	n/a	n/a	200213133322002122001300
13	2784	20	33	0.66	n/a	n/a	123031302201322332122121
14	2273	28	39	0.48	n/a	n/a	111320020320310223232301
15	1928	24	35	0.62	n/a	n/a	210212323022310303233033
16	1376	22	40	0.46	n/a	n/a	323003322112320003322033
17	1219	22	32	0.52	n/a	n/a	323221331212312111231322
18	1797	22	33	0.48	n/a	n/a	130320010000112011330312
19	1761	14	26	0.6	n/a	n/a	023003201011011200332030
20	539	24	39	0.5	0	0	103012211013211131211003
21	1309	24	38	0.62	1	1	303003032210130130122102
22	1012	28	40	0.52	2	2	221233232321030012313000
23	2008	32	39	0.64	3	3	222030231311333021130322
24	1561	24	28	0.66	4	4	033102121311101321130103
25	811	24	47	0.58	5	5	112030032310133322222003
26	1745	34	43	0.46	6	6	013202233201003222331202

Table 2: Variables used in the Simulation

Creat.	Lifetime	Size	Range	Speed	Parent	AncestralID	DNA Sequences
0	1079	18	29	0.38	n/a	n/a	3020332310102232122122011
1	2147	26	40	0.46	n/a	n/a	0201111212223030102311301
2	1254	26	38	0.52	n/a	n/a	0232023023222101033333122
3	974	22	27	0.5	n/a	n/a	3312012220230330221030303
4	2347	24	33	0.64	n/a	n/a	0001111010123321103013002
5	2981	28	39	0.52	n/a	n/a	0313222032312030121300002
6	2281	22	35	0.54	n/a	n/a	2213012203110201133002111
7	2579	34	47	0.44	n/a	n/a	0303122132033213332330122
8	2653	30	37	0.62	n/a	n/a	0130130023032233031213002
9	2137	26	38	0.66	n/a	n/a	3033023311002023020331211
10	1097	20	32	0.46	n/a	n/a	0110133103100030213110011
11	3561	28	41	0.8	n/a	n/a	2213310303201300003312111
12	1657	14	26	0.54	n/a	n/a	3231201002211122212121111
13	2259	24	41	0.52	n/a	n/a	2323033230023320112002233
14	2487	40	50	0.54	n/a	n/a	1300212313333200331313122
15	1219	18	35	0.46	n/a	n/a	3111101300002212210312122
16	1167	24	40	0.5	n/a	n/a	0203313000033202322122333
17	1237	30	39	0.4	n/a	n/a	3221030100012111133303000
18	1149	22	34	0.44	n/a	n/a	1012322300020121322102322
19	2517	32	44	0.54	n/a	n/a	0331013020110311222023300
20	609	18	29	0.38	0	0	3020332310102232122122011
21	1824	26	39	0.48	1	1	0201111212223030102311301
22	504	26	38	0.5	2	2	1102310230232221010333311
23	224	22	28	0.54	3	3	3312012220230330221030303
24	1938	24	33	0.58	4	4	0001111010123321103013002
25	1593	30	41	0.52	5	5	0313212032312030121300002
26	1247	22	35	0.52	6	6	2213012203110201233002111
27	1283	34	47	0.44	7	7	0303122132033213332330122
28	2522	30	37	0.62	8	8	0130130123032233031213002
29	1335	26	38	0.66	9	9	3033023311002023020331211
30	347	20	34	0.46	10	10	0110133103100030213100011
31	2505	28	40	0.8	11	11	2213310303201300003312111
32	364	14	26	0.56	12	12	3231201002211222212121111
33	1745	24	41	0.5	13	13	2323033230023320112002233
34	1631	38	48	0.54	14	14	1300212213333200331313122
35	469	18	34	0.44	15	15	3111101300002212210312122
36	557	24	40	0.5	16	16	0203313000033202322122333

5. DISCUSSIONS

As it can be seen from Table 1 above, the zeroth to the nineteenth creature indicates the first generation of creatures and do not have either parent or ancestor from whom they descended. The twentieth creature (i.e. creature 20) inherited almost all the attributes of its parent (i.e. creature 0). Although, there are some variations in DNA instructions of offspring and parents, the reason for this difference is as a result of mutation on one or more of the DNA bases. Each creature is born with these attributes (i.e. size, range (how far it can see) and speed).

Similarly, there are also variations in the phenotypic attributes of one offspring and its parent; this is due to adaptable behaviour of the offspring. Since, every creature has the ability to think and respond to specific conditions in the environment. They all have capacity to reproduce, based on heritable instructions that are encoded in their DNA. And their DNA instructions do change as a result of mutation.

The process of the cell life is based on its behaviour (i.e. feeding, movement, and replication). As a cell is born, takes in food in each time cycle, and the size of the cell increases by a given amount. Thereafter, if the amount of the food reserve in the cell is less than a given part of its food capacity, the feeding begins. Movement and reproduction take place when their conditions are favourable. This life cycle terminates because of either replication or death. Death, in its turn, may happen through a lack of food in the environment.

Table 2 depict similar scenario but generates different values for the creatures' attributes. Although, there are some instances in the two tables where two or more creatures have the same value for their attributes. It shows the consistence of the process of the simulator.

CONCLUSION

The study of artificial life has help in the simulation of self-replication and natural selection, which invariably assists in eliminating or reducing trivial self-reproduction or un-inheritable mutation.

It has really helped to understand evolutionary processes through computer simulation. It assists in understanding the concept of survival of the fittest through modeling of evolution.

Meanwhile, computer simulation of artificial life is used to study the essential features of living systems (cells) such as evolution, adaptive behaviors etc.

Artificial life simulation has made one to understand the complex information processing that defines living systems. Although, living cells are considered for this work, but represented the whole living systems, due to the fact that, they are the fundamental and basic units of life. This study may offer benefits of reducing red blood cell related problems such as blood cancer, sickle-cell anaemia etc. However, further study will be required to fully implement the study for proffering solution to the related problems of blood cells.

ACKNOWLEDGEMENTS

We thank Dr. Tannell Helland for the permission given to make use of some of his materials for the work. We are very grateful.

REFERENCES

[1] Majd, S.M.A Shariatzadeh, Cell and Molecular Biology, 4thm edition, pp. 23-29 ISBN 964-7006-35-7, Aeeizh, 2004.

[2] Abbas P. D., Bahar K.L., Mozhdeh N., “An artificial life model inspired by the Basic single-cell organism”, International Journal of Computers, Communication and Control, Vol.3, pp. 253-258, 2008.

[3] Ostman, A. Hintze, C. Adami, “Population and Evolution: Critical Properties of Complex Fitness Landscapes”, proceeding of 12th International Conference on Artificial Life, (MIT Press, 2010), pp. 126-132.

[4] Cariani P., “Emergence and artificial Life”. In Artificial Life II Christopher Langton (Ed.). Addison Wesley. pp. 775-797, 1992

[5] Christopher G. Langton, “Studying artificial Life with cellular automata”, Physica D, 22, 120-149, 1986.

[6] Daniel C. Dennett, Darwin’s Dangerous Idea: Evolution and the meaning of Life: New York, Simon and Schuster, London; Allen Lane, 1995.

[7] Eric Bonabeau and Paul Bourgine, “Artificial life as it could be”, World Futures 40(4): pp. 227-249.

[8] Francisco J. Varela, Paul Bourgine; Towards a practice of Autonomous Systems; proceeding of the First European Conference on Artificial Life, MIT Press, pp. 273-280, 1992.

[9] Reynolds C.W: Flocks, herds and schools: A distributed behavioural model. ACM SIGGRAPH Computer Graphics, vol. 21, No. 2, pp. 25-34, 1987.

[10] Richard A. Watson (et al.), “Adaptation without Natural Selection, Natural Systems”. In proceeding of the ALife XII Conference, Odense, Denmark, 2010.