**Cyber-biology-project**

**General description**

The project aims to expand upon the classical cellular automaton model by simulating classical evolutionary principles via addition of genome based sexual reproduction to each ‘cell’ of an organism. Each cell in the proposed simulation will have an arbitrary long genome which would define the way a ‘cell’ would interact with the outside environment. How it would react to its surroundings; compete for resources, interact with other cells and reproduce.

**Implementation**

**1.Environment**

Let our environment be represented as a 2 Dimensional plane, subdivided into pixels of equal size. The plane size is planned to be made parameterizable, it would be possible to select a size before starting a simulation. Another important feature of such a plane would be the possibility to make it loop on the x or/and y axis (Illustration 2). The point of this feature is to avoid system’s impairment around the border region (which is a very common problem for many similar situations).

Each pixel of the environment would be able to host a cell, just like in vanilla Conway's game of life. But unlike CGL it would also have the ability to store energy. See illustration 1.

| 0,c | 0 | 1 | 1 | 4 |
| --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 3,c | c |
| 1 | 8 | 2 | 1 | c |
| 4 | 1 | 2,c | 2 | c |

Illustration 1

Example field. Structure: energy, cell. Each cell of the above table would be a cell in the simulation environment.

| 1 | 2 | 3 | 4 | 5 |
| --- | --- | --- | --- | --- |

Illustration 2

The environment is wrapped around, so neighbours of 3 are 2 and 4. While neighbours of 5 are 4 and 1.

**2.Energy**

Since in the simulation we are making ‘game-of-life’ simulation we will also need to revoke the live/die/reproduce criterion to make it work adequately with improved ‘cell’ mechanics. For that let us define the notion of “Energy”. In my simulation every cell would need energy to live/reproduce, or take any other actions. If a cell runs out of energy it dies. The energy costs would be subdivided into 2 main categories. Maintenance fee, an energetical cost a cell would pay just to exist - it would be based on genome and organism characteristics (stronger ones might need to consume more). The second category would be action fees - how much an organism would pay per action.

In our simulation (at least during early stages of development) the energy distribution would be similar to the one of digital tree evolution projects. Aka, there would exist an energy source, (let us call it ”Sun”) which at the start of a new round would add ‘x’ energy points to every pixel on the plane of our environment. The x would not be constant, contrary to that it would be possible to express x as some periodical function with trigonometric elements, to simulate day/night, or seasonal cycle. In the future by controlling the value of x it would be possible to experiment and see how life adapts to rapid changes to the environment.

Any cell would also have some energy stored to consume when needed. When a cell dies it leaves some energy behind on the tile it died at. That would also allow to implement predatory behaviour into the simulation, together with parasitism and would allow for more complex interactions and more dynamic behaviours of cells in general.

**3.Reproduction**

**a.Sexual vs Assexual - reasonings**

My simulation would permit both types of reproductions, because both are present in the real world as well. The type of reproduction an organism would be practising would be specified in its genome.

From my research/previous experimentation I was able to come to the conclusion that sexual reproduction is in general significantly better - since it allows populations to evolve faster. Imagine the situation, an optimal genome is 12345 (highest adaptability). And in the asexualy reproducing population there appeared (via mutation) two organisms 02000 and 10000, both of them have element of correct sequence, but only 1 of them would win in the population, since in the asexual reproduction there would be no way to combine them together quickly.

| Assexual | Step x (base step) | Step x+1 (mutation) | Step x+n (all others died) |
| --- | --- | --- | --- |
| Sample genome 1 | 00000 | 02000 | 02000 |
| Sample genome 2 | 00000 | 00000 | 02000 |
| Sample genome 3 | 00000 | 10000 | 02000 |

| Sexual | Step x (base step) | Step x+1 (mutation) | Step x+n (descendants of 02000 and 10000 dominate) |
| --- | --- | --- | --- |
| Sample genome 1 | 00000 | 02000 | 12000 |
| Sample genome 2 | 00000 | 00000 | 12000 |
| Sample genome 3 | 00000 | 10000 | 12000 |

One large population sexual reproduction fixes this problem by having the most adaptable cells split their genome evenly to send to the next generation preserving successes of all members of the system, instead of just one most successful. Of course there are problems and disadvantages, like per say the fact that it might happen that “good part of” genome is simply not selected on the next stage of offspring, but these problems are simply covered by the large population size, where by law of big numbers the population would be able to evolve correctly.

**b.Asexual reproduction details**

In the asexual reproduction a ‘cell’ would just create an offspring on one of the sides (where exactly specified by genome). The genome would be taken from it’s parent, with some random small altering (mutation).

Step 1

| Empty cell | Empty cell | Empty cell |
| --- | --- | --- |
| Empty cell | Cell A | Empty cell |
| Empty cell | Empty cell | Empty cell |

Step 2

| Empty cell | Empty cell | Empty cell |
| --- | --- | --- |
| Cell A offspring | Cell A | Empty cell |
| Empty cell | Empty cell | Empty cell |

**c.Sexual reproduction details**

In sexual reproduction a cell would find a partner capable of reproducing. They will split the genome 50/50 and create an offspring in any cell adjacent to parents.

Step 1

| Empty cell | Empty cell | Empty cell |
| --- | --- | --- |
| Cell P1 | Cell P2 | Empty cell |
| Empty cell | Empty cell | Empty cell |

Step 2

| Cell C1 | Empty cell | Empty cell |
| --- | --- | --- |
| Cell P1 | Cell P2 | Empty cell |
| Empty cell | Empty cell | Empty cell |

**d.Lifespan**

A cell would have a life expectancy, the average would be specified in the genome. It would be of a standard Gaussian distribution. So each individual life length would be somewhat unpredictable. For example if life expectancy of an organism is 20 turns (and standard deviation of 1). It could for example live for 18, or 21 years, not just 20.

**4.Simulation order**

The naive solution to the problem would be to make a cell by cell implementation, but it would give cells that start an unfair competitive advantage. Also the newborn cells, they would either be over privileged or placed in an unfortunate position. So classical list stack implementation would be flawed.

a.Approach 1

1->2->3->4->5 … -> 1->2… If new element appears it would be in trouble because all other cells would be able to take ‘action’, whilst it would not be able to do anything.

b.Approach 2

2->3->4->1->5 - an approach would be to select sells randomly, sampling without replacement. But that would be very chaotic on a smaller simulation, and would make the project much more difficult to debug. Plus there is an extremely high risk that the simulation would go out of sinc, killing complicated organisms like colonies and composite nomads.

b.Approach 3

Making a linked list, where every cell would have a next/previous attribute. Each time a new offspring would be born it would be inserted after its parent. Each time a cell dies it is simply removed and the chain is restored. Example.

Cell 1 (Prev 4 Next 2)

Cell 2 (Prev 1 Next 3)

Cell 3 (Prev 2 Next 4)

Cell 4 (Prev 3 Next 1)

Assume Cell 2 Died, we simply restore connection, since due to the chosen data structure we can restore the flow without damaging integrity. Reconnecting cells 1 and 3.

Cell 1 (Prev 4 Next 3)

Cell 3 (Prev 3 Next 4)

Cell 4 (Prev 3 Next 1)

**5.Genome and mutation**

Genome is fundamental for our simulation, in the genome all the characteristics of a cell would be stored. (Reproduction, life expectancy, behavioural preferences). Here is how I am planning to structure it.

Example GENOME

| GENOME | Val 1 | Val 2 | Val 3 | Val 3 | Val 4 | … | Val x |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Gene 1 | 002 | 232 | 212 | 100 | 011 | .. | 071 |
| Gene 2 | 093 | 081 | 122 | 062 | 100 | … | 164 |
| Gene 3 | 051 | 023 | 205 | 111 | 034 | … | 078 |
| Gene 4 | 044 | 082 | 142 | 001 | 044 | … | 195 |
| … | … | … | …. | ….. | ….. | … | … |
| Gene n | 172 | 010 | 029 | 091 | 156 | … | 255 |

a.General info

Genomes would be expressed as a sequence of genes (16/32/64 per genome, have not decided on the exact value yet). Note that by genome I do not mean the classical scientific definition of a “building unit”, but a complete set of characteristics/instructions for the cell. During the simulation it would be possible for the cell to switch between different behaviours.

b.Individual genome structure

The cell will have some physical characteristics

-Health points

-Energy storage capacity

-Livespan

And many behavioural characteristics.

-Movement

-Aggressiveness

-Reproduction

-Energy management

Each val from the table above would represent one of these characteristics. During each new turn the execution might go as follows. Suppose the gene (tiny part of it) is structured like this:

121-001

We first take a random number from 0 to 256. If it is above 121 the specified action would be taken. The genome would be big enough to account for all potentially takeble actions.

c.Active genes

Is an idea I saw on Simullife Youtube channel, which also does cyber biology. There to allow for many different behaviours the gene is slit into a table of many smaller genes responsible for fewer interactions and then the cell picks the active genome that suits it the best depending on the environment.

d.Energy management

To allow for more complicated life forms the cells would have the ability to transfer energy to each other to allow specialisation. Because the simulation would be made in a way in which if a cell is good at extracting resources from the outside world it would have debuffs on a physical level, promoting diversification and formation of more complicated systems.

For example the cell which is specialised at extracting resources would not be able to attack others. But would have an opportunity to give energy to offspring, which in turn might switch to predatory or parasitic behaviour.

The genome would be subdivided into 2 different parts. The first part would concern the physical characteristics of an organism. How fast is it? Can it eat other cells? What size will it have? From there we would derive an energy cost for an organism to last for the next round.

The second part of the genome would be concerning the behavior of an organism. Different biases (Granivorous or Carnivorous) propensity to stay still, propensity to nomad lifestyle. Propensity to symbiosis etc. From a practical standpoint I have further developed my concept, each organism would be similar to a small neural network (1 internal layer). Where input would be environment characteristics, whilst the output would be all possible actions which a being can take during a turn.

This implementation's big advantage is the fact that the creature would have the starting characteristics and biases in its genome (starting weights and balances), whilst would have an opportunity to adapt and change it's weights and balances during the simulation. (Have not decided on how to do this exactly, probably would be some process similar to mutation, but instead of being random it would be guided by immediate reward system)

**6.Observations on simpler systems**

Looking at most basic simulations I have discovered that there are only 3 scenarios that can happen when dealing with trivial simulations.

Scenario 1

Perma-Statis. The simulation reached stability (or loop with step 1), where either all cells died, or all cells stuck at their static places.

Scenario 2

Loop. Conway’s game of life glider is a prime example of cool. Cells engage in repetitive behaviour, each n steps they return to the same configuration (might move as a whole).

Scenario 3

Bifurcation, chaos. The system never stabilises and just acts stochastically.

My final product is expected to end up in one of the following scenarios. But before I expect some time of ‘interesting development’ when different life forms appear and compete for resources. Since the simulation is complex, the time before an appearance of an ultimate gene is expected to be somewhat long.

**7.References Bibliography**

1.Stephen Wolfram - A new kind of science (Book on cellular automatons)

2.Simulife Hub (Youtube channel focusing on making different cyber biology simulations)

**8.Conclusion**

I already have an idea on how I am planning to proceed with the development. The calculations are going to be executed using matrices/tensors, whilst there would exist a separate visualizator program, which would make us able to observe the situation with our own eyes.

My question is whether the project has potential. Since in the lecture you said that you previously worked in a similar discipline. I would like to hear your opinion/critiques on it. Whilst it is still in the draft stage. Thank you very much for your time!

PS When finished I will share the project with you.