Spontaneous emergence of the intelligence in an artificial world

Istvan Elek, Janos Roden, Thai Binh Nguyen

Eotvos Lorand University Faculty of Computer Science, Budapest, Hungary
elek@map.elte.hu,
home page: *http://mapw.elte.hu/elek/*

Abstract. The aim of this paper is to introduce the principles of the evolution units, which are named the digital evolutionary machines. The entities will be constructed based on these principles. Their properties and abilities will be shown with some experimental results. The knowledge base of DEMs was stored in a database, which is the source of the individual knowledge graph. It helps to make decisions of each entity. Their life was observed and analyzed. Some interesting analysis and charts will be shown in order to understand how DEMs work in the artificial world. Finally the DEMs geographical extension will be shown.

1 Introduction

The evolution modeling projects focus on multiplication and the competitive exclusion theory, which is one of the most famous low after Darwin in the evolution biology. The competitiv exclusion theory comes from Georgii Frantsevich Gause, 1932 [12], who was a Russian biologist in the 20th century. An important low of the evolution process, that is irreversible, which was revealed by Dollo [6], and expounded by Gould, 1970 [13]. The key moment of the evolution is the emergence of the intelligence. This early intelligence was really low level, but it can help to survive the fluctuating environment. The changing Earth and its climate were the most remarkable challenge for the ancient organisms.

Many papers (Adami, 2002 [1], 2003 [2], Ofria at al, 2004 [4], Ostrowski, 2007 [15]), dealing with artificial organisms emphasized the importance of complexity. Not only life shows serious complexity. There are human made constructions as well, such as the topology of Internet, that have increasing complex structures with fractal properties (Barabasi, 2002 [5], 2003 [3], Barabasi–Newman, 2006 [14]).

The multiplication and the competition are in the center of works mentioned above. The evolution of organisms‘ body and their intelligence can not be independent. While we are going to model the organisms' evolution, we need to model the development of their intelligence too. There is no right definition of intelligence of the organisms (Turing 's definition does not serve path). Instead of making useless definitions, it is more promising to set up some essential principles that regulate the process of collecting and interpreting data from the surrounding world. There are many research papers and books that describe machines which collect data from the surroundings and they have some kind of remembrance (Russel, 2002 [16]). Their ability of interpretation of the environment is restricted. The limits of these constructions are obvious: their intelligence never becomes similar to that of mammals or octopuses.

This paper is about a computerized approach of the digital evolutionary machines as the elements of the evolution process. The principles of them were established in Elek, 2008 [7], 2009 [8], 2010 [9], [10] and 2011 [11].

2 Knowledge and evolution

Paleontology and geology serve many exciting examples of the one way evolution. The time flows in one direction, forward. Life always tries to adapt itself to the circumstances, mainly to the weather (temperature, seasons, climate). If the climate has changed, the adaptive organisms also change their right properties, skills and manners. If one million years later the climate changed back, the evolution did not remember the previous stage of the environment. The adaptivity also produces new properties that help to survive the changes. The evolution seems to be recursive. It means the regulation is a kind of recursion, like a feedback. The current stage is the basis of the next step. Evolution never gets back. There are no general physical laws in the background where the processes can be computed from.

If an organism did not recognize the enemy, the food and the environment, it had died soon or immediately. Consequently, organisms had to collect data from their surroundings (perception) and interpret them (understanding). The remembrance is probably one of the most important components of understanding the world. The irreversible time, i.e. the serial recording of events and data of the environment produces a huge database. It contains everything that happened to an organism in its lifetime. The key of surviving is the interpretation of the surrounding world and the action.

The weather and climate were the most effective factors of the evolution of organisms. Every organism needs to interpret the measured and stored data of the environment if they wanted to survive. This ability required a huge database containing everything that ever happened to the organism. This is the experience based knowledge base.

3 Principles

Biologists say there is a kind of self organization among protein molecules in vitro. Even if we assume it was the first step toward a real organism, it is evident there is no huge knowledge database and complicated interpretation logic in it. Protein molecules seem to be inclined to form combinations. Their data collection logic also has to be simple. Consequently a knowledge database can be constructed from simple steps of data collection. The algorithm of the knowledge graph search has to be simple.

Let us construct the knowledge-graph. It consists of atoms and their connections. Atoms are the nodes and connections are the edges of knowledge-graph. Let us define a context that includes arbitrary atoms of the tree; consequently, the structure of the graph is not predefined. It depends on the atomic connections that depend on a time series, when events took place in time order one after the other. The general principles declared the equality of atoms and contexts. In other worlds a context can contain simple atoms or other contexts as well.

Let a_{ij} denote the i-th atom in the j-th context which contains N atoms. Let the quantity of the knowledge of a context (k^j) be the sum of the quantity of the knowledge of every atom in it:

$$k^j = \sum_{i=1}^{N} a_i^j \tag{1}$$

The knowledge-base has two basic functions: receiving a question and answering. The key of the problem is to find the path from the questioning node to the answering node in the knowledge-base.

Let l_{ij} denote the strength of the connection between a_i and a_j . Let $l_{ij} = l_{ij} + 1$ if the tour produces good result and $l_{ij} = l_{ij} - 1$ if the result is bad. This logic makes good paths stronger and bad ones weaker. Look at the u and f nodes in the knowledge-graph. The right distance definition depends on the length of the path along branches, so it is graph tour based.

Let a_i and a_j be two nodes of the knowledge-graph where the path includes $m = j - i$ atoms between them. Let d_{ij} be the distance of these two nodes, let the strength of their connection be denoted by l_{ij} , which is the reciprocal of the sum of the strength of the connections between a_i and a_j. The stronger the connection between two nodes, the closer they are.

The goal of the knowledge-graph is to answer a question arising from the circumstance. How to find the right path from the questioning node to the answering one? The fastest is the right path, probably. This logic produces very fast reaction in well known problems and may result fail in unknown cases. The fail means unsuccessful escape, capture or something important for the organism. If it survived the situation, i.e. the reaction was successful, and the path that produced the success became stronger. If it did not survive the situation or the result of the action was failed, i.e. the result of the action was unsuccessful, the organism was knocked out or a path in the knowledge-graph became weaker.

3.1 The knowledge representation

Let $\mathbf{K}$ denote the knowledge-base which consists of n atoms a_i, $a_j \in \mathbf{K}$. Let us name it the knowledge-matrix.

$$\mathbf{K} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \tag{2}$$

Some of the atoms are in touch with other atoms in **K**. Let us describe the links of a_i and a_j atoms with l_{ij}, where

$$l_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \text{ and no link between them} \\ u - v & \text{else where } u \text{ succeeful and } v \text{ unsuccessful} \end{cases}$$

Let us organize the atomic links into a matrix form, and name it a link matrix and denote it by **L**. The elements of the link matrix are l_{ij} that describe the link of the atomic pointpairs. **L** is diagonal ($l_{ii} = 1$) and describes the relationships of the atoms in the knowledge-matrix:

$$\mathbf{L} = \begin{pmatrix} 1 & l_{12} & \dots & l_{1n} \\ l_{21} & 1 & \dots & l_{2n} \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \dots & 1 \end{pmatrix} \quad (3)$$

1. Every atom has one link at least.
2. Let **C** be a context. $\mathbf{C} \subseteq \mathbf{K}$, i.e. **C** consist of any atoms of **K**.
3. Any atoms can be the member of any context.
4. Any contexts can be a member of any contexts. In this case, the knowledge-matrix (**K**) is a hyper matrix where matrix elements can be matrices.
5. In summary an atom can be a
   (a) Simple atom that is really elementary and belongs to one context.
   (b) Multi-member atom that is also elementary but belongs to more than one context.
   (c) Aggregated atom that is a kind of simple atom, but its value is a representation of a context. In other words, its value is a determinant or a spur of the knowledge-matrix of a certain context.
   (d) A complex atom that is a context that includes any kinds of atoms above.

3.2 Properties of the knowledge base

1. The knowledge-base is a continuously increasing database which stores everything that happened to it. This is a one way process. The knowledge-base is different from entity to entity.
2. There is no changing or erasing function in the knowledge-base.
3. Let the data collection be extensive if a DEM-entity perceives the circumstance and stores data. There are some important consequences of the extensive mode:
   (a) Every individual knowledge-base is different. It depends on the career of a certain DEM-entity. There are as many DEM-entities as many kinds of knowledge-base exist.
   (b) If some changes happen that certainly produces the same circumstance in the past, the previously recorded atoms has not been changed, simply a new atom has appended the end of the knowledge-base.

4. Let us have another mode that was named intensive, when there is no perception. It is a meditative stage when the knowledge-base acquires the data came from the extensive stage. The result of the data acquisition in intensive mode may produce new contexts, faster path, more reliable work. This stage is extremely important in learning the circumstance.
5. The feedback is a process when a DEM-entity is informed about the result of its reaction. The result of this process is success or fail. As mentioned previously, the success/fail makes stronger/weaker a certain path in the knowledge-graph of the knowledge-base. The DEM-entity's knowledge-base becomes much stronger/ weaker if the feedback such as rewarding or punishment comes from an external intelligent entity, because its knowledge-base can be considered as an included context.
6. In the history of the Earth there was never only one organism. There were always ensembles. Ensembles make organisms competitive. Competition results in different skills i.e. different knowledge-bases.
7. Different circumstances cause different experiences for the organisms.

The evolution is a process in time and in space dimensions. Regarding the results of paleontology it is known that evolution is recursive, i.e. an evolution step depends on the previous step only and it is a one way process, it can not be turned back. This is known as Dollo's law in paleontology [6]. This law was first stated by Dollo in this way: *An organism is unable to return, even partially, to a previous stage already realized in the ranks of its ancestors.* According to this hypothesis, a structure or organ that has been lost or discarded through the process of evolution will not reappear in that line of organisms.

The development means not only physical but mental changes as well in an organism. Mental means intelligence in this context. The adaptivity comes from its knowledge base. The quality of the knowledge base has to influence the physical properties also. Consequently the physical and mental evolution work collaterally, since the organisms live in various climate, in many challenges, having different experiences with different chances with various knowledge but in many, different places at the same time.

4 Internet is the artificial world

If we are willing to create experienced DEM-entities there is no wizard unfortunately. There is no a recipe to install them from little pieces. How to construct a DEM-entity? Probably, it is impossible to construct only one of it. First we should create an ensemble from many initial DEM-entities and leave them to live in their own world.

We have two access points to this problem. The first one is to construct an artificial circumstance where DEM-entities live in. The second one is to construct many initial DEM entities with simple perception and interpretation functions. Let us look at some of the details:

1. The artificial world (AW) has some essential properties that define the frame of this world.

   (a) Let AW be huge where circumstances have spatial dependencies. Regarding the size of this world, environmental parameters are obviously different. If we leave DEM-entities alone in this world they will have different experiences and different knowledge-bases because of climatic differences.
   (b) If the AW is huge enough, the survival strategies will be different. One of the DEM-entities escapes from the unfriendly circumstances but others try to adapt. Different strategies result different knowledge-bases.
2. If there are many DEM-entities on the same territories, what is the consequence?
   (a) There are many DEM-entities who try to get better strategy in order to be more successful than others. Someone gets advantages but someone gets disadvantages since it fails to answer a certain question.
   (b) Regarding the different DEM-entities and unique knowledge-bases, many different strategies can coexist in the AW. Consequently, many different strategies can be successful at the same time. Someone preferes the escape, but someone the competition.
3. Many DEM-entities will have many different knowledge-bases.

The question is how to construct the prototype of a DEM-entity? Before the construction of the prototype, let us create the artificial world that will be the space for DEM-entities. If AW has been created already many DEM-entities should be available to start up in it. Properties and abilities of DEM-entities were introduced previously, so the task is to make their software representation. Regarding the quantity of DEMs and the huge sized world with different spatial properties may result many formed DEMs, and these entities have different knowledge-bases. An intelligent DEM-entity can exist in ensemble only. There is no lonely intelligence because it is a collective product that is realized in some successful entities' knowledge bases.

4.1 The computerized approach

This paper focused on the appearance of the intelligence and its properties in a digital representation. Digital evolution machines and the Internet, which is the artificial world, where they live in , are in interaction. The purpose of this work is to model this interaction. The task is to store everything sequentially that happened to DEMs, and to construct the knowledge graph, which nodes are the pointers for sequentially stored events.

Briefly the system works as follows:

1. There are a lot of DEMs in the Internet. An individual DEM starts his carrier in a starting URL.
2. This DEM investigates the certain URL, what kind of files are linked to this. Some tags in this file point to another URL-s (< a ref = http://...>). Some of the pointed files are considered as resources which are needed to DEM's life (for instance *doc, pdf, txt, jpg, gif, tif* files are considered resources with various energy content, but *exe, bat* files are considered poisons, and so on).

3. A successful step, when the reached resource produces energy input for a certain DEM, makes path stronger. If the reached node (file) is neutral or contains poison, the path becomes weaker.
4. if a DEM visited a resource, the node (file) runs out for a short period (it needs time for recover itself), so DEM has to move to another URL (for instances, it follows an < a ref = http://...> html tag) for further resources.
5. go to 2

Last but not least, in order to find the rigth URL for the next move, we need an *adviser*, which is one of the most remarkable component of the system. *Adviser* uses the knowledge graph, and give the most promising URL for a DEM. Summarily sequentially stored events and *adviser* functionality figure out a certain DEM's intelligence.

For the technical implementation, there is a MS Sql server, which stores sequentially the every day events, (the table name is *Logbook*). The next table *DemWorkers* stores data for the identification of individual DEMs This table contains the resource requirements also of a certain DEM. The visited web nodes are stored in the table *ArtiworldNodes* A related table, named *UrlsContent* stores the link of an URL (resource files, and potential URLs for further moves). The table *ContentTypes* contains the possible energy sources (food) with energy content units. You can change properties of the artificial world via table *ContentTypes* if you want the energy files to have different energy content. For example a pdf file has higher energy content than a simple txt file. You can adjust the world more, if the energy content depends on the file size too. Do not forget, that the world structure, its topology is a fact, you can not change this property directly. If you define resctricted domains, the topology is changed for the DEMs, because these domains can not be reached as energy source.

DEMs will affect the artificial world, and AW influences DEMs life. This interaction is the target of our research. The system is working, but not finished yet, of course, because this is a changing database. The project info page can be reached on *http://dem.elte.hu* site. Here sooner or later you can reach the database and observe what DEMs do.

4.2 Results

In this section some figures will be shown, which were made from a program package. The purpose of this system is to handle DEMs' life, such as movements, destiny, status and so on. The other task of the system is to register the circumstance, URLs and their contents. This is a kind of projection of the known artificial world, where DEMs have already visited.

Look at the figure series which show the DEMs' life. First, 1000 DEMs were started. The properties of DEMs can be set. In addition we can set up any domain disabled if it is required. This set is obvious, regarding the early evolution stage, when the chemical evolution was typical. It is extremely interesting whether they have any chance to survive.

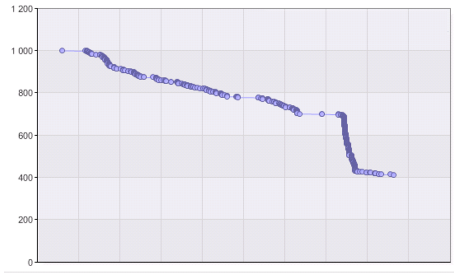

Fig. 1. Alive DEM workers after 5 minutes from start. X, Y axes means time and population (started with 1000 DEMs)

Look at the dem workers' status in the following figures (fig. 1, fig. 2 and fig. 3), which show those workers who alive certain time intervals.

It is surprising that at the beginning DEMs die in large quantities. The most of URLs are not green grass for hunger DEMs, so they have no foods. They obviously perished.

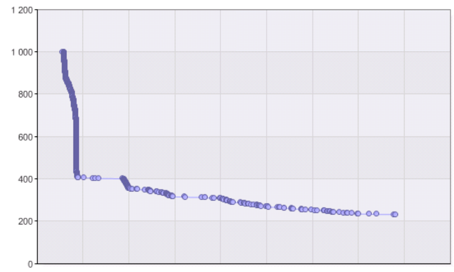

Fig. 2. Alive DEM workers after 120 minutes from start

If you study the figures (fig. 1, fig. 2 and fig. 3) you will find, there is a small population which survived the artificial deserts and all troubles. There were 5 runs with 1000 DEMs, and the results were the same. The consequence is really interesting. Generally the 85 % of DEMs perished after some days, but there remains a little group with 15 % of the starting population, who survived the troubles, and they will be alive for a long time, if the circumstances are steady.

4.3 Geographical extension of DEMs

The geographical extension of DEMs is interesting. They extend almost all over the world, which can be seen on the figures and 4.

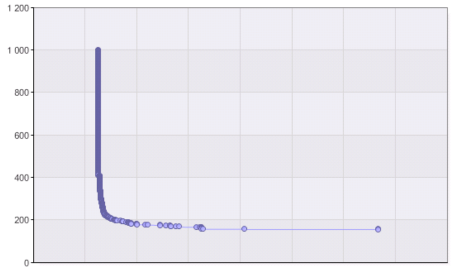

Fig. 3. Alive DEM workers after 4 days from start

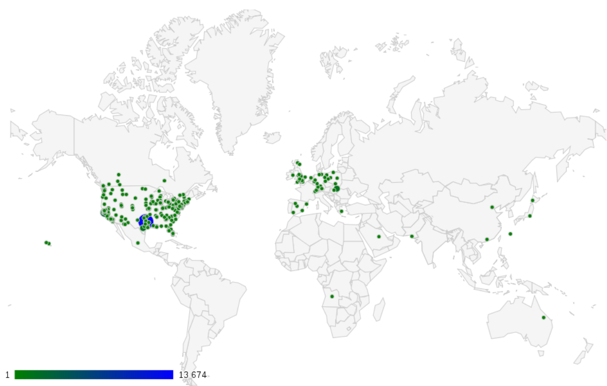

Fig. 4. Cities in the world which are visited by DEMs

References

1. C. Adami: Ab Initio of Ecosystems with Artificial Life, arXiv:physics0209081 v1 22 Sep 2002
2. C. Adami: Sequence Complexity in Darwinian Evolution, Comlexity, 2003, Wiley Periodicals, Vol.8, No.2
3. Albert, Barabasi: Statistical mechanics of complex networks, Reviews of modern physics, VOLUME 74, January 2002
4. Chow, SS, Wilke CO, Ofria C, Lenski RE, Adami C. 2004. Adaptive radiation from resource competition in digital organisms. Science (New York, N.Y.). 305(5680):84-6.
5. Barabasi – Albert – Jeong: Mean-field theory for scale-free random networks, Preprint submitted to Elsevier Preprint, 5 July 2002.
6. L. Dollo: http:/en.wikipedia.org/wiki/Dollo's_law
7. I. Elek: Principles of Digital Evolution Machines, International Conference of Artificial Intelligence and Pattern Recognition, 2008, Orlando, FL
8. I. Elek: Evolutional Aspects of the Construction of Adaptive Knowledge Base, The Sixth International Symposium on Neural Networks (ISNN 2009), Wuhan, China
9. I. Elek: A Computerized Approach of the Knowledge Representation of Digital Evolution Machines in an Artificial World, LECTURE NOTES IN COMPUTER SCIENCE 6145: pp. 533-540. (2010)
10. I. Elek: Digital Evolution Machines in an Artificial World: a Computerized Model, International Conference on Artificial Intelligence and Pattern Recognition (AIPR-10). Orlando, USA, 2010.07.12-2010.07.14., Paper 169.
11. I. Elek – J. Roden – B.T. Nguyen: Spatial Database for Digital Evolutionary Machines in an Artificial World without Knowledge-Base, Fourth International Workshop on Advanced Computational Intelligence (IWACI2011), October 19-21, 2011, Wuhan, China
12. G. F. Gause: Experimental studies on the struggle for existence, Journal of Experimental Biology 9: 389-402, 1932
13. S. J. Gould: Dollo on Dollo's law: irreversibility and the status of evolutionary laws, Journal of the history of biology (Netherlands) 3: 189?212. doi:10.1007BF00137351. PMID 116096510, (1970)
14. M. Newman – A-L. Barabasi – D. J. Watts: The structure and Dynamics of Networks, Princeton University Press, 2006
15. E. A. Ostrowski, C. Ofria, R.E. Lenski: Ecological specialization and adaptive decay in digital organisms, The American Naturalist. 169(1):E1-20-E1-20., 2007
16. S. Russel – P. Norvig: Artificial Intelligence: A Modern Approach, Prentice Hall, 2002