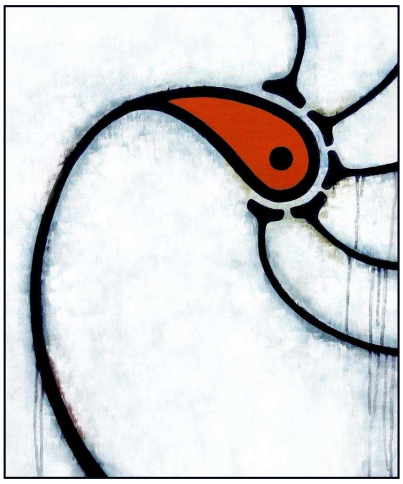

"Epigenetic tracking: an evolutionary-developmental approach to generate very large complex systems"

Alessandro Fontana (Technical University of Gdańsk) and Borys Wróbel (Polish Academy of Sciences)

Abstract—Epigenetic Tracking (ET) is a model of evolution and development of large systems built from cells. ET has been previously used to simulate growth of multicellular bodies with arbitrary 3-dimensional structures. In terms of the number of cells, the complexity of bodies generated with ET is comparable with the complexity of biological organisms. In this short report, we investigate the properties of a mechanism for increasing diversity in the population in the genetic algorithm used with ET. Although the results suggest that this mechanism does not increase the evolvabiltity in ET, their discussion provides us with an opportunity to present our research portfolio and to illustrate the potential of ET for evolution and development of patterned 3D structures. We then present a perspective for the application of ET for development and evolution of large and complex neural networks.

Presented at:

DevLeaNN 2011: A Workshop on Development and Learning in Artificial Neural Networks
Held at the ISC-PIF in Paris, France, on October 27, 28, 2011.
Edited by: T. Kowaliw, N. Bredeche, and R. Doursat
http://devleann.iscpif.fr

Epigenetic Tracking: an Evolutionary-Developmental Approach to Generate Very Large Complex Systems

Alessandro Fontana
IEEE
Vanzago (Milano), Italy, and
Graduate School at the
Technical University of Gdańsk,
Gdańsk, Poland
cnd12001@yahoo.it

Borys Wróbel
Systems Modeling Laboratory,
IO PAS,
Sopot, Poland, and
Evolutionary Systems Laboratory,
Adam Mickiewicz University,
Poznań, Poland
wrobel@evosys.org

Abstract— ***Epigenetic Tracking*** **(ET) is a model of evolution and development of large systems built from cells. ET has been previously used to simulate growth of multicellular bodies with arbitrary 3-dimensional structures. In terms of the number of cells, the complexity of bodies generated with ET is comparable with the complexity of biological organisms. In this short report, we investigate the properties of a mechanism for increasing diversity in the population in the genetic algorithm used with ET. Although the results suggest that this mechanism does not increase the evolvability in ET, their discussion provides us with an opportunity to present our research portfolio and to illustrate the potential of ET for evolution and development of patterned 3D structures. We then present a perspective for the application of ET for development and evolution of large and complex neural networks.**

Keywords- artificial embryology; evo-devo; 3-dimensional pattern formation; 3-dimensional morphogenesis; genetic algorithm

I. INTRODUCTION

Artificial Embryology models can be divided into two categories: grammatical models and cell chemistry models. In the grammatical approach, development is driven by sets of grammatical rewrite rules. Context-free or context-sensitive grammars, instruction trees or directed graphs can be used. An example is L-systems, first introduced by Lindenmayer [1] to describe the complex patterns observed in the structure of plants. The cell chemistry approach aims at simulating cell biology at sub-cellular level, by modeling chemical reactions and gene regulatory networks. The first model of this kind was proposed by Turing [2], who introduced reaction and diffusion equations to explain the striped patterns observed in Nature. Other examples of grammatical models are [3-5]; some cell chemistry models include [6-9].

Epigenetic Tracking (ET), first described in [10], is a model of cellular development that aims to combine some features of both approaches. Once coupled with a genetic algorithm, ET becomes an evo-devo method able to generate arbitrary 3-dimensional cellular structures starting from a single cell. The model has also interesting biological implications, explored in [11].

In this work we propose a mechanism for increasing the diversity within the genetic population, with the ultimate objective to improve the effectiveness of the evolutionary process. Although our results suggest that this mechanism does not influence much the evolvability of the model, discussion of these results provide us with an opportunity to present our research portfolio and to illustrate the potential of ET for evolution and development of large patterned 3D structures.

II. EPIGENETIC TRACKING AS A MODEL OF MULTICELLULAR DEVELOPMENT

Multicellular bodies in ET consist of cube-shaped cells deployed on a grid. Development starts with one cell (zygote) placed in the middle of the grid and unfolds in N steps. Cells belong to two categories: *normal* cells and *driver* cells. Driver cells have an associated variable called *mobile code* (MOC), a set of N integers. While the *genome* (a list of *genes*) is identical for all cells, the value of MOC is different in each driver cell. The MOC values can be seen as the differentiation states of driver cells and allow them to behave differently despite sharing the same genome.

Genes are composed of a left part (called MOS) and of a right part. At each simulation step the MOC values of all driver cells are compared with the left part of all developmental genes. If a match occurs, the right part is executed. Some genes cause the driver cell in which the match occurs to proliferate in the volume around it. Other genes cause cells to be deleted from the grid. Additional fields in the right part may specify the shape of the local structure created during a proliferation event (for example, a sphere or an ellipsoid) and the final differentiation state of the normal cells that will be created (for example, their color). In case of proliferation, both normal cells and driver cells are created. Normal cells fill the volume, driver cells (much fewer in number) are “sprinkled” in the volume. When a new driver cell is formed, it obtains a new and unique MOC value. Whether one of these new cells will become the

Computational resources used for this work were provided by the Polish Ministry of Science and Education (project N519 384236) and the Tri-city Academic Computer Centre (TASK).

center of another proliferation event depends on the presence in the genome of a gene whose left part matches such value.

There are three features that distinguish ET from other models used in the field. The first feature is the presence of two categories of cells: normal cells and driver cells, much fewer in number (by several orders of magnitude). This allows to steer development by acting on a small subset of driver cells and to keep genomes compact, as in Nature.

The second key feature is the presence of the variable MOC in driver cells. The MOC takes different values in different driver cells and represents the source of differentiation during development, leading different driver cells at different times to execute different portions of the genome. This feature represents a key difference with respect to other cellular models, which rely on positional information and chemical micro-environment as basic providers of the information necessary for differentiation.

The third feature is the definition of the events of proliferation and apoptosis in such a way that many cells (instead of one) are created or deleted at once. This increases the power of a single event, allows a reduction of the number of developmental genes needed to generate the shape, and results in a considerable speed up of the morphogenetic process.

This peculiar combination of features translates into a unique power to generate shapes of arbitrary size and complexity (Fig. 1), comparable with the complexity of biological multicellular organisms, by means of relatively compact genomes (for example, creation of about 125 000 was needed to obtain the brain-like structure in Fig. 1).

Finally, it is important to note that although ET has been inspired by known biological mechanisms, additional elements (not necessarily consistent with current biological knowledge) have been added in order to produce interesting behaviors through computer simulations.

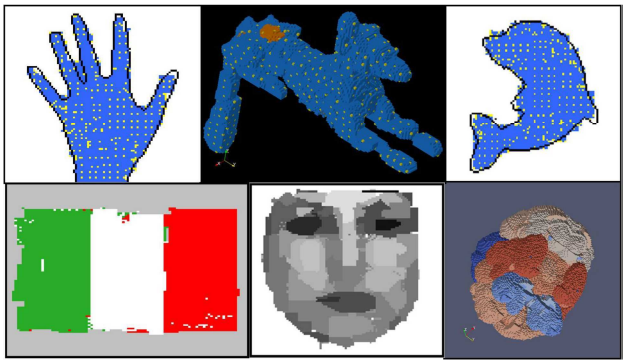

Figure 1. Some 2-D (a hand, a dolphin, Italian flag, a face) and 3-D (a dog, a human brain) shapes generated with ET. Top panels show driver cells (yellow) sprinkled between normal cells (blue), the black line around the hand and the dolphin shows the target. The pattern of the Italian flag and of the face (but not of the brain) is the result of development; only one half of the brain-like shape was developed. All panels show the final developmental stage of the best individual in an evolutionary run.

III. An Attempt to Improve Diversity in Epigenetic Tracking

In this work we explore a novel mechanism to improve the speed of the evolutionary process by increasing the diversity in the population, as a possible alternative to more traditional niching approaches [12]. As a target we have used an asymmetrical shape representing the Egyptian dog-god Anubis patterned with the colors of the Italian flag ("Italian Anubis"), which requires about 80 000 cells.

The mechanism we used in the attempt to improve diversity is based on the concept of the genetic distance between two individuals, calculated as the ratio between the number of different MOS values and the total number of genes in their genomes (a value in the [0,1] interval). For instance, if two genomes G1 and G2 each contain 100 genes, and 20 MOS sequences in G1 are identical to 20 MOS sequences in G2, their distance is (100-20)/100 = 0.80.

The genetic population is subdivided into an ordered set of N subpopulations (Fig. 2). The fitness values of individuals in the first subpopulation are the values of the fitness function: the number of cells inside the target and with correct color minus the number of cells outside the target divided by the number of desired cells (the number of cells required to fill the whole target). The individual with the highest fitness value is taken as the reference element for the next subpopulations. The fitness value of each individual of the *k*th subpopulation is calculated multiplying the value of the fitness function by the genetic distances between this individual and the reference individuals of the previous *k*-1 subpopulations. For instance, the fitness value of each individual in the third subpopulation is the value of the fitness function multiplied by two factors, one measuring the genetic distance between the individual and the reference individual of the first subpopulation, and one measuring the genetic distance between the individual and the reference individual of the second subpopulation.

Unfortunately, the average fitness of the best individuals in 10 independent runs with this mechanism (using 6 subpopulations, 24 individuals each) was 0.8006 (with standard deviation 0.0005), less than the average (0.848, s.d.: 0.001) for 10 runs without it (one population, 144 individuals). In other words, introducing the mechanism actually reduced evolvability; *p*-value for 2-tailed *t*-test: 7E-10 (see also Fig. 3).

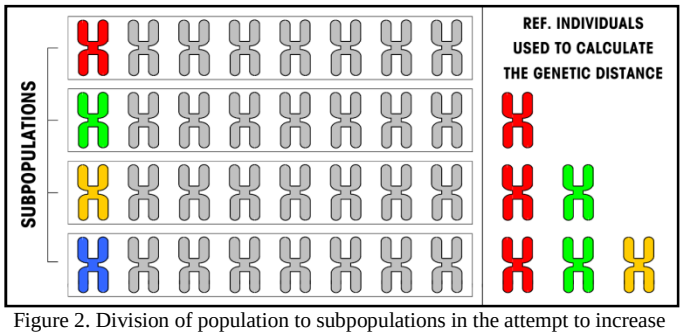

Figure 2. Division of population to subpopulations in the attempt to increase the diversity during the genetic algorithm. The best individuals of each subpopulation, highlighted in color, are used to compute the diversity factor for the individuals in subsequent subpopulations.

IV. Towards Evolution and Development of Neural Networks Using Epigenetic Tracking

Our proposal to increase diversity and thus evolvability in ET is similar to fitness sharing (in which genotypic or phenotypic similarity to other individuals in the population is penalized [12]). Our lack of success thus far may be caused by suboptimal number of subpopulations or their suboptimal size, or by the fact that perhaps comparing developmental programs rather than genomes is more appropriate. We plan to further investigate these possibilities.

Even though our attempt was unsuccessful, our results show that ET allows to generate 3-dimensional multicellular artificial bodies with much larger complexity (Fig. 3) in terms of the number of cells differentiated along the body axis than other approaches (e.g., [8]). This can be considered as a first step towards using ET to grow very large neural networks capable of performing non-trivial computations. Our hope that ET can be used for this purpose is based on already performed experiments in which cells in 3-dimensional structures were evolved to compute a predefined output for a certain input [12]. In this previous work, the input was interpreted as a set of values representing concentrations of chemicals, so cells became metabolic processing units.

In order to extend ET for evolving artificial neuronal networks, the cells would first need to differentiate into artificial neurons. Although more bits will be necessary to specify neuron parameters than are necessary for 3 colors, this is a relatively straight-forward extension of the model. As a first approach, we plan to use stereotypical neurons specified by a formal model (for example, 6 types of neurons with different behaviors in [14] or 20 types shown in [15]), with a stereotypical shape for the EPSP.

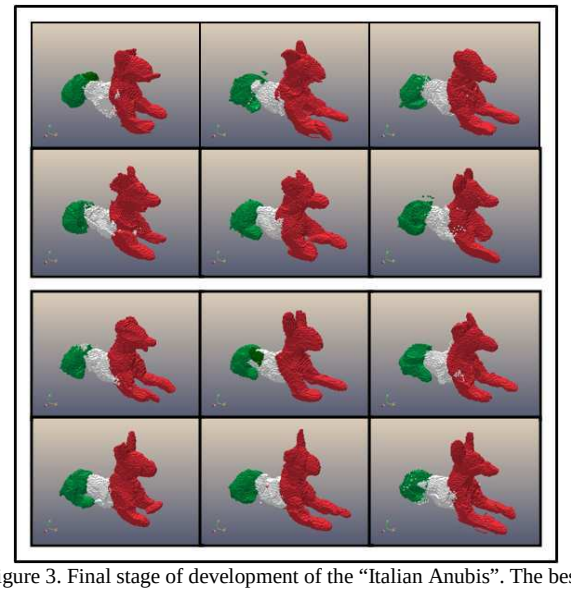

Figure 3. Final stage of development of the "Italian Anubis". The best individuals in 6 independent runs out of 10 performed with the aid of the diversity mechanism described in the text (top) are compared with 6 out of 10 in which the mechanism was not used (bottom).

Specifying the way the connections are made between the neurons will be more challenging. We believe that the first attempt at this could be made by creating a connection when two cells have a matching repertoire of "colors" (several colors can be used for combinatorial matching that could specify much more connections than there are genes in the genome). Additional "colors" may be used to specify the synaptic weight for a given connection or the parameters for learning.

V. Summary

Epigenetic Tracking allows generating multicellular systems with unprecedented complexity in terms of the number of cells and number of detail in the structure. We now plan to build on these foundations to create a system that would allow for evolution and development of very large neural networks.

References

[1] A. Lindenmayer, "Mathematical models for cellular interaction in development," J. Theor. Biol., vol. 18, pp. 280-289, 1968.

[2] A. Turing, "The chemical basis of morphogenesis," Phil. Trans. Roy. Soc. London, vol. B237, pp. 37-72, 1952.

[3] H. De Garis, "Artificial embryology and cellular differentiation," in Evolutionary Design by Computers, P. J. Bentely, Ed. San Franciso: Morgan Kaufmann, 1999, pp. 281-297.

[4] F. Gruau, D. Whitley, and L. Pyeatt, "A comparison between cellular encoding and direct encoding for genetic neural networks," in Genetic Programming 1996: Proceedings of the first annual conference, 1996, pp. 81–89.

[5] G. S. Hornby and J. B. Pollack, "Creating high-level components with a generative representation for body-brain evolution," Artif. Life, vol. 8, pp. 223-246, 2002.

[6] R. Doursat, "Organically grown architectures: creating decentralized, autonomous systems by embryomorphic engineering," in Organic Computing, R. P. Würtz, Ed. Berlin: Springer-Verlag, 2008, pp. 167-200.

[7] M. Joachimczak and B. Wróbel, "Evo-devo in silico: a model of a gene network regulating multicellular development in 3-d space with artificial physics," in Artificial Life XI, S. Bullock, J. Noble, R. A. Watson, and M. A. Bedau, Eds. Boston: MIT Press, 2008, pp. 297-304.

[8] M. Joachimczak and B. Wróbel, "Evolution of the morphology and patterning of artificial embryos: scaling the tricolour problem to the third dimension," LNCS vol. 5777, pp. 33–41, 2009.

[9] J. F. Miller and W. Banzhaf "Evolving the program for a cell: from French flags to boolean circuits," in On Growth, Form and Computers, S. Kumar and P.J. Bentley, Eds. London: Elsevier Academic Press, 2003, pp. 278-301.

[10] A. Fontana, "Epigenetic tracking, a method to generate arbitrary shapes by using evo-devo techniques," in Proceedings of the 8th International Conference on Epigenetic Robotics, 2008.

[11] A. Fontana, "Epigenetic tracking: biological implications," LNCS, vol. 5777, pp. 10–17, 2009.

[12] A. Fontana, "Devo co-evolution of shape and metabolism for an artificial organ," in Artificial Life XII, H. Fellermann et al., Eds. Boston: MIT Press, 2010, pp. 16–23.

[13] O. M. Shir, Niching in Derandomized Evolution Strategies and its Applications in Quantum Control. Universiteit Leiden, 2008.

[14] E.M. Izhikevich, "Simple model of spiking neurons," IEEE Trans. Neural Netw., vol. 14, pp. 1569-1572, 2003.

[15] J. Touboul, "Bifurcation analysis of a general class of non-linear integrate and fire neurons," SIAM J. Appl. Math., vol. 68, pp. 1045-1079, 2008.