An Evolutionary System using Development and Artificial Genetic Regulatory Networks

Song Zhan, Julian F Miller, and Andy M Tyrrell

***Abstract*—Biology presents incomparable, but desirable, characteristics compared to engineered systems. Inspired by biological development, we have devised a multi-layered design architecture that attempts to capture many of the favorable characteristics of biological mechanisms for application to design problems. In this paper we have identified and implemented essential features of Genetic Regulatory Networks (GRNs) and cell signaling so that our system exhibits self-organization which is reminiscent of aspects of biological systems.**

I. Introduction

Biological organisms exhibit characteristics such as robustness, adaptivity, self-reconstruction that are highly desirable in human designed systems. However the exact mechanisms and origins of such characteristics are still far from understood. Recently, the emerging consensus coming from biological research points to the importance of dynamic networks of gene activity, these have come to be known as Genetic Regulatory Networks (GRN) [4]. Such networks are concerned with gene regulation, and protein synthesis. From the system point of view, biological development could be viewed as a multi-layered system with each layer abstracted by different system processes. Bringing these concepts together, one can hypothesize that every multi-cellular organism could be viewed as a multi-layered system that develops from the zygote cell via the mechanisms of gene regulation and cell signaling. In considering biologically inspired artificial designs, one of the important questions concerns the amount of biological realism that should be included. For practical applications, one must bear in mind the limitations of design resources. The work presented in this paper focuses mainly on what biological mechanisms are essential and how they can be implemented. Our aim is to capture an abstraction of biological development and identify the essential aspects of gene regulation and developmental mechanisms required to produce small but useful designs. Section 2 considers the biology that has inspired and informed the work. Section 3 reviews some related work in the application of bio-inspired computational development. Section 4 describes the proposed model. Section 5 is the first analysis of the mechanisms in the model and section 6 summarizes the paper and suggests some possible future work.

S. Zhan is with the Department of Electronics, University of York, UK. (sz503@york.ac.uk)

J. F. Miller is with the Department of Electronics, University of York, UK. (jfm, @ohm.york.ac.uk)

A. M. Tyrrell is the Department of Electronics, University of York, UK. (amt@ohm.york.ac.uk)

II. Biology

Biological development is a differential translation process of genes in different cells and tissues at different time and with different rates [1]. It involves the five processes: cell division, pattern formation, morphogenesis, cell differentiation and growth [19].

A. *Multi-layered view of biology development*

The underlying principle of the development process can be attributed to the function of genetic regulation and the consequent network of interactions that it produces. This is often viewed as a layered construction process from DNA, to mRNA, to proteins and finally organism through protein modification as in figure 1.

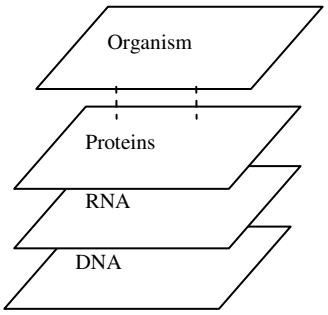


Figure 1: Layered view of biological development

B. *Feedback loops in GRNs*

Two well known regulatory regions; enhancer and inhibitor are incorporated in our model. They have important roles and provide positive and negative feedback control loops in a regulation system. In biology, such feedback loops are thought of as the wheels of dynamical complexity [18]. Negative feedback loops generate homeostasis with or without periodicity. Positive feedback loops generate multiple alternative steady states [2].

C. *Cell Differentiation*

The process of cell differentiation is central to multi-cellular biological development. Even the simplest multi-cellular organisms contain hundreds of cells (for example C. elegans). These cells usually contain the same genetic information and derive from one fertilized cell through mitotic cell division. The differentiation process is controlled by different gene activities in cells and a cascade of cellular processes such as inductive interactions and asymmetries in cytoplasm and these determine division [16].

III. RELATED WORK IN COMPUTATIONAL DEVELOPMENT

Bio-inspired artificial design is a relatively new research field, but the promise from nature is high, and this has motivated research in the last few years. In the fields of computation and engineering, particularly, inspiration from biological development and evolution, has led many researchers to focus on the analysis of such systems so that they can be used to help solve problems concerning pattern formation, robot control and circuit design.

Kumar models an evolutionary developmental system with many natural development processes for 3D multi-cellular development and differentiation [8]. Bowers designs an embryogeny model for pattern formation tasks [3]. His evolved results exhibit canalization in the genotype and phenotype mapping process. He emphasizes the importance of a complex mapping between genotype and phenotype in an evolutionary process and argues that the robustness displayed in the phenotype is the result of functional groupings in the genotype that produce modularity in the mapping process. Roggen and Federici consider a morphogenetic system and a cell chemistry model [13]. In their analysis they conclude that development systems are suitable for large target domains because they reduce the genetic search space and exploit regularities. In addition, they found that the developmental systems exhibit internal temporal dynamics and are able to recover from significant amounts of damage to the phenotype. Others have shown that high levels of self-reconstruction can be achieved by using development mechanisms for both pattern formation and digital circuit tasks in [9][10][11]. Gordon applies a gene regulation network mechanism to problems in evolvable hardware and investigates the scalability of developmental mechanisms for hardware design [7]. Taylor and Quick both use gene regulation mechanisms to model systems for real time control tasks by allowing signal coupling in embodied agents with an environment. They found that the agents are able to self-adapt in the environment and exhibit robustness [12][17]. Eggenberger models a developmental neural controller using a ligand-receptor interaction mechanisms and gene regulation. It is suggested that gene regulation confers adaptivity to a system and that developmental approaches increase evolvability and provide faster convergence towards solutions [5]

IV. THE EMBRYOGENY MODEL

In this section a model of embryogeny is defined which contains fundamental elements that are present in biological development. The development is controlled by gene regulation and protein synthesis.

A. Model Structure

The model consists of three layers as shown in figure 2: a genome layer, a protein layer, and a mapping layer. The genome layer is where the system undergoes its principal development mechanisms: the GRN determines when and where cells in the system grow. The protein layer links the underlying biological mechanisms to the higher level of application design. It synthesizes the basic components that will be utilized for different applications. Finally, the mapping layer utilizes the synthesized functional components for a specific application (i.e. pattern formation, circuits or robot controllers).

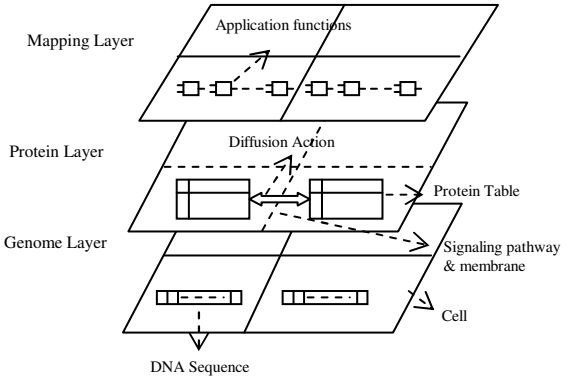


Figure 2: Layered Model Structure. Each layer abstracts different level system processes: The genome layer abstracts the gene regulations, the protein layer is concerned with signaling, and the mapping layer is concerned with the design application.

Genome Layer

In this layer gene regulations are carried out in each developed individual cell, based on the regulatory conditions. Each cell contains a genotype. The system grows from an initial zygote cell to its final structure.

Gene Structure

Each gene has a regulatory region and a product region as illustrated in Figure 3. The regulatory region determines when this gene will be expressed and the product region defines the gene product proteins. When regulatory proteins are bound with the regulatory region in the gene, genes will be expressed and will generate proteins and these proteins in turn will regulate the other genes in the sequence.

En	En_conc	In	In_conc	ProId1	ProId2	----

En: enhancer,
En_conc: enhancer binding concentration,
In: inhibitor,
In_conc: inhibitor binding concentration,
ProIdx: protein Id x,
---: possible additional regulators and products

Figure 3: Structure of a gene

Two regulators are modeled in each gene as in Figure 3: enhancer (En) and inhibitor (In). As the names suggest, the enhancer is used to enhance the expression of the gene and the inhibitor inhibits it. The binding of the regulatory proteins is determined by a binding ID (En and In) and a binding protein concentration threshold (En_conc and In_conc). When the protein's ID matches the regulation ID and its concentration is above a certain threshold, it will express the gene unless the concentration of bound inhibitor protein is larger. When the gene is expressed, it will produce a certain amount of protein

product. The amount is determined by the combination and quantity of regulatory proteins (see Figure 4). Once bound, a process similar to a chemical reaction takes place, proteins will be consumed during regulation.

Assuming there are regulatory proteins bound with genes in the cell the amount of product protein produced is given by

If bound enhancer >bound inhibitor
Protein_produced = enhancer -inhibitor
else
no protein produced (gene inhibited)

Figure 4 Steps in the gene regulation process

Protein Layer

This is the layer where protein synthesis takes place. It is where the underlying proteins are translated into the elements required for the application task. In our model, it is also the place where cell signaling takes place. Proteins in each cell pass through their cell membrane, and via signaling pathways with their neighboring cells are able to exchange cell information among other cells in the system. This negotiation mechanism between cells is important to the development process and helps the system achieve the overall task.

Mapping Layer

The methods used to map the underlying process into application designs is considered in this layer. There are a number of approaches that could be envisaged for this. We could mimic biology and map proteins to application components directly or we could map the regulation process into system states. However, the details of the mapping layer are not presented in this paper.

B. Model Components

Protein

In each cell, all proteins are regulatory proteins and some act as instruction proteins as well. Regulatory proteins bound to regulatory binding sites enhance or inhibit gene action. Instruction proteins instruct the development processes including cell division and cell death. Each different protein has an ID to distinguish itself from other proteins, its quantity is described as floating point number. Each instruction protein is associated with either a division direction or cell death function.

Cell

Cells are the basic computational units. They store all the necessities of the system: gene sequences, proteins, cell membrane and pathway process and are the processing environment for gene regulations and protein synthesis.

Cell Membrane

The cell membrane controls the signaling of proteins between cells. We model the membrane through a threshold implemented as floating point number. Each protein in a cell has a cell membrane threshold that determines the minimum diffusion level for diffusion through the membrane to occur.

Signaling pathway

In biology when proteins bind to receptors on the cell surface, they are transmitted by signal transduction pathways which are a series of chemical events. In our system, we have modelled this via a *pathway* and this and the membrane (via its thresholds) are the same for each cell and are predefined before the system develops; each protein has one membrane and one pathway. When a protein is large enough to pass through its membrane, it will be changed to another protein according to its pathway. For example suppose there are three proteins with membrane thresholds (in parentheses) 0(10.0), 1(15.0), 2(20.0) an *evolved* pathway might be 0→2, 1→0, 2→1. So, for instance, when the concentration of protein 0 is larger than 10.0 part of it will be diffused to neighbors and become protein 2. Following [10] the equation used for chemical diffusion is given by equation (1). This determines the amount of protein that diffuses to neighbors.

$$P^{t+1}_{conc} = 1/2P^{t}_{conc} + 1/(2*\text{NumberofCellBorders})\sum P^{t}_{conc} \qquad (1)$$

C. Development Process – cell division

Cell division is an essential aspect of a developmental process in a multi cellular system. In nature, genes directly or indirectly regulate the cell division process [16]. Here, cell division is determined by active gene product proteins and occurs during the gene regulation process. Division directions are determined by these proteins. Cells will only divide when there is place for daughter cells to be placed (i.e. no cell in the division direction and no predefined system boundary). The proteins in daughter cell are a copied from the mother cell *after* the signaling pathway has taken place in it. We felt that this would make gene regulation differentiation more likely. This is detailed further in the Analysis section.

V. ANALYSIS

Since gene regulation is the fundamental aspect of a development system, our analysis of the system starts from there.

A. Stabilized gene regulations

In order for a self regulatory system to perform reliable actions it needs to arrive into particular regulation states (which will define the cell type). These states normally have a certain degree of homeostasis and present a level noise resistance. In particular, a cellular system invested with GRNs requires firstly, that each cell runs its own gene regulation network and secondly, gene regulations in different cells present regulation of differentiation (i.e. defining when and in what location cells take on particular types).

We made the decision that to keep each gene regulation system balanced the amount of proteins involved in regulation and the amount of proteins produced by it should stay in a comparable balance. Hence when two proteins are involved in regulation, two protein products are involved in the system.

To achieve stable gene regulation genes need to be involved in their own expression. This means that regulation feedback loops are required. This is achieved, in our system, through the use of proteins that are gene enhancers and inhibitors. The role of an enhancer is to exert positive control on regulation. This is necessary for situations where the regulation behavior is required to persist (conferring stable behavior). For example, in figure 5, in the left of the figure, once gene 0 is activated, its own positive feedback loop will cause its own regulation. On the right of the figure, either gene0 or gene1's action will promote the other gene's regulation, and therefore both gene0 and gene1's are regulating each other.

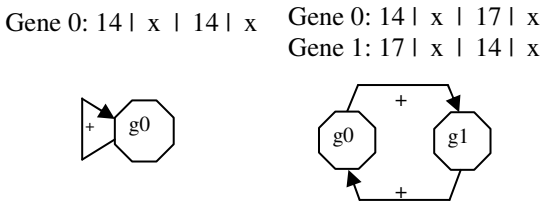


Figure 5: Feedback loops through the use of enhancers in gene regulation (x means 'don't care')

The potentially useful purpose of inhibitors is to exert negative control on regulation and is necessary when the regulation level needs to be reduced. For example, in figure 6, in the left of the figure, gene0 would control its own regulation and the product - protein 8 - will continue to accumulate. In the right of the figure, when gene1's product (0) inhibits gene0's regulation, the production of protein 8 will be reduced.

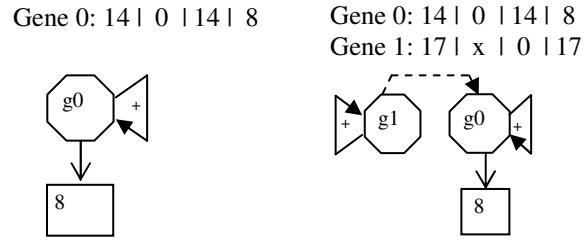


Figure 6: Inhibitory loops for protein expression level control

Gene oscillatory behavior is another possible consequence of inhibition which might be useful to design problems since it can be used to store system information. For example, in figure 7 once gene0 and gene1 are activated from the initial proteins 14 and 0, they provide the regulatory proteins for gene2. However, it needs three development cycles for gene2 to accumulate enough enhance protein 15 to be active. The graph shows the gene actions during development cycles. White represents non-action and black, gene action.

Inhibition is useful for blocking unnecessary gene activation in a regulation network. As a whole, the network serves as a positive loop system. For example, in figure 8, if gene1 is required to be active, gene2's action should be inhibited. A simple cross-inhibition network will regulate gene1's action.

B. Regulation of Cell Differentiation

In biology the differentiation of cells is closely concerned with maternal factors in the cell (i.e. what proteins are present within a cell) [19]. So even if cells have the same gene sequence, the different proteins in each cell determine when and where to regulate which genes and therefore generate cell differentiations.

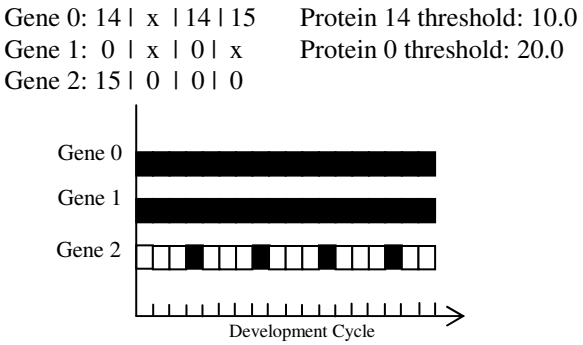


Figure 7: Role of inhibitors for oscillatory gene expression Oscillations

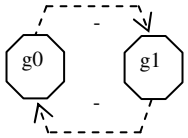


Figure 8: Positive loop system occurring through an even number of inhibitions

In the current system, a development cycle consists of a sequence of genes activity from gene0 through to the last gene. This mimics gene transcription process in biology where a DNA molecule is read from the 5' to the 3' end. Since cell division can occur during this cycle, different cells could receive different amount of proteins. When different gene regulation loops exist in the system, there are chances that genes in different cells are regulated differently. For example in figure 9, we see how gene regulation could generate two separate loops. Consider a two cell environment, which begins with a zygote cell that has been initialized with an amount of protein 1 and suppose that product protein 10 is responsible the north division of the cell. The two resulting cells (after division) will find themselves in different regulation loops.

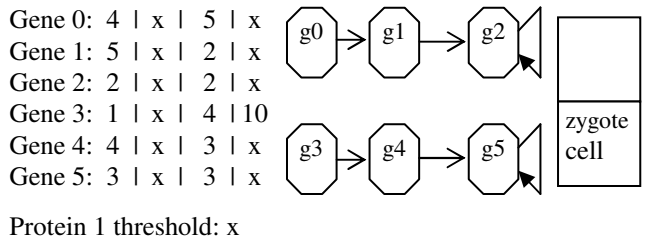


Figure 9 Gene regulation: Differentiations associated with cell division

Signaling pathways between cells also contribute to differential gene regulation. Proteins diffusing to neighboring cells through pathway interactions will react with others so the actual proteins coming into neighbors could be different from those originally present, this also allows differential gene regulation between these cells to happen. For example, in figure 10, there are two regulation loops in the networks. At initialization there is protein 4 in the zygote cell; proteins diffuse to the neighbors in each development cycle after the protein pathway protein 3 becomes protein 1. Therefore, in the zygote cell the gene cycle ends up self-regulating and protein 3 diffuses to its north neighbor but the north cell receives protein 1 and it ends up in a gene4 self-regulation loop.

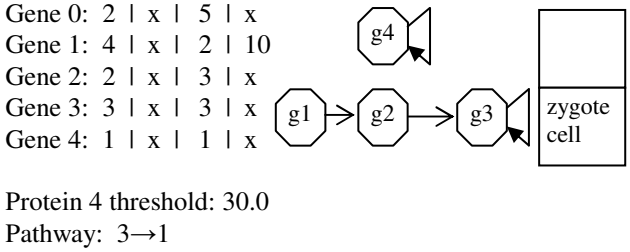


Figure 10 Gene regulation: Differentiation occurring as a result of protein pathway

C. *Evolution of Regulation of Cell Differentiation*

Differentiable regulations in the system are necessary for it to produce phenotypes. This is also one of the interesting areas of investigation within developmental biology [5]. A search algorithm, in this case hill climbing, is used to search some randomly chosen differentiation regulations in cells to demonstrate that this can be achieved. Figure 11 lists the pseudo code of the algorithm we have used and table 1 shows the parameters used for the results in this paper. Fitness is measured by comparing the gene regulation pattern with the desired target pattern at certain development cycles. The number of correct regulations in the developed cells is the fitness. Table 2 illustrates six random target gene regulations in a 3*3 cell structure. Black represents gene action (gene on) and white represents non-action (gene off). The results listed in table 3 are for 10 runs of this system for a number of conditions: (i) evolve only gene regulation in genome layer, (ii) two layer evolution: the gene regulation in genome layer and the cell signaling in protein layer, and (iii) two layer evolution: the gene regulation in genome layer, the cell signaling and membrane threshold in protein layer.

D. *Cell signaling*

Cell signaling interactions are important for achieving an evolvable complex system and as suggested in [14] they play a vital role in achieving system characteristics such as robustness.

In this paper, we have investigated the importance of a presence or absence of cell signaling for a number of targeted gene regulation patterns in a differentiated multi-cellular system. Figure 12 shows the results found for ten evolution runs with 2 genes. It can be seen that higher fitness values were obtained with signaling (right graph) compared with no-signaling (left graph).

1. *Random create n individuals,*
2. *Choose the best individual*
3. *Mutate every gene one by one in sequence in each generation.*
4. *Accept the better ones immediately*
5. *Else if it gets the same fitness accept it with a certain probability*
6. *Else If it becomes worse mutate it m more times with a certain probability*

Figure 11 The Search Algorithm

E. *Membrane*

Membranes act as information filters between cells and can be viewed as part of the cell signaling mechanism. From the development point of view, they sort out the necessary regulatory proteins from the unnecessary for achieving cell differentiation. In Figure 13, we see that higher fitness is obtained when we allow membrane filtering. Figure 14 shows that as the number of genes in a cell grows (from 2 to 3) membranes become even more helpful where it can be seen that membranes with 3 genes provide faster convergence to the target patterns of gene activity than with 2 genes. This suggests that membranes might be more helpful when more complex GRNs are used

VI. SUMMARY AND FUTURE WORK

This paper has presented an evolutionary embryology system that uses genetic regulatory network and development mechanisms. The system organizes itself by its regulation and cell signaling mechanism and develops from a single zygote cell to a multi-cellular system. Cells differentiate directly and indirectly by their internal gene regulations and cell signaling. From the investigation considered so far, it suggests that the biological mechanisms such as, cell signaling and membranes are important processes that assist in achieving cell differentiations and consequently for the development of the system.

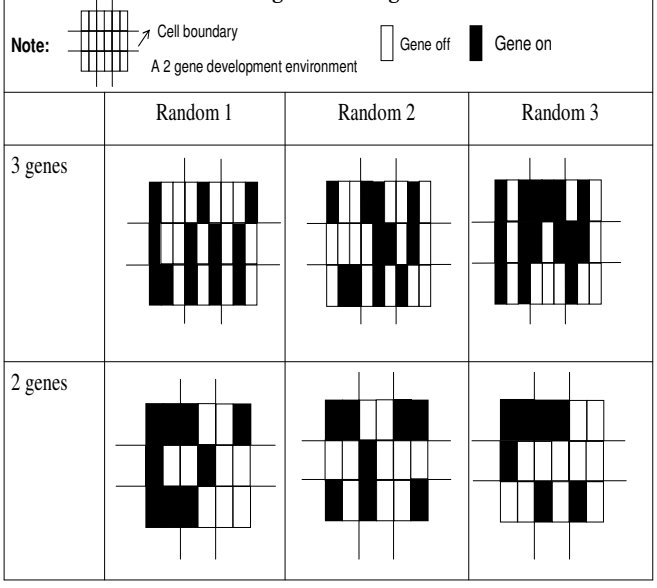
From the model abstraction point of view, many bio-inspired developmental models are based on cellular automata, e.g. in

Miller's system and its implementations neighbor cell coupling is used to achieve the global regulation [10][11]. Similarly the cell chemistry based and morphogenetics system proposed by Roggen and Federici's [13] and Gordon's system [7] also directly apply neighbors' protein states to determine gene regulation in cells. The work presented here has identified and abstracted other biological processes such as signaling pathways and cell membrane thresholds. These can be helpful for tuning signaling between cells so that each cell could either regulate itself alone or strongly or weakly couple with its neighbors.

This paper presents the initial investigation of such underlying processes. We highlight the contribution of gene regulation itself and the assistance of cell signaling to a bio-inspired cellular developmental system and we emphasize two aspects for future work: cell differentiation as a means of determining substructures in a system and stabilisation for conferring a fixed system function.

Future research will investigate further how model parameters and mechanisms influence and confer features such as robustness, fault tolerant and adaptivity. With the mapping layer involved in the designs, additional tasks such as pattern formation, circuits and robot controller as suggested by [15] will be investigated.

Table 2 Random Target Gene Regulations in 3*3 cells



REFERENCES

[1] Banzhaf, W. (2004), On Evolutionary Design, Embodiment, and Artificial Regulatory Networks, Springer-Verlag, pp. 284-292.

[2] Brenner, S., Dove, W., Herskowitz, I. & Thomas, R. (1990). Genetics 126, pp. 479-486

[3] Bowers, C. P. (2005), Formation of Modules in a Computational Model of Embryogeny, IEEE Congress on Computational Evolutionary, pp. 537-542

[4] Davidson, E. (2001). Genomic Regulatory Systems: In Development and Evolution. Academic Press Inc., London, UK.

[5] Eggenberger, P, G´omez, G. and Pfeifer. R, (2002) Evolving the morphology of a neural network for controlling a foveating retina - and its test on a real robot. In Standish, R. K., Bedau, M. A., and Abbass, H. A., editors, Artificial Life VIII: Proceedings of The Eight International Symposium on Artificial Life, pages 243–251, 2002.

[6] Geard, N. L. and Wiles, J. (2003), A Gene Regulatory Network for Cell Differentiation in Caenorhabditis elegans. In Proceedings of The First Australian Conference on Artificial Life, pp. 86-100.

[7] Gordon, T. G. W and Bentley, P. J., (2005), Development Brings Scalability to Hardware Evolution. In Proceedings of the 2005 NASA/DoD Conference on Evolvable Hardware, pp. 272-279, IEEE Computer Society

[8] Kumar, S. (2004), Investigating Computational Models of Development for the Construction of Shape and Form, PhD thesis, University College London.

[9] Liu, H., Miller, J. F., and Tyrrell, A.M. (2004), An Intrinsic Robust Transient Fault-Tolerant Developmental Model for Digital Systems, Workshop on Regeneration and Learning in Developmental Systems, GECCO 2004.

[10] Miller, J. F. (2004), Evolving a self-repairing, self-regulating, French flag organism, Proceedings of GECCO, Part I. LNCS, 3102 Springer 2004 pp. 129-139.

[11] Miller, J. F. (2003), Evolving developmental programs for adaptation, morphogenesis, and self-repair, Seventh European Conference on Artificial Life, LNAI Vol. 2801, pp. 256-265

[12] Quick, T., Nehaniv, C. L. etc (2003), Evolving Embodied Genetic Regulatory Network-driven Control Systems, Seventh European Conference on Artificial Life, LNAI, Vol. 2801, pp. 266-277

[13] Roggen, D. and Federici, D. (2004), Multi-cellular Development: Is There Scalability and Robustness to Gain? 8th Int. Conf. on Parallel Problem Solving in Nature, pp 391-400

[14] Rudge, T. and Geard, N. L. (2005) Evolving Gene Regulatory Networks for Cellular Morphogenesis. In Proceedings of The Second Australian Conference on Artificial Life, pp. 239-252.

[15] Stanley, K. O. and Miikkulainen, R. (2003), A Taxonomy for Artificial Embryogeny, Artificial Life 9: 93-130

[16] Stephen L. Wolfe (1993), Molecular and Cellular Biology, Chapter 22, Wadsworth.

[17] Taylor, T. (2004), A Genetic Regulatory Network-Inspired Real-Time Controller for a Group of Underwater Robots, in Proceedings of Eighth Conference on Intelligent Autonomous Systems (IAS-8), 2004.

[18] Thomas, R. (1991), Regulatory networks seen as asynchronous automata: A logical description. J. Theor. Biol., 153, 1991

[19] Wolpert, L. (1998). Principles of Development. Oxford University Press.

Table 1: Search parameters	
Initial random individuals	15
Number of repeated point mutations	15
Acceptance probability if equal fitness (Step 5, Figure 11)	0.7
Probability of mutation if worse fitness (Step 6, Figure 11)	0.5
Start fitness calculation at development cycle	10
No of protein types	19
development parameters	
Number of instruction proteins	9
Number of state proteins	10
Number of infused zygote proteins	1
Zygote cell position	centre
For the reasons of simplicity threshold bindings are not involved.	

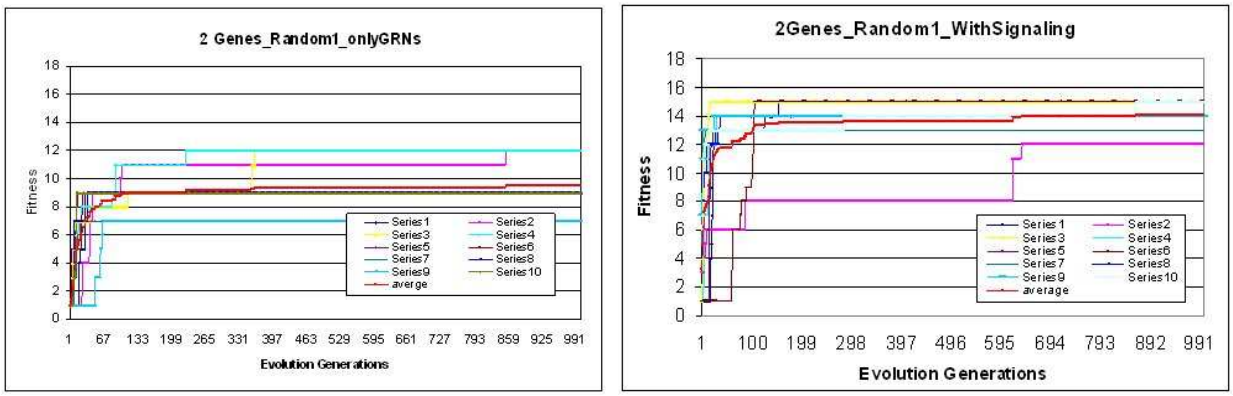


Figure 12: One evolution example with only gene regulation and with cell signalling added

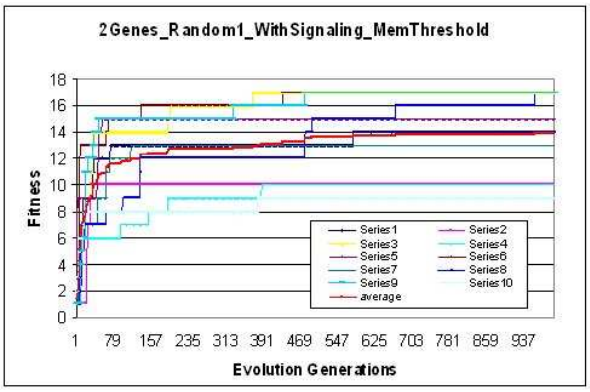


Figure 13: One evolution example with membrane threshold

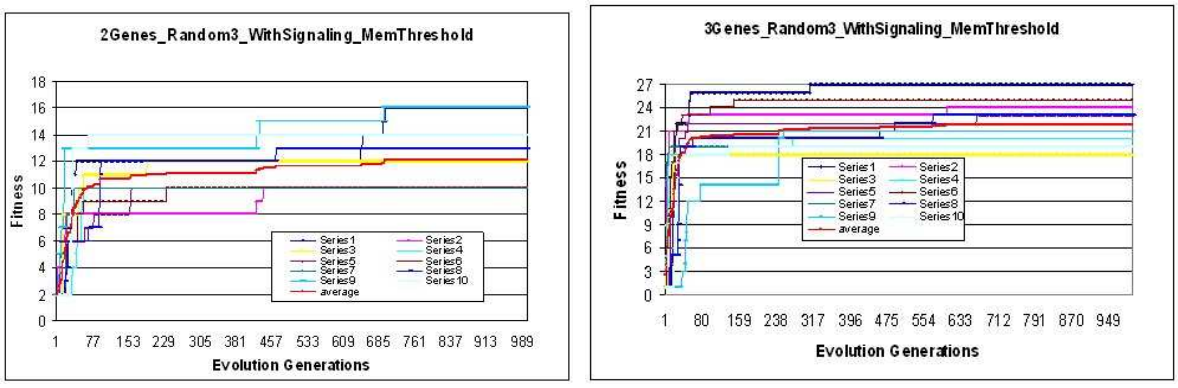


Figure 14: Evolution examples with membrane threshold and increased gene size.

Table 3 Experiment Results										
experiments		Random1			Random2			Random3		
		i	ii	iii	i	ii	iii	i	ii	iii
3 genes	Best Run Fitness	0.85	0.89	1	0.85	0.93	0.93	0.85	0.93	1
	Average 10 run fitness	0.72	0.81	0.87	0.71	0.82	0.85	0.76	0.81	0.81
	Standard deviation of best over 10 runs	0.10	0.09	0.07	0.11	0.08	0.083	0.13	0.16	0.12
2 genes	Best Run Fitness	0.67	0.83	0.94	0.78	0.94	1	0.83	0.83	0.89
	Average 10 run fitness	0.53	0.78	0.77	0.55	0.78	0.73	0.54	0.79	0.67
	Standard deviation of best over 10 runs	0.17	0.08	0.18	0.20	0.17	0.18	0.26	0.12	0.17

Key:

(i) Only gene regulation,
(ii) With gene regulation and cell signaling, and
(iii) With gene regulation, cell signaling and membrane threshold.

From the best run fitness of all experiments we can see that cell signaling and membrane threshold are useful in achieving the target gene regulation patterns. Average runs show that cell membranes work better in 3 gene systems than 2 gene systems. Standard deviation shows the deviations for the best of 10 runs.