Design by Morphogenesis

Cefn Hoile[1,2] and Richard Tateson[2]

[1]COGS, University of Sussex, Brighton BN1 9QH, UK. cefn@newscientist.net

[2]BT Labs, Martlesham Heath, Ipswich IP5 3RE, UK. richard.tateson@bt.com

Abstract

We present a simulation of morphogenesis and cell interaction which allows us to address embryogenesis as an engineering problem. A space of cell control functions is defined using a cybernetics approach by identifying the controlling variables and the degrees of freedom of each individual cell. These functions receive inputs which are dependent on a cell's context, and determine its responses to that context. Multiple cells are coupled through a simulation of a 3d reaction-diffusion fluid matrix in order to generate interactive behaviour. A fitness function is designed which characterises the desired behaviour of the cellular collective. We tune the parameterised control function for the individual cells using a genetic algorithm to maximise the fitness of their collective behaviour.

Introduction

Over the last decade of artificial life research there have been several impressive approaches to simulating morphogenesis (see below). Morphogenesis, literally "creation of shape", is the process by which a fertilised egg changes into the form of the adult organism. It involves cell division and movement. Of course, to achieve the correct pattern of cell movements the cells must change their behaviour. For example genes will be turned on, new proteins will be expressed, perhaps resulting in two neighbouring cells adhering. The understanding of these underlying molecular processes is the realm of developmental biology.

Some of the previous simulations of morphogenesis include simulation of the molecular processes. In the work reported here we aim to show morphogenesis at the cellular level and are not concerned with a biologically plausible set of developmental mechanisms within the simulated cells.

Looking forward to artificial life in the coming decade, we see the simulation of morphogenesis as an approach to system design. As artificial life techniques are increasingly applied to real industrial problems, the issue of how to "engineer" emergent behaviour will become ever more pressing.

In the next section we review earlier work relevant to simulation of morphogenesis. We then explain our motivation for work in this area and give details of our simulation and results to date. Finally we look ahead to future work.

Related Research

Several simulation approaches have been employed to study different classes of cellular behaviour.

Furusawa and Kaneko simulate cell-differentiation using catalytic networks of different chemical species to describe the internal dynamics of cells in a shared fluid (Furusawa and Kaneko 1997). They identify cell-type attractors, and characterise their *potency*. The combinations of cells which can coexist in stability within the same matrix was also found to have attractors. Their catalysis matrix is the basis of the reaction matrix in Design by Morphogenesis (DBM), although the binary values have been replaced by continuous values.

Peter Eggenberger (Eggenberger 1997) uses a model of gene expression to generate symmetrical 3D morphologies from a single seed cell. The interaction of cell products, regulatory molecules and genetic material is prescribed by an affinity function. A genetic algorithm is used to tune the affinities of the genetic material. Triangulation of position is made possible by three sources of diffusible morphogens which are placed in the environment. Forms with bilateral symmetry emerged from the epigenetic interactions of the cells and these diffusible morphogens.

Nicholas Savill and Paulien Hogeweg's (Savill and Hogeweg 1997) models are based on earlier work (Glazier and Graner 1993). Their techniques focus on the energetics of membrane surfaces, and may be extended to explore the generation of tissue types with different mechanical properties, physical behaviour, and complex membrane topologies, such as those found in neural tissues. Their classic work is the simulation of slug formation, cell sorting and slug motility in the slime mould *Dictyostelium*.

Geometrical approaches to the simulation of cell membranes in 2D have been explored (Weliky and Oster 1990). However, these techniques would be challenging to scale to 3D.

Detailed biological models are reported in a manifesto for the computer simulation of cell biology (Kitano et al 1997). The focus of their Virtual Cell Laboratories project is to replicate the known dynamics of individual cells at the biochemical level, as a potential replacement for cell culture. Kitano does not envisage the simulation of multi-cellular tissues in this simulator.

Motivation for Design by Morphogenesis

Multi-cellular organisms have a number of features which are desirable in artificial systems. They can adapt to changing circumstances, through learning or evolution. "Faults" arising from internal errors or injury can often be rectified. Irredeemable faults are seldom the cause of catastrophic failure because the redundancy of the system allows most or all functions to continue.

This contrasts markedly with most current artificial systems. Here the human designer takes on the burden of solving all these problems. The designer must gather a large amount of information about the operating conditions and specification of the system to be designed. Having described the problem, the designer must provide an explicit, practical and comprehensible solution.

This design process usually results in a system which only operates under previously anticipated conditions, and which employs a centralised, hierarchical, non-redundant control system. Such control systems are applauded as transparent and efficient. However, many solutions are unavailable to human designers not because they are poor solutions in terms of performance in the system of interest but merely because they are not comprehensible by humans.

As systems become more complex, by virtue of increasing size, interconnections and dynamism, the shortcomings of rational human design will be exposed. This is already happening in the computation and communication fields. For example there have been failures of large software control systems for air (Lyu 1996) and space travel (Ward and Seligsohn 1990), and of highly complex terrestrial data networks (Andrews 1991).

These facts have led to a great increase of interest over recent years in the potential for non-rational design through evolutionary algorithms (Thompson and Layzell 2000).

The DBM work was motivated by two complementary desires. Firstly to simulate the process of morphogenesis. Secondly to exploit morphogenesis as part of a design process for artificial systems. The importance of the genotype to phenotype mapping is well recognised by researchers who use evolutionary algorithms (Shipman et al 2000). We aim to mimic nature's "mapping" mechanism, and employ its principles in human design problems.

Multi-cellular Dynamics and 'DBM' Dynamics

The dynamics of an isolated cell, or an isolated point in a reaction-diffusion soup can be treated as a closed phase space, in which the rate of change of the system variables is determined entirely by their current values. However, in a multi-cellular organism, these are not independent systems. Each infinitesimal point in a fluid is affected by its neighboring points. Each cell embedded in the fluid is affected by the chemistry taking place around it.

Separating the multi-cellular organism into subsystems with variables and parameters is somewhat artificial given these interdependencies, but the definition of fundamental components is demanded by the formalism of computer simulation. In a discrete time simulation, we can then prescribe how the phase spaces of subsystems intersect, setting the rate of change of variables at intersections to the sum of the contributing vectors of the intersecting spaces.

DBM navigates the multi-cellular repertoire by exploring the range of possible vector fields for each of the coupled sub-units in the simulation. Each sub-system's trajectory through its phase space can be contributed to by other sub-systems, according to their coupling. In each case a parameterised function is employed which can describe a wide range of possible vector fields. The parameters of these functions represent the genome of the individual.

A discrete-time simulation of these coupled units offers a good approximation of the global dynamic qualities of a continuous system. The repeated assessment of interdependent variables over small time segments closely parallels the analytic treatment over infinitely small intervals of a deterministic, fixed-dimensionality, dynamical equation by calculus.

Simulation

The spherical cells are simulated as volumes of fluid with a semi-permeable boundary, embedded at a point in a fluid space.

Genomic response is captured by a feed-forward sigmoidal neural network - theoretically capable of representing any input to output mapping. This has the advantage of representing continuous response functions which cannot be represented by discrete gene-switching models such as boolean networks.

To simulate the chemical medium we employ a 3D cellular automata-style reaction-diffusion model with a fixed set of chemical species.

The parameters specified by the candidate genome are:

- Weights of neural network employed in control of individual cells.
- Reactivity and diffusivity of each chemical species in the system.

Each of these values has an upper and lower bound. The remaining parameters are hard coded to suitable values.

The implementation is broken up into subsystems:

- **Cell Control Mechanism**
  Each cell takes the concentrations of chemicals in its cytoplasm, its membrane, the membranes of cells which it contacts, and its local fluid environment. These values constitute the *input vector* for the neural network whose weights are specified by the genome.
  The resulting *output vector* dictates the cell's behaviour which includes:
  - rate of emission of chemicals into its cytoplasm and membrane
  - rate of active transport of chemicals from its cytoplasm into the local fluid environment
  - change of membrane permeability for each chemical

- chemotaxis (movement in the direction of maximum concentration gradient for each chemical)
- cell division
- cell death

- **Cell Update Mechanism**
  The output states from the Cell Control Mechanism, the positions of each cell, and the concentrations and gradients of each chemical in the local fluid environment form the *input vector* for this subsystem.
  The *output vector* specifies the actual exchanges between the cell and the reservoirs with which it interacts, and the actual movement of the cell within the simulation space.

- **Grid Cube Update Mechanism**
  (Elements of reaction-diffusion CA)
  The *input vector* is the concentration of each chemical within the subcube and within its facing neighbours.
  The *output vector* specifies the chemical exchange with each of its facing subcubes and the rate of transformation from each chemical to another within this subcube.

Simulation Runs

Each simulation run begins with a single cell placed at the centre of a nil concentration 3D space, analogous to the initial state of the zygote. The cell and its daughters then alter and respond to their chemical environment for a fixed period of simulation time. If the number of cells reduces to zero or increases above a prescribed threshold, the simulation is called to an early halt, and the candidate receives a penalty fitness. Otherwise, the candidate is given a fitness which captures its degree of competence at producing a desired 3d form.

Candidates and Evaluation

In our approach, a genome of floating point numbers represents the candidate for assessment. These numbers are used as parameters for the control functions of the coupled subsystems. The dynamics of these coupled systems result, after a number of timesteps, in a multi-cellular form.

The points occupied by cells are read into a matrix. Principal components analysis is undertaken to establish the axes of maximal variance. The original matrix is then re-oriented onto the axes of the principal vectors.

This re-orientation is necessary because cell divisions take place according to a random vector. It is not predetermined in which direction the original spherical symmetry and its descendants will break. Using the arbitrary x, y and z axes of the simulation space would penalise most distributions even if they had the perfect shape, since they would have the wrong orientation. Transforming the axes according to the distribution itself eliminates this problem.

The initial population comprises thirty individuals. In each generation, thirty new individuals are generated by crossover, employing a probabilistically biased rank to select the parents. All sixty individuals are then trialed, and rank selection used to select the thirty individuals which are carried forward to the next generation.

In each simulation run, we employ a fitness function which selects those candidates whose re-oriented distributions best approximate to a target form. Candidates are free to use whatever "affordances" are offered by the simulation dynamics to maximise their fitness.

The strategies they adopt to achieve target structures could offer insights into multi-cellular dynamics in a real fluid environments. They could also indicate design principles for robust self-organising artificial systems.

Results

A two segment target structure was used to test the repertoire of the DBM system. Segmentation is a form achieved very early in the morphogenesis of natural organisms. It is revealed by the organisation of the early embryo into alternating stripes (with the stripes distinguished either by morphology or by some less visible differentiator such as gene expression state). Hence these trials were intended to explore early symmetry breaking and structural formation of an artificial "embryo".

To mimic this process of natural segmentation it was decided to generate segments with differential densities of cells. The ideal candidate would have a high density of cells in specified sections and low densities of cells in between. In a real organism this would correspond to sections of cells of a certain differentiated type, separated by cells of another differentiated type.

Since the transformed co-ordinates always have their greatest variance in the z-axis, this was used to represent a longitudinal axis, along which the segments should form. Individual cells were given credit in proportion to the value of a negative cosine function centred on the mean containing the same number of peaks as the desired number of segments. The wavelength of the function for two segments was, therefore, twice the standard deviation in the z axis. For any size or orientation of distribution, this function will contain the appropriate number of peaks in the longitudinal axis. If the z co-ordinate of an individual cell fell somewhere on the peak of this function, the candidate was credited. If it fell in a trough, it was penalised.

During this run, a penalty score of –20 was assigned to candidates producing more than 40 cells, or whose cells all committed cell suicide. This corresponds to a distribution with the average permissible number, (twenty cells), whose cells are all maximally penalised for incorrect positioning.

The function was used to encourage the development of two segment distributions. Since the distance between any two cells is exactly twice the standard deviation from the mean, both cells are located at the peak of the negative cosine wave. This provides a guaranteed score of 2.0. However, producing such a distribution is not a trivial matter, and it is a simple demonstration that DBM cells can regulate their own activity. Cells during this run were able to divide to produce up to 64 cells within the number of

timesteps allotted. At some point during the simulation, the input to the cell control system must have generated an output to trigger division in order to generate two cells. However, further division must have been inhibited by some change which the cells initiated to themselves or their environment, leading to a change in their input vector.

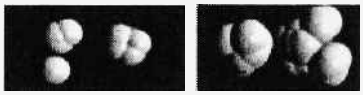
Later solutions in this run were similar, although they contained a larger number of cells in each cluster. The initial division is followed by a migration by the cells away from each other. The daughter cells of these two initial cells, however, did not migrate, maintaining their proximity to the peaks of the negative cosine wave.

The initial migration increases the variance in the z axis of the transformed coordinates. Hence a larger number of cells can share the peaks of the distribution. Below, this strategy is shown for full-term cell distributions of 2, 4, 6, 10, 11, 25 and 37 cells from generation 450.

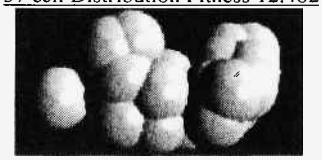
2 and 4 cell distributions Fitness 2.0 and 3.917

6 and 10 Cell Distributions Fitness 5.502 and 6.382

11 and 25 cell Distributions Fitness 6.211 and 11.233



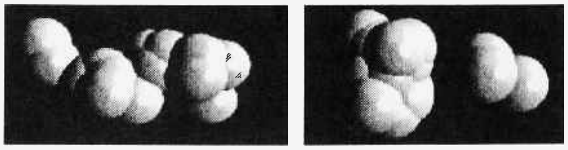
37 cell Distribution Fitness 12.482



Unfortunately, as the number of cells increases, it seems to become harder to maintain the separation of the two clusters and their symmetry. Although originally a candidate may score very highly, the next evaluation introduces an element of noise for which the cell collective is not able to compensate.

The 19 cell distribution below scored only 1.320, through the apparent merging of the two cell clusters.

19 and 25 cell distributions Fitness 1.320 and –4.883



The two clusters of the 25 cell distribution above are asymmetrical. This causes the standard deviation to be significantly reduced, as the mean tends towards the largest cluster. As a result of the repositioned mean the fitness function penalises many of the cells.

Although the genetic algorithm is able to drive the distributions towards large numbers of cells in each segment, the instability that this introduces causes more candidates to achieve negative scores. Where the maximum fitness is driven up, the mean is often driven down. This has the overall result of causing selection to favour cells with stable configurations such as the two-cell system.

Conclusion and Future Work

In this paper we have presented an argument for the use of abstract biological models to explore multi-cellular dynamics. We have detailed a specific approach to the simulation of complex multi-cellular behaviour, and presented an early optimisation run demonstrating that this approach may be used to replicate phenomena in early embryological development.

Future work is to include exploring somatic optimisation by multi-cellular systems, such as neural organisation. The neurons of two cloned rats with identical DNA produce different neural structures during embryological development (Edelman 1992). Nevertheless, both structures "design themselves" in order to process information and participate in the response mechanisms of the rat. In the absence of a single global controller, this self-organisation must take place according to local interactions.

The exploration of more challenging optimisation based on natural systems will be carried out alongside use of the simulation in abstract design problems. For example, an individual's morphology could be evaluated as a solution to a telecommunications network topology problem.

Future implementations may include the use of a spatially distributed genetic algorithm in which only neighbouring individuals in a population space may mate. Innovations may then emerge in particular neighbourhoods, without being diluted or disrupted by genetic material from other solutions. Combinations of benefits can still occur at the borders between competitive neighborhoods.

It is also possible to employ an "evolutionary strategy" algorithm (Bäck et al 1997), in which real-valued genetic information is not mutated randomly within a specific range, but according to variances prescribed at other locations in the genome. This could encourage the gradual variation of network weights and chemical coefficients towards an optimal solution, avoiding large changes in dynamics which can accompany large changes in parameters in complex coupled systems.

Future developments are also planned to improve the performance of the simulator. Parallel processing will permit larger populations, more generations and a finer grained discrete-time simulation, hopefully allowing more complex behaviours to emerge.

Acknowledgments

Many thanks to Brian Goodwin, Inman Harvey, Alan Steventon, Mark Shackleton, Erwin Bonsma, Alex Penn and the reviewers for discussions and critical reading.

A pre-release version of Eos, an evolutionary and ecosystem platform developed at BT Labs, was used to marry the cellular simulation with a genetic algorithm. For Eos information see http://www.labs.bt.com/projects/ftg/

References

Alt, A., Deutsch, A. and Dunn, G. (eds.) 1997. Dynamics of Cell and Tissue Motion. Birkhauser Verlag, Basel.

Andrews, E. L. 1991.Computer Maker Says Tiny Software Flaw Caused Phone Disruptions. *N.Y. Times*. 10th July.

Bäck, T., Hammel, U., and Schwefel, H-P. 1997 Evolutionary Computation: Comments on the History and Current State, IEEE Transactions on Evolutionary Computation, Vol. 1, No. 1

Edelman, G. M. 1992. Bright air, brilliant fire: on the matter of the mind. BasicBooks, New York.

Eggenberger, P. 1997. Evolving Morphologies of Simulated 3D Organisms Based on Differential Gene Expression. In Proc. 4th European Conf. Artificial Life. pp 205-213. Husbands, P. and Harvey, I. (eds), MIT Press.

Furusawa, C. and Kaneko, K. 1997. Emergence of Differentiation Rules leading to Hierarchy and Diversity. In Proceedings of the Fourth European Conference on Artificial Life. pp 172-181. Husbands, P. and Harvey, I. (editors), MIT Press.

Furusawa, C. and Kaneko, K. 1998. Emergence of Multicellular Organisms with Dynamic Differentiation and Spatial Pattern. Artificial Life IV pp79-93 MIT Press.

Glazier, J. A. and Graner, F. 1993. Simulation of the differential adhesion driven rearrangements of biological cells. *Physical Review* E (47) Number 3 (March)

Gleick, J. 1988. Chaos – Making a New Science. Heinemann.

Goodwin, B 1994. How The Leopard Changed Its Spots Weidenfeld and Nicolson

Harvey, I. and Bossomaier, T. Time out of Joint: Attractors in Asynchronous Random Boolean Networks. In Proceedings of the Fourth European Conference on Artificial Life. pp 67-75. Phil Husbands and Inman Harvey (editors), MIT Press.

Kauffman, S. 1993. The Origins of Order. Oxford University Press, New York.

Kitano, H., Hamahashi, S.,Kitazawa, J., Takao, K. and Imai, S. 1997. Virtual Biology Laboratories. In Proceedings of Fourth European Conference on Artificial Life. pp 274-283 Phil Husbands and Inman Harvey (editors), MIT Press.

Lyu, M. R. (ed), 1996. Handbook of Software Reliability Engineering. Institute of Electrical & Electronic Engineers, McGraw-Hill Book Company, New York.

Painter, K. J. 1997 Chemotaxis as a Mechanism for Morphogenesis. DPhil Thesis, University of Utah.

Pittenger M.F., Mackay A.M., Beck S.C., Jaiswal R.K., Douglas R., Mosca J.D., Moorman M.A., Simonetti D.W., Craig S. and Marshak D.R. 1999 Multilineage potential of adult human mesenchymal stem cells. *Science* Vol. 284 pp 143-147

Savill, N. J. and Hogeweg, P. 1997. Modelling morphogenesis: from single cells to crawling slugs. *J. Theor. Biol.* Vol. 184 pp 229-235

Shipman, R., Shackleton, M., Ebner, M. and Watson, R. 2000. Neutral search spaces for artificial evolution: a lesson from life. In Artificial Life VII: Proceedings of the Seventh International Conference. Bedau, M., McCaskill, J., Packard, N. and Rasmussen, S. (editors)

Slack, J. M. W. 1991. From Egg to Embryo. Cambridge University Press.

Slotine, J-J. E. 1994 Stability in adaptation and learning. Animals to Animats 3. pp 30-34 MIT Press.

Thompson, A. and Layzell, P. 2000 Evolution of Robustness in an Electronics Design. In Proc. 3rd Int. Conf. on Evolvable Systems, Springer Verlag. (in press)

Turing, A. The Chemical Basis of Morphogenesis. 1952. *Phil. Trans. R. Soc.* Vol B237 pp 37-72.

Ward, M. and Seligsohn, D. 1990. Missing tests sent Ariane on path to doom. *New Scientist*. p 10, 27th July.

Weliky, M. and G. Oster. 1990. The mechanical basis of cell rearrangement 1. Epithelial Morphogenesis during Fundulus Epiboly. *Development* Vol. 109 pp 373-386.