Asynchronous Dynamics of an Artificial Genetic Regulatory Network

Jennifer Hallinan[1] and Janet Wiles[2]

[1]ARC Centre for Bioinformatics, Institute for Molecular Biosciences and School of ITEE The University of Queensland, AUSTRALIA

[2]School of ITEE and School of Psychology, The University of Queensland, AUSTRALIA

j.hallinan@imb.uq.edu.au

Abstract

The synchrony / asynchrony dichotomy prevalent in models of genetic regulatory networks can be replaced by a parameter, *s*, which is the probability of a node being updated in a single time step. Here we apply the idea of such parameterized synchrony to study the dynamics of the genetic regulatory network extracted from an artificial genome model. We find that the relationship between degree of synchrony and the number of limit cycles is not linear. The number and length of limit cycles peaks at intermediate values of *s*. The proportion of state space explored and the length of transient trajectories also follows this pattern. The richer behavior found at intermediate values of the synchrony parameter is much more characteristic of biological systems than either full synchrony or complete asynchrony.

Introduction

A characteristic property of biological systems is the robustness of their dynamic behaviors despite the inherent stochasticity of the underlying molecular interactions. A current issue in the design of genetic regulatory networks is how such characteristic dynamics can best be modeled while acknowledging the underlying stochasticity of the network components.

Genetic regulatory networks developed using Reil's Artificial Genome (AG) model (Reil, 1999) reflect the behavior of real genetic regulatory networks, implementing and extending random boolean networks by grounding their design in sequence-level models of DNA. They exhibit complex dynamic behavior, ranging from rapid "freezing" at a point attractor in state space, through limit cycles of varying length and complexity, to apparently chaotic dynamics, depending on the connectivity and degree of inhibition of the network. However, these models suffer from several important deviations from biological plausibility. Perhaps the most important of these deviations pertains to the gene update rules.

In the standard AG, as in the classical Random Boolean Network (RBN), all genes are updated simultaneously, on the basis of their inputs from the previous time step. While computationally convenient, such synchronous updating does not occur in biological GRNs; factors such as mRNA and protein synthesis, degradation and transport times mean that the system is replete with delays of varying amounts, and genes are activated or inhibited in a fundamentally asynchronous manner. Unfortunately, with Boolean GRN models, the implementation of asynchronous gene update, whether deterministic or not deterministic, completely alters the network dynamics. Unless connectivity and inhibition levels are very high, most networks quickly move to a point attractor in state space and remain frozen there. The synchrony simplification is therefore widely accepted, largely because synchronous Boolean networks do exhibit dynamic behavior similar to that of biological cells.

The existence of complex dynamic behavior in network models is of interest because biological cells are assumed to function on the "edge of chaos", in the regime between totally frozen and chaotic dynamics. This region is characterized by the presence of limit cycle attractors with wide basins of attraction. Such attractors are widely assumed, following Kauffman (1993), to be models of cell types—each cell type is an attractor in gene expression phase space.

The biological implausibility of synchronous updating is widely recognized, and the effects of synchrony have been examined in a variety of models, including globally coupled logistic maps (Abramson & Zamette, 1998), Conway's Game of Life (Blok & Bergersen, 1999), cellular automata (Schönfisch & de Roos, 1999) and random Boolean networks (Harvey & Bossomaier, 1997; Di Paolo, 2001).

Limit cycles *per se* do not generally exist in an asynchronously updated networks, although networks with such properties can be specifically handcrafted (Nehaniv, 2002). However, several authors have demonstrated the existence of pseudo-periodic "loose" attractors in asynchronously updated networks (Harvey & Bossomaier, 1997; Di Paolo, 2001). Hallinan and Wiles (2004) have demonstrated that networks based upon the AG model can be evolved to exhibit such loose attractors, even when asynchronously updated.

Synchrony / asynchrony is often thought of as a binary condition. An alternative is to consider degree of synchrony as a tunable parameter, implemented as the probability of each node being updated at each time step. Under such an update scheme a network with synchrony 1.0 is the standard synchronous network, while one with synchrony *1/n*, where *n* is the number of nodes in the network, is equivalent to the usual conception of an asynchronous network.

In this study we examine the effects of modifying the degree of synchrony upon the dynamics of AG-generated networks.

Methods

Our model is based on the Artificial Genome model (AG) developed by Reil (1999). A genome is generated at random, using equal proportions of *b* "bases". There are four bases in real DNA – adenine (A), thymine (T), guanine (G) and cytosine (C), so we used four bases, designated 0, 1, 2 and 3. The genome is then searched for instances of a gene marker string of length *l* (we used 0101) analogous to the TATA box to which biological transcription factors bind. The following *g* bases are then designated a gene. The region between the end of a gene and the beginning of the next 0101 marker string becomes the promoter region for the downstream gene.

Each gene is "translated" into a gene product by incrementing each base by 1. A gene with the sequence 012130 will therefore result in a product with the sequence 123201. All of the promoter regions in the genome are searched for matches with each gene product; if a match to the product of gene **A** is found in the promoter region of gene **B**, we say that gene **A** controls gene **B**. This control may be either excitatory—**A** promotes the transcription of **B**—or inhibitory. In this way a genetic regulatory network is constructed from the randomly generated genome (Figure 1).

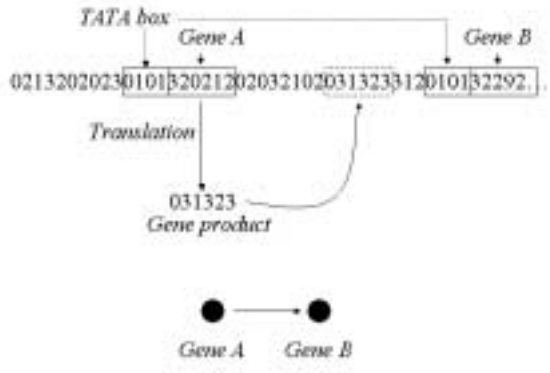

Figure 1. Reil's artificial genome model of a genetic regulatory network.

Ten random genomes and their corresponding networks were generated, with the parameters shown in Table 1. These parameter values were selected because previous experiments indicated that they produce networks which, when updated synchronously, show limit cycle dynamics, but when updated asynchronously rapidly collapse to a point attractor (Hallinan & Wiles, 2004).

Table 1. Parameters used for network generation

Parameter	Value
Chromosome length	14000
Number of bases	4
Gene Marker	0101
Gene length	6
Proportion of inhibitory links	0.4
Maximum number of timesteps	1000

The resulting networks had an average of 57.9 nodes and 172.3 links, giving an average connectivity of 2.97.

Our model differs from that of Reil (1999) in two major ways. One is the manner in which inhibitory links are implemented. In the original model a link was deemed inhibitory if its last base has a particular value. Using this approach a network can have a degree of inhibition of 0, 0.25, 0.50 or 1.0, and the links emanating from a given gene will always be either inhibitory or excitatory, no matter which to gene it links. We designate individual links as inhibitory with a specific probability as they are formed. This scheme allows much finer-grained inhibition, and is more biologically plausible in that it allows a single gene product to participate in some reactions as an inhibitor and in others as an activator.

The other difference, as discussed above, is our update scheme. Instead of synchronous updating we use a synchrony parameter, *s*, which represents the probability of a node being updated at any time step. This approach has been applied to the Game of Life, which shows characteristic phase transitions (Blok and Bergerson, 1999).

Each network was run *n* times, where *n* is the number of nodes in the network, each time with a different initial node activated, for synchrony values ranging from 0.1 to 1.0. With a synchrony value of 1.0 each node has a chance of 1.0 of being updated at each timestep, making it a synchronous updating scheme. In this way much of the state space of each network is explored in a systematic manner, although the stochastic element in the update scheme means that the entire state space is almost certainly not fully explored.

Each network was run for 1000 timesteps and a record kept of the states visited. This state list was then used to construct a state transition diagram for each network run, from which statistics pertaining to the number and length of limit cycles encountered could be compiled.

Results

Typical state transition graphs for networks with low synchrony (0.3) and high synchrony (0.9) are shown in Figure 2.

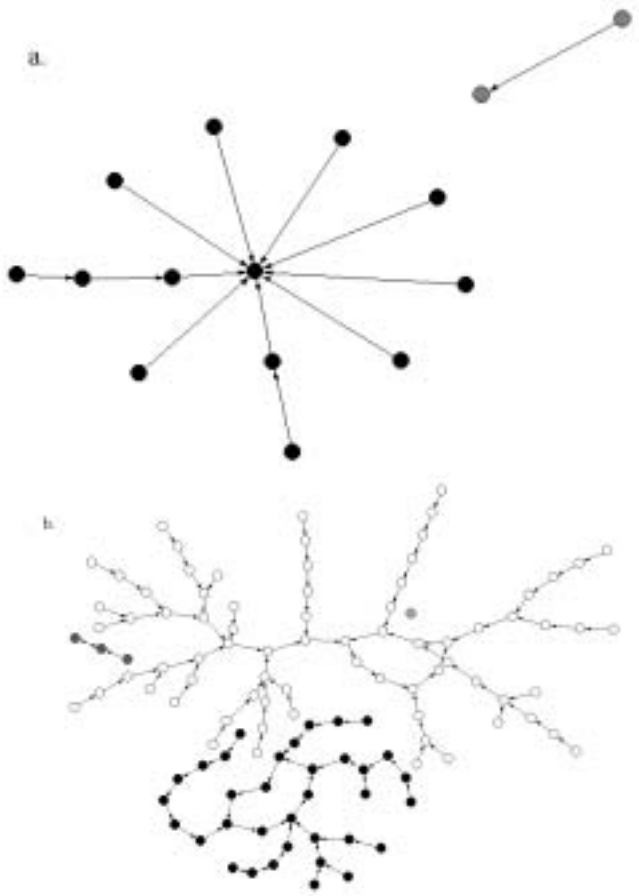

Figure 2. State transition diagrams. a. Network with synchrony 0.3 has two basins of attraction, each leading to a single point attractor and the longest transient in the network is 3. b. Same network with synchrony 0.9 has four basins of attraction, three of which have limit cycle attractors and the longest transient is >15.

The networks with high synchrony visit many more states, and settle to longer attractors than the networks with lower synchrony. This pattern is common to all of the networks; the network dynamics change with the degree of synchrony.

Despite the higher number of basins of attraction in the more synchronous networks, the proportion of all genes in the largest basin of attraction decreases relatively slowly with increasing synchrony (from 0.95 to 0.61), whereas the proportion of genes actually active in each time step increases sharply. In networks with very low synchrony very few genes are active per time step (Figure 3).

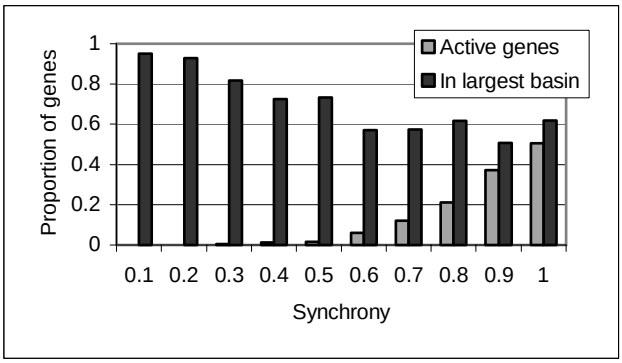

Figure 3. Change in network dynamics with synchrony as evidenced by the proportion of genes active per time step and proportion of states in the largest basin of attraction.

As expected, networks with low synchrony displayed fewer and shorter limit cycles that those with high synchrony. The relationship between synchrony and limit cycle behavior is not, however, linear. For most measures of dynamic behavior a peak occurs a synchrony rate of about 0.9 (Figure 4).

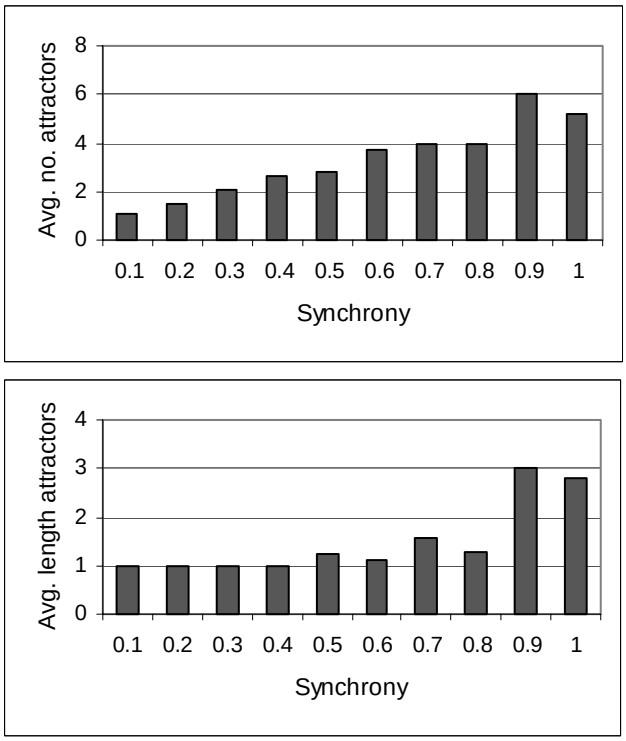

Figure 4. Average length and number of attractors in networks with different levels of synchrony.

Limit cycle length varied from 1 to 14 states, with most limit cycles less than five states long. The longest limit cycles were found, once again, at a synchrony of 0.9 (Figure 5).

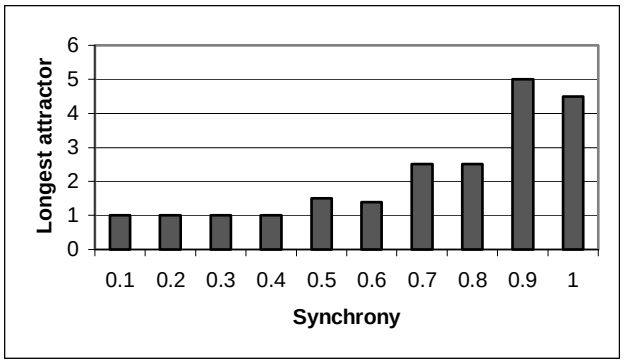

Figure 5. Average length of limit cycles for networks with different levels of synchrony

Although the number of long limit cycles increases with increasing synchrony, all networks have a significant proportion of point attractors (Figure 6).

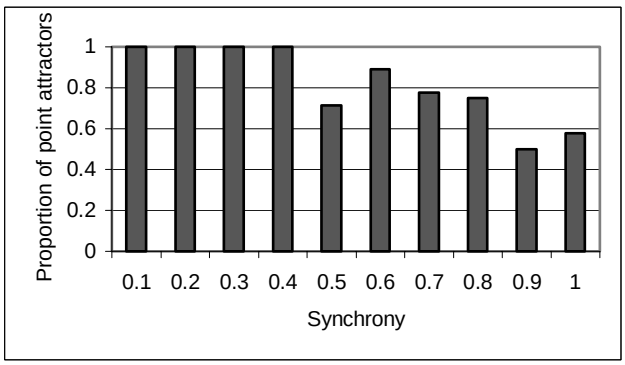

Figure 6. Proportion of attractors which are point attractors, for different levels of synchrony

Discussion

An issue in the design of artificial life simulations is the identification of which features of biological systems are most significant for robust dynamics and what kinds of abstractions provide informative analogues. Computational models of genetic regulatory networks are increasingly taking into account the role of stochasticity, not as an irrelevant detail, but as a core property of the system.

The observation that networks updated asynchronously display limited dynamic behaviour, collapsing rapidly to a stable state, is of concern if these networks are to be used as models of genetic regulatory networks. In previous work we have shown that networks can be evolved to display rich dynamic behaviour under asynchronous updating (Hallinan & Wiles, 2004). The current work builds upon this finding be exploring in more detail the nature of asynchrony in order to make the simplified Boolean model more biologically relevant.

We implemented the degree of asynchrony as a parameter, *s*. Networks with *s* of below about 0.5 have very limited dynamic behavior, tending to move rapidly to a single point attractor from any point in the state space. In contrast, networks with *s* between 0.5 and 1.0 display a range of interesting dynamic behaviors, with the number and length of fuzzy limit cycles increasing with *s* to peak at about 0.9. Interestingly, fully synchronous networks, with an *s* of 1.0, have fewer and shorter limit cycles, on average, than networks with a *s* of 0.9. We are currently testing whether this holds in larger systems, examining factors such as the number of active n nodes, and the way in which basins of attraction change with changes in parameter values.

All of the networks in our study were run repeatedly, with a single node active at timestep 1, a different node each time. This means that each network explored its state space *n* times (where *n* is the number of nodes in the network) per value of *s*, and the order of exploration was the same each time. Although the entire state space of any network is not fully explored, because of the stochasticity in the update rule, this protocol means that the number of trajectories used for state space exploration is equivalent each time, and the state transition diagrams reflect the different dynamics of the networks under different update schemes. As *s* increases, the state spaces of the networks tend to have more basins of attraction and longer limit cycles. The basins of attraction are also larger, as reflected by the longer transients evident in networks of higher *s* (Figure 2).

Although networks with higher *s* have more long limit cycles, all networks have a significant proportion of point attractors. Once again, the proportion of attractors that were point attractors was lowest at an *s* of 0.9 (0.5 compared with 0.57 for $s = 1.0$).

Genetic regulatory networks are characterized by robust temporal dynamics. Kauffman (1993) hypothesizes that an attractor in gene expression space represents a cell type. Models of genetic regulatory networks such as the artificial genome, display limit cycle behavior with large basins of attraction. A large basin of attraction for a limit cycle is necessary for robustness; it implies that a small perturbation in gene expression will leave the network in the same basin of attraction, eventually to return to the same attractor. Cell type is therefore stable, as observed in biological systems.

The observation that both limit cycle length and size of basin of attraction reach a maximum at an *s* of less than 1.0 demonstrates that biologically plausible dynamic behavior in a genetic regulatory network is not dependant upon synchronous node updating. Although the asynchrony parameter, *s*, is still an extreme simplification of the multiplicity of variable temporal delays induced by differential rates of transcription, translation, degradation and other cellular processes, it illustrates that stochasticity, feedback and delay are fundamental to network robustness in both biological and computational networks.

References

Abramson, G. & Zanette, D. H. 1998. Globally coupled maps with asynchronous updating. Physical Review e 58(4): 4454 - 4460.

Blok, H. J. & Bergersen, B. 1999. Synchronous versus asynchronous updating in the "game of Life". Physical Review e 59(4): 3876 - 3879.

Di Paolo, E. A. 2001. Rhythmic and non-rhythmic attractors in asynchronous random Boolean networks. Bio systems 59: 185 - 195.

Hallinan, J. & Wiles, J. 2004. Evolving genetic regulatory networks using an artificial genome. In Chen, Y.-P. (ed.) 2nd Asia-Pacific Bioinformatics Conference (APBC2004), Dunedin, New Zealand. Conferences in Research and Practice in Information Technology, Vol. 29.Australian Computer Society, Inc.

Harvey, I. & Bossomaier, T. 1997. Time out of joint: Attractors in random Boolean networks. In P. Husbands & I. Harvey (eds.) Proceedings of the Fourth European Conference on Artificial Life (ECAL97). Cambridge, MA: MIT Press: 67 - 75.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabasi, A.-L. 2000. The large-scale organization of metabolic networks. Nature 407: 651 - 654.

Kauffman, S. A. 1993. The Origins of Order: Self-Organization and Selection in Evolution. New York: Oxford University Press.

Nehaniv, C. L. 2002. Evolution in Asynchronous Cellular Automata, In R.K. Standish, M.A. Bedau, and H.A. Abbass (eds.) Proceedings of the Eighth International Conference on Artificial Life, MIT Press, pp. 65-78.

Reil, T. 1999. Dynamics of gene expression in an artificial genome: Implications for biological and artificial ontogeny. In Floreano, D., Mondada, F. & Nicoud, J. D. (eds.) Proceedings of the 5th European Conference on Artificial Life. New York, NY: Springer Verlag: 457 - 466.

Schonfisch, B. & de Roos, A. 1999. Synchronous and asynchronous updating in cellular automata. Bio systems 51: 123 - 143.

Sole, R. & Pastor-Santorros, R. 2002. Complex networks in genomics and proteomics. Santa Fe Institute Working Paper 02-06-026.

Wagner, A. & Fell, D. 2001. The small world inside large metabolic networks. Proceedings of the Royal Society of London, Series B 268: 1803 - 1810.