

# Is there a Liquid State Machine in the Bacterium *Escherichia Coli*?

Ben Jones, Dov Stekel<sup>o</sup>, Jon Rowe, Chrisantha Fernando\*  
School of Computer Science, School of Biosciences<sup>o</sup>  
University of Birmingham  
Birmingham, UK  
Email: C.T.Fernando@cs.bham.ac.uk

**Abstract**—The bacterium *Escherichia coli* has the capacity to respond to a wide range of environmental inputs, which have the potential to change suddenly and rapidly. Although the functions of many of its signal transduction and gene regulation networks have been identified, *E.Coli*'s capacity for perceptual categorization, especially for discrimination between complex temporal patterns of chemical inputs, has been experimentally neglected. Real-time computations on time-varying inputs can be undertaken by a system possessing a high dimensional analog fading memory, i.e. a liquid-state machine (LSM). For example, the cortical microcolumn is hypothesized to be a LSM. A model of the gene regulation network (GRN) of *E.Coli* was assessed for its LSM properties for a range of increasingly complex stimuli. Cooperativity between transcription factors (TFs) is necessary for complex temporal discriminations. However, the low recurrence within the GRNs autonomous dynamics decreases its capacity for a rich fading memory, and hence for integrating temporal sequence information. We conclude that coupling of the GRN with signal transduction networks possessing cross-talk, and with metabolic networks is expected to increase the extent of non-autonomous recurrence and hence to facilitate enhanced LSM properties.

## I. INTRODUCTION

Maass et al [1] produced a model called the Liquid State Machine (LSM). They hypothesize that a high dimensional recurrent dynamical network designed such that it can act as a universal analog fading memory can possess transient internal states. In comparison to a traditional neural network approach whereby the circuit represents a 'decision' by reaching one of many alternate stable attractors, the transient states of an LSM are capable of being observed by linear readout elements to produce stable outputs. This allows robust real-time processing of time-varying inputs. All that needs to be trained, is a readout module consisting of linear perceptrons. Relatively few parameters are required to specify the LSM, provided that a separation property can be achieved<sup>1</sup>. Training (by evolution or lifetime learning) would only need to take place on a layer of linear readout elements instead of on the recurrent neural network itself.

LSMs have been successfully implemented in a wide range of media, from spiking neurons in cortical microcolumns [2],

<sup>1</sup>The separation property requires that the distance between the trajectories of internal states of the system is roughly proportional to the distance between two different input streams that caused them.

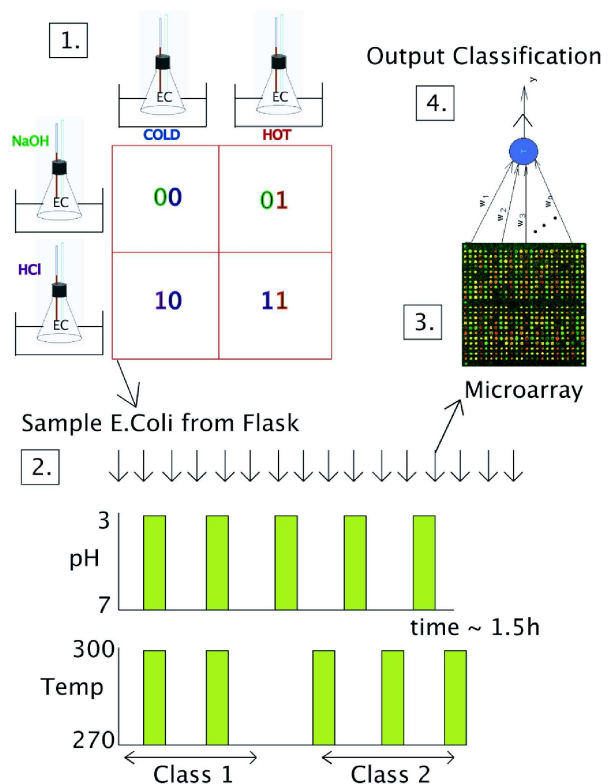


Fig. 1. The proposed sequence discrimination experiment with *E.Coli*. See text for details.

to the surface of water in a bucket [4]. They have been shown to be capable of robust speech recognition and real-time robot control [3]. Furthermore, multiple readout elements can be trained to make different discriminations by observing the same LSM.

We ask whether the GRN of *E.Coli* can act as a LSM and subsequently whether *E.Coli* is capable of making complex perceptual discriminations between temporal sequences of chemical inputs. We test the hypothesis that the GRN acts as the liquid, and that a structural intra-cellular protein or a DNA sequence acts as the linear readout element. Within

this framework, the free concentration of the DNA is taken as a function of its evolvable binding affinities to the various proteins and mRNAs whose production is controlled by the GRN.

The model we build here is a first step at simulating the wet experiment shown in figure 1. A population of E.Coli with mean replication period greater than 1.5 hours are placed in a flask, and exposed to the four possible combinations of high and low temperature, and low and neutral pH. The pairs are labelled using the binary coding scheme shown in part 1 of figure 1. Various sequences of combinations are administered, and perceptual categories are defined, e.g. as shown in part 2 of figure 1, 'in phase' is considered to be class 1 and 'out of phase' is considered to be class 2. At regular intervals, samples of E.Coli are removed and a microarray is then taken to sample the state of the GRN; see part 3. Finally, a parallel perceptron is trained to classify the set of microarray data into the desired classes; see part 4. Our model of this experiment predicts that a perceptron may indeed be expected to be able to discriminate between complex environmental sequences to which E.Coli were exposed, merely by observing the mRNA concentration vectors, and therefore that E.Coli is capable in theory of discriminating between temporal sequences of inputs.

## II. METHODS

### A. Topology

The gene transcription network of E.Coli is available from various sources such as RegulonDB and Ecocyc, [5], but we chose Uri Alon's [6] modification of Shen-Orr's data [7] due to its convenient format. It consists of 577 interactions between 116 TFs and 419 operons. The topological properties of the network [8] and the dynamical properties of species operons [10] of the network have been investigated [9], and it is found that the GRN is arranged largely in a feed-forward hierarchy but with some recurrence [11]. The dataset only gives the topology and sign of interactions, i.e. whether a TF is an activator, a repressor or a dual, at a particular Operon. It does not provide  $K_m$  values for binding of TFs to the operator, nor does it provide information about cooperative interactions between TFs at the same Operon, or between operators at a distance by DNA looping [12]. In the absence of such thermodynamic data for most TFs, we were forced to generate the parameters from 'reasonable' statistical distributions.

### B. Transfer Functions

The formalism of Bintu et al [12] is used to describe the transfer function at each node (Operon) due to the interaction of TFs with the RNAP (RNA Polymerase). The probability that an RNAP is bound to an Operon is given by the equation,

$$P_{bound} = \frac{1}{1 + \frac{N_{NS}}{P F_{reg}} e^{\Delta\epsilon_{pd}/k_B T}}, \quad (1)$$

where  $N_{NS}$  is the number of non-specific binding sites for RNAP ( $5 \times 10^6$ ).  $P$  is the number of RNAP molecules (2000, [15]).  $F_{reg}$  is a regulation factor that describes the effective increase ( $F_{reg} > 1$ ) or decrease ( $F_{reg} < 1$ ) of the number of RNAP molecules available to bind to the promoter.  $\Delta\epsilon_{pd} = k_B T \ln(K_{pd}^S/K_{pd}^{NS})$ , with  $k_B = 1.38 \times 10^{-23}$  joules/K (Boltzman's constant),  $T = 280K$ ,  $K_{pd}^S$  and  $K_{pd}^{NS} = 10000nM$  are the equilibrium dissociation constants for specific binding and non-specific binding of RNAP to this promoter. Values of  $K_{pd}^S$  are generated randomly according to a log normal distribution with the mean 544nM, and variance 267nM. For example, the lac promoter has  $K_{pd}^S = 550nM$  and the T7 promoter has  $K_{pd}^S = 3nM$ .

TFs stabilize or destabilize the binding of the RNAP to the promoter in a manner described by  $F_{reg}$  (see Table 1 from [12]).  $F_{reg}$  for a simple repressor, R, is

$$F_{reg} = \left(1 + \frac{[R]}{K_R}\right)^{-1} \quad (2)$$

with  $K_R$  being the equilibrium dissociation constant of the repressor, generated from the log normal distribution with mean 0.1nM and variance 0.04nM. Note the tendency for repressors to bind more strongly than RNAP molecules on the whole.  $F_{reg}$  for a simple activator, A, is,

$$F_{reg} = \frac{1 + \frac{[A]}{K_A} f}{1 + \frac{[A]}{K_A}} \quad (3)$$

where  $f = e^{-\epsilon_{xp}/k_B T}$ , with  $\epsilon_{xp}/k_B T$  being the 'glue-like' interaction between TF (x) and RNAP (p). It is randomly generated according to the log normal distribution with mean -3.14, and variance 0.93.  $K_A$  is generated from the same distribution as  $K_R$ .  $F_{reg}$  for dual repressors, R1 and R2, cooperatively interacting is,

$$F_{reg} = \left(1 + \frac{[R_1]}{K_{R_1}} + \frac{[R_2]}{K_{R_2}} + \frac{[R_1][R_2]}{K_{R_1}K_{R_2}}\omega\right)^{-1} \quad (4)$$

where  $\omega = e^{-\epsilon_{x_1x_2}/k_B T}$  is the cooperativity factor,  $\epsilon_{x_1x_2}/k_B T$  being drawn from a log normal distribution with mean 1.6 and variance 1.13.  $F_{reg}$  for dual activators, A1 and A2, cooperatively interacting is,

$$F_{reg} = \frac{1 + \frac{[A_1]}{K_{A_1}} f_1 + \frac{[A_2]}{K_{A_2}} f_2 + \frac{[A_1][A_2]}{K_{A_1}K_{A_2}} f_1 f_2 \omega}{1 + \frac{[A_1]}{K_{A_1}} + \frac{[A_2]}{K_{A_2}} + \frac{[A_1][A_2]}{K_{A_1}K_{A_2}}\omega} \quad (5)$$

Where a promoter is controlled by a single TF, then either equation (2) or (3) is used to calculate  $F_{reg}$ . Where there is more than one TF acting on a promoter, for each possible pair of TFs, a cooperative interaction is generated with probability  $P_c$ , typically set to 0.3. Any TFs not included in cooperative

interactions are assumed to act independently on the RNAP according to equation (2) or (3). On a complex promoter with more than one TF, the total  $F_{reg}$  is given by the product of component  $F_{reg}$ s calculated for each cooperative pair and each independently acting TF.

The value of  $P_{bound}$  is used to calculate the rate of mRNA production at an Operon by assuming that the RNAP produces mRNA at 50 nucleotides per second and that the average length of mRNA is 1000 nucleotides. This gives an mRNA production in Moles of mRNA per second for a cell of volume  $1.5 \times 10^{-15}$  of  $3.6 \times 10^{-8} Msec^{-1} Operon^{-1}$ . Each operon is assumed to be present in copy number 1. Protein production rate is first order in [mRNA] and is assumed to be 15 amino acids per second. This gives a value of protein production of  $apx 0.0001 Msec^{-1}$  per Mole of mRNA in an E.Coli sized cell.

The rate of decay of mRNA is first order and is given by the log normal distribution with mean 4.85 minutes and variance 4.08 minutes. The rate of decay of protein is also first order and is given by the log normal distribution with mean 4.05 minutes and variance 28.25 minutes. These values are scaled by  $4.3 \times 10^{-3}$  to give a decay rate with units  $Msec^{-1}$  per Mole of Protein.

The  $K_R$  or  $K_A$  of a transcription factor can be altered by external signal molecules (inhibitors), which bind to the TF and alter its chances of binding to the operator. The extent of inhibition by a signal molecule is given by calculating a modified K, given by  $K_{mod} = K/(1 + [I]/K_I)$ , with  $K_I$  drawn from a log normal distribution with mean  $0.09 \times 10^{-3}M$ , and variance  $0.07 \times 10^{-3}M$ . Typically in experiments where the E.Coli is stimulated, external signal molecule concentrations are administered in a square wave with baseline 0M and plateau 0.001M.

The network is initialized with low concentrations of TFs, e.g.  $10^{-8}M$ , and allowed to settle to its steady-state. The volume of the cell is assumed to be  $1.5 \times 10^{-15}$  Liters. A C++ program is written to run the ordinary differential equations by Euler Integration with a fixed time-step of 0.003 seconds, for 1800000 time-steps. The entire stimulus duration is 1.5 hours,  $apx 3$  times the minimum replication period of E.Coli. If much shorter stimulus durations are used, the transcription network does not have time to display significant differences in mRNA or protein concentrations.

### C. Discrimination Tasks

1) *The XOR Problem:* The XOR problem is a linearly non-separable problem, i.e. a single-layer perceptron cannot solve it. The inputs  $\{0,0\}$  and  $\{1,1\}$  must be classified as class 0, and  $\{0,1\}$  and  $\{1,0\}$  must be classified as class 1. Therefore a simple threshold that could be used to solve the OR or AND tasks cannot be used to solve the XOR task. Each digit of the XOR input is represented as a randomly

chosen 50% of inhibitor concentrations. '1' is represented as a high concentration, and '0' as a zero concentration of inhibitors. It is assumed that instantaneous changes in inhibitor concentration can be induced, although it is our intention to add a much more realistic signal transduction model in further work. The stimuli are presented in the order shown in figure 2. The perceptron is then trained to discriminate class 1 from class 0 on the basis of the mRNA and protein concentrations alone.

2) *Distinguishing the phase of two signals: A and B:* Imagine that the first digit of the XOR input represents some chemical A in the environment, and the second digit represents some chemical B. Can a perceptron be trained to distinguish between an environment in which A and B occur together at the same time compared to an environment in which A and B occur in-between each other? This is effectively a special case of the XOR task, produced by clustering the presentations of classes; see figure 3. When the 00 and 11 inputs are presented repeatedly, this is equivalent to A and B occurring together. When the 10 and 01 inputs are presented repeatedly, this is equivalent to A and B occurring in-between each other, i.e. in phase but shifted by half a period. In fact the same class definition as the XOR problem is used since a chain of 00 and 11 inputs will be in class 0, and a chain of 01 and 10 inputs will be in class 1.

### D. Linear Readout by Multiple Perceptrons

A variant of the P-delta rule [13] is used to train multiple perceptrons to classify the stimuli into their appropriate classes. Protein and mRNA concentrations are used (separately) as inputs to each perceptron making up the collection of parallel perceptrons. Summation of the outputs of the perceptrons gives an indication of the certainty of the classification. If the classification is incorrect, then the perceptrons that failed to classify correctly are updated. The perceptron weight update equation for a single perceptron is,

$$w_{i+1}(t+1) = w_{i+1}(t) + \eta(d(t) - y(t))x_i(t) \quad (6)$$

The bias weight,  $w_0$  is set to -1 for all perceptrons;  $d$  and  $y$  are the target and computed outputs respectively for the whole input vector  $\mathbf{x}$ , and  $\eta$  is the learning parameter. In line with the P-delta rule, the weight values are also stabilized to yield a parameter  $\mu$ , which can either be positive or negative (equation 7). The value  $\varepsilon$  is set to 0.1 throughout and the parameter  $\gamma$ , which specifies the margin around zero, is set to 0.01, also throughout.

$$\mu = \begin{cases} +\varepsilon & \text{if } \gamma \geq \mathbf{w} \cdot \mathbf{x} \geq 0 \\ -\varepsilon & \text{if } -\gamma \leq \mathbf{w} \cdot \mathbf{x} < 0 \end{cases} \quad (7)$$

Depending on the output of the dot product  $\mathbf{w} \cdot \mathbf{x}$ , therefore, the value of  $\mu$  is either positive or negative. It is then factored in to the weight update equation,

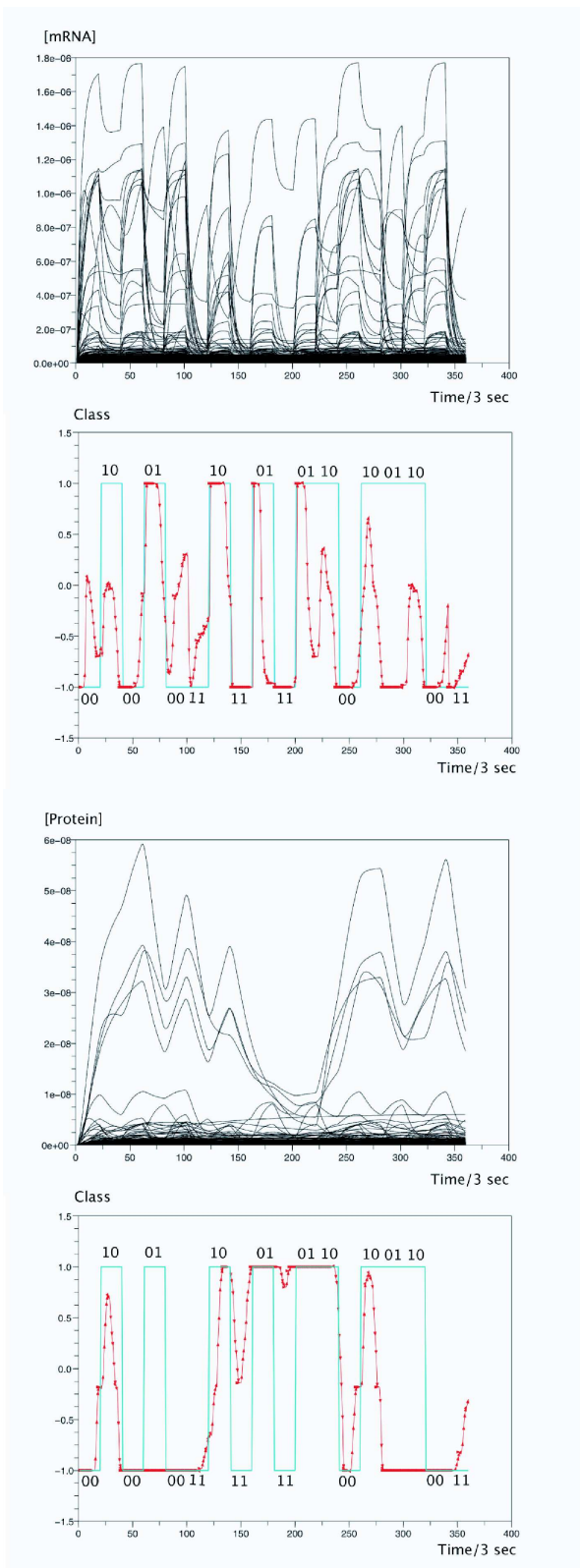


Fig. 2. Perceptrons trained on random presentation of 00,10,01,11 cases, requiring XOR classification. *Top 2*: The top graph shows mRNA concentrations over time, followed below by the target classes superimposed upon the summed output of 50 perceptrons (Dotted line), with  $MSE = 0.07$ . The binary two digit number indicates which of the combinations of the two inputs were presented at that period. *Bottom 2*: As above, but perceptrons were trained on protein concentration rather than mRNA concentration.  $MSE = 0.08$

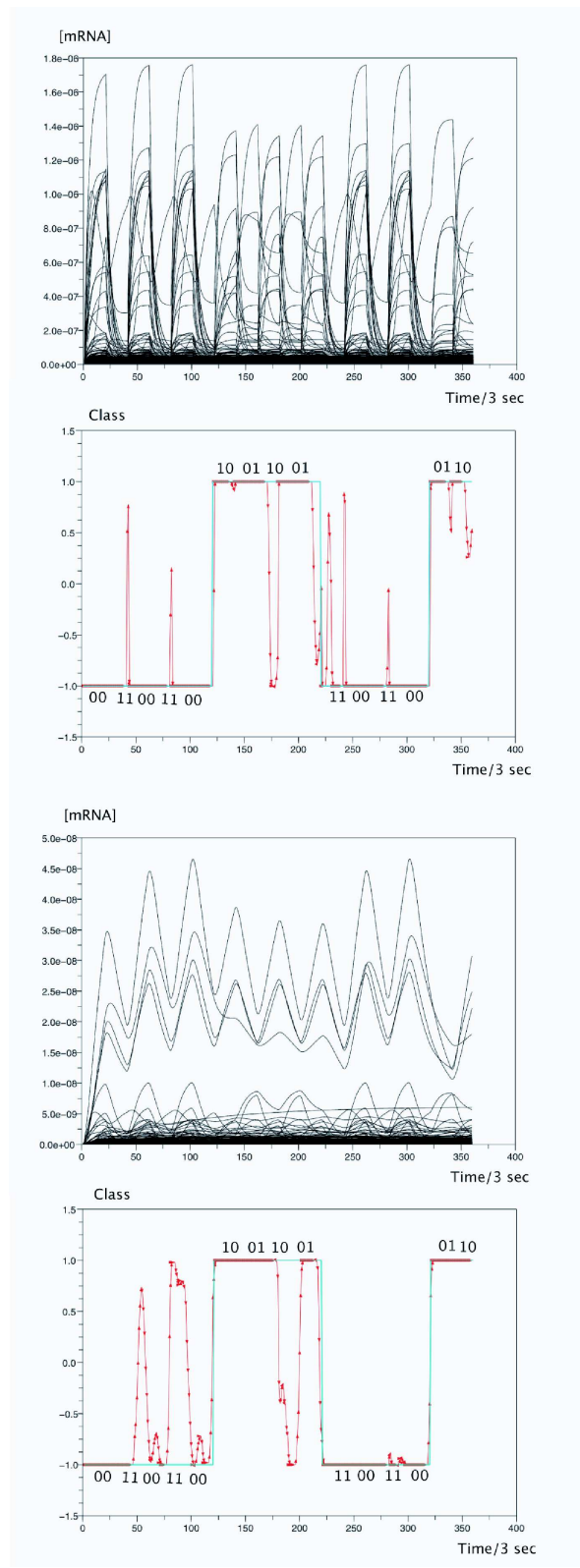


Fig. 3. Parallel perceptrons were trained to discriminate between in-phase and out-of-phase chemical 'spikes'. *Top 2*: The top graph shows mRNA concentrations over time, followed below by the target classes superimposed upon the summed output of 50 perceptrons (Dotted line),  $MSE = 0.02$ . *Bottom 2*: As above, with perceptrons trained on protein concentrations,  $MSE = 0.03$ .

XOR mRNA	acs, dps, flhE, fucAO, glpACB, <b>narGHJI</b> , nhaA tdcABCDEFG, tsx.
XOR Protein	narGHJI, clpP, focA_pflB, glnALG, hycABCDEFGH, <b>narGHJI</b> , nfo, ompF, osmC
AB mRNA	acs,soda, tdcABCDEFG
AB Protein	flhE, <b>focA pflB</b>

Fig. 4. Most important TFs and mRNAs involved in the classifications for the XOR, and AB tasks.

$$w_{i+1}(t+1) = w_{i+1}(t) + \eta(d(t) - y(t)) \mu x_i(t) \quad (8)$$

In effect, the value of  $\mu$  establishes a margin around the dot product  $w \cdot x$ , preventing the dot product from getting too close to zero.

### III. RESULTS

Figure 2 and 3 show the outputs of a trained parallel perceptron on the XOR and the AB task. In both cases the parallel perceptron imperfectly, but significantly above chance, identifies the target classes correctly. The mean squared errors are shown in the figure captions, and are all less than 0.1. Both mRNA and protein concentrations can be used to make the classification. Performance is in fact better on the AB task in figure 3 than on the randomized presentations of the inputs in figure 2. Further experiments are required to test the capacity of the trained perceptron to generalize to stochastic noise in the data, and to reordering of class presentations. A perceptron capable of generalization from the XOR classification task should be able to exhibit good performance in the AB task.

The most important mRNAs and proteins involved in the classification (i.e. those with the largest perceptron weights) were examined for all 4 cases and shown in table 4. These may differ given different random initializations of K values.

### IV. CONCLUSION

Neural network metaphors have been used to describe the information processing taking place within GRNs [14]. We undertook to assess the capacity of the GRN of E.Coli, at least the version found in [6], to behave as a liquid state machine. Only partial success in solving the XOR problem was achieved, although good performance in the AB task was observed. Further experiments are required to systematically test whether other intra-cellular networks in E.Coli could contribute to produce a richer LSM than is possible using the GRN model alone, which is of relatively low dimensionality, and contains limited recurrence. For example, signal transduction networks and metabolic networks will be added to the model, and their contributions to the LSM will be examined. The protein decay rates in the current model are somewhat too rapid. In real life many proteins

have decay rates much longer than the lifetime of an E.Coli. This suggests the possibility that perceptual discriminations could actually take place over many generations. Our model does not consider the effect of cell division. We hope that an experimentalist may be encouraged to conduct a similar experiment in a wet lab.

### ACKNOWLEDGMENT

Jon Rowe and the ESIGNET 6th Framework grant for research on the evolution of cell signaling networks.

### REFERENCES

- [1] W. Maass, T. Natschlagler, and H. Markram. *Real-time computing without stable states: A new framework for neural computation based on perturbations*. Neural Computation, 14(11):2531-2560, 2002.
- [2] W. Maass, T. Natschlagler, and H. Markram. *Computational models for generic cortical microcircuits*. In J. Feng, editor, Computational Neuroscience: A Comprehensive Approach. CRC-Press, 2002.
- [3] W. Maass, R. A. Legenstein, and H. Markram. *A new approach towards vision suggested by biologically realistic neural microcircuit models*. In Proc. of the 2nd Workshop on Biologically Motivated Computer Vision, Lecture Notes in Computer Science. Springer, Nov. 2002.
- [4] C. Fernando, and S. Sojakka. *Pattern recognition in a Bucket*. LNCS. Advances in Artificial Life. 2801:588-587. 2003.
- [5] <http://ecocyc.org/>
- [6] <http://www.weizmann.ac.il/mcb/UriAlon/>
- [7] S Shen-Orr, R Milo, S Mangan and U Alon, *Network motifs in the transcriptional regulation network of Escherichia coli*. Nature Genetics, 31:64-68, 2002.
- [8] R Milo, S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii and U Alon, *Network Motifs: Simple Building Blocks of Complex Networks Science*, 298:824-827, 2002
- [9] U. Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall, 2007.
- [10] N. Yildirim, and M. Mackey. *Feedback Regulation in the Lactose Operon: A Mathematical Modeling Study and Comparison with Experimental Data* Biophysical Journal, 84:2841-285, 2003
- [11] H. Ma, J. Buer, and A. Zeng. *Hierarchical Structure and modules in the Escherichia Coli transcriptional regulatory network revealed by a new top-down approach*. BMC Bioinformatics. 5:199, 2005.
- [12] L. Bintu et al. *Transcriptional regulation by the numbers: models*. Curr Opinion. in Genetics and Development. 15:116-124, 2005
- [13] P. Auer, H. Burgsteiner, and W. Maass. *Reduced Communication for Distributed Learning in Neural Networks* LNCS. ICANN 2002
- [14] N.E. Buchler, U. Gerland, T. Hwa. *On schemes of combinatorial transcription logic* Proc. Nat. Acad. Sci. 100(9): 5136-5141, 2003.
- [15] D.C. Grainger, D. Hurd, M. Hudson, J. Holdstock, and S.J.W. Busby. *Studies of the distribution of Escherichia coli cAMP-receptor protein and RNA polymerase along the E.Coli chromosome* Proc. Natl. Acad. Sci. 102(49): 17693-17698, 2005.