Title: Reinforcement Learning in Synthetic Gene Circuits

Adrian Racovita[1] and Alfonso Jaramillo[*,1,2,#]

[1]Warwick Integrative Synthetic Biology Centre (WISB) and School of Life Sciences, University of Warwick, CV4 7AL, Coventry, UK

[2]Institute for Integrative Systems Biology (I2SysBio), University of Valencia-CSIC, 46980 Paterna, Spain.

[*] On leave of absence from CNRS, France.

[#] To whom correspondence should be addressed: Prof. Alfonso Jaramillo. School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK, Alfonso.Jaramillo@synth-bio.org, Tel: +44 (0)24 765 7343

Abstract:

Synthetic gene circuits allow programming in DNA the expression of a phenotype at a given environmental condition. The recent integration of memory systems with gene circuits opens the door to their adaptation to new conditions and their re-programming. This lays the foundation to emulate neuromorphic behaviour and solve complex problems similarly to artificial neural networks. Cellular products such as DNA or proteins can be used to store memory in both digital and analog formats, allowing cells to be turned into living computing devices able to record information regarding their previous states. In particular, synthetic gene circuits with memory can be engineered into living systems to allow their adaptation through reinforcement learning. The development of gene circuits able to adapt through reinforcement learning moves Sciences towards the ambitious goal: the bottom-up creation of a fully-fledged living artificial intelligence.

Introduction

Living systems are able to efficiently adapt to the environment by adjusting the regulation of their encoded genes. A cell may adapt either by modifying its genome or by regulation of gene expression. Cells learn to solve problems by adapting the regulation which controls alternative decisions. Re-engineering this regulation could allow for the creation of systems able to attain expertise at a given task. In a way, this approach merges the concepts of Darwinian evolution with those of neuroscience in order to outline a synthetic circuit capable of adaptation through learning. The challenge lies in defining the aspects of cell memory and learning while satisfying the constraints imposed by the experimental implementation. Any attempt at outlining a machine learning model in biology would be hindered by the usual difficulties associated with this field: what would the optimal structure for such a system be, how can the meta parameters be defined and ultimately how can novel genetic devices capable of storing and reading memory be engineered? With this in mind, researchers have collectively aimed towards the creation of a biological artificial intelligence[1-5]. Engineered cells able to make decisions would allow for numerous applications in the fields of personalized medicine or environmental protection.

Memory can be regarded as the ability of a given system to store information about its previous states[6]. In synthetic biology, memory allows for the generation of a desired output based on receiving and amplifying a specific external input in a reproducible manner. *Analog memory* encodes information as a gradual response, usually defined as ranging from none to complete activation, an aspect which makes it sensitive to regulation and commonly altered by noise. *Digital memory* on the other hand offers a more discrete

implementation as the output can only be null or fully induced. This approach is less sensitive to noise but is accompanied by the limitation of producing complex adaptive landscapes that difficult optimisation[7,8]. *Learning* in computation systems defines the ability to modify memory in response to a set of inputs, either supervised or unsupervised. Multiple types of learning can be outlined and implemented through synthetic gene circuits. *Supervised learning* assumes availability of correct input-output pairs as training examples[9,10], that the algorithm tries to generalise to a larger input space. In biological systems, the association between the environmental cues and the produced change in the memory generates an altered phenotype. This association can be adjusted to generate a desired phenotype. On the other hand, *unsupervised learning* is achieved without the need of known targeted phenotypes and with minimal external guidance. In this case, learning is based on successive alternative phenotypes. In living organisms, such phenotypes are produced in response to multiple environmental cues to adjust the memory. Learning shares characteristics of evolution as it infers acquiring new or improved functions, while adaptation refers to carrying out the same tasks but in different conditions. Since developing new traits leads to adaption, one may infer that the mechanisms behind learning and Darwinian adaptation can share some similar characteristics[6]. *Reinforcement learning* differs from other types of learning as it is achieved through maximising a reward signal rather than using the relation between inputs and outputs for acquiring new functions[9]. It assumes the availability of a long-term reward function that maps the history of the system to the reward. Reinforcement learning algorithms then adapt the memory to maximise the long-term reward. This innate ability to adapt without having to change the memory through external intervention is what suggests that reinforcement learning is the most suitable option for learning with synthetic gene circuits.

In multicellular organisms, memory is recorded through changes in cellular connections[11]. The engineering of learning with synthetic gene circuits aims to find ways to emulate this mechanism at singular cell level, usually employing newly engineered gene circuits or natural ones exploited to fulfil a new role[3]At the base of any system capable of learning there is an internal memory and a regulatory mechanism through which the memory of its states can be changed. The way learning is achieved is strongly related to the type of memory the system carries. Thus, digital-type memories are only able to record information regarding their ON and OFF states and have been considered the most viable for a number of successful applications[12,13]. On the other hand, devices with analog memory are able to store information in relation to the strength of the signal applied, allowing for intermediate states to be recorded. Throughout the history of computer science, the trend has been to switch from analog systems to digital to solve the problem of intrinsic noise of analog information. However, in living systems, analog mechanisms already provide adaptable systems able to give robust outputs in spite of noise.

The following sections will present some of the most recent developments in the engineering of memory systems with synthetic gene circuits. These include the *in vivo* development of digital memory devices, the transition to more tuneable analog systems as well as the breakthrough innovations towards creating a biological memory reprogrammable through reinforcement learning, the focus of this paper. Moreover, we will analyse the concept of hybrid systems able to implement both digital and analog characteristics[8]. Thus, the aim of this mini-review article is to analyse the current state of the art as well as propose novel approaches for reinforcement learning with synthetic gene circuits.

Digital memory

DNA encodes the memory necessary for evolution. DNA can be used as a means of writing and reading digital information *in vivo,* owing to the intrinsic storage capability of the genetic code[4,14-16]. Moreover, with the recent advances and cost reduction of DNA synthesis[17], researchers have developed approaches of how one can encode data into living cells and then retrieve it without damaging the engineered cells. A suitable example of how this is achieved is the use of a gene circuit that is constitutively inactive unless recombinases are used to flip the promoter sequence in the right orientation, thus switching on transcription[18-24]. Placing the activation of a gene under the control of an external input offers significant control to writing into the cellular memory. Through this type of synthetic learning, cells become able of conducting biological, noise-free binary computation, as well as passing the newly recorded information to future generations by means of permanent genome modifications. Moreover, multiple elements with these characteristics could be linked in circuits and networks, enabling more advanced computation with predictable outcomes.

A more advanced way of achieving *in vivo* binary computation is the use of a systems that cause site-directed mutations[17,25,26]. This offers increased accuracy as the mutations would only occur at the desired site. In this case, learning is manifested as a prescribed change in the DNA sequence, with the cell recording the memory of previous exposure to the mutagenesis system. Thus, the engineered cells carry memory able to be written and read in a predictable manner similar to electronic storage devices. Memory systems that do not rely on mutating the DNA is similar to a Lamarckian evolution[17]. This implementation enables

the analysis of the genome in different conditions, both as a function of time as well as signal strength[12]. More recent work used nucleotide-level resolution in order to transform cell genomes into digital memory devices[14].

Analog memory

Gene circuits could be engineered to exploit the analog features of molecules[8,27]. Analog behaviour started to be employed in the most recent innovation such as the CAMERA system[28]. The aim was to create a reliable genetic device meant to record the history of cellular events. The CAMERA system supported writing and erasing analog cellular memory via a controllable signal. The information was encoded as the ratio between two co-transformed plasmids and learning was achieved using antibiotic selection by introducing genes encoding for antibiotic resistance. A variation of the same mechanism involved the cleavage of one plasmid using Cas9 nuclease in order to change the plasmid ratio and thus the memory. The memory becomes accessible through sequencing, enabling the analysis of the mutations to infer the time profile of the variations in the environment. In another implementation, researchers engineered cells able to respond to both chemical inputs as well as light[29]. The changes induced by stimuli in the genomic DNA are permanent and can be re-traced to produce the timescale of their occurrence. A compelling proposal would be the use of hybrid systems able to conduct analog-digital computation[8]. This would imply the integration of an analog memory storage with digital computation through synthetic gene circuits. The analog aspect would allow to weigh two alternative digital gene circuits according to learning and adapt in order to decide towards the preferential option.

Adaptable Regulatory Systems

In the field of electronic computation, memristors are devices able to store memory related to their previous states[30]. The unique characteristic to maintain their resistance in absence of electrical charge allows for the interpretation of memristors as equivalents of synapses[31]. This property makes memristive systems viable candidates for biological computers. One way to achieve this would be using the functional aspects of memristors in order to link synthetic neuron-like components. Although not precise proxies of real neurons, these novel devices could still be linked via synthetic synapses to form artificial neural networks. Thus, a signal fired by a presynaptic neuron could be passed through a memristor which would then trigger an action potential in the postsynaptic neuron. The variable resistance of the memristor would serve as the synaptic weight which could be modified in order to use the system for learning. However, given the usual dynamic voltage range of memristors (0.2-2V) which are significantly above the biological voltage thresholds of approx. 100mV[31], there is a need to engineer a biological equivalent that would offer more realistic dynamics when emulating cognitive networks. This is currently the case with other implementations of machine learning, which have the challenge of an increasing demand of processing power.

Regardless of the application, the behaviour of memristive devices depends on their previous history. In biochemistry, the recent work of several groups has led to the construction of molecular memristors able to mimic cognitive functions[32-34]. One example of this is the diffusive memristor functioned using Silver nanoparticles. With this, it was possible to build a synthetic neuron which presented the characteristic integrate-and-fire dynamics. Through coupling multiple of these neurons, a fully functional memristive

artificial neural network could be built. However, this achievement was still hindered by the functional discrepancies between memristors and real synapses, but it could be employed in experiments of encoding information in the relative synaptic strengths. Other implementations involved using oxide-based memristors to emulate the Ca^{2+} dynamics which govern biological synapses[35]. This biology-driven approach, in addition to the previously described memory-storing capacities of memristors in electronics strengthens the potential of such devices to be used in biological computing.

One first step towards a biological memristive device involved the creation of bioinspired bio-voltage memristors[31]. This system exploited the active protein nanowires of *Geobacter sulfurreducens* which have previously been proven to catalyse Silver ion reductions[36]. This property enabled the creation of biochemical memristors, with the flux of Silver ions serving as the signalling flow of artificial synapses. Following this, a successful implementation included the analysis of the electrical properties of fungi belonging to *Pleurotus* spp.[37], a significant step taken towards building a biological memory-storing device. In a previous study[38], it was shown that fungi are able to record the action of exogenous stimuli as fluctuations in the electrical signalling between distinct members of the same population, in a similar fashion to how information is passed from sensory neurons to the central nervous system and then motor neurons. This analogy enabled the outline of a model of a fungal automaton able to implement logical functions. With these findings, the gap between organic and inorganic computation decreases, possibly paving the way towards a hybrid of the two.

Regulatory Systems Capable of Reinforcement Learning

The implementation of molecular reinforcement learning would become possible via the creation of a genetic memristor. Using such a genetic device would enable researchers to modulate its transcription rate by means of a promoter with tuneable strength which would function as the equivalent of a potentiometer. The advantage of a genetic memristor would be the tuning of gene regulation according to its history which would offer the cells unsupervised adaptation based on the output. We call memregulon (as a contraction of memory regulon) a genetic memristor, which operates through reinforcement learning and a population-based memory. Two types of experimental memregulons implementations are illustrated in Figure 1A,C and Figure 1B,D. The first set-up is illustrated in Figure 1A, where two types of cells (A and B) are mixed in a co-culture. The memory of the resulting population is given as the A:B ratio. Cells A carry a genetic memristor built as an operon (regulon) with transcription rate switchable through a regulator (R). Cells B derive from cells A through mutations (either a knockout or knockdown of their functions) in the reporter gene X and the learning gene L. The product of gene X can be a reporter, an enzyme or a regulator of a downstream gene circuit. The product of gene L is a protein able to manipulate the A:B ratio by either altering the number of cells or by transmuting A into B or *vice versa* (Figure 1C). Several mutagenesis systems already exist that could be employed to achieve this [14,39,40]. In the absence of regulator R, both A and B cells will show similar mean growth rates, offering stability to the co-culture memory. Thus, changing the memory (learning) at the population level requires the presence of regulator R that will enable the expression of the L gene. In addition, the product of the gene L is switched on through an independent regulatory mechanism under the control of a suitable chemical compound (I).

An example of gene L is an antibiotic resistance gene, in which case learning would occur at a slower rate in co-transformed cells (Figure 1B,D) as it would require the death of the cells which spontaneously lost a plasmid. The second implementation of the memregulon (Figure 1B,D) combines the CAMERA analog memory system described before with gene regulation to produce non-Darwinian adaptations that change the phenotype according to the memory state (Figure 1E,F). The promoter output increases the level of gene L expression. Therefore, when gene L is activated it will lead to a greater shift in memory if the promoter is more active (Figure 1E). The activation of gene L acts as a positive/negative reinforcement of the behaviour (gene X expression by the population) if L increases/decreases the A cells (Figure 1A,C) or plasmids (Figure 1B,D) ratios respectively. Figure 1E,F shows the positive reinforcement case. Leaky transcription (Figure 1E) leads to a basal learning rate which produces a continuous memory loss. In the second type of memory because learning occurs at a slower rate, this memory loss will be smaller. In absence of leakage, both types of memory show similar behaviour.

Recently, in an oral presentation at the Synthetic Biology UK 2019 meeting held in Warwick on December 2019, we have shown for the first time our experimental validation of a library of 25 memregulons of the type of Figure 1B,D reprogramming a microbial consortia of *E. coli* to learn to play the game of tic-tac-toe. As natural multiple-copy plasmids already contain promoters with combinatorial regulations and antibiotic resistance genes, we may conjecture the existence of non-Darwinian adaptive mechanisms exploited by microbial consortia to optimize decision making to changing environments.

Perspective

- Gene circuits with analog memory are allowing the development of gene circuits able to adapt to the environment, usually by supervised learning. The introduction of reinforcement learning will allow for an unsupervised adaptation of gene circuits towards complex algorithms.
- Darwinian evolution of gene circuits could be recast as a reinforcement learning problem when an effective analog memory could be achieved. Current approaches involving analog memory systems usually require an *in vitro* memory access (*e.g.* sequencing), which prevents their use in gene circuits.
- Future directions could involve the search for new mechanisms for analog memory that could regulate gene expression (i.e. implementing a memregulon) that will already exist in living systems. In this way, by understanding the combinatorial regulation of their natural memregulons (understanding their "language") we could imagine training those microbial ecosystems new tricks and even complex algorithms.

Author contributions

A.R and A.J. wrote the manuscript.

Funding

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) grant EVO-ENGINE BB/P020615/1 and EPSRC-BBSRC grant BB/M017982/1 (WISB center).

Figure caption.

Figure 1 Proposal for a memregulon: A minimal distributed gene circuit able of unsupervised adaptation through reinforcement learning. (**A**) Co-culture memory implementation. We show a population of two type of cells: cells A carry an operon that is activated in presence of a regulator R and cells B being a copy of cells A with a knockout or knockdown of a selection gene L. At the population level, the gene L controls the A:B ratio through one of the following mechanisms: it can hinder the multiplication of either A or B, promote the multiplication of A or B or transforms A into B through a mutation or recombination of the operon with gene L. Transmuting A into B would allow for the implementation of a single-cell level memory. (**B**) Two-plasmid memory implementation. In this setup, memregulons are built with operons like those of the previously described cells A and B which would be co-transformed as Plasmid A and Plasmid B respectively in a single cell. (**C**) Representation of population memory experiment. Reinforcement learning occurs when the activity of L (in presence of I) changes the A:B ratio (which acts as an analog memory). (**D**) Representation of co-transformation memory. Reinforcement learning is manifested as a change in the total level of X inside the population. (**E**) Memory changes based on increased gene L expression. Because the activity of L increases learning, the memregulon system implements the simplest form of reinforcement learning. Leakage causes a residual learning and, therefore, memory loss. (**F**) The total amount of X produced by the population is tuned by changing the memory state in the presence of I.

References

[1] Macia J, Vidiella B, Sole RV. (2017) Synthetic associative learning in engineered multicellular consortia. Journal of The Royal Society Interface. **14**, 20170158.

[2] Banda P, Teuscher C, Stefanovic D. (2014) Training an asymmetric signal perceptron through reinforcement in an artificial chemistry. Journal of The Royal Society Interface. **11**, 20131100.

[3] Didovyk A, Kanakov OI, Ivanchenko MV, Hasty J, Huerta R, Tsimring L. (2015) Distributed Classifier Based on Genetically Engineered Bacterial Cell Cultures. ACS Synth. Biol. **4**, 72–82.

[4] Genot AJ, Fujii T, Rondelez Y. (2013) Scaling down DNA circuits with competitive neural networks. Journal of The Royal Society Interface. **10**, 20130212.

[5] Nesbeth DN, Zaikin A, Saka Y, Romano MC, Giuraniuc CV, Kanakov O, et al. (2016) Synthetic biology routes to bio-artificial intelligence. Essays Biochem. **60**, 381–391.

[6] Gandhi N, Ashkenasy G, Tannenbaum E. (2007) Associative learning in biochemical networks. J Theor Biol. **249**, 58–66.

[7] Kauffman S. (1993) The Origins of Order: Self-Organization and Selection in Evolution New York: Oxford Univ. 1993.

[8] Sarpeshkar R. (2014) Analog synthetic biology. Philos Trans A Math Phys Eng Sci. **372**, 20130110–22.

[9] Sutton RS, Barto AG. (1998) Reinforcement Learning. MIT Press; 1998.

[10] Russell S, Norvig P. (2002) Artificial intelligence: a modern approach.

[11] Fernando CT, Liekens AML, Bingle LEH, Beck C, Lenser T, Stekel DJ, et al. (2009) Molecular circuits for associative learning in single-celled organisms. Journal of The Royal Society Interface. **6**, 463–469.

[12] Perli S, Perli SD, Cui CH, Cui C, Lu TK. (2016) Continuous genetic recording with self-targeting CRISPR-Cas in human cells. Science. **353**.

[13] Rodrigo G, Jaramillo A. (2007) Computational design of digital and memory biological devices. Syst Synth Biol. **1**, 183–195.

[14] Farzadfard F, Gharaei N, Higashikuni Y, Jung G, Cao J, Lu TK. (2019) Single-Nucleotide-Resolution Computing and Memory in Living Cells. Mol Cell. **75**, 769–780.e4.

[15] Aubert N, Dinh Q, Hagiya M, Fujii T, Iba H, Bredeche N, et al. Evolving cheating DNA networks: a case study with the rock–paper–scissors game. MIT Press

.

[16] Church GM, Gao Y, Kosuri S. (2012) Next-Generation Digital Information Storage in DNA. Science. **337**, 1628–1628.

[17] Farzadfard F, Lu TK. (2018) Emerging applications for DNA writers and molecular recorders. Science. **361**, 870–875.

[18] Siuti P, Yazbek J, Lu TK. (2014) Engineering genetic circuits that compute and remember. Nat Protoc. **9**, 1292–1300.

[19] Shipman SL, Nivala J, Macklis JD, Church GM. (2016) Molecular recordings by directed CRISPR spacer acquisition. Science. **353**, aaf1175–aaf1175.

[20] Ham TS, Lee SK, Keasling JD, Arkin AP. (2008) Design and Construction of a Double Inversion Recombination Switch for Heritable Sequential Genetic Memory. PLoS ONE. **3**, e2815.

[21] Moon TS, Clarke EJ, Groban ES, Tamsir A, Clark RM, Eames M, et al. (2011) Construction of a Genetic Multiplexer to Toggle between Chemosensory Pathways in Escherichia coli. J Mol Biol. **406**, 215–227.

[22] Bonnet J, Subsoontorn P, Endy D. (2012) Rewritable digital data storage in live cells via engineered control of recombination directionality. Proc. Natl. Acad. Sci. U.S.A. **109**, 8884–8889.

[23] Siuti P, Yazbek J, Lu TK. (2013) Synthetic circuits integrating logic and memory in living cells. Nat. Biotechnol. **31**, 448–452.

[24] Yang L, Nielsen AAK, Fernandez-Rodriguez J, McClune CJ, Laub MT, Lu TK, et al. (2014) Permanent genetic memory with >1-byte capacity. Nat Meth. **11**, 1261–1266.

[25] Farzadfard F, Gharaei N, Citorik RJ, Lu TK. (2020) Efficient Retroelement-Mediated DNA Writing in Bacteria. bioRxiv. **576**, 149–62.

[26] Schubert MG, Schubert MG, Goodman DB, Wannier TM, Kaur D, Kaur D, et al. (2020) High throughput functional variant screens via in-vivo production of single-stranded DNA. **110**, 491.

[27] Daniel R, Rubens JR, Sarpeshkar R, Lu TK. (2013) Synthetic analog computation in living cells. Nature. **497**, 619–623.

[28] Tang W, Liu DR. (2018) Rewritable multi-event analog recording in bacterial and mammalian cells. Science [Internet]. **360**, eaap8992. Available from:

http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=29449507&retmode=ref&cmd=prlinks

[29] Farzadfard F, Lu TK. (2014) Synthetic biology. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. Science. **346**, 1256272.

[30] Chua L. (1971) Memristor-The missing circuit element. IEEE Transactions on Circuit Theory. **18**, 507–519.

[31] Fu T, Liu X, Gao H, Ward JE, Liu X, Yin B, et al. (2020) Bioinspired bio-voltage memristors. Nat Comms. **11**, 1861.

[32] Wang Z, Joshi S, Savel'ev S, Song W, Midya R, Li Y, et al. (2018) Fully memristive neural networks for pattern classification with unsupervised learning. Nat Electron. **1**, 137–145.

[33] Wang Z, Joshi S, Savel'ev SE, Jiang H, Midya R, Lin P, et al. (2017) Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. Nat Mater. **16**, 101–108.

[34] Wang Z, Rao M, Han J-W, Zhang J, Lin P, Li Y, et al. (2018) Capacitive neural network with neuro-transistors. Nat Comms. **9**, 3208.

[35] Kim S, Du C, Sheridan P, Ma W, Choi S, Lu WD. (2015) Experimental Demonstration of a Second-Order Memristor and Its Ability to Biorealistically Implement Synaptic Plasticity. Nano letters. **15**, 2203–2211.

[36] Law N, Ansari S, Livens FR, Renshaw JC, Lloyd JR. (2008) Formation of nanoscale elemental silver particles via enzymatic reduction by Geobacter sulfurreducens. Appl Environ Microbiol. **74**, 7090–7093.

[37] Beasley AE, Abdelouahab M-S, Lozi R, Powell AL, Adamatzky A. (2020) Mem-fractive Properties of Mushrooms.

[38] Adamatzky A. (2018) Towards fungal computer. Interface Focus. **8**, 20180029.

[39] Nyerges Á, Csorgő B, Draskovits G, Kintses B, Szili P, Ferenc G, et al. (2018) Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance. Proc. Natl. Acad. Sci. U.S.A. **115**, E5726–E5735.

[40] Zheng K, Wang Y, Li N, Jiang F-F, Wu C-X, Liu F, et al. (2018) Highly efficient base editing in bacteria using a Cas9-cytidine deaminase fusion. Commun Biol. **1**, 32.

Memregulons: Adaptive minimal gene circuits