Evolution and Analysis of the Dynamics of Learning Behaviour and Memory
2nd Year Progress Report

Eduardo Izquierdo-Torres

May 22, 2006

This report comprises a short-version of the motivation for my PhD work, a brief overview of the current outlook of the contents of the [future] thesis, a summary of the progress in the research so far and a strategy for getting the rest of it done (in time).

1 Project Motivation

One of the greatest challenges faced by autonomous robotics today is generating (and understanding) robots that can adaptively modify their behaviour according to changes in their environment. The motivation for my work is to evolve and analyse agents that learn during their lifetime without introducing learning mechanisms a priori into the agent's structure or internal mechanisms. In my work, that which is to be learned is a feature within a continuum from the environment. The tasks are loosely inspired on imprinting behaviour (in for example *C. elegans*, birds, etc.) and they are cognitively relevant because they involve memorization and decision-making in very simple models. The interest in the synthesis of such agents is in the analysis of the evolutionary dynamics as well as the evolved mechanisms and interactions using the language of dynamical systems theory. The main outcome of this thesis will be to provide an understanding of learning and memory behaviour that helps shift the focus from 'accurately representing' the environment to dynamically engaging it (with a body) so as to generate coordinated patterns of behaviour.

2 Current Thesis Outlook

1. Introduction: where I describe what the thesis of the work is, its motivation and the layout.

2. Background and Related Work: where I give an overview of the study of learning and memory. From neurosciences (e.g. Marder et al., 1996; Kandel et al., 2000; Major & Tank, 2004), psychology (e.g. Thorndike, 1898), AI and evolutionary robotics (e.g. Yamauchi and Beer, 1994; Tuci et al., 2002; Phattanasri et al., in press) to philosophy and other areas such as cybernetics (e.g. Ashby, 1952; Maturana, 1970; Maturana & Varela 1973). This list of references is not meant to cover the whole spectrum but only a taste.

Part I. Abstract Learning Scenario: the thesis will be divided into two parts: in the first a disembodied and non-situated dynamical system will be studied in a task that requires the agent to learn a feature from a continuum and long after the signal has disappeared to be able to recognise whether a presented feature is the same or different than the first learned one. The system has to be able to do this for a number of successive tests. This behaviour is loosely inspired in filial imprinting in, for example, birds, as studied by Konrad Lorenz. This corresponds to a form of irreversible learning. However, the same situation will be extended to include the case when the system has also to re-learn new signals, all during its lifetime (reversible learning). The interest in this abstract dynamical systems and discretely tokenised learning scenario is to explore the internal dynamics as in-depth as possible in scenarios that closely resemble conventional psychologists' learning tasks.

3. Evolution of Abstract Learning Agents: section where I describe the experimental set-up, including the evolutionary technique and all of the 'tricks' that eventually helped, also where I layout the evolutionary runs that were carried out and an analysis of the evolutionary dynamics, including which ‘tricks’ helped and which didn’t, what parameters gave better results and why.

4. Dynamical Systems Analysis of Abstract Learning Agent: section where I analyse one best evolved agent that performs the learning task in the abstract scenario in-depth. This will prove to be hard for at least two reasons: (1) to avoid falling into simple correlations and representationalist-style-vague talk and (2) because of the sheer size of the circuits that currently solve the task (e.g. 20 nodes).

Part II. Situated Learning Scenario: The second big experimental part of the thesis deals with the simplest embodied and situated version of the same task. The agent/environment/task in this case is loosely inspired on behavioural plasticity observed in *C. elegans* – which is a similar form of learning to imprinting (i.e. animals that were cultivated normally with food at temperatures ranging from 15C to 25C migrate to the cultivation temperature on a temperature gradient and move isothermally at that temperature. By contrast, the animals migrate away from the temperature at which they were previously starved). They don't call it imprinting, but it is a very appropriate behavioural paradigm to continue to study learning and memory in the embodied version.

5. Evolution of Situated Learning Agents: similar to chapter 3 but with the embodied experimental set-up. Emphasis here will be made on the similarities and the differences of these two experiments. One difference is the continuity of the sensory stimulation (as opposed to the experimenter tokenised stimuli in the abstract case) and another one is the active role in choosing the stimuli to be sensed arising from sensori-motor coordination.

6. Dynamical Systems Analysis of Situated Learning Agent: similar to chapter 4 but with a different agent and environment altogether. Although the circuits needed here may be slightly smaller than in the abstract scenario the analysis will include complex sensori-motor loops.

7. The role of the agent’s embodiment and situatedness for learning behavior: where I discuss in-depth the differences between the ‘solutions’ to learning and memorisation for the different scenarios from both an evolutionary perspective (e.g. how easy were they to synthesise?) and from a dynamical systems perspective on the evolved mechanisms (e.g. how complex do they need to be?).

8. Philosophical considerations: learning as the basis for homeostasis, autopoiesis and adaptativity. In both of these scenarios, but particularly the embodied one, it is very clear that successfully evolved agents are capable of maintaining certain internal organisation in the form of the self-regulation of some internal variables. The original idea of studying dynamical systems that can learn during their lifetime was to study systems with ultrastability such as Ashby’s Homeostat – but without the need to incorporate the ‘adaptive rules’ a priori, In this chapter I hope to come around to this, my original motivation, and discuss the philosophical implications to the understanding of situated, embodied and dynamical system agents that can learn, in the light of very inter-related ideas such as the ones in the title of the chapter.

9. Physical implementation of learning agent: where I show a simple ‘hardware implementation’ of the agent described in chapter 6 – this is desirable for my overall doctoral training but it currently seems to contribute little to the actual thesis.

10. Concluding Remarks: where the contributions of my work are summarised and made concise. Also where an outlook of the work that my thesis opens up to explore is overviewed.

3 Strategy

Most of the work in the PhD so far has been in obtaining the experimental results (described in the outline above). The strategy has been the following:

1. Implement and debug abstract and embodied scenarios. Play around with parameters and evolutionary and set-up tricks until I can consistently evolve agents for abstract and embodied scenarios. This task has been taken incrementally: evolving simple versions of the tasks first and growing in complexity as I manage to evolve them consistently. The path taken is described in the table bellow:

- *1* Abstract Scenario
  - *1/1* Irreversible learning (where the agent learns only once during the beginning)
    - *1/1/1* Discrete features
      - *1/1/1/1* One test (where the agent is tested only once to see whether it remembers the learned feature).
      - *1/1/1/2* Several successive tests (where the agent is tested several times for how well it remembers the learned feature).
    - *1/1/2* Features on a Continuum
      - *1/1/2/1* One test (similar to 1/1/1/1)
      - *1/1/2/2* Several successive tests (similar to 1/1/1/2)
  - *1/2* Reversible Learning (only with features on a continuum and several successive tests).
- *2* Embodied and situated scenario (only features on a continuum)
  - *2/1* Irreversible learning
    - *2/1/1* One test
    - *2/1/2* Several successive tests
  - *2/2* Reversible learning (only for several successive tests).

The idea of these incremental stages is to serve as an exercise and testing bed for each generally harder task to follow. Ideally for the thesis I will analyse in-depth *only* the last 'most capable' of evolved agents in both the abstract and embodied scenarios. Ideally, the analysed agents will be capable of reversible learning of features on a continuum with random delays and 'good/decent' memorisation performance over several successive tests.

2. Analyse the best evolved agents in the most complex task reached for each of the abstract and embodied scenarios. The plan here is to get as much help from Dr. Randall Beer as possible. This is the main reason for spending the following 3 months in Indiana University (from June to Ausgust 2006).

3. If everything goes according to plan then I should have the third year of the PhD to write each of the chapters of the thesis based on the experimental results from this current year of research. In order to accomplish this, I hope to treat the thesis as the writing of 5 journal papers. I will describe ahead how they are initially arranged:
   a. The combination of chapters 3 and 4, about the evolution and analysis of the dynamical systems for the abstract learning scenario should make for an in-depth Journal of Adaptive Behavior paper.
   b. I would like to aim the combination of chapters 5 and 6, about the evolution and analysis of embodied agents in a learning scenario, towards a more biological audience. Although I know this could prove harder than it sounds.
   c. I would like to aim chapter 7, about the role of embodiment, situatedness, and sensori-motor coordinations in learning and memory towards the cognitive science and psychology audience.
   d. Chapter 8 is one of the most important for me. I am not sure where it would be appropriate to aim at, but I know I would like to work with Ezequiel in this general endeavour.

e. If there were such a thing as chapter 9 then its obvious place would be a Robotics/Hardware oriented journal.

4. Finally, I would ‘copy and paste’ the papers into the thesis and write up the Introduction, Related work and Conclusion chapters to submit for September 2007 (fingers crossed).

4 Progress

In relation to the previously outlined strategy, I have spent the last 3 months working on [1], evolving agents on both abstract and embodied scenarios. Fortunately, by now I have managed to evolve agents towards the last stages for each of the versions. I will describe briefly ahead:

4.1 Abstract Learning Scenario

The overall idea is that of an agent who is exposed to its parent during the beginning of its life, the parent is in-charge of bringing the food during this initial growth/imprinting phase. After a certain period the parent goes away, and after some time, an individual approaches the agent, the agent has the option to approach this individual and ultimately open its mouth to receive potential food from this individual for the cases when this individual is the parent (the same individual who brought the food during the initial stage), or to maintain the mouth closed and avoid this individual in the case that it is any other individual who is not the parent. Whenever the agent chooses correctly its response it is given a reward, which can be thought of as either nutritious food or avoid being eaten in both of the scenarios above respectively. Whenever the agent responds incorrectly, either by accepting food from a stranger or by refusing food from its parent it receives a punishment, which can be thought of as either eating rotten/poisonous food or starving, respectively. As an extension of this scenario, we will deal with the case when the parent goes off to holidays or dies and a formerly stranger individual becomes the new carer and food bringer for the agent. In this case, the agent has to be able to re-learn a new parent and accordingly, associate food with the new individual.

As regards the details of the experimental set-up. The feature that the agent has to remember is a signal between [1, 2] provided for a fixed length of time (10 units of time). There is a second signal sensed by the agent that is related to food (i.e. reward/punishment) that can be -1, 1 depending on the case. There are random delays between presentations of different individuals

The agent is a fully connected CTRNN. The feature of the individual and the food are fed to all nodes in the CTRNN via a set of weights that are evolved along with the rest of the CTRNN parameters. The agent has one output node. The response of a circuit is interpreted as ‘approach and open mouth’ when the average activation of the output node over the evaluation period exceeds 0.5 and as ‘avoid and close mouth’ when it is bellow 0.5. The agent is evaluated after each test individual is presented and a successful agent must produce the correct output for any number of individuals presented after the ‘parent’.

A minimal microbial genetic algorithm is employed. An agent is evaluated every time it is chosen for a tournament in the artificial evolution algorithm. A fitness evaluation consists of a rather large number of trials. In each trial the parent signal is drawn randomly. As for the test individuals to follow, there’s a 50% chance to present parent or non-parent test individual. In the case of a different signal a random signal is chosen. The fitness of an agent is its average performance on all presented trials.

There's a range of tricks and that have made evolution of the latter (more difficult) tasks possible I will mention them here briefly in categories:

1. Task set-up
    a. Gaussian-weighted evaluation of agent's output
    b. Mouth opening mechanism is binary and depends on the accumulation of activation of the output node over a period of time.
2. GA and CTRNN parameters:
    a. Genotype-phenotype mapping:
        i. Weights: mapped linearly from [0,1] to something like [-6, 6].
        ii. Time-parameters: Map exponentially from [0,1] to [e^0,e^3].
        iii. Biases: mapped linearly from [0,1] to [-10,10] .
    b. Number of trials per fitness (this is related to the inter-trial variability): 100-500.
    c. Evolutionary operators: Gaussian vector mutation and no recombination.
    d. Number of nodes in the CTRNN: from 2 to 20.
3. Evolutionary 'tricks'
    a. Gradual increment from noise-less to noise-full simulations.
    b. Gradual increment from one test individual to N test individuals.
    c. Gradual increment from one learning scenario to N learning scenarios (reversibility).
    d. Non-parent test individuals that are 'too similar' to the actual parents are not used during evolutionary training.

The trajectory from simple to harder learning tasks in this abstract scenario has been very interesting. The simplest version of the task managed to evolve using only 3 node circuits while the current hardest task seems to require between 15 and 20 nodes.

There's a range of interesting questions that arise from the evolution of these agents that I have explored with the evolved agents for the previous versions and hope to do so more in-depth for the last evolved one.

1. What is the role of slow and fast components (e.g. neurons, nodes) in the dynamics of learning? Preliminary results show that best evolved agents are comprised of around 30% slow acting components versus 70% fast acting ones.
2. For how long does the 'memory' of a feature last on the best evolved agent? What affects how good an agent will remember the feature? Will seeing the feature often during the tests improve the agent's chances of remembering it? Preliminary results point towards the imminent decay of the memory of the feature as well as improved remembering performance when the previous test individuals are the parent.
3. What helps evolve agents that can perform this task? Although full experiments testing which of the 'tricks' being used to evolve successful agents helps and which doesn't, the intuition is that several of the ones mentioned previously will prove to be quite helpful. Particularly helpful seems to be the mouth opening mechanism as a binary choice, as opposed to trying to maximise the response of the output node towards 1 or 0 for the evaluation period – this seems to prevent the agent from overspecialising on some part of the task while doing the rest of it poorly.

4.2 Embodied & Situated Learning Scenario

The overall idea of these experiments is inspired on a form of learning observed in *C. elegans*. It has been studied that the animals that were cultivated normally with food at temperatures ranging from 15C to 25C migrate to the cultivation temperature on a temperature gradient and move isothermally at that temperature. By contrast, the animals migrate away from the temperature at which they were previously starved. More or less thermal-imprinting. Also the can re-learn new preferred temperatures after a long time of starvation in the current preferred temperature. I think this is a very appropriate behavioural paradigm to continue to study learning and memory in a lightly embodied and situated version of the previous abstract experiments. I have designed the experiments for the situated (and lightly embodied) version of this memorisation task loosely inspired on such extraordinary behavioural plasticity observed in *C. elegans*. There are other advantages of taking inspiration from observations in *C. elegans*, but mainly the simplicity and current understanding of their nervous systems and internal mechanisms in general.

The details of my experiment are as follows. The scenario is based in a 2D world. The agent is a fully connected CTRNN. The agent has a body and it can move forward at fixed speeds () and can turn its head to the sides. The direction in which the head is facing determines the direction of movement of the agent. Two opposing motors determine the position of the head. An arbitrarily determined node in the circuit controls each of the motors. The agent has two sensory perturbations from the outside world (much as the previous abstract version). One is the temperature of its surroundings and the other is whether there is food or not. The feature that the agent has to 'remember' is the temperature that is on a continuum between [0,1].

The task is as follows. The agent spends the first units of time in an environment where the temperature is the same throughout. This temperature is picked at random between the range [0,1]. There is food in this environment. After some time, the agent is taken to a different environment and placed at random. In this environment there is no food and the temperature is no longer uniform but has a linear thermal gradient. A successful agent will 'swim' towards the gradient in search for its 'growth temperature' and then navigate isothermally. If physically displaced at any time (or what's the same if the thermal gradient moves) the agent has to find its way towards its growth temperature again. This signals the successful learning/imprinting of the feature on the continuum (the temperature). If placed in an environment with food in a different temperature, then the agent should re-learn this new preferred temperature and swim towards it later on.

Similar evolutionary techniques are being used for this case and so far I have managed to evolve agents that can perform successfully the thermal imprinting, swimming several times towards the preferred temperature after physical displacements. I have not yet evolved agents towards the re-learning ability. Interestingly enough, less than half of the nodes in the CTRNN circuit are needed to perform this task so far.

5. Concluding Remarks

This report has been deliberately over-focused on the strategy for generating a thesis. Nevertheless, my PhD experience has encompassed much more than just that. The thing that I am learning the most from is undeniably teaching. Very fruitful has been the organising of scientific meetings such as the *activate.d* sessions, the ECAL workshop and the workshop here in Sussex. Also nourishing has been the involvement in the editing of a journal issue. Presenting papers at conferences has been very demanding but useful and also has been the making of a scientific poster. Finally, participating in the hiring of new faculty for the Cognitive Science program in Indiana University as one of three candidates, including travelling to Bloomington for a talk and over 10 separate interviews with faculty members has been a very rewarding experience, with much to be learnt and improved for future applications.

References

Ashby, W. R. (1952). Design for a brain. New York: Wiley.

Maturana, H. R. (1970). Biology of cognition. In H. R. Maturana and F. J. Varela (Eds.), Autopoiesis and cognition (pp. 1–58). Boston: Reidel.

Maturana, H. R., & Varela, F. J. (1973). Autopoiesis: The organization of the living. In H. R. Maturana and F. J. Varela (Eds.), Autopoiesis and cognition (pp. 59–138). Boston: Reidel.

Kandel, E.R., Kupfermann, I., and Iverson, S. (2000). Learning and memory. In E.R. Kandel, J.H. Schwartz, and T.M. Jessell (Eds.), Principles of Neural Science, 4th edition (pp. 1227-1246). McGraw-Hill.

Major, G. and Tank, D. (2004). Persistent neural activity: prevalence and mechanisms. Current Opinion in Neurobiology 14:675-684.

Marder, E., Abbott, L.F., Turrigiano, G.G., Liu, Z. and Golowasch, J. (1996). Memory from the dynamics of intrinsic membrane currents. Proc. Nat. Acad. Sci. USA 93:13481-13486.

Thorndike, E.L. (1898). Animal intelligence: An experimental study of the associative processes in animals. The Psychological Review, Series of Monographs Supplements (Supplement 2) 4:1-109.

Tuci, E., Quinn, M. and Harvey, I. (2002). An evolutionary ecological approach to the study of learning behavior using a robot-based model. Adaptive Behavior 10(3-4):201-221.

Yamauchi, B. and Beer, R.D. (1994). Sequential behavior and learning in evolved dynamical neural networks. Adaptive Behavior 2(3):219-246.

Phattanasri, P. Chiel, H., and Beer, R.D. (submitted). The Dynamics of Associative Learning in Evolved Model Circuits. Submitted to the Journal of Adaptive Behavior.