Modelling mental representation as evolved second order learning

Solvi Arnold[1*], Reiji Suzuki[1], and Arita Takaya[1]

[1]Graduate School of Information Science, Nagoya University, Aichi 464-8601, Japan
(Tel: 81-52-789-4716, Fax: 81-52-789-4800)

[*]solvi@alife.cs.is.nagoya-u.ac.jp

Abstract: Mental representation is a fundamental aspect of advanced cognition. An understanding of the evolution of mental representation is essential to an understanding of the evolution of mind. However, being a decidedly mental phenomenon, its evolution is difficult to study. This research addresses the question of how representation ability may emerge from non-representational cognition. We reformulate cognitive map ability, a paradigm case of mental representation, in terms of second order learning. We provide a neural network species with neural mechanisms for second order change and evolve them in an environment of randomly generated Tolman mazes, known to require mental representation. Some runs of this model evolve near-optimal performance, providing support for the hypothesis that mental representation is evolved second order learning.

Keywords: artificial life, cognitive map, evolution of mind, learning, mental representation, neural networks.

1 INTRODUCTION

Mental representation (MR for short) is, abstractly put, the ability to simulate or reconstruct in the mind aspects of the environment that lie outside the scope of one's current perception. The type of MR we focus on in this paper is the ability to navigate using "cognitive maps": mental representations of the layout of an environment (see Tolman [1]). Much of an environment may lie outside one's view, but mental representation of these parts makes it possible to take those parts into account nonetheless. Other types of MR are "mental time-travel" and "theory of mind" (see Takano & Arita [2], Minoya, Arita & Omori [3] for computational approaches to the latter). There too, inaccessible aspects of the environment (respectively: future and past, other minds) are mentally simulated or reconstructed.

The evolution of MR is not well-understood. MR is a highly structured and organized form of cognition, and already in the early decades of connectionism, it has become clear (contrary to common intuition) that adaptive processes such as evolution or learning do not, in general, produce such structured or organized AI (see e.g. Fodor & Pylyshyn [4]). This raises the question how MR can have evolved in biological cognitions. We propose that the representational nature of MR emerged from the interaction between adaptation processes, specifically: first and second order learning. We provide a proof of concept for this hypothesis in the form of a computational model in which a neural network species with the basic elements for second order change is evolved in an environment composed of maze tasks generally believed to demand cognitive map ability. If second order learning can be evolved into MR, this may improve our understanding of the evolutionary transition from non-representational to representational cognition.

2 TOLMAN MAZES

In experimental psychology, MR ability in biological species is often studied using Tolman mazes [1]. A Tolman maze has multiple paths from its start to its goal. Exploring the maze, subjects latently learn to take the shortest path.

Subsequently, the shortest path is blocked, meaning they have to switch to a different path. However, the maze is set up so that after blockage of the shortest path, the new optimal choice of path depends on the location of the blockage (Fig. 1). Many species need to re-learn on the altered layout, while some species find the new optimal path without additional exploration (needing merely to observe the position of the blockage). The standard explanation is that the former category of species does not acquire a mental representation (cognitive map) of the maze during initial exploration, while the latter does. Animals that have acquired a representation of the maze can exploit that representation to infer the new optimal path without additional exploration.

Fig. 1. Generated Tolman maze on 7x7 grid.
S: Start G: Goal. Dot colours indicate path lengths.
Blue circles indicate possible locations for the blockage. Blockage on the left blue ring obstructs only the short path, while blockage on the right blue ring blocks both the short and medium path.

3 HYPOTHESIS

Here we propose a different explanation, in terms of second order learning. MR can be viewed as acquired isomorphism between a mind's content and its environment. It has previously been established that innate isomorphism can emerge from interactions between evolution and learning (Arnold [5]). Acquired isomorphism might similarly emerge from interactions between first and second order learning.

Let us first give a working definition of first order learning: "advantageous change in behaviour, caused by exposure to pertinent information". Second order learning, then, must be advantageous change in (first-order) learning, caused by exposure to pertinent information. When after blockage of the shortest path in a Tolman maze a subject infers the new optimal path without additional exploration, we can view this inference as a split-second learning process: exposure to pertinent information (observation of the location of the blockage) produced an advantageous change in behaviour (the subject abandons the blocked path and switches to the new optimal path). Viewing it as such requires that we explain how this (first order) learning process can produce such a fast and effective behaviour-update (the information in the observation alone does not suffice to explain the update). Our explanation is that during exploration, not just the animal's behaviour has adapted to the maze (via a learning process), but that that learning process itself has been adapted to the maze as well. In other words, a second-order learning process optimized the (first-order) learning process to the current environment: the optimal change in behaviour (from selecting the blocked path to selecting the optimal path) has come to be causable by minimal information: mere observation of the location of the blockage has come to trigger this behaviour-update.
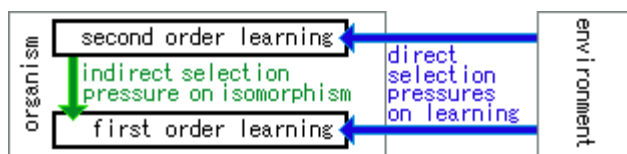

Fig. 2. Learning and isomorphism (MR).

A second order learning process associates possible future observations with suitable behaviour updates. This may sound infeasible, and for almost all first order learning schemes, it is. But if the first order learning scheme produces behaviour-implementations that are isomorphic with the environment, then second order learning suddenly becomes quite feasible: simply keeping the behaviour-implementation isomorphic with the environment (i.e. updating it continuously with present perception) keeps the behaviour optimal with respect to that environment. So when there is selection pressure on second order learning, then indirectly there is selection pressure on isomorphism-based first order learning schemes. Mental representation, then, could be this isomorphism.

The proposed reformulation might seem highly specific to the case of Tolman mazes, but equivalent reformulations can be given for many or all other situations that involve MR. We omit detailed examples here, but the general form is as follows. Consider an environmental object X to be represented. We (should) perceive our subject as representing X if and only if it can pre-emptively adjusts its behaviour so as to avoid or bring about specific unseen situations involving X after some period of observation of X. In all such cases, observation affects future changes in behaviour. To the extent such pre-emptive adaptation characterizes MR, explanation in terms of second-order learning should be applicable. In the focal case of cognitive maps, X is the maze, but the general scheme may equally well describe a scenario of spontaneous novel tool use (a scenario generally recognized as involving MR), with X being the tool.

4 MODEL

We test the hypothesis using a model in which neural nets with the basic elements for second order neural change are evolved in an environment composed of Tolman mazes.

Each individual's fitness is assessed on multiple Tolman mazes (randomly generated continuously over the course of the experiment). In each maze, the agent first gets 150 time steps to explore the maze (exploration phase), is then placed back at the start position and given exactly enough time steps to reach the target over the shortest path (exploitation phase). On a portion of the mazes, the shortest path is blocked at the start of the exploitation phase. When the agent arrives at the blockage, its remaining time steps are again set so as to make it only just possible to reach the target. Fitness is awarded for reaching the target. A population of size 100 is evolved, using a genetic algorithm with mutation but no crossover.

In addition to standard (activatory) connections, the network species has connection types that can be wired up to create second order changes in behaviour. These connection types are (1) activatory connections with lag, and (2) neuromodulatory connections. That two mechanisms are provided is merely to improve flexibility so as to facilitate evolution. Theoretically, one mechanism should suffice.

(1) Activatory connections with lag. These transmit activation signals much like standard connections, except with a short delay. This makes it possible for the nets to retain activation patterns over time. This sort of connection is common in recurrent neural networks. Retention of activation makes it possible for perception of a stimulus to af-

fect behaviour at some later point in time. For long-term change, one needs loops (in the simplest case, a single reflexive connection). Trivially, second order behaviour change can be construed too: if a retained local activation pattern A somewhere in a net affects behaviour, and another retained local activation pattern B affects A, then B has a second order effect on behaviour.

(2) Neuromodulatory connections. These connections convert activation into "modulation": the pre-synaptic neuron's activation arrives at the post-synaptic neuron as modulation. Typically, a neuron's modulation determines its connections' plasticity, in a simple (e.g. Hebbian) weight update rule (see Soltoggio et al. [6]). We let a gene pick from among a number of update rules (various combinations of pre- and post-synaptic activation and modulation values). Connection weight changes can change behaviour, and again quite trivially, they can change change in behaviour, too. Say change in connection X affects behaviour. Then weight change in connections that control change in X has a second order effect on behaviour.

Naturally these mechanisms can also be combined to create second order dynamics. Fig. 3 gives a schematic representation of the four possible circuits that produce second order dynamics (provided the connections are non-zero).
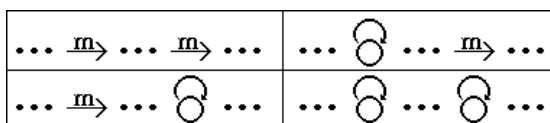

Fig. 3. Genotypic circuits for second order change. Any of these combinations occurring along a path from input to output neurons can cause second order changes. Arrows marked with M are modulatory connections. The lagged reflexive connections may produce local instead of reflexive connectivity in the phenotype, but either can support sustained change.

Informed by what's known about the neurology of spatial representation (See Moser et al. [7] for a review), we let the genotype encode not simple neurons, but neuron grids. We use square grids of three sizes: 1x1, 3x3, and WxW, where W is the size of the world (7 for our 7x7 world). Given the setup of the model, sizes larger than W offer no additional functionality (i.e. WxW is functionally equivalent to an infinite grid). The peculiar 3x3 size is included because it can concisely encode directional information. Connectivity is defined on two levels: inter-grid and intra-grid.

Intra-grid all connections are of the lagged activatory type. They can be reflexive or local. If a grid has a reflexive connection to itself, this means that all neurons in the grid get a reflexive connection to themselves. A local connection means that each neuron has connections to its four neighbours in the grid. So, connectivity within a grid is

uniform (the genotype essentially defines connectivity for a single neuron, and the number of copies to make of that neuron). Propagation within a grid runs at a faster (24x) timescale than propagation among grids. Observe that reflexive connections allow retention of an activation pattern on a grid, while local connections lead to diffusion of activation patterns over time.

Grids can be connected by any connection type other than lagged activatory connections. If the genotype defines a connection between two grids, then the phenotype gets uniform connectivity between the neurons in the two grids. If the grids are equal in size, connectivity is one-to-one, otherwise all-to-all. This leads to a highly symmetrical connectivity, which by itself would cause the activation within a grid to remain uniform and redundant. This symmetry is broken by our neurotransmitter logic.

There are two global "neurotransmitter" values, nt-Bx and nt-By. These dynamically control (in two dimensions, as the neuron grids are 2D) which connection subsets of an all-to-all projection can transmit activation. When both are zero, then this set comprises connections linking corresponding neurons in the grids (e.g. the centre neuron in the pre-synaptic grid to the centre neuron in the post-synaptic grid). Non-zero nt-B values cause simple offsets, as illustrated in Fig. 4. Currently, nt-Bx and nt-By values are hardwired to reflect the agent's current x-coordinate and y-coordinate, so signal transfer can shift along with position in space. This makes it relatively easy for evolution to devise nets that store information in different locations in a grid depending on their own position in space: if a smaller grid projects to a larger grid, then the activation pattern on the smaller grid affects only a sub-region of the larger grid. We will call this sub-region the *focal area* of the smaller grid on the larger grid. Nt-B does not correspond directly to any biological neurotransmitter, but can be reduced to a biologically plausible neurotransmitter via a trivial network transformation (which increases network size dramatically).

Inclusion of coordinates in the input is unnatural, but preliminary experiments with a simple spatial memory task have shown it quite possible with our model to evolve agents that keep track of their own coordinates. Cognitively interpreted, the coordinates in the input and their linkage to the nt-B values make it trivially easy to evolve an innate sense of space as an extended medium in which movement predictably changes one's position. Construction of the ability to represent the volatile and non-uniform *contents* of space, however, is left to evolution.

The nets have one 3x3 grid and a number of 1x1 grids (i.e. single neurons) receiving input. The 3x3 grid encodes for each of the four cardinal directions whether there is a wall in that direction (on the 4 neurons adjacent to the middle neuron). The 1x1 grids encode whether the current posi-

tion is the start position, whether the current position is the goal position, and the current phase (exploration or exploitation). Additionally, there are input neurons for bias (always 1.0) and noise (random real numbers from [0,1]). Output is read from two 3x3 grids. From the four neurons corresponding to the cardinal directions, the one with the highest activation is selected, and movement in that direction is performed (if possible). One set is read during exploration and the other during exploitation (so that the nets can easily evolve specialized behaviour per phase).
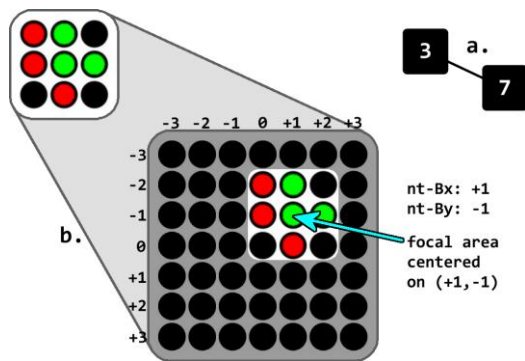

Fig. 4. Neural grids and nt-B. **a.** Genotype encoding a 3x3 grid, a 7x7 grid, and their connection. **b.** The corresponding phenotype. The 3x3 grid projects into the focal area of the 7x7 grid. The position of focal areas for projections between unequally sized grids is dynamically controlled by the global neurotransmitter values nt-Bx and nt-By. This mechanism lets the nets conveniently allocate neurons to spatial locations.

5 RESULTS

While success rates are very low, the model does occasionally produce networks with optimal performance on our Tolman mazes. We concisely describe an evolved solution here.

As the agent moves through the maze, the 3x3 input grid projects onto different regions of a 7x7 grid. Modulation working on the 7x7 grid records the paths travelled into the connectivity of this grid, replicating the maze layout. Activation diffusion in the grid then comes to follow this layout. The T-splits in the layout receive high activation due to their high connectivity, and consequently activation in other neurons on the grid comes to depend on proximity (along the replicated paths) to these T-splits. Effectively, local activation levels in the 7x7 grid come to indicate path length, and can thus be used to select the shortest path. Blockages are handled simply by updating (weakening) the connectivity in the grid at the position corresponding to the blockage. The activation pattern then settles into a new state with low activation on the blocked paths, as positions on these paths lose their proximity to one of the T-splits.

This solution is representational in nature: the connectivity pattern that forms on the grid during exploration is isomorphic with the maze, and the activation pattern on it identifies the paths and orders them by length. Also, the solution shows second order dynamics on the neural level. The neural circuit are of the bottom-left type in Fig. 3: a modulatory connection projects to a grid with activation retention (in the form of local connectivity). This second order circuit is crucial to the ability to solve the Tolman mazes (breaking either connection removes this ability).

6 CONCLUSION & FUTURE WORK

We proposed the hypothesis that mental representation is evolved second order learning, and tested this hypothesis using an artificial life model in which cognitive map ability, a paradigm case of mental representation, is evolved from the neural elements for second order neural change. Although success rates are very low, successful runs provide a proof of principle for our hypothesis. Future directions for our research are extension of this approach to other domains of representation, such as temporal or social forms.

REFERENCES

[1] Tolman E C (1948), Cognitive maps in rats and men, Psychological Review 55(4):189–208
doi: 10.1037/h0061626

[2] Takano M, Arita T (2006), Asymmetry between Even and Odd Levels of Recursion in a Theory of Mind, Proceedings of ALIFE X, 2006, pp.405-411

[3] Minoya K, Arita T, Omori T (2011), An artificial life approach for investigating the emergence of a Theory of Mind based on a functional model of the brain, Proceedings of the 2011 IEEE Symposium on Artificial Life, 2011, pp.108-115

[4] Fodor J, Pylyshyn Z (1988), Connectionism and Cognitive Architecture: a Critical Analysis. Cognition, 28(1988):3-71

[5] Arnold S (2011), Neuro-cognitive Organization as a Side-effect of the Evolution of Learning Ability, Proceedings of the IEEE Symposium on Artificial Life, 2011, pp.100-107

[6] Soltoggio A, Bullinaria J A, Mattiussi C et al. (2008), Evolutionary Advantages of Neuromodulated Plasticity in Dynamic, Reward-based Scenarios, Proceedings of Artificial Life XI, MIT Press, 2008.

[7] Moser E I, Kropff E, Moser M (2008), Place Cells, Grid Cells, and the Brain's Spatial Representation System Annual Review of Neuroscience 31:69-89
doi: 10.1146/annurev.neuro.31.061307.09072