3J3-OS-20b-2

Evolution of Social Representation in Neural Networks

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

[*1] Graduate School of Information Science, Nagoya University, Japan

This paper outlines an Artificial Life approach to Theory of Mind (ToM), the ability to employ mental models of other minds in order to understand or anticipate the behaviour of others. We designed a model in which a population of neural network (NN) agents evolve the ability to predict, on basis of observation of past behaviour, others' future behaviour in novel circumstances. As agent behaviour is guided by private mental states, correct prediction of others' future actions requires that the agents learn to recognize others' mental states from observation of their behaviour. Such learning ability cannot be captured with conventional learning algorithms, but we find that NNs equipped with neuromodulation mechanisms can be evolved to perform favourably on this task. The resulting networks are seen to behave as though they have a primitive form of first order ToM.

1. Introduction

Theory of Mind (ToM) is the ability to employ mental models (representations) of other minds, in order to understand or anticipate the behaviour of others [Dennett 1987]. The adaptive advantages of ToM are likely to be a driving factor in the evolution of recognition of others as well as the self as intentional agents. Understanding or anticipating the actions of others typically involves placing oneself in the other's shoes, and reasoning or intuiting what one would do in their position. This mental operation involves a generalization over perspectives: one has a first-person perspective on one's own mind but a third-person perspective on others' minds. Recognition of oneself as an instance of a class of agents one observes in the outside world provides a basis for a third-person concept of the self, as one intentional agent among others. When we think of recursive ToM, the necessity of a third-person concept of the self becomes quite obvious. In order for X to think about what Y is thinking about X, X must understand Y's (third-person) perspective on X. Given that ToM involves a third-person concept of the self, it is likely to have played a crucial role in the evolution of the understanding of the self. At present, this research focuses on non-recursive (i.e. first-order) ToM, with recursive ToM as a future goal.

2. Research Goal

Representation is a tough issue, but other minds presents a special challenge in that they are (1) invisible (we cannot see the minds of others, we can only guess at the existence of other minds via observation of behaviour) and (2) themselves capable of representing, leading to recursive and reflexive scenarios such as mind X representing mind Y that represents mind X. Point (1) has implications for learning about other minds: One might learn about another's behaviour via direct observation of that behaviour, but for learning about another's mind one needs forms of learning ability that incorporate inference from externally visible behaviour to (invisible) mental states. This sort of learning is difficult to capture with traditional AI conceptualizations of learning. Indeed, while computational work on ToM exists, the mechanisms for representing other minds are usually explicitly given [Takano 2006] [Noble 2010]. In this research we instead aim to let such mechanisms evolve from scratch, using a minimalistic evolutionary neural network model.

Contact: Solvi Arnold, Graduate School for Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601, Tel: 052-789-4716, Fax: 052-789-4800, E-mail: solvi@alife.cs.is.nagoya-u.ac.jp

3. Model outline

Agents are implemented as neural networks (NNs). Network architecture is evolved using a basic Genetic Algorithm, which we will not detail here. Agents interact in pairs. During its lifetime, each agent is part of multiple pairings. In each pair, there is a fixed role division: one agent acts at zero-order ToM (L_0), and one at first-order ToM (L_1). Each pair interacts for a set number of time-steps.

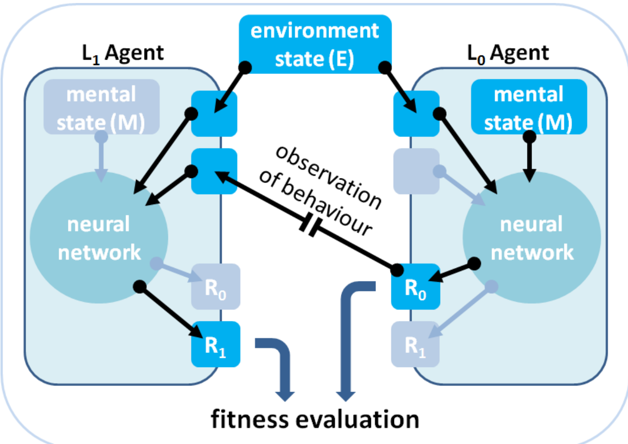

Figure 1. Agent interaction. The L_0 agent computes its action (R_0) from the (shared) environmental state and its (private) mental state. The L_1 agent computes a prediction (R_1) of this action. After the prediction is made, R_0 is revealed to the L_1 agent as data to drive its learning process.

Our model is not intended to capture any specific social interaction scenario in particular. Instead we take a more abstract approach, in which the logic that determines the fitness effect of performing a given action in a given state is generated randomly for each experiment (i.e. the fitness function is randomly generated for each run of the model). The idea is that if arbitrary fitness functions can be handled successfully, then the model has generality. Thus there is no concrete "task" to solve, there are merely *environmental states*, *mental states*, *actions*, and a randomly generated *base logic* that relates these elements.

Environmental state: bit-string of length N_e. The environmental state is shared between interacting agents (i.e. both agents see the same state). The environmental state changes every time-step, so each pair of agents will always interact under a number of environmental states.

Mental state: bit-string of length N_m. Each agent has a private mental state, invisible to its interaction partner. Mental states remain constant over the course of the interaction of an agent pair.

Action: bit-string of length N_a. At each time-step, each agent outputs an action.

Base logic: generates the optimal action choice for each (environmental state, mental state) pair. The base logic abstractly represents social scenarios.

Fitness scores for the action choices of the agent performing the L_0 role are calculated as proximity to the optimal action as generated by the base logic. Meanwhile, fitness scores for the action choices of the agent performing the L_1 role are calculated as proximity to the action choice of the L_0 agent. As such, the L_1 agent must try to predict the action of the L_0 agent, but the action choice of the L_0 agent depends on the L_0 agent's mental state, which is invisible to the L_1 agent. Herein lies the challenge: in order for the L_1 agent to be able to predict the L_0 agent's future moves under future environmental states, the L_1 agent must infer the L_0 agent's mental state from the L_0 agent's action choices under the current and preceding environmental states. The goal of the model is to let this learning ability evolve. Evolution of learning ability is made possible using neuromodulation [Soltoggio 2008]. This technique has previously been employed to evolve spatial representation ability in NNs [Arnold 2012] [Arnold 2013]. We omit detailed explanation here, but the basic idea is to introduce a special connection type that lets neurons send modulatory signals to one another, and to let these signals control connection weight change. This allows for evolution to shape the weight update dynamics of the networks by shaping the modulatory connectivity. This provides a basis for endogenously controlled behaviour change, i.e. a basis for learning ability. Additionally, our NN logic allows for recurrent connectivity, so activation can be retained over time. This can serve as a basis for memory mechanisms. After each time-step of an interaction, the actual action choice of the L_0 agent is revealed to the L_1 agent, and connection weight updates are performed. By observing both the L_0 agent's action choice and the environmental state that led the L_0 agent to choose that action, the L_1 agent has the necessary information to infer the L_0 agent's mental state, and from there correctly predict its behaviour under other environmental states. This prediction ability is what we aim to let evolve in our agent population.

4. Results

We have experimented with this model with the N_e, N_m and N_a parameters all set to 3. Some runs fail, but we found the model quite capable of producing agents that can predict other agents' L_0-action choices under unseen environmental states at far better than chance performance, or even near-optimal performance, after observing their behaviour under only a subset of environmental states. Thus the agents are successfully learning how their interaction partner maps environmental states to actions. In that mapping, the partner's mental state plays a central role. As such, it seems that by observing their partner's behaviour, the agents learn their partner's mental state, suggesting that these agents have evolved a primitive form of first-order Theory of Mind. Figure 2 shows the evolution process of a representative successful run.

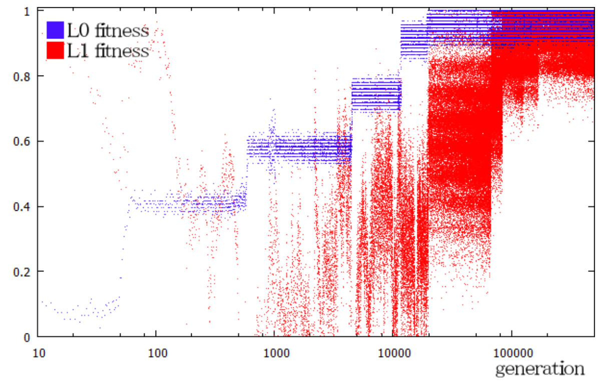

Figure 2. Evolution process of an example successful run of 500000 generations. Note the log scale on the x-axis. L_0 fitness scores are averages over individuals that have been copied from the preceding generation without mutation (i.e. offspring of the elite from the previous generation). L_1 fitness scores are averages over interactions between such individuals. Theoretical maximum fitness is 1. The best possible fitness for an individual without learning ability was computed to be 0.12 for this run's base logic. The average L_0 and L_1 scores over the final 1000 generations of this run were 0.971 and 0.974, respectively.

5. Future work

Future goals in this research are: 1) To investigate where and how the partner's mental state is represented in the agents' neural structures, and the neural circuitry used to compute L_0 responses is co-opted to compute the L_1 responses (as this would constitute a sort of identification with the partner agent). 2) To experiment with more complex scenarios for the L_1 response (i.e. not mere prediction, but acting in anticipation of the L_0 agent's action). 3) To extend the model to higher (recursive) orders of ToM. 4) Once the model works for mental and environmental states of sufficient size, we aim to replace the randomly generated base logic with simple games or cognitive psychology experiments that involve ToM.

References

[Arnold 2012] Arnold, S., Suzuki R. and Arita, T. "Second Order Learning and the Evolution of Mental Representation," in *Artificial Life XIII: Proceedings of the Tenth International Conference on Artificial Life*, MIT press, pp. 301-308. (2012)

[Arnold 2013] Arnold, S., Suzuki, R., and Arita, T. “Selection for Reinforcement-Free Learning Ability as an Organizing Factor in the Evolution of Cognition,” *Advances in Artificial Intelligence*, vol. 2013, Article ID 841646, 13 pages. (2013)

[Dennett 1987] Dennett, D. C. The Intentional Stance. MIT Press / Bradford Books, Cambridge, MA. (1987)

[Noble 2010] Noble, J., Hebbron, T., Van Der Horst, J., Mills, R., Powers, S. T. and Watson, R. "Selection pressures for a theory-of-mind faculty in artificial agents," in *Artificial Life XII: Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems*, MIT Press, 615-615. (2010)

[Soltoggio 2008] A. Soltoggio, J. A. Bullinaria, C. Mattiussi, P. Dürr, and D. Floreano, "Evolutionary Advantages of Neuromodulated Plasticity in Dynamic, Reward-based Scenarios," in *Artificial Life XI: Proceedings of the Tenth International Conference on Artificial Life*, MIT Press, pp. 569-576. (2008)

[Takano 2006] Takano, R. and Arita, T. "Asymmetry between even and odd levels of recursion in a theory of mind," in *Artificial Life X: Proceedings of the Tenth International Conference on Artificial Life*, pages 405–411. MIT Press, Cambridge, MA. (2006)