(Proceedings of the Third International Conference on Cognitive Neurodynamics, pp. 615-611.)

Synthetic approach to understanding meta-level cognition of predictability in generating cooperative behavior

Jun Namikawa, Ryunosuke Nishimoto, Hiroaki Arie and Jun Tani
Lab. for Behavior and Dynamic Cognition
RIKEN Brain Science Institute, Saitama 351-0198, Japan
e-mail: jnamika@brain.riken.jp

Summary

We propose that "predictability" is a meta-level cognitive function that accounts for cooperative behaviors and describe this from a dynamical systems perspective based on a neuro-robotic experiment. In order to bring about cooperative behaviors among individuals, individuals should attempt to predict the behavior of their partners by making internal models of them. However, the behaviors of partners are often unpredictable because individuals possess free will to generate their own independent actions. Thus, acquiring internal models which attempt to completely predict the actions of others seems to be intractable. In the current study we suggest that, when learning internal models for interacting with the partners, cooperative agents should maintain predictability monitoring mechanisms by which attention is oriented more toward predictable segments in spatio-temporal sensory input space .

1 Introduction

This study represents an attempt to understand the nature of cooperative behaviors in terms of abstract models of perceptual inference and learning. The concept of cooperative behavior has been widely investigated in the fields of distributed intelligence [1] [2] [3] [4] and social psychology [5] [6]. Except for interactions among individuals having no internal states, the individuals should make internal models of them in order to predict behaviors of partners. However, acquiring internal models to be able to completely predict behaviors of others seems to be intractable because individuals are capable of generating voluntary actions. An important point is that when achieving cooperative behaviors, an individual only needs to predict the behaviors of their partner that relate to their own actions. Therefore, we propose an artificial neural network model possessing predictability monitoring mechanisms by which they can focus on predictable segments in spatio-temporal input sequences. We hope to show that predictability as the meta-level cognition is indispensable in achieving cooperative behaviors.

In addition, we try to describe that the problems of how attention works can be resolved using exactly the same principle. Attention, defined broadly as the cognitive process in which processing resources are allocated to one aspect by ignoring other things, is a ubiquitous feature in information processing [7] [8] [9] [10]. For example, learning performance of a mixture of RNN experts was improved by adding a mechanism by which each expert network selectively concentrates on a primitive pattern in spatio-temporal time series [11]. It has been suggested that attention can be understood as inferring the level of precision during hierarchical perception in a Bayesian fashion [12]. We pursue these attempts to understand attention in terms of dynamical systems perspective.

The work in the current paper is related to temporal sequence learning problems for artificial neural networks. Artificial neural networks have been widely applied to learning problems for various kinds of temporal sequences [13] [14] [15] [16] [17]. However, in spite of the considerable accounts carried out since the mid-1980s, it has been thought that neural networks could not be scaled so as to be capable of learning complex sequence patterns, especially when the sequence patterns to be learned contain long-term dependencies. This is due to the fact that the error signal cannot be propagated effectively in long time windows of sequences using the gradient descent method, because of the potential nonlinearity of the neural dynamics [17]. This paper will claim that the predictability monitoring mechanism improves learning performance because it avoids the unreasonable interference of the error signal corresponding to unpredictable parts in spatio-temporal sequences.

2 Methods

This section explains an artificial neural network model together with how the model is applied to a specific robotic experiment. The network received two different modality inputs, proprioceptive somato-sensory input and vision input. These different modality sensations came together in the network to generate predictions of the future states. The next visuo-proprioceptive states which were predicted from the current states were used to control the robot. In addition, the network also predicted the prediction errors between visuo-proprioceptive inputs and predicted values, as "prediction of prediction errors". The dynamics of the network is described by the following differential equation:

$$\tau\dot{\boldsymbol{u}}(t) = -\boldsymbol{u}(t) + W^u \langle \hat{\boldsymbol{x}}(t-\xi), f(\boldsymbol{u}(t)) \rangle + \boldsymbol{I}^u, \quad (1)$$

$$\boldsymbol{x}(t) = f(W^x f(\boldsymbol{u}(t)) + \boldsymbol{I}^x), \quad (2)$$

$$\boldsymbol{v}(t) = g(W^v f(\boldsymbol{u}(t)) + \boldsymbol{I}^v), \quad (3)$$

where $\boldsymbol{u}$ is membrane potential of internal context neurons (in experiments the number of neurons is 50), $\hat{\boldsymbol{x}}$ is sensory-motor input, $\boldsymbol{x}$ is sensory-motor output, $\boldsymbol{v}$ is variance as described below, and ξ is the feedback time delay of the controlled robot. The functions f and g denote component-wise application of tanh and exp, respectively. $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$ denotes the concatenation of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.

To perform given tasks of cooperative behaviors, the network learned to predict sensory feedback for the next time step through training processes. The network was trained by means of supervised learning using teaching sequences obtained using the robots. The training of the network is defined as maximizing (or integrating over) the likelihood P as follows:

$$P = \prod_t \prod_i \frac{1}{\sqrt{2\pi v_i(t)}} \exp\left(-\frac{(x_i(t) - \hat{x}_i(t))^2}{2v_i(t)}\right), \quad (4)$$

where $x_i(t)$ and $v_i(t)$ are generated by the network, and $\hat{x}_i(t)$ is training data representing visuo-proprioception. A notable point in this scheme is

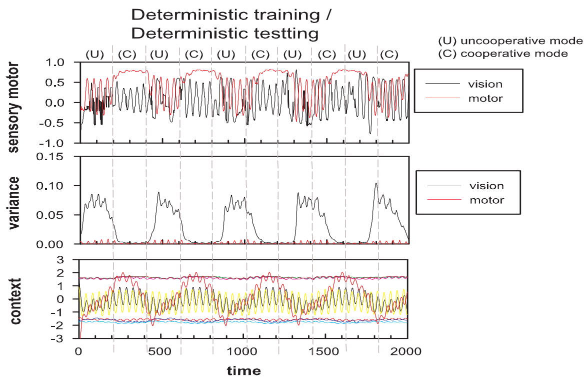

Figure 1: Example of behavior sequence generated by a trained network. The network learned behaviors in which cooperative actions and uncooperative actions are switched in a deterministic manner. In the case of sensory-motor, 2 out of a total of 8 dimensions were plotted (vision: vertical axis of object position captured by the vision sensor, motor: left arm pronation). In variance, two lines corresponding to the relative sensory-motor were depicted. In context, first 8 neural units were plotted (a total of dimensions is 50).

that the network generates the prediction of prediction error as the variance vector $\boldsymbol{v}$. Since the variance $v_i(t)$ works to be a weighting factor for the mean square error $(x_i(t) - \hat{x}_i(t))^2$ [11], the network is able to control the importance of sensations via the variances and, as such, this control might play a role of the "attention". Maximizing P is achieved using the back propagation through time (BPTT) method [15].

3 Results

Two small humanoid robots, A and B, interact with each other in a physical environment. In cooperative mode, robot B will periodically move the object it holds and robot A must attempt to track it with its hand. Robot B will enter uncooperative mode 50% of the time and will randomly move the object, in defiance of its partner. The task for robot A is to learn at which times it is possible to cooperate with robot B and to so when feasible. In the experiment, robot B is controlled by the experimenter and robot A is controlled by the neural network.

Three experiments were carried out, each con-

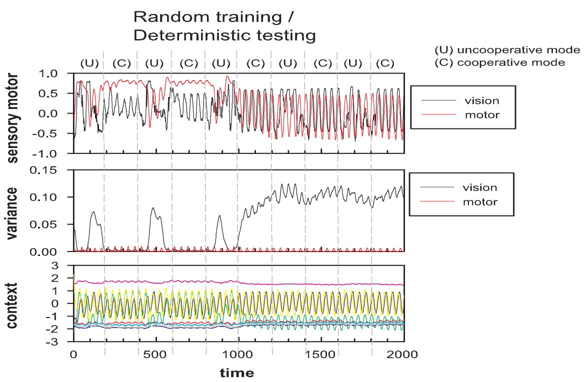

Figure 2: This figure uses the same format as Figure 1 but shows the result when cooperative/uncooperative actions are switched randomly in training.

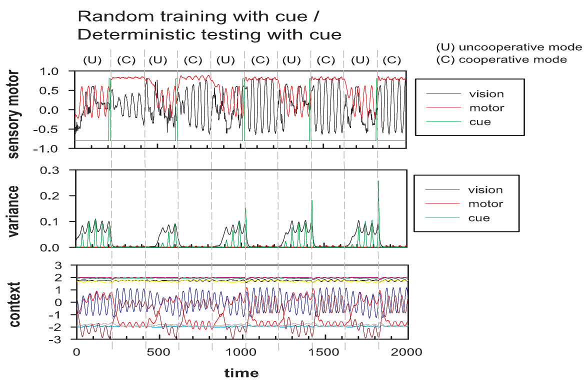

Figure 3: This figure uses the same format as Figure 1 and Figure 2 but reports the result when both sensory-motor and cue are presented simultaneously.

sisting of a training and a testing phase. During training, the switching of robot B's cooperative/uncooperative behaviors was either deterministic or random, depending on the experiment. During testing, the switching of robot B's behavior was deterministic for *all* experiments.

In experiment 1, during training and testing, robot B switched its behavior in a deterministic fashion. Figure 1 displays an example of network dynamics during testing, where robot A learned to predict robot B's behaviors that changed from cooperative to uncooperative actions in a deterministic manner. Through training, the robots were able to reproduce the object manipulation cooperatively, and they also generated uncooperative actions. It can be seen here that the variance corresponding to the vision sensor increased when the partner moved the object randomly.

In experiment 2, robot B switched its cooperative/uncooperative behaviors randomly during training. Figure 2 depicts robot A's network dynamics during testing when robot B switched its behaviors deterministically. Although, the variance plot initially appears to show the correct switching between cooperative/uncooperative modes, ultimately robot A always migrates towards an uncooperative behavior.

The difference between these experiments is attributed to the difference of predictability in behaviors during training. In the case of Figure 1, switching between cooperative/uncooperative behaviors is periodic, and so the network was able to attend to this temporal cue even though ignoring sensations of the uncooperative behavior. In experiment 2, there was no temporal cue available during training, therefore during testing the network could not exploit this during the absence of predictable input. Thus, if some cues which indicate change of partner's behaviors are presented, the network will be able to generate both cooperative/uncooperative actions.

In experiment 3, again robot B switched its cooperative/uncooperative behaviors randomly during training but this time an additional cue was added to indicate the switching to cooperative behavior. The cue took the form of a key press from the experimenter. Figure 3 shows the network dynamics of robot A during deterministic testing containing both sensory-motor and cue inputs.

In this case, the trained network was able to switch from uncooperative behaviors to cooperative one by means of cueing. However, even though the timing of the cue was not predictable, the network could recognize when robot B is in uncooperative mode and activities of variance were initiated before cueing in Figure 4. It can be inferred that the trained robot A prepared mentally for the coming cue by predicting prediction errors.

4 Conclusion

As shown by the robotics experiments, cooperative/uncooperative behaviors of the robots appeared to be controlled by the variances, the so-

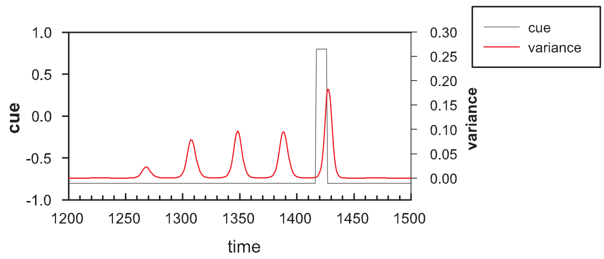

Figure 4: This figure demonstrates how networks generate variances corresponding to unpredictable inputs.

called "prediction of prediction errors". The results revealed that the predictability in terms of "prediction of prediction errors" enables the attention of agents to be focused on predictable parts in the sensory sequences through learning of prediction models for partners' actions. It is also suggested that "prediction of prediction errors" as the meta-level cognition is indispensable in achieving autonomous mechanisms of joint attention in cooperative behaviors.

Acknowledgments: The present study was supported in part by a Grant-in-Aid for Scientific Research on Innovative Areas "The study on the neural dynamics for understanding communication in terms of complex hetero systems" from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

[1] Reynolds, C.W.: Flocks, herds and schools: A distributed behavioral model. ACM SIGGRAPH Comp. Graph. **21(4)** (1987) 25–34

[2] Asada, M., Uchibe, E., Hosoda, K.: Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development. Artif. Intell. **110(2)** (1999) 275–292

[3] Yamaguchi, H.: A cooperative hunting behavior by mobile-robot troops. Int. J. Robot. Res. **18(9)** (1999) 931

[4] Kube, C.R., Bonabeau, E.: Cooperative transport by ants and robots. Robot. Auton. Syst. **30(1-2)** (2000) 85–101

[5] Brothers, L.: The social brain: a project for integrating primate behavior and neurophysiology in a new domain. Concept Neurosci. **1** (1990) 27–51

[6] Frith, U., Frith, C.: The social brain: allowing humans to boldly go where no other species has been. Philos. T. Roy. Soc. B **365(1537)** (2010) 165–175

[7] Posner, M.I.: Orienting of attention. Q. J. Exp. Psychol. **32(1)** (1980) 3–25

[8] Treisman, A.M., Gelade, G.: A feature-integration theory of attention. Cognitive Psychol. **12(1)** (1980) 97–136

[9] Duncan, J., Humphreys, G.W.: Visual search and stimulus similarity. Psychol. Rev. **96(3)** (1989) 433

[10] Spratling, M.W.: Predictive coding as a model of biased competition in visual attention. Vision Res. **48(12)** (2008) 1391–1408

[11] Namikawa, J., Tani, J.: A model for learning to segment temporal sequences, utilizing a mixture of RNN experts together with adaptive variance. Neural Netw. **21(10)** (2008) 1466–1475

[12] Feldman, H., Friston, K.J.: Attention, uncertainty, and free-energy. Frontiers in Human Neurosci. **4** (2010)

[13] Jordan, M.I.: Indeterminate motor skill learning problems. Attention and Performance XIII (1989)

[14] Elman, J.L.: Finding structure in time. Cognitive Sci. **14(2)** (1990) 179–211

[15] Rumelhart, D.E.: Learning internal representations by error propagation. DTIC Document (1985)

[16] Jaeger, H., Haas, H.: Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science **304(5667)** (2004) 78

[17] Schmidhuber, J., Gers, F., Eck, D.: Learning nonregular languages: A comparison of simple recurrent networks and LSTM. Neural Comput. **14(9)** (2002) 2039–2041