Development of a Dividual Model Using a Modular Neural Network for Human-Robot Interaction

Toshiyuki Tanaka [1] and Kunikazu Kobayashi [1*]

[1] Aichi Prefectural University,
1522-3 Ibaragabasama,
Nagakute, Aichi 480-1198, Japan
[*] E-mail: kobayashi@ist.aichi-pu.ac.jp

Abstract

Currently, in the field of human-robot interaction (HRI), robots have a problem that can only interact the same at all times with humans. In this paper, therefore, we introduce the concept called a dividual and build a model of the dividual to grow through interactions with others. In addition, using a modular neural network and reinforcement learning (actor-critic), we confirmed process to choose an appropriate dividual out of plural dividuals.

Keywords: Model of dividual, Human-robot interaction, Robot, Modular neural network, Reinforcement learning

1. Introduction

In recent years, many types of robots have been developed and successfully applied to a variety of fields, such as medical care and disaster relief. Currently, in the field of human-robot interaction (HRI), robots have a problem that they can only interact with humans in a stereotypical way. Humans can however change correspondence depending on a human to be interacted, and realize a variety of interactions (communications). On conventional interaction between humans and robots, robots receive unilateral orders by performing prearranged movement and utterance from humans, and perform given tasks[1,2,3,4,5]. This cannot however realize robots which live together and support humans.

As one of researches on HRI, Kojima tried to have a mind like human for building a social relationship and proposed a model of development to obtain communication skills through social interactions[6]. Shibata made robots play a role in human society. Then he developed a seal-shaped robot and reported that elderly people get pleasure and spiritual comfort through physical interactions such as touching and petting[7]. Naya et al. also presented a system dealing with haptic interactions between humans and robots[8].

In the present paper, in the field of HRI, we develop a dividual model to grow through interactions with others. Then, we introduce a concept of dividual that it is formed into a self with respect to another human through repetitive communications with others[9,10]. Individual cannot divide anymore whereas dividual can divide into plural ones. We use two machine learning techniques to construct the dividual model. Then, we confirm a process to choose an appropriate dividual out of plural dividuals when we appropriately prepare an input set for a dividual model defined by category elements and action ones.

2. Dividual

A concept of dividual is proposed by Japanese novelist Keiichiro Hirano to interact properly with another human[9]. Individual cannot divide anymore whereas dividual can divide into plural ones. Based on the concept of dividual, a human can change to an another self according to environments and human relations. Dividual is formed into a self with respect to another human through repetitive communications with others. In addition, dividual is strictly divided into three types as shown in Fig.1.

Figure 1: Conceptual diagram of dividual.

The first one is a social dividual. This is a standard dividual to interact with a stranger or an unfamiliar person. The second one is a group-oriented dividual. This corresponds to a dividual for a specific group such as a school class or a tennis club. The third one is an individual-oriented dividual. This is a dividual for a specific person such as family members or a close friend. In the present paper, the last one, i.e. individual-oriented dividual is treated.

3. Dividual Model

3.1. Design of Dividual Model

We introduce a modular neural network as a general framework of the dividual model[11,12]. One module corresponds to one dividual as show in Fig.2. Firstly, a dividual model creates a dividual as a module by giving inputs and learning. This corresponds to interaction between humans and robots. Secondly, the model chooses an appropriate dividual out of plural ones according to input information by an output value from each modules. This corresponds robots could change interaction according to humans.

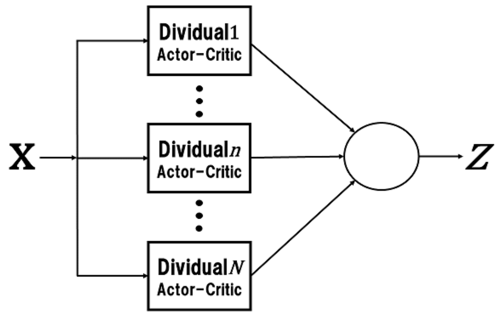

Figure 2: Configuration of a dividual model.

3.2. Learning of Dividual Model

We introduce an actor-critic reinforcement learning method to train each modular neural network in Fig.3[13]. Reinforcement learning is a learning framework that acquires the desirable output and time sequence with a trial and error by being maximized value for result of interaction with environment, that is reword of expectation. One of the representative technique is actor-critic. The actor-critic method consist of an actor to select an action and a critic to evaluate its action.

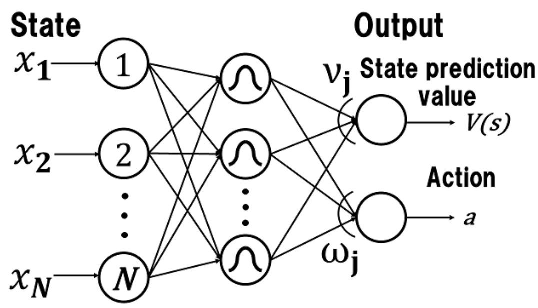

Figure 3: Structure of a neural network based on actor-critic.

The parameters are repeatedly modified as follows.

TD error: Critic observes the rewords r and the next state $\boldsymbol{s}' = (x_1, x_2, \cdots, x_N)^T$, and calculate temporal difference error (TD error) as learning index of critic and actor. TD error δ at time t is expressed by Eq.(1).

$$\delta = r + \gamma V(\boldsymbol{s}') - V(\boldsymbol{s}), \quad (1)$$

where γ is a discount rate, $0 \leqslant \gamma \leqslant 1$.

Learning of middle unit: The output function of middle unit is assumed as Gaussian function with an average $\boldsymbol{c}_j$ and a variance σ_j^2 in Eq.(2).

$$y_j = \exp\left(-\frac{\|\boldsymbol{s} - \boldsymbol{c}_j\|^2}{2\sigma_j^2}\right), \quad (2)$$

The average $\boldsymbol{c}_j$ becomes close to input $\boldsymbol{s}$ as Eq.(3).

$$\boldsymbol{c}_j \leftarrow \boldsymbol{c}_j + \zeta\, \delta\, v_j \frac{\boldsymbol{s} - \boldsymbol{c}_j}{\sigma_j^2} y_j, \quad (3)$$

where ζ is a learning coefficient, $0 \leqslant \zeta \leqslant 1$. In the present paper, the variance σ_j^2 of Gaussian function is assumed to be fixed.

Learning of critic: A weight v_j is updated by Eq.(4) so that TD error becomes zero.

$$v_j \leftarrow v_j + \eta\, \delta\, y_j, \quad (4)$$

where η is a learning coefficient, $0 \leqslant \eta \leqslant 1$.

Learning of actor: A weight ω_j is updated so as to take a higher state value $V(s)$ using Eq.(5).

$$\omega_j \leftarrow \omega_j + \rho\, \delta, \quad (5)$$

where ρ is a learning coefficient, $0 \leqslant \rho \leqslant 1$.

Selection probability of modules: We introduce a soft-max technique based on Gibbs distribution in Eq.(6) so as to adjust the selection probability of modules.

$$\pi(\boldsymbol{s}) = \frac{\exp(p(\boldsymbol{s}))}{\sum_{\boldsymbol{s}' \in X} \exp(p(\boldsymbol{s}'))}, \quad (6)$$

where X refers to an input set and the parameter $p(\boldsymbol{s})$ is updated by Eq.(7).

$$p(\boldsymbol{s}) \leftarrow p(\boldsymbol{s}) + \beta\, \delta, \quad (7)$$

where β is a learning coefficient, $0 \leqslant \beta \leqslant 1$.

4. Computer Simulation

Through computer simulations, we confirm a process to choose an appropriate dividual out of plural dividuals.

4.1. Design of Input Information

We prepare an input set for the dividual model in advance. It basically consists of two kinds of information. The first one is category information to identify a person, e.g. name, sex, age, nationality, hobby and etc. The second one is action information to express interactions, e.g. question, answer and etc. For convenience of identifying a person, we should keep the input dimensions of action information as small as possible. Further, we design four patterns for each person whose category elements are identical but action elements are not.

4.2. Simulation Results

When the proposed model newly gets an input for an unfamiliar person, it creates a novel dividual module for him/her. We designed that each dividual module gets a positive reward for an input containing good action and a negative reward for the other inputs so as to form an appropriate dividual in a module. The accumulated rewards for three dividual modules is illustrated in Fig.4. This figure shows that each module learns correctly because the accumulated rewards is monotonically increasing.

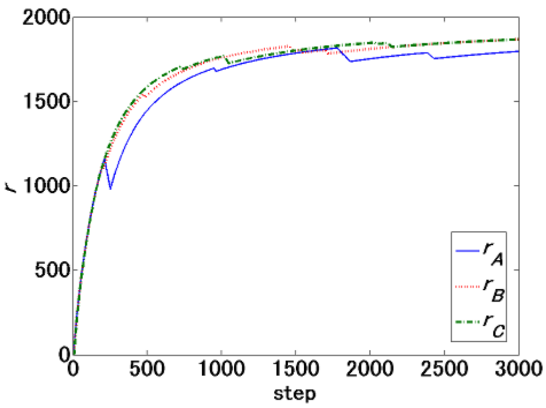

Figure 4: Accumulated reward for three persons.

Figure 5 illustrates the transition of output $V(\boldsymbol{s})$. As shown in this figure, the output of each module takes higher value than other two modules after 3,000 learning steps. We can therefore identify the person by observing $V(\boldsymbol{s})$. Thus, we can confirm the process to choose an appropriate dividual out of plural dividuals for input information.

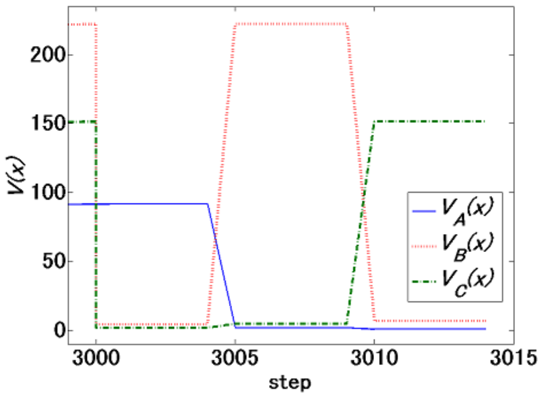

Figure 5: Output value $V(x)$ from each module.

5. Conclusion

In the present paper, in the field of HRI, we developed a dividual model to grow through interactions with others. Through computer simulations, we confirmed a process to form an appropriate individual-oriented dividual.

As a future problem, we should implement social dividual and group-oriented dividual. Then we should verify the validity of a dividual model.

In the near future, when a home robot is introduced to our house, we hope that it should be developed to learn and grow through interactions with family members, and can separately correspond to each family member. In order to realize such a home robot, we must develop a system to enable smooth communication between humans and robots. Once a dividual based human-robot interaction system is developed, robots can take a role in a family member and a doctor and then relationship between humans and robots will be dramatically changed.

Acknowledgments

This work was partly supported by President's Special Research Fund from Aichi Prefectural University, Japan.

References

1. T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A Survey of Socially Interactive Robots," Robotics and Autonomous Systems, **42**, 143–166 (2003).
2. M. A. Goodrich and A. C. Schultz, "Human-Robot Interaction: A Survey," Human-Computer Interaction, **1**(3), 203–275 (2007).
3. T. Kanda, "Research Trends towards Social Robots in HRI," Journal of the Robotics Society of Japan, **29**(1), 2–5 (2011) (in Japanese).
4. T. Ono, M. Imai, T. Etani, and R. Nakatsu, "Construction of Relationship between Humans and Robots," Journal of Information Processing, **41**(1) 158–166 (2000) (in Japanese).
5. K. Kobayashi and S. Yamada, "Human-Robot Cooperation with Commands Embedded in Actions," Transactions of the Japanese Society for Artificial Intelligence, **21**(1), 63–72 (2006) (in Japanese).
6. H. Kozima, "Epigenesis of social intelligence and"theory of mind," IPSJ SIG Notes. ICS, ICS122-3, 13–18 (2000) (in Japanese).
7. T. Shibata, "Research on Interaction between Human and Seal Robot, PARO," Journal of the Robotics Society of Japan, **29**(1), 31–34 (2011) (in Japanese).
8. F. Naya, K. Shinozawa, and K. Kogure, "Human-Robot Tactile Interaction," IPSJ Magazine, **44**(12), 1227–1232 (2003) (in Japanese).
9. K. Hirano, "Who am I? : From Individual to Dividual," Kodansha Shinsho (2012) (in Japanese).
10. M. Kawamura and K. Kobayashi, "An Action Selection Method Using Degree of Cooperation in a Multi-agent Reinforcement Learning System," Journal of Robotics, Networking and Artificial Life, **1**(3), 231–236 (2014).
11. T. Tanaka and K. Kobayashi, "Construction of a Dividual Model Using a Modular Neural Network for Human-Robot Interaction," Proceedings of Tokai-Section Joint Conference on Electrical and Related Engineering, N1-6 (2014) (in Japanese).
12. R. Anand, K. Mehrotra, C. K. Mohan, and S. Ranka, "Efficient Classification for Multiclass Problems Using Modular Neural Network," IEEE Trans. Neural Networks, **6**(1), 117–124 (1995).
13. R. S. Sutton and A. G. Barto, "Reinforcement Learning: An introduction," MIT press (1998).