

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

Toshiyuki Tanaka

*Aichi Prefectural University, 1522-3 Ibaragabasama
Nagakute, Aichi 480-1198, Japan*

Kunikazu Kobayashi

*Aichi Prefectural University, 1522-3 Ibaragabasama
Nagakute, Aichi 480-1198, Japan*

E-mail: kobayashi@ist.aichi-pu.ac.jp

<http://www.ist.aichi-pu.ac.jp/~koba/>

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 the process to choose an appropriate dividual out of plural dividuals.

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

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. And robots are getting into life place without realizing^{1,2,3}. For example, a voice recognition agent using artificial intelligence technique on a smart phone, and a humanoid communication robot manufactured by a major communication company. In Japan, Prime Minister Abe added “robot revolution” to economic growth strategy⁴.

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^{5,6,7,8,9}. 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 interactions¹⁰. 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 petting¹¹. Naya et al. also presented a system dealing with haptic interactions between humans and robots¹².

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^{13,14,15}. 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 the 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.

The outline of this paper is as follows. Section 2 introduces a concept of dividual. Section 3 proposes a dividual model using a modular neural network and actor-critic reinforcement learning. Section 4 explains the results of computer simulations to verify the performance of the proposed model. Finally, Section 5 summarizes the present paper and states future work.

2. Dividual

A concept of dividual is proposed by Japanese novelist Keiichiro Hirano to interact properly with another human¹³. Individual cannot divide anymore whereas dividual can divide into plural ones. The dividual is formed for every partner of communication. Based on the concept of dividual, a human can change to 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.



Fig. 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

We imagine a new HRI system which learn and grow through interactions with family members, and develop

a dividual model to grow through interactions with others.

3.1. Design of Dividual Model

We introduce a modular neural network as a general framework of the dividual model^{16,17}. This is because a dividual is represented by not a centralized network but a distributed network. Then we suppose that a human may switch dividuals not consciously but automatically. One module corresponds to one dividual as show in Fig.2.

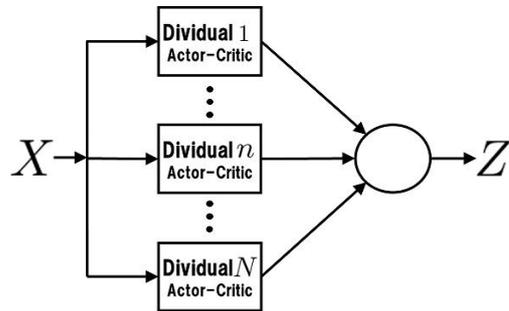


Fig. 2: Configuration of a dividual model.

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 inputs by an output value from each module. This corresponds robots could change interaction according to humans.

3.2. Learning of Dividual Model

We introduce an actor-critic reinforcement learning method to train each module in Fig.3¹⁸. This is because a dividual is formed under the influence of external factors such as environment and interactions with humans. 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 reward of expectation. One of the representative technique is actor-critic. The actor-critic method consists of an actor to select an action and a critic to evaluate its action.

The parameters are repeatedly modified as follows.

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

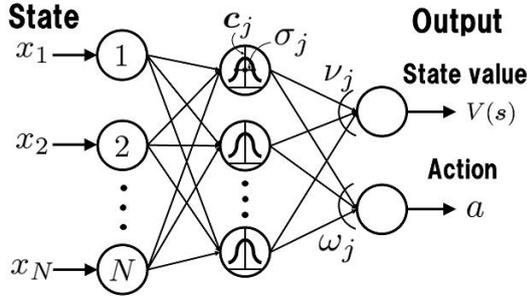


Fig. 3: Structure of a neural network based on actor-critic.

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

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

Learning of middle unit: The output function of middle unit is assumed as Gaussian function with an average c_j and a variance σ_j^2 in Eq.(2).

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

The average c_j becomes close to input s as Eq.(3).

$$c_j \leftarrow c_j + \zeta \delta \nu_j \frac{s - c_j}{\sigma_j^2} y_j, \quad (3)$$

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

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

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

where η is a discount rate, $0 \leq \eta \leq 1$.

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 discount rate, $0 \leq \rho \leq 1$.

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.

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

where X refers to an input set and $q(s)$ is updated by Eq.(7).

$$q(s) \leftarrow q(s) + \beta \delta, \quad (7)$$

where β is a discount rate, $0 \leq \beta \leq 1$.

4. Computer Simulation

Through computer simulations, we confirmed the process to form dividuals and 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. Then we perform the following manipulations sequentially for learning in a neural network.

Labeling elements: We label input elements with appropriate words and values.

Conversion 1: We convert labeled inputs into binary digits. This manipulation digitizes personal information.

Conversion 2: We convert '0' into '-1', and '-1' into '0'. The manipulation solves a problem that 0 has no

information amount when updating weights in a neural network.

Thus, an input is given to the dividual model as a vector consisting of ‘-1’, ‘0’ and ‘1’. In actuality, we define input information as Table 1.

Table 1: Input information

name	sex	age	nationality	hobby
A	man	21	Asia	sports
B	man	40	America	reading
C	woman	30	Asia	cooking
D	woman	10	Asia	game
E	man	19	Europe	music

Next, we label input elements for input sets using Table 2. For example, if a person’s name is A, we label as ‘1’. Then if it’s sex is man, we label as ‘1’. After that, if his age is 21, we label as ‘2’. Similarly, we label other input elements for input sets. ‘-1’ means no information in action information.

Table 2: Labeling elements for input sets

		-1	1	2	3	4	5
category	name (3bit)	-	A	B	C	D	E
	sex (2bit)	-	man	woman	-	-	-
	age (2bit)	-	1~20	21~30	31~40	-	-
	nationality (2bit)	-	Asia	Europe	America	-	-
	hobby (3bit)	-	reading	sports	cooking	music	game
action	question (2bit)	nothing	Where will you go for the vacation? etc.			-	-
	answer (3bit)	nothing	I’m going shopping, etc.			-	-

Next, we perform two types of conversions as shown in Table 3. Conversion 1 converts labeled information into binary digits. Accordingly, we can express “001 01 10 01 010 -1-1 -1-1”. Conversion 2 converts 0 into -1 and -1 into 0. Thus, we can express “-1-11 -11 1-1 -11 -11-1 00 000”. In Table 3, ‘*’ can take any values from {-1, 0, 1}.

Table 3: Conversion of input information

person	labeling	conversion 1	conversion 2
A	11212 **	001 01 10 01 010 *****	-1-11-11 1-1 -11 -11-1 *****
B	21331 **	010 01 10 10 001 *****	-11-1-11 1-1 1-1 -1-11 *****
C	32313 **	011 10 11 10 011 *****	-111 1-1 11 1-1 -111 *****
D	42115 **	010 01 11 10 101 *****	-11-1-11 11 1-1 1-11 *****
E	51124 **	010 01 01 10 100 *****	-11-1-11 -11 1-1 1-1-1 *****

4.2. Simulation Settings

We imagine interaction between humans and robots as follows. On one day, a robot gets information about a human, and performs appropriate behavior towards him/her. On next day, the robot thinks about him/her, and performs both different and appropriate behavior towards him/her. Thus we think that a robot can perform communication with him/her. So we prepare four patterns for each person, that is, category elements are identical but action elements are not. Then we define both positive and negative rewards. The positive rewards denote input sets having good action for each module. On the other hand, the negative rewards express other inputs.

We define simulation settings as shown in Table 4. So we use parameters of updating equations in an actor-critic neural network as shown in Table 5.

Table 4: Simulation settings

parameter	value
Learning times	3000
Number of middle units	5
Number of Learning person	3
Number of input pattern	4
Positive reward	20.0 or 1.0
Negative reward	-1.0

Table 5: Parameter settings for update equations

parameter	value
Discount rate of TD error γ	0.09
Learning coefficient of middle units ζ	0.001
variance σ^2	1.0
Learning coefficient of critic η	0.01
Learning coefficient of actor ρ	0.01
Learning coefficient of soft-max β	0.4

4.3. 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 are illustrated in Fig.4. This figure shows that each module learns correctly because the accumulated rewards are globally monotonically increasing.

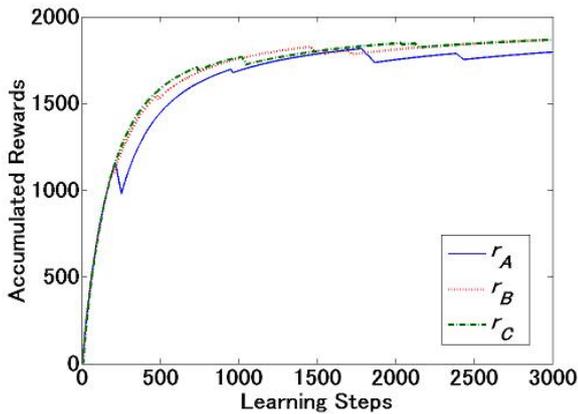


Fig. 4: Accumulated reward for each module.

Figure 5 illustrates the transition of $V(s)$. 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 $V(s)$. Thus, we can confirm the process to choose an appropriate dividual out of plural dividuals for input information.

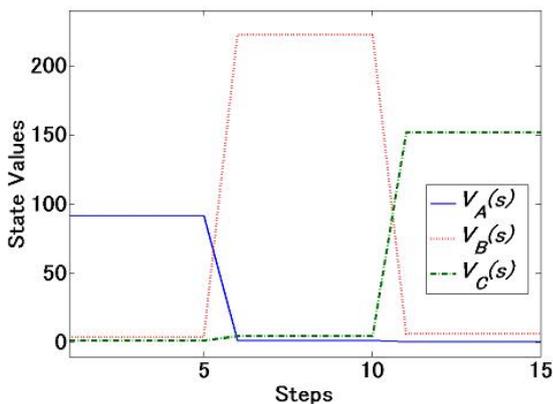


Fig. 5: State value $V(s)$ from each module.

5. Conclusion

In the present paper, in the field of HRI, we introduced a concept of dividual. In addition, we developed a dividual model to grow through interactions with others using a modular neural network and reinforcement learning (actor-critic). Then we confirmed the process to form dividuals and to choose an appropriate dividual out of plural dividuals.

As future work, we should implement social dividual and group-oriented dividual. And we should extend the dividual model considering actually interaction. Finally, we should verify the validity of the 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.

Acknowledgements

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

References

1. H. Inoue, T. Kanade, Y. Anzai, and H. Sena, “*Robotics Creation*,” Iwanami Shoten, (2005) (in Japanese).
2. I. Shimoyama, T. Shibata, K. Ikuta, N. Hata, J. Tani, T. Mathui, and H. Hirukawa, “*Robots Frontier*,” Iwanami Shoten, (2005) (in Japanese).
3. M. Inaba, A. Kagami and K. Nishiwaki, “*Robots Anatomy*,” Iwanami Shoten, (2005) (in Japanese).
4. Y. Honda, “*Robots Revolution : Why do Google and Amazon invest it?*,” Shodensha (2014) (in Japanese).
5. T. Fong, I. Nourbakhsh, and K. Dautenhahn, “*A Survey of Socially Interactive Robots*,” *Robotics and Autonomous Systems*, **42**, 143–166 (2003).
6. M. A. Goodrich and A. C. Schultz, “*Human-Robot Interaction: A Survey*,” *Human-Computer Interaction*, **1**(3), 203– 275 (2007).

7. T. Kanda, "Research Trends towards Social Robots in HRI," *Journal of the Robotics Society of Japan*, **29**(1), 2–5 (2011) (in Japanese).
8. 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).
9. 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).
10. H. Kozima, "Epigenesis of social intelligence and "theory of mind"," *PSJ SIG Notes. ICS, ICS122-3*, 13–18 (2000) (in Japanese).
11. 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).
12. F. Naya, K. Shinozawa, and K. Kogure, "Human-Robot Tac-tile Interaction," *IP SJ Magazine*, **44**(12), 1227–1232 (2003) (in Japanese).
13. K. Hirano, "Who am I? : From Individual to Dividual," *Kodansha Shinsho* (2012) (in Japanese).
14. K. Suzuki, "Smooth Society and the Enemy," *Keisoshobo* (2013) (in Japanese).
15. M. Kawamura and K. Kobayashi, "An Action Selection Method Using Degree of Cooperation in a Multi-agent Reinforcement Learning System," *Journal of Robotics, Net-working and Artificial Life*, **1**(3), 231–236 (2014).
16. T. Tanaka and K. Kobayashi, "Development of a Dividual Model Using a Modular Neural Network for Human-Robot Interaction," *Proceedings of International Conference on Artificial Life and Robotics (ICAROB2015)*, pp.107–110 (2015).
17. 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).
18. R. S. Sutton and A. G. Barto, "Reinforcement Learning: An introduction," *MIT press* (1998).