Analyzing the Goal Finding Process of Human’s Continuous Learning with the Reflection Subtask

Tomohiro YAMAGUCHI *, Yuki TAMAI *, Yoshimichi HONMA *, and Keiki TAKADAMA **

Abstract: This paper reports our experimental results on analyzing a human’s goal finding process in continuous learning. The objective of this research is to make clear the mechanism of continuous learning. To fill in the missing piece of reinforcement learning framework for a learning robot, we focus on two human mental learning processes, awareness as pre-learning process and reflection as post-learning process. To observe mental learning processes of a human, we propose a new method for visualizing them by the reflection subtask for human to be aware of the goal finding process in continuous learning with invisible mazes. Two-layered task is introduced. The first layer is the main task of continuous learning designing the environmental mastery task to accomplish the goal for any environment. The second layer is the reflection subtask to make clear the goal finding process in continuous learning. The reflection cost is evaluated to analyze it.

Key Words: reinforcement learning, learning process, awareness, reflection, continuous learning.

1. Introduction

Researches on learning process are divided into two fields. One is a learning algorithm for robot in robotics and artificial intelligence [1], and the other is learning of a human in psychology [2]. For the learning robot or agent, reinforcement learning is the major framework since it automatically learns after a learning goal is set in the learning environment. The main feature of reinforcement learning is that the learning goal is given by the human designer. On the other hand, researches on human learning ability have been performed in various research fields such as psychology, education, business, and so on. One of the main features of human learning ability is that it covers a vast territory of learning ability including discovery of learning goals, awareness, reflection, self-regulated learning [3], or continuous learning.

The objective of this research is to analyze the continuous learning ability of a human. The question is how to observe mental learning processes of a human. Previous methods of human learning researches mostly depend on observable behaviors or activities. On the other hand, a learning process of a human has a major difficulty in observing since it is a mental process. Then a human learning process is yet-to-be-defined. So it is necessary to add a new twist to observe the human learning process.

To solve these problems, we propose a new method for visualizing two kinds of mental learning processes, called awareness and reflection. Awareness process plays an important part to trigger behavior change. To visualize awareness process, we introduce invisible walls in the learning environment to observe the behavior, whether wall encountering action or a wall avoiding action. If a learner is aware of a presence of an invisible wall, the learner may behave to avoid it. Reflection process is the conscious re-evaluation of experience for the purpose of guiding future behavior. To visualize reflection process, we design the reflection subtask in order to assist to work through the main learning task. We focus on continuous learning and aim for modeling the unified continuous learning process model based on reinforcement learning framework for both a human and a learning agent [4]. Our new approaches are as follows:

1. To fill in the missing piece of reinforcement learning whose learning process is behavior change, we add two mental learning processes, awareness as pre-learning process and reflection as post-learning process.

2. The learning environment is an invisible maze consisting of invisible walls which are perceived as a sign that suggests the number of walls in the neighborhood.

3. We add the reflection subtask for expressing and summarizing something to be aware of learning from mistake or to be reflected on learning from experience. A learner can mark up his/her traces of the actions and signs for future success. It turns out to visualize his/her mental learning processes.

4. To examine the reason why the non-continuous learner stops learning the task, we analyze the learner’s performance of both the main learning task by achievement cost and the reflection subtask by reflection cost in each learning stage.

This paper reports our experimental results on analyzing human’s continuous learning ability with the reflection cost. In our previous work, the experimental results showed that there is a strong negative correlation (the correlation factor is $-0.78$) between the number of discovered solutions and the reflection cost [5],[6]. The learners who found many solutions learn continuously and work smart with a small reflection cost. The main
factor whether they learn continuously or not is to find new learning goals. To examine the reason why the non-continuous learners stop learning the task and why the continuous learners keep learning, each learner’s performance of the reflection sub task in each learning stage is analyzed. In the next section, we give a theoretical background on continuous learning and the continuous learning process model.

2. Background

This section describes a theoretical background of this research, continuous learning, awareness and reflection. The concepts awareness and reflection are viewed differently across the disciplines among computer-supported cooperative work (CSCW), computer-supported collaborative learning (CSCL), psychology, business, educational sciences, computer science, and so on. First we summarize an overview of continuous learning and its learning process which is compared with that of reinforcement learning, then our usage of awareness and reflection is introduced. For more information, refer [5], [6].

2.1 An Overview of Continuous Learning

The concept of continuous learning comes from industrial and organizational psychology. Smita and Trey [7] review research on continuous learning. There are three levels, individual, group, and organizational level. This paper focuses on the individual level of continuous learning since we try to analyze the goal finding process of continuous learning. At the individual level it is self-directed or autonomic. It means that the learning goal of the continuous learning is not given by others but is decided by the learner oneself. One of conceptual definitions of continuous learning is according to Sessa and London [8], “Continuous learning at the individual level is regularly changing behavior based on a deepening and broadening of one’s skills, knowledge, and worldview.”

2.2 Learning Process Model to Achieve Continuous Learning

A learning process model to achieve continuous improvement has been proposed [9]. Learning process is defined as a process that results in changed behavior. Learning process consists of several mental processes. This model assumes a learner and the leader. The role of this leader is mentor or coacher rather than teacher [10]. Continuous learning seeks learner-centered rather than teacher-centered.

2.3 Comparison of Continuous Learning with Reinforcement Learning

Figure 1 compares learning process between continuous learning and reinforcement learning. Figure 1(a) shows the learning process of continuous learning in the field of psychology [9]. The major feature of it is that there are 1) awareness and 5) reflection processes in order to circle each process endlessly. The reason is that continuous learning is designed for an adult learner who is expected to spiral up along the endless learning cycle toward mastery of some professional skill. Thus it has commitment process to commit a learning goal. However, each process is mental process. In contrast with it, Fig. 1(b) shows the learning process of reinforcement learning in the field of machine learning. Major distinction is that there is no corresponding continuous learning process of 1) awareness and 5) reflection since its process has a start and an end. Since the objective of existing reinforcement learning is to find an optimal solution for the learning goal given by a human designer. Note that there is no previous research on continuous learning except our researches [4]–[6] in artificial intelligence field. The next subsection reviews awareness and reflection as important mental processes of continuous learning process.

2.4 Awareness and Reflection on Learning Processes

The common feature of awareness and reflection is focusing experience of a learner on some information for future improvement. The main differences between them are that awareness relates to the perception, reflection mainly relates to the action or the behavior consisting of perception and action.

2.4.1 Awareness as pre-learning process

As pre-learning process, awareness plays an important part to trigger behavior change. Then we focus on the meaning of awareness as the need for distinction of indistinguishable perceptions or experience between future success and future failure. We assume that these indistinguishable perceptions occur by partially observable states [11].

2.4.2 Reflection as post-learning process

There are several meanings of reflection on learning. One of them, learning through reflection is defined as “those intellectual and affective activities in which individuals engage to explore their experiences in order to lead to new understandings and appreciations” [12]. We take to the concept of reflective learning [13] which is “the conscious re-evaluation of experience for the purpose of guiding future behavior.” As post-learning process, reflection plays a part to create some meaning out of behavior change or learning result. Then we focus on the meaning of reflection as creating the becoming explanation or interpretation toward the rule which can lead the learner not to future failure but to future success.

3. Analyzing the Goal Finding Process by the Two-Layered Task

This section describes designing the two-layered task to make clear the goal finding process in continuous learning. The goal finding process is a process of meta-learning to discover learning goals while learning the main learning task. First, the environmental mastery task to accomplish the goal for any environment is introduced, then the meta-level learning task is described.

3.1 The Environmental Mastery Task

The first layer is the main task of continuous learning. We define the environmental mastery task which is to accomplish the goal for any environment as one of the continuous learning tasks. Since the objective of the continuous learning is to
learn something new endlessly, we design the goal of the task as the maze sweeping task. It is to find the different longest paths from the start state to the goal state in the maze as many as possible. As the learning environment for the environmental mastery task, we design the sequence of the learning stages in ascending order of difficulty such as from a simple rectangle maze to complex shape one whose boundary has a concavo-convex shape or which has inner walls. To find out the solution of the maze sweeping task, it is necessary to explore the boundary of the maze area to grasp the shape of the entire maze. Then they are the subgoals for the continuous learning in the maze sweeping task. So the action of first encounter with an invisible wall consisting of walls in-line is rewarded as the exploration to be aware of discovering the subgoal. Note that it is not instructed to a learner. If the learner is aware of this rewarding rule on first contact event, at first he/she is encouraged that the encounter with an invisible wall is a positive event, then he/she may become aware of efficient exploration of an invisible wall and an invisible side of the boundary of the maze. The essence of this idea is the potential for the automatic sub-reward generation according to the complexity of a learning environment. In the maze environment of our system, the boundary shape and inner obstacles of a maze represents the complexity of the learning environment.

3.2 The Meta-Level Learning Task to Find the Learning Goals

The second layer is the meta-level learning task. We assign the reflection subtask for human to be aware of the goal finding process in each learning stage to observe the process. We design the learning environment with an invisible goal state, and each maze is surrounded by invisible walls as the boundary of the maze, and some mazes in the later stages may have inner walls as invisible obstacles. To perform the maze sweeping task, it is necessary for a learner to identify them. There are two kinds of objectives of the reflection subtask. The first one is that learners are instructed to make out the reflection map of each maze by placing color markers on it on the current maze at the trial of the maze sweeping task in order to give a new learner a clue to perform the stage.

The second objective is to find learning goals which are not instructed directly. The human learners are expected to become to see the meaning of the reflection map and the reflection subtask as their learning stages progressed. In the learning experiment, the reflection cost of each learner during the learning stages is analyzed, whether stable or not. Note that the reflection cost is the time a learner worked through the reflection subtask per the number of found solutions. If the reflection cost of a learner is stable during the learning stages, it suggests that the learner can perform the reflection subtask in a certain amount of time and can perform a very good job of it, even though it becomes more difficult as the learning stage progressed.

To make clear the reflection mental process of the continuous learners, we report the experimental results. We have a hypothesis that an intelligent learner utilizes information acquired from the current trial and errors for the forward trials or achievement in the way of the reflection subtask. For example, it is important to be aware of finding the invisible walls through the reflection subtask since the distribution of the markers placed on found invisible walls may suggest the blank area to search as making out the reflection map.

4. Designing the Continuous Learning Process with the Invisible Maze Model

This section summarizes the design concepts for the continuous learning process, then describes the way to design the continuous learning support system based on the reinforcement learning process model to guide a human to achieve continuous learning. For modeling the continuous learning process as shown in Fig. 1, this paper formalizes them by a maze model with invisible walls and an invisible goal and designs the maze sweeping task which requires discovering and mastering invisible various learning goals of a learner.

4.1 The Flow of the Continuous Learning Process

Figure 2 shows the flow of the continuous learning process. This process consists of a triple cycle. The innermost cycle is called a trial. A trial is defined as a transition sequence from a start state to encountering either a goal state or a wall. In this cycle, a learner repeats an action and his/her mental process including awareness until he/she results in either success or fail of the task. The second cycle is called an achievement.

An achievement is defined as the learning of a maze sweeping task with the fixed start and the invisible goal. In this cycle, the learner reflects the trial by the reflection subtask when the trial terminates by the encounter with a wall or a goal. If current trial is not accomplished, he/she restarts the trial from the start state. The outmost cycle is the continuous learning cycle. When the learner accomplished current achievement, he/she can challenge next new achievement.

4.2 Designing the Learning Environment by an Invisible Maze

Designing a learning environment for a human learner, we adopt a grid maze model from start to goal since it is a familiar example to find the path through a trial and error process [4]. A maze model is defined by five elements, state set, a sign of walls, transitions and walls, action set, and meta-actions.

Figure 3 shows the structure of a 2D grid maze. The \( n \times m \) grid maze with four neighbors consists of \( n \times m \) number of 1 \( \times 1 \) squares. It is called a simple maze which is surrounded by walls in a rectangle shape. Figure 3(a) shows a \( 3 \times 2 \) simple maze with a start and a goal. In a grid maze, every square that touches one of their edges except a wall is connected. Each square in a maze model is called a state.

Transitions in a maze model is defined between two connected states. They are represented as the labeled directed graph as shown in Fig. 3(b). Action set is defined as a set of

![Fig. 2 The flow of the continuous learning process.](image-url)
labels to distinguish the possible transitions of a state. In a grid maze model with four neighbors, a learner can execute four kinds of absolute actions: \{up, right, left, down\} or relative actions: \{forward, turn-right, turn-left, reverse\}. The action toward a wall results in the transition to the start state to restart the trial. Transition to a goal state results in automatic restart to the start state with the judgment whether the maze sweeping task is accomplished or not. Next we describe a maze sweeping task which is performed in an achievement. It is defined as to find paths from a fixed start state $S$ to the invisible goal state $G$ which visit all states only at once in the maze.

### 4.3 Visualizing the Awareness Process by a Sign

A sign is defined as the number of walls of four neighbors in each state; $\{0, 1, 2, 3, 4\}$. Figure 4 shows an illustrated example of a non-simple maze with invisible walls. Figure 4(a) shows a $4 \times 4 - 2$ invisible maze and an example of the transition sequence as if wall-following. Dashed lines are invisible walls which are invisible for a learner during trials. Figure 4(b) shows the distribution of signs among visited states in the maze. In our previous experiment, a number was displayed as a sign of surrounding walls in each visited state. In this case, the sign ‘2’ means a corner [5],[6].

Making the learning tasks more difficult to learn continuously than previous research, we introduce a marked sign $[5,6]$. Figure 4(c) shows the current version of the distribution of the marked signs which the learner perceives. The marked sign is displayed at the visited state to assist awareness of invisible walls until the end of the trial. A marked sign is defined by the sign sequence. A sign sequence between two neighbor states are described as follows:

- (the sign of previous state, the sign of current state)

In Fig. 4(c), the marked sign * which comes from (0, 1) suggests the presence of a wall. A state whose sign is 2 suggests whether a corner or a passageway. The marked sign + with light color which comes from (0/1/2, 3) suggests the presence of a corner (or a passage). The marked sign + with dark blue which comes from (2, 2) suggests the presence of cascade corners, cascade passages or cascade corner and passage. A state which sign is 3 suggests the dead end. The marked sign with light/normal/dark green which comes from (0/1/2, 3) suggests the presence of a deadend.

### 4.4 Visualizing the Reflection Process by the Reflection Subtask

This subsection describes the way to visualize reflection process. We introduce the reflection subtask which is used to describe the awareness from a sign of the invisible walls to visualize his/her reflection process. The reflection subtask is defined to attach a marker on a state in the reflection map for expressing and summarizing a learner’s mental learning processes. The markers are one yellow marker and many RGB markers. Figure 5 shows the initial reflection map of each stage with ten kinds of markers. RGB markers are red, green, and blue markers, these color density is in three tones, light, normal and dark in each color. Therefore there are nine kinds of many RGB markers and one yellow marker. Note that right hand of diamond shape marker is the eraser icon to erase a marker in the reflection map.

At the end of a trial, markers can be attached on any state in the reflection map. Figure 6 shows an example of the reflection subtask of Stage 1. Figure 6(a) shows an example of the trial of Stage 1. In each state, whether visited or not is distinguished by its background color of the square, either gray (visited) or white (not visited). A star mark is a sign whose color suggests that the wall surrounding situation is different. If the goal state is found, its background color is yellow and small G is displayed. Figure 6(b) shows an example of the reflection map of Stage 1. In each square, a maximum of four kinds of markers are placeable. Typical usage of them is to memorize the special event such as the encounter with an invisible wall or a goal, or to summarize the sign distribution in the searched maze. The reflection map is displayed during a train of trials.
4.5 Designing the Sequence of Learning Goals

Designing the learning goals, even the case of simple mazes may result in the difficulty of a partially observable state \[11\], it can be worked out by slightly broadening a learner’s perceptual states to distinguish partially observable states. The case of non-simple mazes results in more difficulty of partially observable states since they have various non-distinguishable states. This paper describes learning goals for a simple maze with a partially observable state. Under simple and invisible mazes, the standard steps to master them are the following steps:

1. Find a goal state if it is invisible.
2. If walls are invisible, estimate the borders of the current maze such like wall-following behaviors.
3. Search all maze sweeping paths from the start state.

We focus on learning goals for the second step to estimate the invisible borders of a simple maze. To solve this, the learner should try to consider the sequence of signs toward neighbor state. If the learner encounters either a wall or a sign 0 state, he/she can decide the direction of walls. In this case, an important learning goal to prevision an invisible wall is being aware of sign sequences as (0, 1) or (1, 2). Note that these increasing patterns for provisioning an invisible wall hold good only under simple mazes.

5. Experiment

To examine the reason why the non-continuous learner stops learning the task, we conducted the more difficult continuous learning experiment than previous research by eight human subjects. In our previous work, the experimental results showed that there is a strong negative correlation (the correlation factor is \[-0.78\]) between the number of discovered solutions and the reflection cost \[5],[6\]. The learners who found many solutions learn continuously and work smart with a small reflection cost.

The main factor whether they learn continuously or not is to find new learning goals. To examine it, this paper analyzes each learner’s reflection cost which is the performance of the reflection subtask in each learning stage. We designed the continuous learning task as 30 stages in ascending order of difficulty. Note that there is no limit on the time of the experiment nor the number of encounters with a wall. Main difficulties are as follows:

1. The number of stages is reduced from 54 to 30. So the degree of increasing difficulty per stage is augmented than our previous experiment.
2. The perceived sign is more difficult to understand as we described before. If the learner becomes aware of the meaning of the color of the signs, the learner can estimate the presence of walls so as to continue achieving more difficult second-half learning stages.

5.1 Experimental Setup

5.1.1 The learning environment for the continuous learning

We prepare the sequence of invisible mazes consists of total 30 stages. In each stage, there is one maze with two or more solutions to be found, excepting the 18th, 19th and 26th stages with only one solution. The sequence of stages is arranged in order of increasing difficulty. The first three stages are practice stages, and the first nine stages are simple mazes, the 10th stage or later are all non-simple mazes. Note that a non-simple maze is surrounded by outside walls in a concavo-convex shape, or it has an internal dividing wall.

5.1.2 Visualizing reflection process by the reflection subtask

We prepared the reflection subtask with ten kinds of color markers as we described. The method of operation is instructed, but the role of each color marker is not assigned.

5.1.3 The first instruction for subjects

The first instruction is instructed at the beginning of the experiment.

- A solution is a path from the start state to the goal state.
- The number of discovered solutions is limited by a ceiling in each stage.
- The objective of the task is to find more sweeping solutions in each maze.
- A sweeping solution is the path to visit all states from the start state to the goal state.
- The first three stages are practice stages.

5.1.4 The second instruction for subjects

The second instruction is instructed after first three stages are finished.

- The real purpose of the task is to learn the way to get better by challenging the maze task one after another.
- In the reflection subtask, place the minimum markers in the reflection map to hint the way to solve the maze required by a new challenger of this maze task.
- You may decide freely what to get better.
- If you are hard to get better, you may terminate the maze task.

5.1.5 The measurement item

The measurements of the experiment are as follows:

(a) The total time of the experiment
(b) The number of challenged stages
(c) The number of discovered sweeping solutions
(d) The number of encounters with a wall
(e) The total time of the achievements
(f) The total time of the reflection subtask
(g) The reflection cost
(h) The achievement cost

Note that the total time of the experiment (a) is defined by Eq. (1). The reflection cost (g) is defined by Eq. (2); it evaluates the quality of the reflection process. The achievement cost (h) is defined by Eq. (3); it evaluates the quality of the achievement process.
Fig. 7 The reflection cost in each stage of a continuous learner.

\[ a = (e) + (f), \]  
\[ g = (f)/(c), \]  
\[ h = (e)/(c). \]

5.2 Summary of the Experimental Results

The experimental results are as follows:

(a) The total time of the experiment
   Average time is 42 min, maximum is 91 min, minimum is 20 min.

(b) The number of challenged stages
   Average is 17.4, maximum is 30, minimum is 9.

(c) The number of discovered sweeping solutions
   Average is 29.3, maximum is 64, minimum is 17.

The reason that the standard deviation in (c) is bigger than that of (b) is that there are about two sweeping solutions in each stage. Two subjects within eight continued to perform the experiment more than 90% of the stages (more than 27 stages within total 30 stages). Considering the number of stages as the continuous learning ability, we call these subjects as continuous learners. Subjects who perform the experiment less than 90% of the stages are called non-continuous learners.

5.3 Analyzing the Experimental Results

To examine the reason why the non-continuous learners stop learning the task, we analyze the learners’ performance of both the main learning task and the reflection subtask in each learning stage. As the experimental results, there are three cases whether the human learning is continuous or not. Figures 7, 8, 9, and 10 show the relationship between the reflection cost and the progression of stages. The horizontal axis is the progression of stages, the vertical axis is the reflection cost in each stage (second/solution). Figure 7 shows an example of the continuous learning case in which the reflection cost of a continuous learner is stable during the learning stages. It suggests that the continuous learner can perform the reflection subtask in a certain amount of time, even though it becomes more difficult as the learning stage progressed.

On the other hand, Fig. 8 shows two kinds of non-continuous learners. Figure 8(a) shows an example of increasing the reflection cost case that it becomes too large to continue the task. In contrast to this, Fig. 8(b) shows an example of decreasing case, it suggests that the motivation of the learner becomes too small to continue the task.

Interestingly, one non-continuous learner is a compound case. Figure 9 shows the reflection cost and the achievement cost of the compound non-continuous learner, which increases in the first half stages, then it decreases in the second half stages. The reason why the reflection cost of this learner decreases after the 10th stage is that it becomes difficult from the 10th stage on since each maze is the non-simple maze from the 10th stage on. The 11th stage is at the peak of both the reflection cost and the achievement cost is 11th stage as shown in Fig. 9 since the learner struggles against the non-simple mazes. After the 11th stage, these costs decrease and then fade out.

Figure 10 shows the reflection cost in each stage among eight subjects. Each line segment shows the linear approximation of this relationship of each subject. The right side’s end point of each line segment is the last stage that the subject achieved. Figure 11 shows the achievement cost in each stage among them. As the experimental results, both the reflection cost and
the achievement cost of the continuous learner are stable during the learning stages as compared to non-continuous learners. It suggests that the continuous learner can perform continuously and can work smart with small and stable reflection costs.

5.4 Discussions

First, the meaning of the reflection cost and its usage are discussed. We regard the reflection cost as the level of proficiency on the reflection skill. If a learner gets more proficient at reflection, the reflection cost comes down then the learner becomes to solve invisible mazes more easily. Therefore, the learner can achieve more difficult learning stages within the range of achievement cost.

Then, as the future direction of this research, we discuss the way to generalize our proposed methods. Since the objective of the continuous learning task is to learn the way to learn continuously. One of good ways is to find a new interpretation of the task. One of the applications of the continuous learning task is a human robot interaction task. A learning stage is a context of interactions between a human and a robot. The maze in the stage is a situation of the interaction, surrounding walls are the frame to set the interaction in it. Sign from the human in the interactions suggests that the action of the robot will be out of the situation.

6. Conclusion

We described the way to design the continuous learning support system and reported our experimental results on analyzing human’s goal finding process for continuous learning with the reflection cost. As the experimental results, both the reflection cost and the achievement cost of the continuous learner are stable during the learning stages as compared to non-continuous learners. It suggests that the continuous learner can perform continuously and can work the reflection subtask smart with small and stable reflection costs in a certain amount of time, even though it becomes more difficult as the learning stage progressed. Future work is to design the collaborative continuous learning system between a human learner and a learning robot.

Acknowledgments

We would like to thank Prof. Habib and Prof. Shimohara for offering a good opportunity to present this research. This work was supported by JSPS KAKENHI (Grant-in-Aid for Scientific Research) Grant Number 16K00317.

References

[1] S. Russell and P. Norvig: Artificial Intelligence: A Modern Approach, 3rd edition, pp. 830–857, Prentice Hall, 2009.

[2] F. Marton and S. Booth: Learning and Awareness, pp. 33–55, Routledge, 1997.

[3] D.H. Schunk and B.J. Zimmerman: Motivation and Self-Regulated Learning: Theory, Research, and Applications, pp. 1–30, Routledge, 2007.

[4] T. Yamaguchi, K. Takemori, and K. Takadama: Modeling a human’s learning processes toward continuous learning support system, M.K. Habib, J. Paulo Davim, (eds.), Mechatronics Engineering, pp. 69–94, Wiley-ISTE, 2013.

[5] T. Yamaguchi, K. Takemori, Y. Tamai, and K. Takadama: Analyzing human’s continuous learning ability toward the intelligent robotics, Proceedings of the Twentieth International Symposium on Artificial Life and Robotics 2015 (AROB 20th 2015), pp. 752–757, 2015.

[6] T. Yamaguchi, K. Takemori, Y. Tamai, and K. Takadama: Analyzing human’s continuous learning processes with the reflection sub task, Journal of Communication and Computer, Vol. 12, No. 1, pp. 20–27, 2015.

[7] J. Smita and M. Trey: Facilitating continuous learning: Review of research on individual learning capabilities and organizational learning environments, Paper presented at the annual meeting of the AECT International Convention, Louisville, http://www.memphis.edu/icl/id/crcl/crcl-smita-research.pdf, 2012.

[8] V.I. Sessa and M. London: Continuous Learning in Organizations: Individual, Group, and Organizational Perspectives, pp. 17–36, Psychology Press, 2006.

[9] B.A. Buckler: Learning process model to achieve continuous improvement, The Learning Organization, Vol. 3, No. 3, pp. 31–39, 1996.

[10] K.B. Dorval, S.G. Isaeks, and R.B. Noller: Leadership for learning: Tips for effective mentoring and coaching, K.W. McCluskey (ed.), Mentoring for Talent Development Unknown Binding, Reclaiming Youth International, http://www.cpsb.com/resources/downloads/public/Mentoring_Coaching.pdf, 2003.

[11] L.P. Kaeldbling, M.L. Littman, and A.R. Cassandra: Planning and acting in partially observable stochastic domains, Artificial Intelligence, Vol. 101, pp. 99–134, 1998.

[12] D. Boud, R. Keogh, and D. Walker: Reflection: Turning Experience into Learning, pp. 7–18, Routledge, 1985.

[13] B. Krogstie, M. Prilla, and V. Pammer: Understanding and supporting reflective learning processes in the workplace, Proceedings of the 8th European Conference on Technology Enhanced Learning (EC-TEL 2013), pp. 151–164, LNCS 8095, Springer, 2013.

Tomohiro Yamaguchi (Member)

He received his B.E, M.E. and Ph.D degrees from Osaka University, Japan, in 1985, 1987, and 1996 respectively. He joined Mitsubishi Electric Corporation in 1987 and moved to Matsushita Electric Industrial in 1988. He worked at Osaka University from 1991 to 1998 as a research associate. He moved to Nara National College of Technology (currently National Institute of Technology, Nara College) as an associate professor in 1998 and is currently a professor of the Department of Information Engineering from 2007. His research interests include interactive recommender system, music information retrieval, multiagent reinforcement learning, autonomous learning agent, human-agent interaction, learning support system, human learning process and mastery process. He is a member of the Japanese Society for Artificial Intelligence.
Yuki Tamai  
He graduated the Faculty of Advanced Engineering, National Institute of Technology, Nara College, Japan and received his B.E. degree from the National Institution for Academic Degrees and University Evaluation, Japan in 2017. Currently he joined NTT-Neomeit from 2017. His research interests include learning support system, motivating human, learning process, human-agent interaction and human learning process.

Yoshimichi Honma  
He received his B.E. and M.E. degrees from Toyohashi University of Technology, Japan, in 1992 and 1994 respectively. In 1997, he joined Nara National College of Technology, where he is currently a lecturer of the Department of Information Engineering from 2000. His research interests include learning support system, network management, computer security and computer architecture.

Keiki Takadama (Member)  
He received his B.E. degree from Ritsumeikan University, Japan, in 1993, M.E. degree from Kyoto University, Japan, in 1995 and got Ph.D. Degree from The University of Tokyo, Japan, in 1998, respectively. He joined Advanced Telecommunications Research Institute (ATR) International from 1998 to 2002 as a visiting researcher and worked at Tokyo Institute of Technology from 2002 to 2006 as a lecturer. He moved to The University of Electro-Communications as an associate professor in 2006 and is currently a professor of the Department of Informatics from 2011. His research interests include multiagent system, autonomous system, human-agent interaction, reinforcement learning, learning classifier system, and emergent computation. He is a member of IEEE and a member of major AI- and informatics-related academic societies in Japan.