Interactive robot learning system based on human operation

D. Katagami

Tokyo Institute of Technology
4259Nagatsuta midori-ku
Yokohama, 226-8502, Japan
katagami@ntt.dis.titech.ac.jp

S. Yamada

National Institute of Informatics
2-1-2 Hitotsubashi Chiyoda-ku
Tokyo, 101-8430, Japan
seiji@nii.ac.jp

Abstract

Conventionally, in the system which contains a human element in an evaluation system, a user's physical and mental load was a big problem. Many of conventional researches have mainly corresponded by the three improvement methods to this problem. In this research, the mechanism in which a robot's action is automatically gained by human operation is mounted. The purpose of this research is reducing a user's cognitive load. They are the improvement of an input interface, the improvement of a presentation interface, and improvement in the speed of EC convergence. On the other hand, in this research, we attention to an operated type robot as an object of research. It aim at the construction of a system design in consideration of interaction between a robot and an operator in this research. A user operates a robot, in order to attain a task. When a system uses a user's operation information for the instruction information on a robot's action, a system gains a robot's behavior automatically by managing a task, without a user being conscious of instruction. Therefore, it is possible to reduce a user's cognitive load. In this research, we implement such a mechanism and verify by some experiments that it is possible to reduce a user's cognitive load.

1 Introduction

In recent years, various agent robots, such as a pet type robot and a robot for welfare, have come to play an active part in the same environment as man. Many studies of the approach using interaction with the human who exists in environment has been proposed. Particularly for the robots that do not have a priori knowledge or commit trial and error in the initial stage, human instruction is the very effective acquisition technique of autonomous behavior. However, in a certain level of autonomous robot, it is not necessary to follow instruction from human all the time. In the stage which does not need instruction, robot should demonstrate its autonomy based on the instruction rules stored by interaction with human without putting a burden on human. Therefore, we need to the technique of establishing a robot's autonomy from through interaction between human and a robot is required.

Asoh et al.[1] proposed the framework that a mobile robot built the map information of the unknown environment, called Jijo-2 which performs a communication by voice conversation with human. Ishiguro et al. [5] built the state space of the mobile robot by reinforcement learning. Horiguchi et al. [3] used the idea of the mutual leadership pattern interaction as the design of the interaction of the robot with the human and realized the cooperation behavior of the automation process of a mobile robot and human operations by using power feedback. Inamura et al. [4] indicate acquirement behavior of a robot using Bayesian Network based on a dialog with a user.

However, in the research which contains such a human element in an evaluation system, a user's load has been pointed out as a big problem. Many of conventional researches have mainly corresponded by the three improvement methods to this problem. They are (1) the improvement of an input interface, (2) the improvement of a presentation interface, and (3) improvement in the speed of EC convergence. About two of the former can be called what aimed at reduction of a user's physical load by the improvement of each interface. And the latter can be called what aimed at reduction of a user's physical load by shortening of the study time in improvement in the speed of EC. However, the question of how to reduce the user's load the problem effectively is still open.

Then, in this research, we pay attention to an operated type robot. We aim at reduction a user's load by construction of a system design using the interaction between a robot and a user and the timing of operation of a robot. A user operates a robot, in order to attain a task. In this system a user operates a robot, in order to attain a task. By using the user's operation information for the instruction information for a robot's action, a system gains a robot's behavior automatically only by a user managing a task. Moreover, the timing of operation prepares and considers some setup, in order to reduce the load to a user. Thereby, a user can create instruction information for a robot, without being not much conscious of instruction. Therefore, it is possible to reduce a user's cognitive load.

In this research, we verify by some experiments that it is possible to reduce a user's cognitive load by mounting such a mechanism.

2 Interactive Teaching

2.1 Teacher's Load

Generally, in interactive evolutionary learning, the more it is taught, the better the performance is. However, human labor is not unlimited. It is clear that it is trade-off like it is better as instruction cost lowers. Human's labor has a limit in cooperating with a machine without tiredness, carrying out comparison evaluation of many individuals (or rules) for every generation, and inputting an evaluation value. This has been a serious practical problem. Moreover, as the second problem, the number of individuals and the number of search generations must be lessened as compared with the usual EC search in order to reduce physical and mental load in case human evaluates individuals. It makes convergence worse. As a result, it is difficult to reduce the number of times of instruction (physical load).

Then, we pay attention to a user's cognitive load. In the case of study using the interaction with a user like this research, if a user's load is divided into two kinds, physical load and cognitive load, it can be said that the conventional research has aimed at mainly reduction of physical load. However, since there is the above-mentioned problem, it is difficult to mitigate this.

On the other hand, in this research, a user operates a robot directly with input equipment. It considers performing automatic generation of a rule from the operation information and the environment information at that time as teaching. A user can performs teaching without consciousness of performing the input of comparison or an evaluation value for many individuals like the conventional interactive technique. Moreover, when there is no input of instruction information from a user, a robot can act autonomously based on the rule accumulated by experience of an interaction with the user till then. It is expected that this load problem is sharply mitigated in the following three points by this method.

(1) Consciousness of teaching: A rule is automatically created by intuitive teaching

(2) Timing of teaching: Operate a robot (teaching) at the time of a user's favorite.

(3) Cooperation with an Autonomic System: A system learns autonomously.

In this paper, it verifies by experiment about (1) and (2).

2.2 Timing of Teaching

We think that the timing of teaching is greatly concerned with the above-mentioned teacher's load. Generally, timing of teaching is performed beforehand (Off-line Teaching), or has much what is performed at the time of the demand of a system (Passive Teaching). Since these techniques have left the timing which instruction performs to the system side, in order to teach, a man side must stand by. Not to mention the experiment in a simulation, a teacher's load increases further in the real environmental learning that needs more time for an experiment.

Then, we propose the following Active Teaching methods. We perform the experiment which measures a user's cognitive load as compared with the conventional Off-line Teaching method and the Passive Teaching method, and verify by psychological evaluation. Each technique is as follows.

2.2.1 Active Teaching

In this teaching method, it is possible that a teacher gives instructs to a robot at favorite timing. In this research, this is called Active Teaching method. Seeing a robot perform autonomous action, a teacher operates a robot to favorite timing and makes a task. Thereby, teacher can instruct to a robot being unconscious of teaching, without worrying about whether he teaches by seeing a robot's action. Thereby, a teacher can teach without worrying about whether being conscious of teaching, whether it teaches, or not when he/she saw a learner's all actions. However, it is difficult to include such specification in a system side.

2.2.2 Off-line Teaching

Off-line teaching is the method of performing exploration by instruction at Teaching Mode beforehand, and performing exploitation at Autonomous Behavior Mode.

2.2.3 Passive Teaching

We define passive teaching method as the method of directing teaching at the time of the demand of a system to a user. Mishima and Asada el al. have improved that the efficiency of learning gets worse by Passive Teaching for a gap (Cross Perceptual Aliasing) of the environmental recognition produced between a teacher and a learner [8]. In study efficiency, Passive Teaching has little futility of teaching and is considered to be a good method. However, the teacher has to be supervising until a system requires action. Moreover, since it does not know when the timing comes, it is thought that a mental load becomes large to the number of instruction.

3 Teaching based on Interactive Classifier System

3.1 Interactive Classifier System

We introduced the above-mentioned technique into the method of Interactive Classifier System (ICS) [6] developed so far. ICS is the robot study model

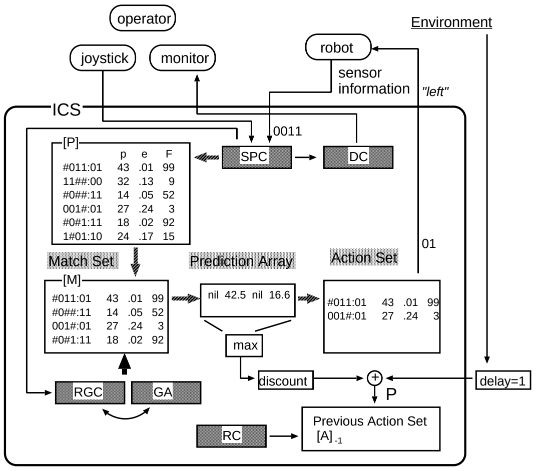

Figure 1 Overview of Interactive Classifier System

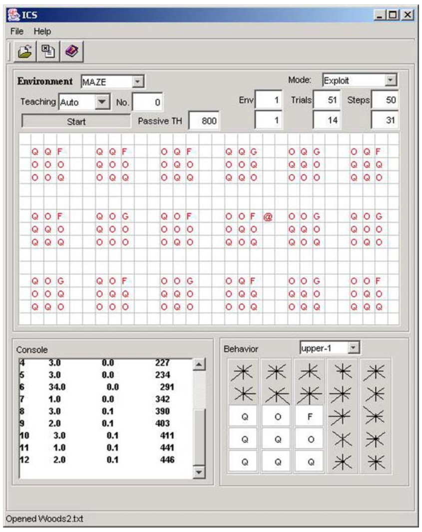

Figure 2 User Interface

that can also perform study by instruction in addition to autonomous study. It included the interactive function of Interactive Evolutionary Computation in (Learning Classifier System (LCS). ICS uses XCS which is one of the LCS as study algorithm. XCS[9] which Wilson proposed is used for LCS that is study algorithm. XCS adds the parameter that is what improved ZCS and is called accuracy. The framework figure of the built system is shown in Fig.1.

ICS consists of a rule generation component (RGC), a sensor processing component (SPC), a display component (DC) and a reinforcement component (RC). Each module is explained below.

yRGCz Rule Generation Component creates the rule by instruction. A teacher operates it using an input equipment, looking at the information displayed on an interface in a robot. A sensor processing part (SPC) receives the operation history of a there, and the sensor information of the robot at that time, RGC creates a rule newly from it, and it adds to a rule list. The creation procedure of a rule was improved so that a rule could be created from instruction information (the action to which operator operated the robot) on the basis of XCS[9].

1. ICS receives a robot's sensor information X and instruction information a_t from SPC.
2. Some classifiers that matched X is moved from a group [P] to a match set [M]. ICS turns regularly the Prediction value of classifier which supports each act a_i in [M] with a Fitness value, and creates $P(a_i)$. The value of $P(a_i)$ is put on Prediction Array, and the act of classifier chosen by $P(a_i)$ is chosen by act selection methods. Act selection methods are performed by deterministic selection method or roulette wheel selection method.
3. If $a_j \neq a_t$ to compare act a_j chosen by act selection methods and act a_t obtained by teaching, the action part of the rule which has a_j in an action part in [M] will be rewritten to a_t. A change will not be made if $a_j = a_t$.
4. The action set [A] which consists of classifiers in [M] which supports selected act a_j is created. When act a_j or a_t is sent to an effect machine, and in case of a_t, reward r_{teach} is given immediately. When there is no input of a_t, remuneration r_{imm} is returned from environment.

yRCz Reinforcement Component is a reinforcement learning part in classifier system. It learns by updating the parameter of classifiers chosen last time step. When there is no operation of a teacher, a robot can act autonomously from the rule created by then.

yDCz Display Component takes charge of the display of the data processed by SPC. The developed interface is shown in Fig.2.

ySPCz Sensor Processing Component performs processing of a robot's various sensors and processing of teaching information. It is sent to DC and RGC and the processed data is displayed and ICS creates classifiers from them.

3.2 Procedure of Learning

ICS performs two modes: a teaching mode and an autonomous behavior mode by turns. The procedures of the two modes are shown in the following.

Teaching Mode

1. Prepare the robot's state space.
2. It teaches depending on any of the procedure of the timing of three kinds of instruction they are.

3. An operator creates a rule by instruction information and environmental information at the time.
4. If there is no rule belonging to the same cluster, it will add as a rule newly.
5. If there is a rule belonging to the same cluster, a strength value will be updated by reward.

Autonomous Behavior Mode

1. The robot behaves by conforming to stored rules in Rule List.
2. If the average of the number of the time steps from GA of just before in a match set exceeds a threshold, GA will be performed to the match set.

3.3 Procedure of the Timing of Teaching

The timing of teaching has three timing described in Chapter 2.2. Each procedure is shown below. Each is performed in **Step 2** in teaching mode.

Off-line Teaching

1. A teacher directs action to state space.

Passive Teaching

1. Act A will be performed if there is effective action A to state space.
2. If there are no directions, directions will be requested to a teacher.

Active Teaching

1. To state space, if there are directions from a teacher, it will perform.
2. If there are no directions, a robot will perform exploration autonomously.

4 Experiment

4.1 Experimental Settings

We test a preliminary experiment to evaluate the effectiveness our ICS. This is a very simple domain. We use Woods2 environment which one of Wood-like environments[9] as an environment in the experience. It used as a test-bed in several works based on classifier system. Fig.3 shows Woods2 environment. This environment is markovian multi-step problem. The left and right edges of Woods2 are connected, as are the top and bottom. Woods2 has two kind of "food" and two kind of "rocks". F and G are the two kind of food, with sensor codes 110 and 111, respectively. O and Q are the two kind of objects, with sensor codes 010 and 011, respectively. Blanks have sensor code 000. The system, here regarded as an animat or artificial animal, is represented by *. To sense its environment, * is capable of detecting the sensor codes of objects occupying the eight nearest cells. The encoding of a classifier is as follows. A classifier, for example,

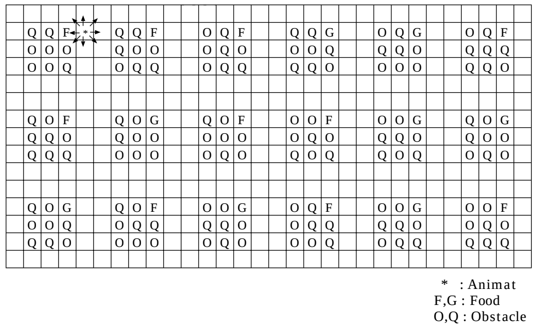

Figure 3 Woods2 Environment

is the 24-bit string 000000000000000010010110. The left-hand three bits are always those due to the object occupying the cell directly north of *, with the remainder corresponding to cells proceeding clockwise around it. The animat's available actions consist of the eight one-step moves into adjacent cells, with the move directions similarly coded from 0 for north clockwise to 7 for northwest. If a cell is blank, * simply moves there. If the cell contains food, * moves to the cell, "eats" the food, and receives a reward($r_{imm} = 1000$). ICS used a population size, N, of 800 classifiers. Parameters were set as follows: $\alpha = 0.1$, $\beta = 0.2$, $\gamma = 0.95$, $\theta = 25$, $\epsilon_0 = 0.01$, $\chi = 0.8$ and $\mu = 0.04$

4.2 Experiment Description

We conducted the comparison experiment with Active Teaching, Passive Teaching or Off-line Teaching. It is one trial, when it arrives at the goal or 50step movement is carried out. Seven graduate students were experimented on the subject by considering 50 trial as one experiment.

In the experiment with the cognitive load of human being like this research, in order to investigate the load, the method of preparing another task has been performed. For example, the method of measuring a participant's cognitive load by performing another task, while the participant in an experiment performs an original task is performed [2].

In this research, a primary task is that a user gives an agent instruction information as an agent arrives at Food. In order to measure a user's cognitive load in teaching, a user must solve two digits addition problems as much as possible while performing an agent's instruction task.

4.3 Effectiveness of Teaching

In this work, we investigated performance in average Step to Food and average of generated Population Size Fig.4 shows the Steps to Foods. And, Fig.5 shows Population Size.

In each technique, there were few differences about the instruction effect. In fact, since study algorithm is the same, although the effect is based on whether

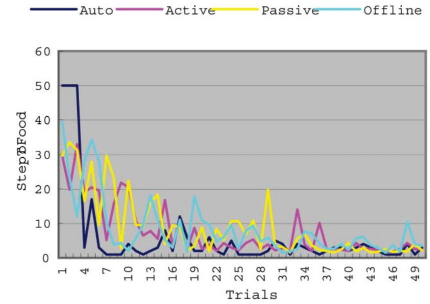

Figure 4 Step to Food

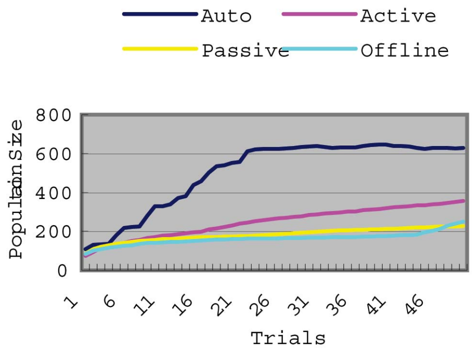

Figure 5 Population Size

a user's instruction mistake changes with the timing of instruction, it can be said that there were few the differences. Although each user does not grow familiar with teaching yet but performance is low in the initial stage, it turns out that it is getting used by about 30 trial.

In Fig.5, the method which based on Autonomous Behavior Mode is converged with the about 600 classifiers. On the other hand, each method which based on Teaching Mode is converged with the about 200 classifiers. This is because the rule which attached importance not only to efficiency but to the intention by judgment of man is created. When there are two or more solution methods efficiency is the same, it is shown that it is difficult to create a rule with many variations. To the surprising thing, as for the Active Teaching method, population size is over the 400 classifiers. The following things can be considered as this reason. When the Active Teaching method teaches to a user's favorite timing, teaching will be performed for every moment. Therefore, a user will do different instruction even if environmental conditions is the same. It is necessary to analyze a rule in detail and to examine it.

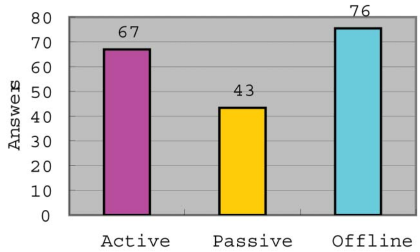

Figure 6 The number of Answers

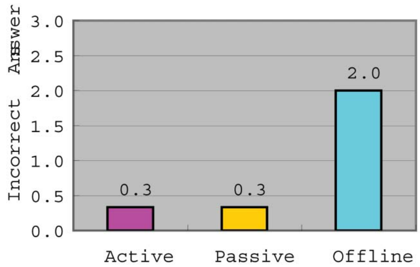

Figure 7 The number of Incorrect Answers

4.4 Discussion of Teaching Load

In order to measure the load of instruction, we measured the number of answers and the number of wrong answers of two digits addition problems in question in a secondary task. Fig.6 shows the number of answers of two addition problems solved while the user performed the primary task. In the Active Teaching method, each subject's number of average answers is 67 questions. Each user is teaching by light load to the extent that it is the same as teaching by Offline. On the other hand, the Passive Teaching method has the low number of answers. This system is determining the timing of teaching and has given a user cognitive load.

Fig.7 shows the number of the mistakes of the answer of the secondary task which the user solved during the experiment of a primary task. In the Offline Teaching method, the mistake of two questions occurred in about 80 questions on the average. In Offline Teaching method, although the speed of task achievement goes up when a user teaches continuously, accuracy is lost simultaneously.

Fig.8 shows the result of the questionnaire of cognitive load. The subject evaluated by the seven-point method about validity, efficiency, and satisfaction to the three teaching methods based on evaluation of usability, respectively.

- Effectiveness:
  Accuracy and completeness with which users achieve specified goals

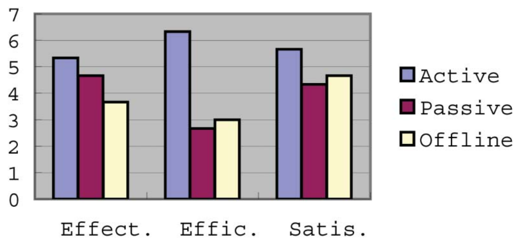

Figure 8 Usability

- Efficiency:
  Resources expended in relation to the accuracy and completeness with which users achieve goals
- Satisfaction:
  Freedom from discomfort, and positive attitudes towards the use of the product

Although there was no statically significant about Effectiveness and satisfaction, statically significant about Efficiency was seen ($p < .05$). When the questionnaire survey of free description of a user was conducted, in the Passive method, there was opinion "I cannot be concentrated on a secondary task since timing of instruction is restricted (specification)", "there is much check work and I cannot be concentrated on a task", etc. Since the system side has determined timing, it turns out that cognitive load is increasing. As a result of a secondary task, although the effect of cognitive load mitigation was acquired, since the user did not recognize the effect, it was not in Satisfaction, or Effectiveness with conclusion directly. On the other hand, in the Active Teaching method, the opinion "the motion of an agent needed to be observed and it took time", "O, Q, F, G, etc. waver for a moment since it is a character", etc. was acquired. On the other hand, in the Active Teaching method, the opinion "I needed to observe the motion of an agent and required time", "I mixed up the difference among characters, such as O, Q, F, and G," etc. was acquired. When a user teaches to his/her timing, a user is enabled to grasp a motion of the instantaneous agent, and the judgment becomes exact. Therefore, these will be improved.

5 Conclusion

We proposed the learning system which mitigates the load of teaching of a user by the timing of the operation using the interaction between a user and an operated type robot. We performed the comparison experiment with the conventional teaching method by the simulation, and verified about (1) Consciousness of teaching and (2) Timing of teaching. In order to measure the load of instruction, we measured the number of answers and the number of wrong answers of two digits addition problems in question in a secondary task. Moreover, we investigated the effect by the questionnaire of the cognitive load based on evaluation of usability. In Active Teaching method, we were able to mitigate cognitive load, without lowering the performance (Step to Food) of a task. It is under implementing this system to a pet robot AIBO in order to experiment in a real environment now.

References

[1] H. Asoh and Y. Motomura and I. Hara and S. Akaho and S. Hayamizu and T. Matsui: Combining probabilistic map and dialog for robust life-long offifce navigation; *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 807–812 (1996)

[2] J. Crandall and M. A. Goodrich: Characterizing Efficiency of Human Robot Interaction: A Case Study of Shared-Control Teleoperation; *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1290–1295 (2002)

[3] Y. Horiguchi and T. Sawaragi and G. Akashi: Naturalistic Human-Robot Collaboration Based upon Mixed-Initiative Interactions in Teleoperating Environment; *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 876–881 (2000)

[4] T. Inamura, M. Inaba, and H. Inoue: User Adaptation of Human-Robot Interaction Model based on Bayesian Network and Introspection of Interaction Experience; *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2139–2144 (2000)

[5] H. Ishiguro and R. Sato and T. Ishida: Robot Oriented State Space Construction; *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1496–1501 (1996)

[6] D. Katagami and S. Yamada: Interactive Classifier System for Real Robot Learning; *IEEE International Workshop on Robot and Human Interaction*, pp. 258–263 (2000)

[7] D. Katagami and S. Yamada: Interactive Evolutionary Robotics from Different Viewpoints of Observation; *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1108–1113 (2002)

[8] C. Mishima and M. Asada: Active Learning from Cross Perceptual Aliasing Caoused by Direct Teaching; *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1420–1425 (1999)

[9] S. W. Wilson: Classifier fitness based on accuracy; *Evolutionary Computation*, Vol. 3, No. 2, pp. 149–175 (1995)

[10] S. W. Wilson: ZCS: a zeroth order classifier system; *Evolutionary Computation*, Vol. 2, pp. 1–18 (1994)