A Decision Making System of Robots Introducing a Re-construction of Emotions Based on Their Own Experiences

Shogo Watada, Masanao Obayashi[*], Takashi Kuremoto, Shingo Mabu
Graduate School of Science and Engineering, Yamaguchi University,
Ube, Yamaguchi, 755-8611, Japan

Kunikazu Kobayashi
School of Information Science and Technology, Aichi Prefectural University,
Nagakute, Aichi, 480-1198, Japan

[*]Corresponding author. E-mail: m.obayas@yamaguchi-u.ac.jp
Tel: +81-83-933-5982, Fax: +81-83-933-5988

Abstract

We already proposed a decision making system for autonomous robots, which is based on a Markovian emotional model proposed by Banik. In our proposed system, decision making of the robot is designed automatically by two processes that are the construction of emotions and optimization of behavior selection policy for emotions. In this study, we introduce an experiences-based reconstruction of emotions to our decision making system as an approach to an automatic construction of decision making online learning.

Keywords: Emotion model, Decision-making, Optimization problem.

1. Introduction

There exists much literature on the study which focuses on the concept of emotions that animals have, for developing autonomous systems in various fields[1-5]. We have already proposed a decision making system for autonomous robots, which is based on a Markovian emotional model proposed by Banik[5,6]. The Markovian emotional model is a probabilistic model that consists of a finite state machine[5].

In our proposed system, a decision making system of the robot is designed automatically by two processes. The first is a construction of emotions based on learning of input stimuli. We proposed use of a SOM (Self-organizing map) to associate input stimuli with internal parameters that represent emotional states. The second is optimization of behavior selection policy for emotions by optimization of system parameters which define behavior selection probability of each emotion.

In our previous studies, a construction of emotions based on learning of the input stimuli was designed only using learning-samples of input stimuli on offline in pre-learning. Therefore, it made useless emotional areas corresponding to the unrealistic input stimuli in the SOM.

In this study, we introduce an experiences-based reconstruction of emotions in our decision making system as an approach to an automatic construction of decision making online learning. That is, we thought that more effective associations with input stimuli and

parameters decided by considering actual frequency and measurement values of input stimuli were achieved by online learning.

We confirmed the usefulness of our method by computer simulation for an E-puck mobile robot. The task for the robot is to maximize floor areas which the robot has passed. In the simulation, the result showed that additional learning of input stimuli makes unevenness of frequency of use of each node in the SOM decreases and it also makes more efficient associations between input stimuli and emotional states.

This paper is organized as follows. Section 2 describes the Markovian emotional model proposed by Banik. In Section 3, the proposed behavior selection system is described. In Section 4, the simulation results of the proposed method are presented. Finally, Section 5 summarizes the effectiveness of the proposed method and future works.

2. The Markovian emotional model

The Markovian emotional model referenced in this paper is a probabilistic model that consists of a finite state machine[5]. It consists of four basic emotions: joy, anger, fear and sadness. Emotional states Y are presented as shown in Eq. (1).

$$Y = [y_{joy} \quad y_{anger} \quad y_{sadness} \quad y_{fear}] \tag{1}$$

The emotional state transition at time k is given by Eq. (2).

$$Y_{k+1} = CY_k \tag{2}$$

Here C is the emotional state transition matrix, which can be expressed as Eq. (3).

$$C = \begin{vmatrix} P_{joy/joy} & P_{joy/anger} & P_{joy/sad} & P_{joy/fear} \\ P_{anger/joy} & P_{anger/anger} & P_{anger/sad} & P_{anger/fear} \\ P_{fear/joy} & P_{fear/anger} & P_{fear/sad} & P_{fear/fear} \\ P_{sad/joy} & P_{sad/anger} & P_{sad/sad} & P_{sad/fear} \end{vmatrix} \tag{3}$$

$P_{A/B}$ is the probability of transition from state B to state A. $q_{*/*}$ is a prior probability of emotion transition, which should be set in advance (See Eq. (4)). These values $P_{*/*}$ can be changed online by the influence of emotion-inducing factors: α, β, γ and δ for joy, anger, fear, and sadness, respectively. These factors vary depending on the internal state and information given by certain other robots and have the role to stimulate each

emotion. For example, probabilities of state transition from joy to other states are expressed by Eq. (4).

$$\left. \begin{aligned} P_{anger/joy} &= q_{anger/joy} + (\beta - \alpha)q_{anger/joy} \\ P_{sad/joy} &= q_{sad/joy} + (\delta - \alpha)q_{sad/joy} \\ P_{fear/joy} &= q_{fear/joy} + (\beta - \alpha)q_{fear/joy} \\ P_{joy/joy} &= 1.0 - (P_{anger/joy} + P_{sad/joy} + P_{fear/joy}) \end{aligned} \right\} \tag{4}$$

3. Proposed behavior selection system

We have already proposed a decision making system of a robot, which uses a Markovian emotional model proposed by Banik. The overall structure of the system can be shown in Fig.1. The system consists of four modules: Cognition, Emotion, Behavior-selection, and Behavioral-making. Z is input stimuli, α, β, γ and δ are emotion-inducing factors, X_k is a behavior selection probability vector at time k, s is a robot state and u is a control input for the robot.

3.1. *The function of each module*

The function of each module is as follows; the Cognition module transforms input stimuli Z such as sensor values into emotion-inducing factors α, β, γ and δ. The Emotion module is a Markovian emotional model proposed by Banik, and this module updates emotional states Y using α, β, γ and δ. The Behavior-selection module determines the behavior selection probability vector X using Y by Eq. (5).
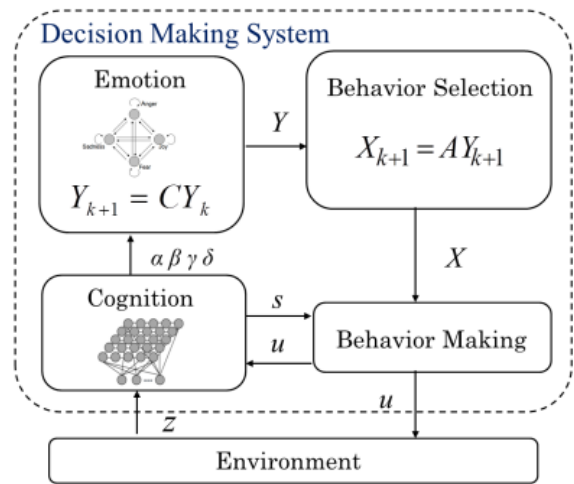
Z: Measurement   Y: Emotional state
C: Emotional state transition matrix
A: Behavioral state transition matrix
s: Situation Behavior state α, β, γ, δ: Emotion-inducing factors

Fig.1 A proposed behavior selection system.

$$X_{k+1} = AY_{k+1} \qquad (5)$$

Elements of X are selection probability of each behavior of the robot, for example, X is defined such as Eq. (6). T means transpose of a vector.

$$X = \begin{bmatrix} x_{drive} & x_{left} & x_{right} & x_{Back} \end{bmatrix}^T \qquad (6)$$

A is a transition matrix consists of probabilities $P_{behavior/emotion}$ which means a probability of transition from *emotion* to *behavior*, and represented by such as Eq. (7).

$$A = \begin{vmatrix} P_{Drive/joy} & P_{Drive/anger} & P_{Drive/sad} & P_{Drive/fear} \\ P_{Left/joy} & P_{Left/anger} & P_{Left/sad} & P_{Left/fear} \\ P_{Right/joy} & P_{Right/anger} & P_{Right/sad} & P_{Right/fear} \\ P_{Back/joy} & P_{Back/anger} & P_{Back/sad} & P_{Back/fear} \end{vmatrix} \qquad (7)$$

Finally, the Behavior-making module generates the control input u based on X taking into account states of the robot.

3.2. Clustering of input stimuli in Cognition module

In our proposed system, we used a SOM in the Cognition module in order to cluster input stimuli such that similar input stimuli correspond with similar emotion-inducing factors. In this method, we regard the position of each node in the SOM as value of emotion-inducing factors. Since it is considered that groups of similar input vectors are compressed in the output layer as weight vectors after learning, each value of emotion-inducing factors is uniquely determined to a stimulation input vector if we consider that the two-dimensional output layer is a two-dimensional map consisting of emotional inducing factors.

The structure of the Cognition module in this study is shown in Fig.2. Each node in the input layer of the SOM corresponds to N input stimuli. The Cognition module constitutes the association between emotions and input stimuli by clustering of input stimuli to the output layer. Then, after learning, the Cognition module can convert input stimuli into emotion-inducing factors by determining a winner node and normalizing the position of it into range from 0 to 1.

In previous studies, learning of the SOM was processed only using learning-samples of input stimuli in advance. Therefore, it generated useless emotional areas corresponding to the unrealistic input stimuli in the SOM. In this study, we propose the method that the
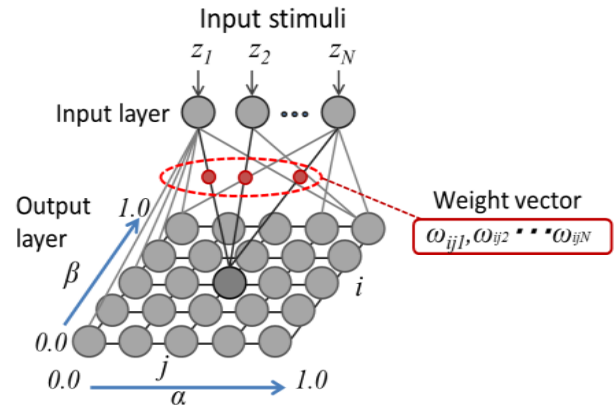

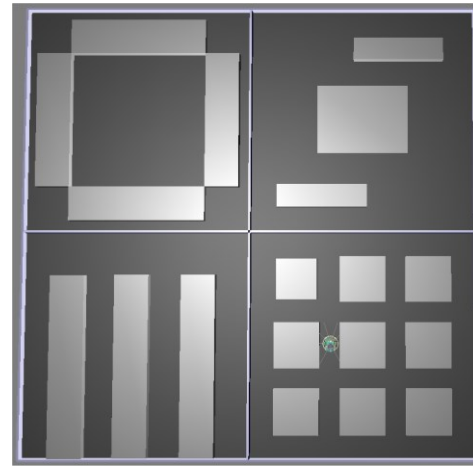Fig.2 New cognition module proposed in this paper.
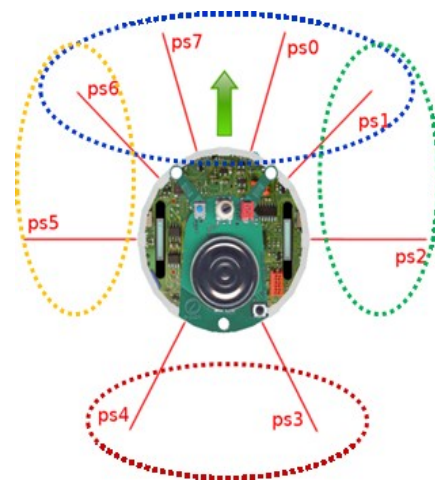

Fig.3 Environments used in simulations.

Fig.4 Sensors of e-puck.

association between input stimuli and emotion-inducing factors becomes more efficient by additional learning of realistic input stimuli on online during the task.

3.3. *Optimization of behavior policy by GA*

In our Behavior-selection module, the policy of behavior selection depends on transition matrix *A*. So, it becomes possible to make various methods for behavior selection by adjusting parameters being included in *A*. In this work, we use GA for setting *A*. GA is a meta-heuristic algorithms for searching the approximate solution. The detail explanation about GA is omitted.

4. Simulation

For evaluating our system, we simulated the problem of maximizing areas that the robot passed in predefined time. In such as house cleaning robots, it is demanded that avoiding obstacles using own sensors, robots pass more large areas in the environment. In this work, we showed that our proposed system generate a behavior policy automatically that a two-wheel mobile robot passes larger areas in fixed time. The simulation which used an e-puck robot was carried out using a Webots simulator which make possible to simulate in a very realistic way.

4.1. *Simulation setting*

The simulation was executed through four environments with different structures which are shown in Fig.3. Each environment is a square of 1 m on a side, and there are walls or obstacles (brighter areas in Fig.3). The e-puck is equipped with 8 infrared sensors which measure proximity of obstacles in a range of 4 cm. We defined four behaviors of the robot; straight, left, right and back, and defined the motor torques of left and right wheels on each behavior as follows: straight: [100, 100]; left [-50,50]; right: [50, -50]; back: [-100,-100]. The Behavior-making module calculates output torques by weighted syntheses of torques with rate of each behavior probability in the probability vector *X*.

The number of input stimuli is 6, and there are shown in Table1. The pre-learning of SOM used learning samples which were generated randomly based on prior knowledge of devices such as the range of the sensor.

On the basic setting of simulation, the pre-learning of the input stimuli runs at first. Then, the decision

making policy is generated by optimization of the transition matrix *A* using GA. In the GA, each individual executes the task in all 4 environments, and the fitness value of the individual is determined as the sum of above 4 areas the robot passed. The setting of the SOM and the GA are shown in Table 2.

Table 1 Input stimuli of the robot.

z_0	The average distance value of four sensors on front (see Fig.4).
z_1	The average distance value of two sensors on right (see Fig.4).
z_2	The average distance value of two sensors on left (see Fig.4).
z_3	The average distance value of two sensors on rear (see Fig.4).
z_4	The total moving distance of robot from 12 seconds ago.
z_5	The total rotation angle of robot from 12 seconds ago.

Table 2 Parameters setting of SOM and GA.

SOM		GA	
Number of learning	500	Number of Individuals	24
Number of nodes	30×30	Final Generation	50
Learning coefficient η	0.0001	The limited time in one environment	600 (s)
$\sigma(0)$*	24.0		

* The neighborhood function: $\sigma(step)=\sigma(0)(1.0\text{-}step/maxStep)$

4.2. *Simulation result*

In this simulation, pre-learning of the Cognition module was performed 1000 steps using random learning samples of input stimuli that their ranges were set as follow; z_0, z_1, z_2, z_3: [0.0:4.0]; z_4, z_5: [-4.0:4.0]. Then, the additional learning runs while young generations in the process of GA after pre-learning. In additional learning, the SOM is updated on online using input stimuli which occur in the task in the first 10 generations of 50 generations. Since updating of the SOM is running every steps, the total number of updates is 576,000.

Since the position of each node of the SOM corresponds to one of values of emotion-inducing factors, it is easy to understand the relationship between input stimuli and emotions by making the emotional map. Fig.5 shows the emotional map made from the SOM after the pre-learning. On this map, the input stimuli are represented in each circle. A shorter distance to obstacle is represented as a bigger size of an ellipse, and magnitudes of a moving distance and a rotation angle are represented as the length of the vertical direction and the horizontal direction of arrows, respectively. For example, in Fig.5, Joy corresponds to a situation of input stimuli such that there are no obstacles

around the robot, and the robot moves forward straight. In the pre-learning, SOM learns equally all situations such as that the robot is surrounded by wall, or the robot moves backward. On the other hand, Fig.6 shows the emotional map after the additional learning. In this map, we can see many input stimuli that occurred frequently during the task such as that corresponding to the situation that the robot moves forward or turns left. These are because unnecessary input stimuli that the robot did not experience during the simulation were excluded. This result shows that the SOM reflects realistic input stimuli by additional learning.

Observing behaviors of the robot using final parameters in GA, the behavior policy were such as follows: go straight bending slightly the left when there are no obstacles around the robot; turn left when there're obstacles in front or right side of the robot; turn right when there're obstacles in left side of the robot. The trajectories of the robot in four environments are shown in Fig.7. In our method, the behaviors mentioned above are able to be explained by parameters of A shown in Eq. (8). For example, in probabilities of transition from Joy, which correspond to no existence of wall around the robot according to Fig.6, $P_{dri/joy}$, 0.64, and $P_{lef/joy}$, 0.23, are higher than others. Therefore, the robot went straight bending slightly the left when the robot is in the open space .

$$A = \begin{bmatrix} P_{Dri/joy} & P_{Dri/ang} & P_{Dri/sad} & P_{Dri/fea} \\ P_{Lef/joy} & P_{Lef/ang} & P_{Lef/sad} & P_{Lef/fea} \\ P_{Rig/joy} & P_{Rig/ang} & P_{Rig/sad} & P_{Rig/fea} \\ P_{Bac/joy} & P_{Bac/ang} & P_{Bac/sad} & P_{Bac/fea} \end{bmatrix} = \begin{bmatrix} 0.64 & 0.44 & 0.00 & 0.34 \\ 0.23 & 0.16 & 0.80 & 0.12 \\ 0.12 & 0.21 & 0.00 & 0.42 \\ 0.01 & 0.19 & 0.20 & 0.10 \end{bmatrix}$$

(8)

Next, the result using additional learning was compared with the result using only pre-learning. Fig.8 shows transitions of fitness values at two results in GA, and the vertical dashed line on 10 steps means generation which additional learning finishes. It shows that the fitness value of the system using additional learning is higher than the other.

Fig.9 shows histograms of the number of each node referenced as a winner node in each system, where white nodes were referenced frequently. In the case of only prior learning, areas of white nodes are concentrated, and it means that emotion-inducing factors changes little. On the other hand, in the case of additional learning, areas of white nodes are spread because that useless node areas of SOM decrease by additional learning. The performance of the decision making process of the robot is improved by the more
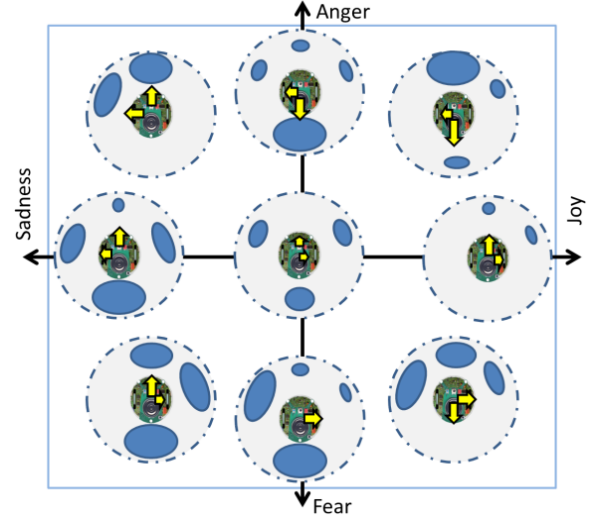

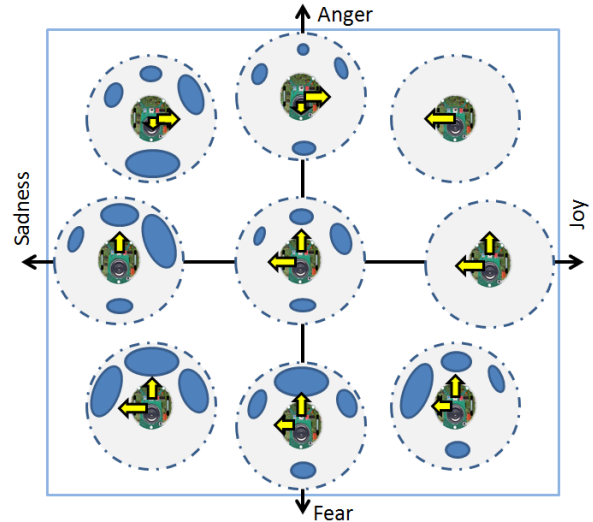Fig.5 The emotional map after pre-learning.

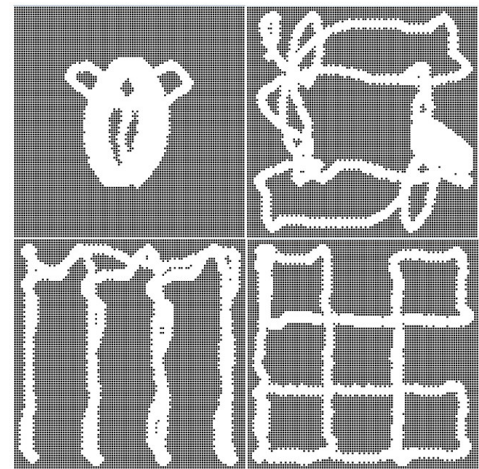Fig.6 The emotional map after additional learning.

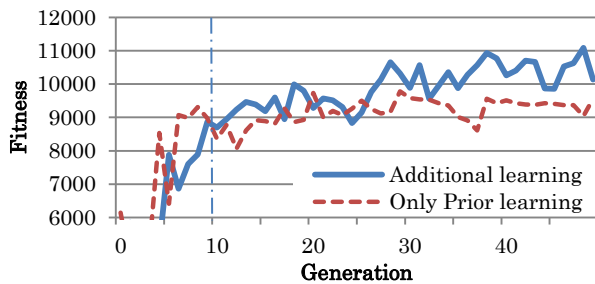Fig.7 Trajectories of the robot in four environments.
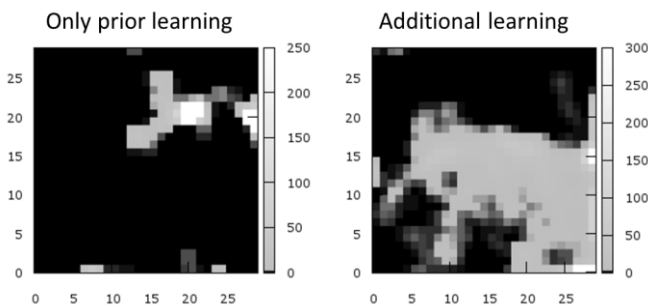
Fig.8 Transition of fitness values.

Fig.9 The histogram of the number of times each node is selected.

efficient association between emotion-inducing factors and input stimuli.

5. CONCLUSION

In our previous studies, a construction of emotions was processed only using learning-samples of input stimuli on offline. Therefore, it generated useless emotional areas in the SOM.

In this work, we proposed reconstruction of emotions by additional learning input stimuli obtained during the task. Then, more effective associations between input stimuli and parameters by additional learning using actual input stimuli were made.

The proposed method was evaluated by a simulation of a problem which maximizes areas a robot passed in a predefined time. The result showed that since an additional learning decreases useless emotional areas, it improved the decision making system of the robot. In addition, it was impossible to explain the decision-making policy of robots by analyzing the generated SOM and parameters of *A*.

For further works, it is necessary to improve our system for online learning of decision making. In current work, it is difficult to determine system parameters on a real robot because of using GA for parameter setting. Therefore, we will attempt to consider a method which optimizes parameter setting on online and robot tasks that make possible to evaluate the system on online instead of additive learning of using GA.

References

1. J.Moren, C.Balkenius, *A computational Model of Emotional Learning in the Amygdla* (Cybernetics and System32(6), 2001) , pp.611-636.
2. M. Obayashi, T. Takuno, T. Kuremoto, K. Kobayashi, *An Emotional Model Embedded Reinforcement learning System* (Proc. of IEEE International Conference on Systems, Man, and Cybernetics, 2012) , pp.1058-1063.
3. Even Daglarli, Hakan Tameltas, Murat Yesiloglu, *Behavioral task processing for cognitive robots using artificial emotions* (Neurocomputing 72, 2009), pp.2835-2844.
4. Christopher P. Lee-Johnson and Dale A. Carnegie, *Senior Member, IEEE, Mobile Robot Navigation Modulated by Artificial Emotions* (CYBERNETICS, VOL. 40, NO. 2, 2010), pp.469-480.
5. Sajal Chandra, Banik, Keigo Watanabe, Kiyotaka Izumi, *Improvement of group performance of job distributed mobile robots by an emotionally biased control system*, (Artificial Life Robotics(12), 2008), pp.245-249.
6. Shogo Watada, Masanao Obayashi, Takashi Kuremoto, Kunikazu Kobayashi, Shingo Mabu, *Behavior selection method of robots based on a Markovian emotional model (in Japanese)*. (IEEJ Transactions on Electronics, C, Vol.134, No.1, 2013)