Fuzzy Inference Model for Learning from Experiences and Its Application to Robot Navigation

Manabu Gouko, Yoshihiro Sugaya and Hirotomo Aso
Department of Electrical and Communication Engineering,
Graduate School of Engineering, Tohoku University
Aoba 6-6-05, Aramaki, Aoba-ku, Sendai, 980-8579, Japan.
gouko@aso.ecei.tohoku.ac.jp

Abstract

A fuzzy inference model for learning from experiences (FILE) is proposed. The model can learn from experience data obtained by trial-and-error of a task and it can stably learn from both experiences of success and failure of a trial. The learning of the model is executed after each of trial of the task. Hence, it is expected that the achievement rate increases with repetition of the trials, and that the model adapts to change of environment. In this paper, we confirm performance of the model by applying the model to a robot navigation task simulation and investigate the knowledge acquired by the learning.

1. Introduction

Over the past few decades, a considerable number of studies have been conducted on the intelligence system as typified by a robot. In recent years, the concern with the system that acquires knowledge by learning has been growing. In such a research, the system acquires knowledge based on interaction in the environment. If a designer easily interprets the knowledge which the system acquired, he or she can make use of the information to design the system, which facilitates the construction of a more flexible system. Additionally if prior knowledge that the designer has is fed to the system, the system can learn more effectively using it as a bias. That is to say, it brings many advantages that humans interpret the knowledge of the system.

Fuzzy inference model can use knowledge which is easily interpreted by humans. The model can conduct advanced inference like humans and has been applied to many intelligent systems. In the model, the knowledge is described in if-then rule form.

Recent studies on learning fuzzy inference model can be classified into three main groups according to learning method. In the first group, models learn by using supervised learning method[1, 6, 7]. This is an efficient method when input-output training data (teacher data) are available, but it is difficult to determine teacher data in a changing environment. Moreover, it cannot learn from an evaluation value which means the success or failure of a trial of a task; note that the evaluation is not teacher data. In the second group, models learn by using genetic algorithm[2, 4]. It is possible to learn from the evaluation value, but the method requires much calculation for learning. Therefore, it is difficult for the model to adapt to environment changes.

In the third group, models learn by using reinforcement learning (RL)[5, 8]. RL is on-line learning through interactions with a dynamic environment and it is possible to learn from an evaluation value (reward). Many conventional models using RL learn the optimum behavior by a searching the environment, but it requires large number of trials-and-errors[5]. On the other hand, there are some models to learn based on experiences[8]. The knowledge learned by these models is not always optimum, but the learning requires relatively small number of trials-and-errors. Several studies have been made on the fuzzy inference model that learns from experiences, however, many conventional models which learn from experiences learn only from experience of either success or failure. Additionally little attention has been given to learning in a changing environment and investigation into the knowledge acquired by learning.

We focus on a fuzzy inference model for learning from experiences (FILE) which is proposed by the authors[3]. FILE can stably learn from both experiences of success and failure. In the model, the learning executes after each end of a trial. Hence, it is expected that an achievement rate of a task increases with repetition of the trials. In this paper, we confirm performance of the model by applying the model to a robot navigation task simulation, and we investigate change of the knowledge caused by an environmental change.

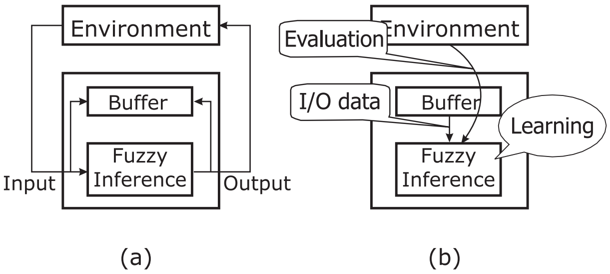

Figure 1. FILE. (a) Trial mode. (b) Learning mode.

2. Proposed model : FILE

Fig.1 shows the structure of FILE. FILE consists of a fuzzy inference unit and a buffer. FILE tries to achieve a task by using own knowledge. One trial is defined as a period from the start of a task until the end. FILE reasons and decides actions several times during a trial (Fig.1(a)). For example a sensor information as an input is given to the fuzzy inference unit, and the unit decides an output. Such input/output (I/O) data by the fuzzy inference unit are stored in the buffer during a trial. The stored data, which mean experiences, are exploited in learning mode with the evaluation value (Fig.1(b)). Evaluation value E, which is an evaluation for the trial, is fed to FILE at the end of trial. FILE updates own knowledge by learning. As a result, the model acquires the knowledge which is suitable for the environment.

2.1. Trial mode

In the trial mode, FILE tries to achieve a task by using own knowledge. The inference is executed at the fuzzy inference unit and that result is outputted. The unit has n fuzzy rules described in if-then form. $Rule^i$ represents ith fuzzy rule is written as follows:

$$Rule^i : \text{if } x_1 \text{ is } A_{i1} \cdots \text{ and } x_j \text{ is } A_{ij} \text{ and} \cdots \cdots \text{and } x_m \text{ is } A_{im} \text{ then } y = b_i \quad (i = 1, 2, \cdots, n), \tag{1}$$

where $x_1, x_2, \cdots, x_m$ are input variables and y is an output variable. $A_{i1}, A_{i2}, \cdots, A_{im}$ are linguistic labels which represent fuzzy sets and b_i is a constant output value. y^* is a result of inference and calculated by:

$$\mu_i = \textstyle\prod_{j=1}^{m} A_{ij}(x_j) \quad (i = 1, 2, \cdots, n), \tag{2}$$

$$y^* = \textstyle\sum_{i=1}^{n} \mu_i b_i / \sum_{i=1}^{n} \mu_i, \tag{3}$$

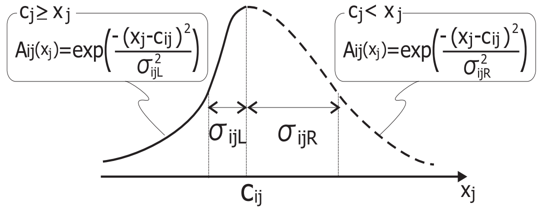

Figure 2. Membership function.

where μ_i is a firing strength of the ith rule and $A_{ij}(x_j)$ is a membership function of x_j (fig.2) and calculated by:

$$A_{ij}(x_j) = \exp\Big(-(x_j - c_{ij})^2/\sigma_{ijk}^2\Big), \quad (i = 1, 2, \cdots, n, j = 1, 2, \cdots, m), \tag{4}$$

where c_{ij} and σ_{ijk} are parameters to give center and width, respectively. The membership function used in the model has asymmetric widths σ_{ijL} and σ_{ijR} in the left and right. k is an index to represent which side of the function. If $c_{ij} \geq x_j$, we set $k = L$, otherwise we set $k = R$. I/O data ($x_1, x_2, \cdots, x_m$ and y^*) of the fuzzy inference unit are stored in the buffer during a trial, which are called learning data.

2.2. Learning mode

In the learning mode, FILE learns from the learning data and the evaluation value (Fig.1(b)). Each parameter of the membership function is updated by learning. The evaluation value $E(-1 \leq E \leq 1)$ has a positive value when a trial succeeded. When a trial failed, E has a negative value. When E is positive, the model learns to reinforce I/O relationships of the learning data. When E is negative, the model learns the repulsive relationships. The learning process is given below.

(1) A set of learning data $(\boldsymbol{x}^l, y^l) = (x_1^l, x_2^l, \cdots, x_m^l, y^l)$ is picked from the buffer.

(2) μ_i^l which is the firing strength of ith rule for $\boldsymbol{x}^l$ is calculated from eq.(2) and i_s is calculated by $i_s = \text{augmax}_i\ \mu_i^l$.

(3) $c_{i_s j}$ and b_{i_s} are updated as follows:

$$c_{i_s j}^{new} = c_{i_s j}^{old} + \alpha E A_{i_s j}(x_j^l)(x_j^l - c_{i_s j}^{old}) \quad \text{if } E \geq 0 \quad (j = 1, 2, \cdots, m), \tag{5}$$

$$b_{i_s}^{new} = b_{i_s}^{old} + \beta \mu_{i_s}^l E(y^l - b_{i_s}^{old}) \quad \text{if } E \geq 0, \tag{6}$$

where α and β are learning rates.

(4) The widths of the membership functions are updated as follows:

$$\sigma_{ijp}^{new} = \begin{cases} \sigma_{ijp}^{old} - \gamma_1 E A_{i_s j}(x_j^l)\mu_i^l & \text{if } E \geq 0, \\ \sigma_{ijp}^{old} - \gamma_2 E(1 - A_{i_s j}(x_j^l))(1 - A_{ij}(x_j^l)) & \\ & \text{if } E < 0, \end{cases}$$
$$(i = 1, 2, \ldots, n, \; j = 1, 2, \ldots, m), \quad (7)$$

where γ_1, γ_2 are learning rates. The index p is determined as the following:

$$p = \begin{cases} L & \text{if } c_{i_s j} < x_j^l \leq c_{ij} \;\; \& \;\; i \neq i_s \\ & \text{or } \;\; x_j^l \leq c_{ij} \leq c_{i_s j} \;\; \& \;\; i \neq i_s, \\ R & \text{if } c_{ij} \leq x_j^l \leq c_{i_s j} \;\; \& \;\; i \neq i_s \\ & \text{or } \;\; c_{i_s j} \leq c_{ij} \leq x_j^l \;\; \& \;\; i \neq i_s, \\ \phi & \text{otherwise.} \end{cases} \quad (8)$$

If $p = \phi$, neither width is updated. The width of the rule except for the i_sth rule is decreased corresponding to μ_i^l, when E has a positive value. On the whole, the firing strength μ_{i_s} is increased. On the other hand, the firing strength μ_{i_s} is decreased when E has a negative value.

(5) After the above operation is completed, the data used for the learning is deleted from the buffer, and then the procedure (1) - (4) is repeated. The learning mode is finished when there is no data in the buffer.

From the viewpoint of the interpreting of the rule by humans, it is more desirable that the model has few rules. However, learning often becomes unstable by an influence of dispersion of the learning data when a model has a few rules. FILE can execute stably learning even if there are few rules[3]. In the eq.(5), $A_{i_s j}(x_j^l)$ prevents the learning from becoming unstable due to the dispersion of the learning data. A learning data which is apart from the center of membership function $A_{i_s j}(x_j^l)$ exerts small influence on the learning, thus the learning becomes stable. In the eq.(7), the role of $A_{i_s j}(x_j^l)$, $1 - A_{i_s j}(x_j^l)$ and $1 - A_{ij}(x_j^l)$ is stabilization of the learning, as we said earlier[3].

Naturally, the model has the limits of the adaptability by using only a few rules without addition of the new rules. Therefore, this model should be improved in order that it can add a rule, but in this paper, we are not concerned with that.

3. Experimental results and discussion

We apply FILE to a robot navigation task simulation[4] and confirm the performance of the model. Additionally, we investigate change of the knowledge caused by an environmental change. The robot navigation task aims at moving of a mobile robot from start to goal without any collisions with a wall. Fig.3(a) shows the top view of the robot. The diameter of the robot is 20cm. The robot has a sensor which can detect obstacles. The sensing area is ± 30 degrees to the traveling direction and the maximum depth of the sensing area is 200cm. The robot moves based on the information which is observed from the environment. If there is an obstacle in the area, the model executes the inference and decides the traveling direction. When there is no obstacle in the area, the robot goes straight without the inference. One step is defined as that the robot moves by an observation. The robot moves forward L[cm] every step. Fig.3(b) shows input variables, x_1 is the angle between the traveling direction and the nearest detected wall. x_2 is the shortest distance between the robot and the detected wall. x_3 and x_4 are the distances between the robot and the walls detected at the right and left edge of the sensing area. The fuzzy inference unit infers a steering angle from the input variables and puts out it. The maximum steering angle is ± 10 degrees. In the fuzzy inference unit, each input variable and output value are normalized.

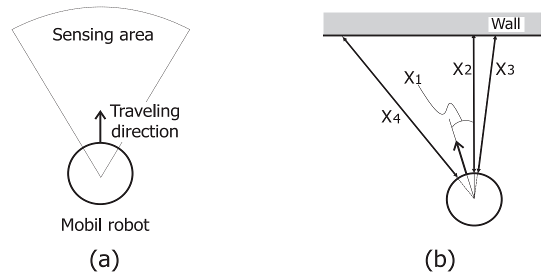

Figure 3. (a) Top view of a mobile robot. (b) Input variables. x_1 is the angle between the traveling direction and the wall. x_2, x_3 and x_4 are the distances between the robot and the wall, respectively.

The experiment is carried out by using the course as shown in fig.4(a). The robot moves to the goal from the start by repeating the inference. One trial is defined as that the robot reaches the goal or collides with the wall after start. In fig.4(a), a broken line represents the range of start positions of the robot. The start position is chosen at random within the range at every trial. E is fed to the model at the end of every trial. When the robot reaches the goal, we set $E = 1$. When the robot collides with the wall, we set $E = -1$. We confirm performance of learning by using the same navigation task without learning (test task). In the experiment, whenever a learning mode finished, the trial of the test task is executed 100 times. When the trial of the test task finished, the robot starts a new trial for learning mode.

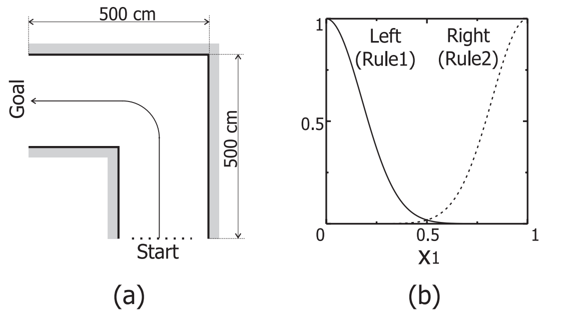

Figure 4. (a) Course of robot navigation. (b) Membership functions of x_1 of each rule which are given as the prior knowledge.

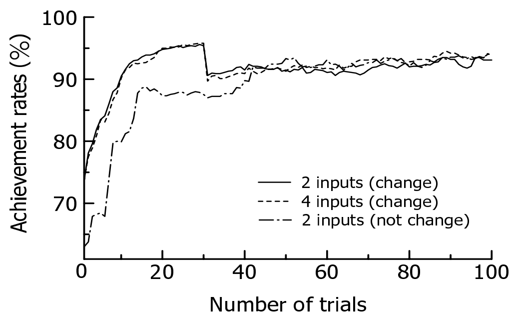

Figure 5. Relation between the number of trials and average of achievement rates.

The achievement rate of the test task is expected to increase as learning time increases.

In the experiment, parameters are $\alpha = 0.004, \beta = 0.0001, \gamma_1 = 110, \gamma_2 = 0.004$. We give two simple rules as a prior knowledge to the model. When there is the nearest wall in the left of the robot ($x_1 < 0.5$), the robot turns the steering wheel to the right by $Rule^1$. When there is the nearest wall in the right of the robot ($x_1 > 0.5$), the robot turns the steering wheel to the left by $Rule^2$. Fig.4(b) shows the membership functions of x_1 of each rule which are given as the prior knowledge. FILE cannot execute the learning, if a designer cannot give any prior knowledge for it. We assume that FILE is used under the environment where a designer can give it some knowledge. In both rules, the same membership function was given for x_2, x_3 and x_4. The center and width of the membership function of x_2 are 0.7 and 0.5. The center and width of the membership functions of x_3 and x_4 are both 0.5.

3.1. Change of the speed

In this experiment, the speed of the robot is changed from $L = 10$ to 25. We set $L = 10$ until the 30th trial, and we set $L = 25$ afterwards. The experiment is carried out by using a 2-input model and a 4-input model. The 2-input model infers only from input variables x_1, x_2, and the 4-input model infers from input variables x_1, x_2, x_3, x_4.

Fig.5 is the result of the experiment, and indicates the relation between the number of trials and average of achievement rates. A solid line indicates the 2-input model, and broken line indicates the 4-input model. A chain dash line indicates the case using the 2-input model although the speed is fixed to 25 from the beginning of the task. The achievement rates of 2-input and 4-input models increase to 95.4% and 95.8% until speed is changed. Each rate in the 31st trial drops to 90.6% and 89.7%, however, until the 100th trial, each rate is increased to 93.1% and 94.1% by the learning after the change of the speed.

The acquired knowledge is investigated by observing the membership function (fig.6). Firstly, we investigate the change of membership function of x_1(fig.6(a)). Before the learning ($t = 0$), the cross point of the functions is 0.5. In the 30th and 100th trial ($t = 30, 100$), the cross point moves to 0.39 and 0.37, respectively. Even if x_1 is a little smaller than 0.5, the firing strength of $Rule^2$ is higher than $Rule^1$, therefore the robot steers to the left. This means that the knowledge that tends to turn to the left was acquired by repetition of the trial and the learning. Such tendency becomes stronger as the speed increases. It was found from the result that the robot adapts to the increase of the speed by changing own knowledge.

We also investigate the change of the membership functions of x_3 and x_4 of the 4-input model. Fig.6(b) shows the membership functions of x_3 and x_4 in $Rule^2$ at the 100th trial. $Rule^2$ means turning the steering wheel to the left. The model learned the knowledge that the robot turns to the left, if x_4 is larger than x_3. This knowledge is suitable, because this relation between x_3 and x_4 is observed when the wall is on the right of the robot. As a result, FILE learned the suitable membership function of x_3 and x_4 in $Rule^2$.

Furthermore, we confirm that the knowledge which uses x_3 and x_4 changes to the effective knowledge in each environment by learning. Fig.7 shows the trajectories of the robot of 4-input model in the test task. These trajectories were obtained by inference that used only x_3 and x_4. The robot starts from the same position and direction in each trajectory. In fig.7, the small circle represents the robot. Before the learning ($t = 0$) the robot goes straight on the wall and collides. After 30 trials ($t = 30$), the robot can reach the goal, however, the robot becomes unable to reach the goal again owing to increase in speed ($t = 31$). Finally the robot reaches the goal by the learning in the changed environment ($t = 100$). The result shows that the knowledge

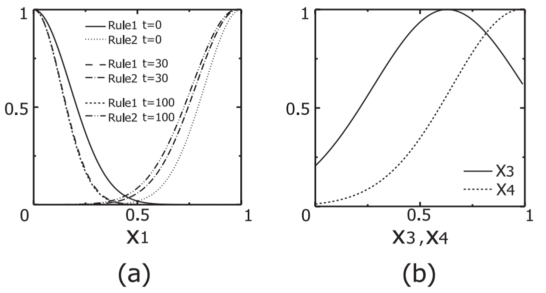

Figure 6. (a) Membership functions of x_1 of each rule. (b) Membership functions of x_3 and x_4 in $Rule^2$.

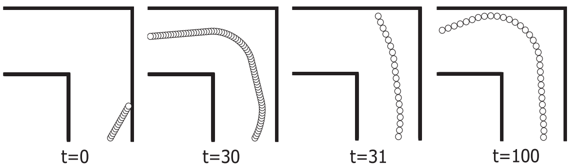

Figure 7. Trajectories of the robot.

which uses x_3 and x_4 changes to the effective knowledge for each environment by learning.

3.2. Change of the course

By another experiment, the course is changed during the experiment. Until the 30th trial, left turn course as shown in fig.4(a) is used, and right turn course is used after that. Only a direction of a corner varies and width and length are the same. The speed of the robot is constant ($L = 10$). Fig.8 is the result of the experiment. A solid line indicates the case using 2-input model, and broken line indicates the case using 4-input model. The achievement rates of 2-input and 4-input models increase to 93.6% and 94.3% until the course is changed. Each rate in the 31st trial drops to 34.6% and 31.3%. Until the 100th trial, each rate is increased to 91.6% and 89.3% by the learning after the change of the course.

Fig.9(a) shows the membership functions of x_1. In the 30th and 100th trial ($t = 30, 100$), the cross point is 0.41 and 0.54. It can be said that the robot acquired the knowledge that tends to turn to the direction of each course.

Fig.9(b) shows the membership functions of x_3 and x_4 in $Rule^1$ at the 100th trial. In this figure, relation between x_3 and x_4 is opposite to fig.6(b). $Rule^1$ means turning the steering wheel to the right. It can be said that the model learned the suitable membership functions of x_3 and x_4 in $Rule^1$.

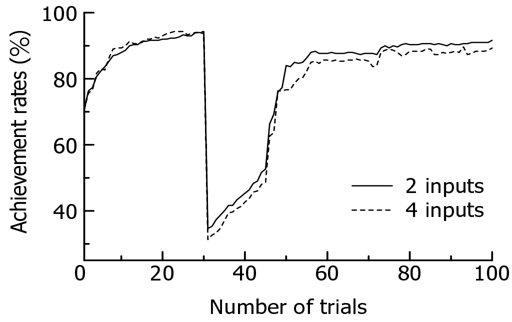

Figure 8. Relation between the number of trials and average of achievement rates.

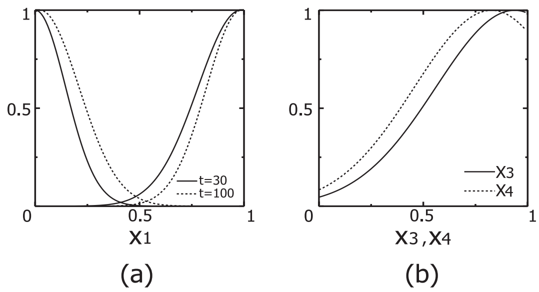

Figure 9. (a) Membership functions of x_1 of each rule. (b) Membership functions of x_3 and x_4 in $Rule^1$.

The difference in the center of each membership function in fig.9(b) is small in comparison with the previous experiment(fig.6(b)). This difference results from the difference in the speed of the robot. Fig.10 shows behavior of the robot which avoids colliding. In Fig.10(a) and (b), the start positions of the robots are the same. The case that the robot moves forward at low speed is shown in fig.10(a). The robot can avoid colliding without approaching the wall very much. The case that the robot moves forward at high speed is shown in fig.10(b). The robot goes to the wall closer in comparison with fig.10(a), because the speed of the robot is fast. Fig.10 shows that though the robots turn in the same way to the left, observed x_3 is smaller when the speed is faster. It can be said that the difference of the center of the membership functions in each speed results from such difference of the observed variables.

4. Conclusion

In this study, we focus on a fuzzy inference model for learning from experiences (FILE) which is proposed by the

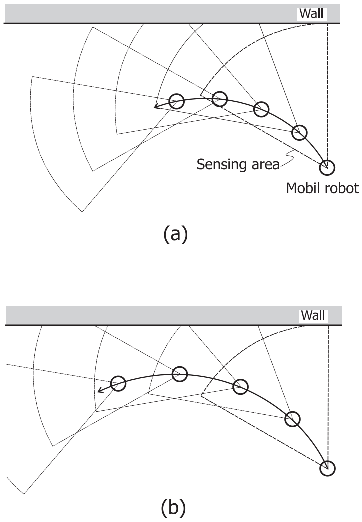

Figure 10. Behavior of the robot which avoids colliding. (a) Low speed. (b) High speed.

authors[3]. FILE can stably learn from both experiences of success and failure. In the model, the learning executes after each end of a trial. In this paper, we applied the model to an environment which changes and confirmed performance of the model by using a robot navigation task simulation. Additionally we investigated change of the knowledge caused by an environmental change. In the experiment, we showed that the robot adapts to a changing environment and acquires knowledge which is suitable for each environment by the learning.

A further direction of this study will be that we add mechanism which generate new rule to the model and adapt the model to more complicated tasks.

Acknowledgment

This work is supported in part by the 21st Century COE Program (Grant No.14170017).

References

[1] M. Er, T. P. Tan, and S. Y. Loh. Control of a mobile robot using generalized dynamic fuzzy neural networks. *Microprocessors and Microsystems*, 28:491–498, 2004.

[2] T. Furuhashi, K. Nakaoka, K. Morikawa, H. Maeda, and Y. Uchikawa. A study on knowledge finding using fuzzy classifier system. *Journal of Japan Society for Fuzzy Theory and Systems*, 7(4):839–848, 1995. in Japanese.

[3] M. Gouko, Y. Sugaya, and H. Aso. Learning fuzzy inference model and acquisition of knowledge. *Proceedings of the 3rd student-organizing international mini-conference on information electronics system (SOIM-COE05)*, pages 285–288, 2005.

[4] H. Itani and T. Furuhashi. A study on teaching information understanding by autonomous mobile robot. *Trans. of SICE*, 38(11):966–973, 2002. in Japanese.

[5] L. Jouffe. Fuzzy inference system learning by reinforcement method. *IEEE Trans. Syst., Man, Cybern., part C*, 28(3):338–355, 1998.

[6] G. Leng, T. M. McGinnity, and G. Prasad. An approach for on-line extraction of fuzzy rules using a self-organising fuzzy neural network. *Fuzzy Sets and Systems*, 150:211–243, 2005.

[7] T. Nishina and M. Hagiwara. Fuzzy inference neural network. *Neurocomputing*, 14:223–239, 1997.

[8] T. Takahama, S. Sakai, H. Ogura, and M. Nakamura. Learning fuzzy rules for bang-bang control by reinforcement learning method. *Journal of Japan Society for Fuzzy Theory and Systems*, 8(1):115–122, 1996. in Japanese.