# Learning fuzzy inference model and investigation of acquired knowledge

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.
Phone:+81+22-795-7088, Fax:+81-22-263-9419,
Email:gouko@aso.ecei.tohoku.ac.jp

*Abstract*– **Fuzzy inference models can conduct advanced inference using knowledge which is easily understood by humans. In this paper, we propose a leaning fuzzy inference model. The model can learn with experience data obtained by trial-and-error of a task. The learning of the model is executed after each 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. We confirm the performance of the model by experiences and the validity of learning by investigation of the knowledge acquired by the learning.**

## I. Introduction

Over the past few decades, a considerable number of studies have been conducted on the intelligence system as typified by a robot. The systems are expected to behave autonomously using their knowledge in an environment.

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. In this paper, we focus on the learning fuzzy inference model.

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]. 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 based on 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, 3]. It is possible to learn with 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)[4, 5]. RL is on-line learning through interactions with a dynamic environment and it is possible to learn based on 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[4]. On the other hand, there are some models to learn based on experiences[5]. 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 based on experiences, however, little attention has been given to investigation into the knowledge acquired by learning.

In this study we propose a new fuzzy inference model which learns based on experiences and investigate the knowledge acquired by learning. In proposed 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. Many conventional models which learn based on experiences learn only from either experience of success or failure, but the proposed model can learn from both experiences. In the model, the knowledge is described in if-then rule form and it is easy to understand by humans. We confirm performance of the model by using a robot navigation task simulation. We investigate the knowledge acquired by the learning.
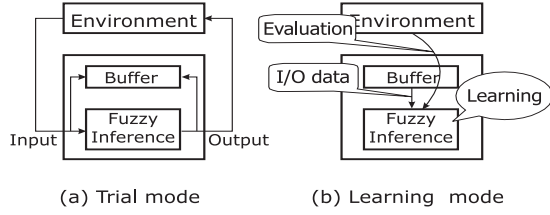
Figure 1. Proposed model.



Figure 2. Membership function.

## II. The Proposed Model

Fig.1 shows the structure of the proposed model. The model consists of a fuzzy inference unit and a buffer. The model tries to achieve a task by using own knowledge. One trial is defined as a period from the start of the task until the end. The model reasons and decides actions several times during a trial(Fig.1(a)). Input/output (I/O) data of the fuzzy inference unit are stored in the buffer during a trial, and 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 the model at the end of trial. The model updates own knowledge by learning. As a result, the model acquires the knowledge which is suitable for the environment. Regardless of the success or failure of the trial, learning is executed sequentially at every trial.

### A. Fuzzy inference unit

The unit has $n$ fuzzy rules described in if-then form. $Rule^i$ represents $i$th fuzzy rule is written as follows:

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

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

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

$$y^* = \sum_{i=1}^{n} \mu_i y_i / \sum_{i=1}^{n} \mu_i, \quad (3)$$

where $\mu_i$ is a firing strength of the $i$th rule and $A_{ij}(x_j)$ is a membership function of $x_j$ shown as fig.2. In fig.2, $c_{ij}$ and $\sigma_{ijk}$ are center and width of the membership function $A_{ij}(x_j)$, respectively. The membership function used in the model has two different widths in the left and right. $k$ is an index to represent which width to use. 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 and called learning data.
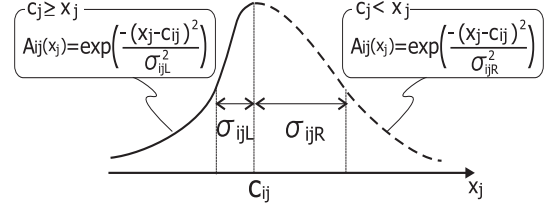
### B. Learning mode

In the learning mode, the model learns with 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 positive value when a trial succeeded. When a trial failed, $E$ has negative value. When $E$ is positive, the model learns to reinforce I/O relationships of the learning data, that is success learning. When $E$ is negative value, the model learns the repulsive relationships, that is failure learning.

We will now discuss a problem which is caused by using the learning scheme. Generally, a tactic to achieve a task is not only one. In the same task, inputs to the same output are dispersed, because different tactics choose the same output for different inputs. The success learning occasionally becomes unstable by such dispersion, because the rule which learns from the data obtained by a tactic is influenced by the learning from the data obtained by other tactic. It also has to be considered that the learning becomes unstable in the failure leaning. The cause of the instability is that all of the learning data obtained by the failure trial is not necessarily an improper relationship.

If a learning rate is very reduced, this problem in the success learning can be avoided, however, that causes decrease of the learning speed. In the failure leaning, the schemes that use information about a time have been proposed [5, 6]. In those schemes, the I/O data is stored with the time stamp. The data which is observed in the far past from the time when a task failed is considered as correct data and used to learn. However, this scheme depends on a task, because the inference just before the failure is not necessarily causes of the failure in all task. We cope with the problem in other scheme.

We will take a close look at a change of the fuzzy rule. The rule represents the I/O relationships by an if-part and a then-part. Each part is changed by the learning. The if-part shows an input area which is expressed by the membership functions. The cause of the instability in the success learning is that the changing of the input area is sensitive to the dispersion of the learning data. The cause of the instability in the failure learning is that the proper relationship represented

by the rule is changed for the worse by learning. We expect that learning data which represents proper relationships exists near the center of input area. Therefore, we propose the learning scheme which refers the distance between the center of input area and learning data. In the scheme, the learning data which is near the center of input area exerts big influence on the learning, in success learning. On the other hand, the data which is apart from the center exerts a small influence. In failure learning, the learning data which is near the center of input area exerts small influence. On the other hand, the data which is apart from the input area exerts a big influence. It is expected to prevent that the learning becomes unstable by using the scheme.

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

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

$$c_{i_s j}^{new} = c_{i_s j}^{old} + \alpha E A_{i_s j}(x_j^l)(x_j^l - c_{i_s j}^{old}) \quad \text{if } E \geq 0$$
$$(j = 1, 2, \cdots, m), \quad (4)$$

$$b_{i_s}^{new} = b_{i_s}^{old} + \beta \mu_{i_s}^l E(y^l - b_{i_s}^{old}) \quad \text{if } E \geq 0, \quad (5)$$

where $\alpha$ and $\beta$ are learning rates. In the eq.(4), $A_{i_s j}(x_j^l)$, which is a membership function, indicates the influence of the distance between the center of input area and the learning data. Consequently, as the distance becomes small, the influence on the learning grows.

(4) The width is updated as follows:

$$\sigma_{ijp}^{new} = \begin{cases} \sigma_{ijp}^{old} - \gamma_1 E A_{i_s j}(x_j^l)\mu_i^l & \text{if } E \geq 0, \\ \sigma_{ijp}^{old} - \gamma_2 E(1 - A_{i_s j}(x_j^l))(1 - A_{ij}(x_j^l)) & \\ & \text{if } E < 0, \\ (i = 1, 2, \ldots, n, \ j = 1, 2, \ldots, m), & (6) \end{cases}$$

where $\gamma_1, \gamma_2$ are learning rates. A parameter $p$ represents which width of left-right is updated and it is shown in the following:

$$p = \begin{cases} L & \text{if } c_{i_s j} < x_j^l \leq c_{ij} \ \& \ i \neq i_s \\ & \text{or } x_j^l \leq c_{ij} \leq c_{i_s j} \ \& \ i \neq i_s, \\ R & \text{if } c_{ij} \leq x_j^l \leq c_{i_s j} \ \& \ i \neq i_s \\ & \text{or } c_{i_s j} \leq c_{ij} \leq x_j^l \ \& \ i \neq i_s, \\ \phi & \text{otherwise}, \end{cases} \quad (7)$$

if $p = \phi$, neither width is updated. The width of the rule except for the $i_s$th rule is decreased corresponding to $\mu_i^l$, when $E$ has positive value. On the whole, the firing strength $\mu_{i_s}^l$ is increased in the
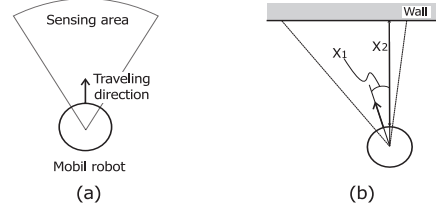


Figure 3. (a) Mobile robot. (b) Input variables.

success learning. When $E$ has negative value, the width of the rule except for the $i_s$th rule is increased by eq.(6). On the whole, the firing strength $\mu_{i_s}^l$ is decreased in the failure learning. In the eq.(6), $A_{i_s j}(x_j^l)$, $1 - A_{i_s j}(x_j^l)$ and $1 - A_{ij}(x_j^l)$ indicate the influence of the distance.

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

## III. EXPERIMENTAL RESULTS AND DISCUSSION

We carry out a mobile robot simulation[3], and we confirm the performance of the model and the validity of the learning by investigation of the knowledge acquired by the learning. 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. A diameter of the robot is 20cm. The robot has a sensor which can detect an obstacle. The sensing area is ±30 degree to the traveling direction and 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 during the trial. Fig.3(b) shows input variables, $x_1$ is the angle formed by the nearest detected wall and the traveling direction. The $x_2$ is the shortest distance between the robot and the detected wall. The fuzzy inference unit infers a steering angle from the input variables and puts out it. The maximum steering angle is ± 10 degree. In the fuzzy inference unit, each input variable and output value are normalized.

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 started. In fig.4(a), a broken line represents the range of start positions of the robot. The start position is decided at random within the range at every trial. $E$ is fed to the model at the end of every trial. When the
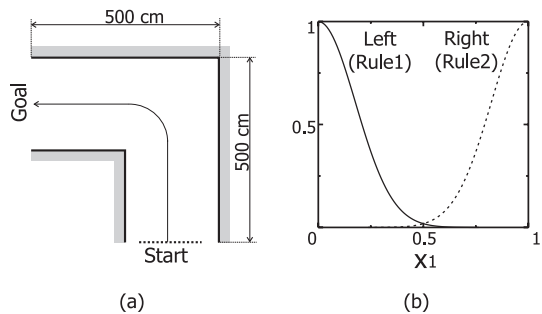
Figure 4: (a) Course of robot navigation. (b) Membership function.



Figure 5: (a)Achievement rates. (b)Membership function.

robot reaches the goal, we set $E = 1$. When the robot collides with the wall, we set $E = -1$. We confirm the performance of learning by using the same navigation task without learning (test task). In the experiment, whenever the 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 in learning mode. The achievement rate of the test task is expected to increase as learning time increases.

In the experiment, parameters are $\alpha = 0.004$, $\beta = 0.0001$, $\gamma_1 = 110$, $\gamma_2 = 0.004$ and $L = 10$. 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 the rule1. 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 the rule2. Fig.4(b) shows the membership functions of $x_1$ of each rule which are given as the prior knowledge. In both rules, same membership function was given for $x_2$. The center and of the membership function of $x_2$ are 0.75, respectively.

Fig.5(a) 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 case using the proposed learning method, and broken line indicates the case using the learning method without considering the distance between the input area and the learning data. In the learning method without considering the distance, we used the learning laws eq.(4) and (6) where $A_{i_s j}(x_j^l)$, $1 - A_{i_s j}(x_j^l)$ and $1 - A_{ij}(x_j^l)$ are replaced with 1. In the case using proposed learning method, achievement rate is increased stably to 95.5% from 74%, and kept. In the other case, achievement rate is unstable as shown by the error bar which represents dispersion. It was found from the result that the model can stably learn by using the proposed learning method.

The acquired knowledge is investigated by observing the membership function in after 60th trial (fig.5(b)). We investigate the change of membership function of $x_1$. In fig.5(b), before the learning ($t = 0$), the point of intersection of each membership function is 0.5. After 60th trial ($t = 60$), the point changes to 0.39. Therefore,
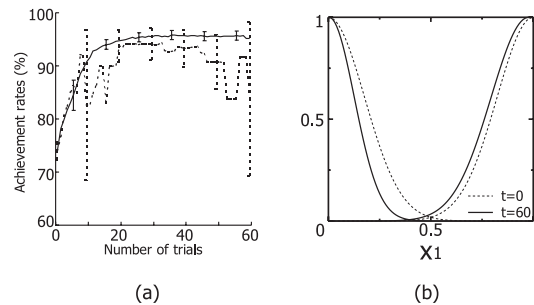
even if $x_1$ is a little smaller than 0.5, the firing strength of rule 2 is higher than rule1, and consequently the robot steers to the left. As a result, the knowledge that tends to turn to the left was acquired by repetition of the trial and the learning.

In this experiment, we also confirmed that the achievement rate is raised to 1.76% of average by the failure learning at a time.

## IV. Conclusion

In this paper, we propose the fuzzy inference model that learns based on experiences and investigate the knowledge acquired by the learning. The model can learn from both experiences of success and failure. In the model, the knowledge is described in if-then rule form and it is easy to interpret by humans. We confirmed performance of the model by using the robot navigation task simulation, and investigated the knowledge acquired by the learning. A further direction of this study will be that we adapt the model to an environment which changes.

## References

[1] Nishina, T. and Hagiwara, M.: Fuzzy inference neural network, *Neurocomputing*, Vol. 14, pp. 223–239 (1997).

[2] Furuhashi, T., Nakaoka, K., Morikawa, K., Maeda, H. and Uchikawa, Y.: A Study on Knowledge Finding Using Fuzzy Classifier System, *Journal of Japan Society for Fuzzy Theory and Systems*, Vol. 7, No. 4, pp. 839–848 (1995). in Japanese.

[3] Itani, H. and Furuhashi, T.: A Study on Teaching Information Understanding by Autonomous Mobile Robot, *Trans. of SICE*, Vol. 38, No. 11, pp. 966–973 (2002). in Japanese.

[4] Jouffe, L.: Fuzzy Inference System Learning by Reinforcement Method, *IEEE Trans. Syst., Man, Cybern., part C*, Vol. 28, No. 3, pp. 338–355 (1998).

[5] Takahama, T., Sakai, S., Ogura, H. and NAKAMURA, M.: Learning Fuzzy Rules for Bang-Bang Control by Reinforcement Learning Method, *Journal of Japan Society for Fuzzy Theory and Systems*, Vol. 8, No. 1, pp. 115–122 (1996). in Japanese.

[6] Unemi, T.: Learning not to Fail by Instance-based Reinforcement Learning Method, *Journal of Japanese Sociaty for Artificial Intelligence*, Vol. 7, No. 6, pp. 1001–1008 (1992). in Japanese.