1. INTRODUCTION

Robots are becoming increasingly familiar. A robot must behave with sensitivity to communicate smoothly with a person. This sensibility is fundamental to communication. Even a person who does not understand a language can feel emotion. However, it is difficult to achieve natural communication even though the algorithm for emotion generation is simple and suitable for more conventional robot communication. Multiple robots are expected to coexist in the future because the number of robots is increasing. Therefore, human-like expressiveness is important in not only robot–human interactions but also robot–robot interactions. In group communication, robots are necessary social behavior in order to communicate naturally [1]. In robot–robot communication, robots are able to communicate by telepathy. However, in human communities, it is important for robot–robot communication to be intuitively understood by humans [2, 3].

In a study of a robot with the emotion function, Harata proposed an emotion growth model of a robot based on affective psychology [4]. In addition, Fukumoto proposed parent–child robots that grow through the interaction of robots using this growth model [5]. Fukumoto also aimed at community formation using parent–child robots, in which the number of parents corresponded to that of the children [6]. However, these researchers did not consider group communication because the model of the previous study assumes one-to-one communication with a robot.

A robot that is not directly related to communication in a group is in the situation of a bystander (hereafter called the “bystander robot”). The study of bystander robots addressed mainstream physical movements such as the appropriate body placement of the robot considering the distance between others [7, 8], changing the robot’s gaze target as the speaker changes [9], and the presence of a bystander robot [10]. However, there have been no studies addressing the expressiveness of a bystander robot. Smooth group communication requires expressiveness in bystander robots.

Consequently, in this study, we propose a sympathy expression model to enable a bystander robot to acquire cooperative expressiveness by learning the emotional display of others. The object of this model is expressiveness of a robot related to communication as a bystander rather than as a speaker or addressee. We aim to construct social communication from the interaction of two robots using a neural network (NN)-based emotion generation model and build a cooperative expressiveness model for the bystander robot.

2. PSYCHOLOGICAL MODEL

We construct the internal system of the robot to communicate using the emotion function. A communicating robot must act similar to a human; this requires appropriate emotion functions. Therefore, we use affective and social psychology models to deal with emotions here.
2.1 Emotion dimension theory

In emotion dimension theory, emotions are expressed as vectors along a dimension without being identified discretely. For example, the emotion “anger” labels the state of the vector as “anger,” not as a systematized state. Therefore, we recognize “anger” as consisting of emotions different from “fear”, for example, by using situational context.

The Russell circle model is one of the foundational expressions of emotion dimension theory [11]. Russell proposed a model in which all emotions are placed in a circle on a plane along the two dimensions of “pleasure-misery” and “arousal-sleepiness,” as shown in Figure 1. Emotions locate a synonym at a neighboring position on the circle and an antonym at the opposite position. Each emotion is displayed with the direction and size of a vector on the two-dimensional coordinate axes, and the difference in direction of the vector of each emotion expresses a correlation coefficient. In addition, the strength of emotions is shown by the size of the vector from the origin.

2.2 Balance theory

Balance theory, proposed by psychologist Fritz Heider [12], assumes that the feelings among three people tend to enter a balanced state through interaction. In this model, P is oneself, O is the other, and X is the object. We label “the impression for O of P,” “the impression for X of P” and “the impression for X of O judging from P” with a positive (+) or negative (−) impression. This theory shows that the product of these three impressions tends to become the balance state of positive (+). Figure 2 shows illustrations of an imbalance state and a balance state of balance theory. Either impression changes to the balance state from the imbalance state by changing “positive (+) to negative (−)” or “negative (−) to positive (+).”

2.3 The theory of moral sentiments

“The Theory of Moral Sentiments,” proposed by Adam Smith, states that society is organized in a particular order for “sympathy,” and lets you choose a moral judgment because “the fair observer” in oneself predicts the reactions of others [13]. To make a moral judgment, it is necessary to have a function for understanding objective emotions that are different from one’s own emotions from the viewpoint of a third party. Humans have an internal “fair observer” to realize this function. In other words, the emotion of the fair observer is “moral emotion.” Moral emotion objectively captures the emotions that we generally expect based on past experience.

In addition, Smith explains sympathy as follows: “When the sympathy feelings of the observer completely accord with the primary emotions of the person concerned, the observer necessarily thinks the primary emotions to be appropriate for this person concerned.” Sympathy emotions are formed when the observer perceives agreement between his or her own emotions and those of the participant. In other words, sympathy is the state in which the expressed emotions of both the person concerned and an observer match. The observer approves the person concerned by creating a state of sympathy. Therefore, the observer predicts the feelings of the person by consulting with the inner fair observer. This fair observer is formed by experiencing the emotional display of oneself and others. We can thus form a fairer observer if we experience the emotional display of more people. Conversely, the formed fair observer becomes a partial emotional display as a result of experiencing the emotional display of only a few others. As a result, the observer with little experience might exhibit egocentric emotional displays. In this study, we treat this observer as “the bystander.”
3. SYMPATHY EXPRESSION MODEL OF THE BYSTANDER ROBOT

In this study, we propose a sympathy expression model of a bystander robot with three robots, including a bystander, used as the minimum communication flow. In the proposed model, the bystander robot generates emotions in two networks, emotion and moral emotion generation. The bystander robot judges sympathy with the participant based on the cosine similarity of two emotion vectors on the Russell circle model and decides which emotion, if any, it should express.

3.1 Summary of the proposed model

The existing robot emotion generation model is designed assuming one-to-one communication with the robot with no consideration of group communication. A robot in such a situation must return some kind of reaction to be used as an input by the other robot. However, a robot in the situation of a bystander not fully involved in the communication between a speaker and an addressee would not return a reaction for all input in group communication.

Our proposed model can express precise sympathy by a bystander robot that learns general emotions in group communication. We prepare for a bystander robot and two other robots, which communicate regardless of a bystander robot and take the role of speaker and addressee. A bystander robot generates emotions by receiving input resembling that received by the addressee robot in the conversation.

Each of three robots has an NN of three layers of perceptrons as an emotion generation function. This NN outputs emotions as input for the action that is the outside stimulation from others. The emotions are expressed as a point in the Russell circle model at a coordinate in the unit circle. Moreover, each of the three robots expresses different emotions for the same action. The combined load of three-layer perceptron in these emotion generation functions is determined by prior learning of emotion training signals using backpropagation (BP). Subsequently, the NN of the emotion generation function learns no further emotions.

However, the bystander robot has another three-layer perceptron as a moral emotion generation function in addition to the emotion generation function. This perceptron outputs emotions for the input of the similar emotion generation function. Figure 3 shows a diagrammatical view of the function of each robot, speaker, addressee, and bystander. The NN generating moral emotion learns emotion output from a teacher signal from the NN generating the emotions of the addressee or the bystander using BP according to the distance between the emotion points of the addressee and the bystander on the Russell circle model. When the bystander robot compares the emotions provided by a moral emotion generation function with those provided by its own emotion generation function to judge when it can sympathize with an addressee, it expresses emotions provided by the emotion generation function. The bystander robot controls emotions provided by its emotion generation function without expressing them, when it judges that it cannot sympathize with an addressee.

In other words, the proposed model is the moral emotion generation function of the bystander, used to control emotional expression. Using the proposed model, a bystander robot controls emotions as needed and thereby takes part in cooperative communication.

3.2 Learning of moral emotion

Generally, we form moral emotions over a long period by communicating with many others in our social lives. However, it is difficult for many robots to communicate using long-distance movement, as a result of being active in limited domains. Therefore, a robot must form a moral emotion by learning from a small number of others. In this study, a bystander robot forms a moral emotion generation function in which it learns stochastically according to the Euclidean distance between each emotion on the Russell circle model to form this moral emotion from a few robots. This learning uses BP. Because BP is supervised learning, a training signal is necessary. The training signal is determined according to a two-dimensional probability distribution.
Figure 4: Two-dimensional probability distribution

\[ f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} \exp\left[ \frac{-1}{2(1-\rho_{xy}^2)} \left( \frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{2\rho_{xy}(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right) \right] \]

This is expressed as and takes the maximum in 
\((x, y) = (\mu_x, \mu_y)\).

When \(x\) and \(y\) do not have correlation, and the values of the dispersion are equal, then
\[ \rho_{xy} = 0 \]
\[ \sigma_x^2 = \sigma_y^2 = \sigma^2 \]

We substitute these expressions into Eq. 1.

\[ f(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[ \frac{-1}{2\sigma^2} \left( (x-\mu_x)^2 + (y-\mu_y)^2 \right) \right] \]

In this study, we treat emotions as points on the Russell circle model. The emotion of the bystander and addressee are \((x, y)\) and \((\mu_x, \mu_y)\), respectively, and we determine the probability using Eq. 4. Depending on the probability, the robot decides whether to use the emotions of the addressee or its own bystander emotions as a training signal. We update the combination load of the perceptron of the moral emotion generation function using the training signal the robot selected using BP.

3.3 Judgment of sympathy emotion

We treat emotions as points in the Russell circle model; using the vector from the origin to a point, we judge sympathy according to cosine similarity. Cosine similarity indicates that the direction of two vectors is similar, with the value being close to one. When this value is greater than 0.9, two robots are in a state of sympathy. Figure 5 shows an example of sympathy and antipathy. In the figure, the angle expressed by two vectors is large in the state of antipathy and small in the state of sympathy. Two emotions are similar if the angle expressed by the two corresponding vectors is small on the Russell circle model. In this study, we treat the state of two emotions being close as “sympathy” and the opposite state as “antipathy.”

When a bystander judges that two emotions generated by the emotion generation function and the moral emotion generation function are in the state of sympathy, it expresses an emotion. Figure 6 shows the expressiveness process of a bystander robot. The action input into the emotion generation function and the moral emotion generation function of the bystander robot is the same as the action that an addressee receives. Two emotions generated by each function are used to judge the state of sympathy. In that case, the bystander robot expresses one of the two generated emotions. In other words, the bystander expresses emotions when a person would be likely to sympathize with an addressee.

4. SIMULATION

We confirm that the moral emotion generation function of the proposed model enables appropriate sympathy expression through simulation. In the evaluation, we use balance theory, a basic principle that produces stable
tripartite relations. In the case of expressiveness of the addressee and the bystander, we express the impression that the speaker has of the bystander using an “unpleasantness degree” parameter. We check whether this unpleasantness degree decreases as the bystander’s learning of moral emotion advances.

4.1 Simulation summary
In this simulation, to evaluate the proposed model, we compare from the perspective of the speaker the expressiveness of a bystander robot with the moral emotion generation function and the expressiveness of the bystander robot with a function to reproduce addressee emotions. We call the model built by the emotion generation function and the moral emotion generation function “the proposed model” and the model built by the emotion generation function and the function to reproduce emotions of addressee “the comparison model.” This comparison model can completely restrain emotions. Therefore, by comparing with the comparison model, the proposed model can be confirmed to give a good impression to the speaker by expressiveness, excluding restraint. The ratio by which an addressee and a bystander sympathize is set to half by prior learning of the training signal. In one step, the speaker provides the action to an addressee, and the addressee outputs the emotion. The speaker replaces the addressee for each step. Figure 7 shows a situation in which the speaker changes for each step. Table 1 shows the conditions of the simulation. First, we determine the combined load of the emotion generation function according to prior learning in this simulation. At each step after prior learning, we choose one among 32 kinds of action at random and input the same action into the proposed model and the comparison model. The input signal action is a symbolic signal; no actual movement is assigned. We compare the output of the addressee with the provided bystander output and determine an unpleasantness degree based on an impression to be determined by a sympathy judgment.

Here, we explain the details of the unpleasantness degree. It is used for checking the impression that the bystander’s reaction to the addressee makes on a speaker, and it is determined by the tendency to strive for the balance state of balance theory. At each step, when the impression on the bystander by the speaker is “+,” the unpleasantness degree decreases, and when the impression is “−,” the unpleasantness degree increases. When we apply balance theory, we see the impression that the speaker receives in each state, shown in Figure 8. In the figure, A is the speaker, B is the addressee, and C is the bystander. In addition, Table 2 shows the value of the unpleasantness degree in each state. As the learning of the moral emotion generation function advances, the state of antipathy changes to the state of restraint. Therefore, it is expected that the unpleasantness degree decreases as the steps advance. We examine the unpleasantness degree for each speaker, and observe how it changes as the steps increase. In addition, we examine the ratio of whether the bystander was able to sympathize substantially with the addressee.

![Figure 8: Impression that bystander C makes on speaker A](image)

<table>
<thead>
<tr>
<th>Table 2: Setting of the unpleasantness degree</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State</strong></td>
</tr>
<tr>
<td>A bystander expresses emotions,</td>
</tr>
<tr>
<td>the impression for a speaker of a bystander</td>
</tr>
<tr>
<td>is “+” (Sympathy)</td>
</tr>
<tr>
<td>A bystander expresses emotions,</td>
</tr>
<tr>
<td>the impression for a speaker of a bystander</td>
</tr>
<tr>
<td>is “−” (Antipathy)</td>
</tr>
<tr>
<td>A bystander doesn’t express emotions,</td>
</tr>
<tr>
<td>the impression for a speaker of a bystander</td>
</tr>
<tr>
<td>is none (Restraint)</td>
</tr>
</tbody>
</table>

![Table 1: Simulation conditions](image)
4.2 Simulation process

In this simulation, the role of the robot is changed for each step. Therefore, we show the simulation procedure below.

I. Three robots exist, A, B, and C, which are respectively assigned the role of speaker, addressee, and bystander.

II. The three robots have an emotion generation function composed by a NN to create emotions of their own. The bystander has another emotion generation function composed by a NN to reproduce the emotion of the addressee by creating a moral emotion besides its emotion generation function. As the step progresses, the learning of this moral emotion generation function proceeds.

At each step of the simulation process, we see the following:

III. The speaker provides a randomly selected action to the addressee and bystander (the same input is given to the addressee and bystander).

IV. The addressee outputs its emotion (as a response to the input from the speaker).

V. The bystander outputs its emotion if two internal NN results are similar.

VI. The speaker provides the action given to the others to themselves. (The speaker does not output this emotion.)

VII. The speaker compares their emotions himself with that of the addressee, and the emotions of the addressee and bystander. The speaker determines the unpleasantness degree based on the balance theory.

VIII. The speaker and addressee, i.e., A and B, are exchanged and go to the next step (return to III) The bystander, C, does not change roles.

After the simulation, the following are considered:

IX. If the unpleasantness degree has declined, we can verify that the bystander can properly communicate in the group communication.

X. We compare the expressiveness ratio of the proposed model with the comparison model. If the ratio of restraint in the proposed model is less than the comparison model, the proposed model is verified to be effective.

In accordance with the above procedure, we run a simulation to verify the proposed model. However, as robots A and B change roles at each step, we evaluated the unpleasantness degree and expressiveness ratio by picking up the speaker’s step of each robot.

4.3 Simulation result

Figures 9 and 11 show the results of the simulation regarding the change in the unpleasantness degree when the speaker is A and B, respectively. The horizontal axis indicates the number of steps, and the vertical axis indicates the unpleasantness degree. The dotted line shows the change in the unpleasantness degree in the case of exposed emotions without a moral emotion. These figures show that the unpleasantness degree decreases as the steps advance in both models. In both figures, the proposed model reduced the unpleasantness degree more than the comparison model. Figures 10 and 12 show the expressiveness ratio for the addressee of the bystander in the last 100 steps, when the speaker is A and B, respectively. These figures show that the proposed model has a slightly higher sympathy ratio than the comparison model. In both figures, the proposed model has a lower restraint ratio than the comparison model. In addition, the proposed model has a greater antipathy ratio than the comparison model. That is, the proposed model ratio of antipathy is high, even though the proposed model has reduced more the unpleasantness degree than the comparison model.

4.4 Consideration

In this simulation, the unpleasantness degree tended to decrease in both models as the steps advanced. In cases where the bystander has no moral emotion indicated by the dotted line, the unpleasantness degree has remained near zero. Therefore, we consider that the unpleasantness degree is decreased in both models by having the moral emotion. Consequently, we consider that the bystander robot came to be able to control the expressiveness that gives the speaker a bad impression by having the function as a standard for expressiveness. In addition, we believe that the bystander robot came to be able to express cooperative emotions in order to not give a negative impression to the speaker. Even if the speaker has changed in a simulation, from the fact that the unpleasantness degree has decreased in both cases, we consider that one moral emotion generation function of the bystander is functioning effectively against two different robots.

On the other hand, we were not able to explain why the proposed model has a higher ratio of antipathy than the comparison model. The proposed model includes the step of learning emotions of their own in the moral emotion generation function. Therefore, it is difficult to predict accurately the emotions of the addressee for individuals using moral emotion.
However, the proposed model can determine the appropriate emotions for overall communication rather than for individuals. Consequently, the proposed model has a lower unpleasantness degree than the comparison model. In addition, the proposed model has a lower restraint ratio than the comparison model; the communication is more active in order to increase the proportion of expressiveness in the proposed model. Figure 13 shows antipathy to obtain a good impression from the speaker. The bystander robot is able to predict this situation using moral emotion. The proposed model aims not only to avoid a bad impression on the speaker, but also enables a good impression in some situations. Therefore, using this model, we think that the bystander robot avoids giving an unpleasant impression to the speaker, and can show cooperative expressiveness. In other words, the proposed model was able to increase expressiveness opportunities by approximately 10%. We can expect to activate group communication by bystanders from the results.

In summary, we show from this simulation result that the proposed model has achieved a balance between the increasing expressiveness opportunities and expressing emotions so as not to give an unfavorable impression to the speaker.

5. CONCLUSION

In this study, we proposed a model in which a bystander robot with a function for controlling emotions using moral emotion sympathized cooperatively to achieve a social, emotional creature-like display of itself. This model not only avoids making bad impressions on the speaker, but also makes a good impression in some situations. We consider that robots must behave cooperatively with each other. Therefore, we proposed a sympathy expression model of a bystander robot that can express appropriate sympathy emotions by learning moral emotion from the difference in expressiveness of neighboring robots. We also confirmed its effectiveness.

However, using this model, it is difficult to detect the difference in the sensitivity between robots having similar sensitivity levels, because this model is based on the assumption that the sensitivity of each robot is different. Hence, examining the behavior in the case of robots with
similar sensitivity is necessary. In addition, we assume communication among three robots in this study, but examining effectiveness with more robots might be necessary. We aim to implement the proposed model on a deployed robot, and it will be necessary to examine these remaining problems in the future.

REFERENCES
5. A. Fukumoto, H. Nagano and M. Tokumaru; Care action generation model for robot as parent child interaction with emotion growth functions, 14th International Symposium on Advanced Intelligent Systems-ISIS2013, T1c-4, 2013.