Empathic Active Inference: Active Inference with Empathy Mechanism for Socially Behaved Artificial Agent

Tadayuki Matsumura[1], Kanako Esaki[1] and Hiroyuki Mizuno[1]

[1]Research & Development Group, Hitachi, Ltd., 1-280, Higashi-koigakubo, Kokubunji-shi, Tokyo 185-8601, Japan.

tadayuki.matsumura.bh@hitachi.com

Abstract

This paper proposes a method for an artificial agent to behave socially by controlling it by active inference with an empathy mechanism. Active inference is a Bayesian hypothesis for understanding the mechanism of a biological agent's cognitive activities and is basically defined for single-agent cases. We extended active inference to the case of an agent surrounded by other agents. These other agents are not only objects of recognition but also sources of social perceptions and actions. An agent controlled with the proposed method infers the others' expectations toward itself on the basis of an empathy mechanism and tries to act in response to the expectations. Although defining proper sociality for a given situation is difficult since it differs by situation, we define sociality as an agent behaving as others expect. Accordingly, the others surrounding the agent are teachers for the agent to learn proper sociality; thus, an agent controlled with the proposed method can learn proper sociality in a variety of situations in a unified manner. We evaluated the proposed method regarding the controlling of autonomous mobile robots (AMRs) and evaluated sociality from the trajectory of the AMRs. From the evaluation results, an agent controlled with the proposed method could behave more socially than an agent controlled by standard active inference. In two agents case, the agent controlled with the proposed method behaved in a social way that decreased travel distance of another by 13.7% and increased margin between the agents by 25.8%, even if it increased travel distance of the agent by 8.2%. They also indicate that an agent controlled with the proposed method behaves more socially when it is surrounded by altruistic others but less socially when surrounded by selfish others.

Introduction

We humans are social animals. We form a community and establish explicit or non-explicit rules in the community. It is essential for the normal functioning of a society that each individual follows these rules. Therefore, artificial agents operating in our daily spaces are also required to follow such rules. In other words, an artificial agent is required to behave socially. Having sociality helps artificial agents operate in our complicated real environment. Our social abilities are acquired through a long evolutionary process and fundamental for our cognitive systems. We do not live alone, and there are always others around us. We can make appropriate cognition and take proper actions due to the help of the presence of others. Others surrounding us are not only objects of recognition but also sources of social recognition and social action.

On the basis of these ideas, we propose a method for an artificial agent to behave in a socially desirable manner with the help of others surrounding it. The proposed method is based on active inference (Friston et al., 2011; Adams et al., 2013; Friston et al., 2016; Friston et al., 2017;). Active inference was proposed under the context of the free energy principle (FEP), which is a hypothesis for understanding the mechanism of a biological agent's cognitive activities (Friston et al., 2006; Friston, 2010a). In FEP, the brain is viewed as a device performing variational Bayes inference. The human brain is explained as always predicting the future and works to decrease the uncertainty of predictions. Similar ideas were widely studied with certain contexts such as Bayesian brain hypothesis (Knill and Pouget, 2004), predictive coding (Rao and Ballard, 1999), and Helmholtz machine (Dayan et al., 1995). The unique point of active inference is that it explains actions as well as perception with only one principle: minimization of variational free energy. It is assumed that there is an internal model for predicting external environments in the human brain. The process of perceptions is explained as the process of minimizing free energy of the internal model by updating the parameters of that model. The process of taking actions is also explained as the process of minimizing expected free energy for the future under a certain action. Active inference is studied in a variety of environments (Pio-Lopez et al., 2016; Parr and Friston, 2017; Friston et al., 2018). It has been combined with deep learning for applying it to more complicated environments such as robot control (Ueltzhöffer, 2018; Millidge, 2020; Fountas, 2020; Tschantz, 2020; Catal, 2020; Catal, 2021). Although active inference is basically defined for single-agent cases, some works recently extended it to multi-agent cases (Friedman, 2021; Kaufmann, 2021; Albarracin, 2022). We also extended active inference to include free energy of others surrounding an agent. More specifically, the action of the agent is determined on the basis of two type of uncertainty, i.e., (1) the agent's uncertainty of others and (2) others' uncertainty of the agent. By assuming the second type of uncertainty, an agent controlled with the proposed method (hereafter, empathetic agent) attempts to act based on the other's expectations, which makes it behave in a socially desirable manner. This is achieved by estimating others' predictions regarding the agent on the basis of the idea of an empathy mechanism called a mirror system (Rizzolatti and Craighero, 2004; Cattaneo and Rizzolatti, 2009). With this mirror system, the empathetic agent estimates the prediction of others using the same model used to predict the future of its external environment. When the model is used to predict others' predictions, the input data of the model is changed from the observation of the agent to the observation of the others, which is also estimated by the agent.

We evaluated the proposed method in a situation of controlling autonomous mobile robots (AMRs). AMRs should not only avoid collisions but should also have sociality when they are operating in human spaces. For example, the sociality for AMRs can be defined as the margin of distance from others during their movement. Defining the proper sociality is difficult since it varies from situation to situation. Therefore, it is difficult to learn the proper sociality in standard reinforcement learning in which the proper margin of distance from others has to be manually encoded as the reward in the training for each situation. In this paper, this problem is solved by defining sociality as behaving as others expect. Namely, proper sociality in a certain environment is expressed as the behavior of others in the environment. Others are not obstacles but help to generate social behavior in our agent for controlling the AMRs. The sociality of the empathetic agent is acquired through the learning of the internal models for predicting the external environment including others' behaviors. In the evaluations, the empathetic agent learned different sociality according to the difference in the surrounding environments. When it learned in the environment in which others were selfish, it tended to behave selfishly. On the other hand, when it learned in the environment in which others were altruistic, it tended to behave altruistically.

Empathic Active Inference

Active Inference

As mentioned above, active inference is a hypothesis for understanding the mechanism of actions of biological agents (Friston et al., 2011; Adams et al., 2013; Friston et al., 2016; Friston et al., 2017;) and was proposed under the context of the FEP (Friston et al., 2006; Friston, 2010a). The FEP is widely studied and is applied to explain many cognitive abilities/phenomena such as behavior (Friston et al., 2010b), planning (Kaplan and Friston, 2018), autism (Quattrocki and Friston, 2014), and attention (Feldman and Friston, 2010). The most basic cognitive mechanisms, perceptions, and actions are also explained as the process for minimizing variational free energy. It seems natural to explain perception as inference minimizing the uncertainty of the internal model for the inference. More interestingly, an action is also explained with the same principle. In the FEP, an action is explained as the process of inference in which biological agents actively act to decrease uncertainty, and the best action is that expected decrease uncertainty the most. This process for performing actions is called active inference.

We now mathematically describe the FEP and active inference. As illustrated in Figure 1, there is an agent that receives observations (o_t) from an environment at time t. There are hidden states (s_t) behind the process of generating the observation. The agent takes an action (a_t) each t then receives the next observation (o_{t+1}) from the environment. The agent always infers the hidden state and action at the current state from the current observation as the following posterior,

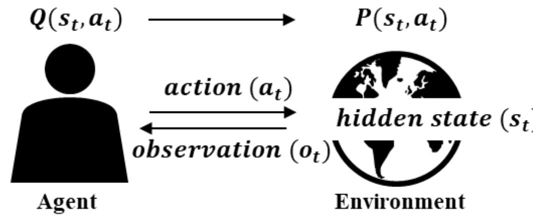

Figure. 1: Overview of free energy principle

$$P(s_t, a_t|o_t). \quad (1)$$

A variational density, $Q(s_t, a_t)$, is assumed to infer the posterior by variation methods. Under this context, variational free energy (F) is expressed as

$$F = -logP(o_t) + KL[Q(s_t, a_t)||P(s_t, a_t|o_t)], \quad (2)$$

where KL is a Kullback–Leibler divergence. Variational free energy is the same as evidence lower bound (ELBO) in machine learning (Blei et al., 2017). In the FEP, it is important idea that the action (a_t) is also inferred as like as the hidden state (o_t). On the basis of this idea, the process of minimizing free energy can be two types of processes. The first type of minimization process is related to the inference for hidden state. This is related to perception, i.e., it can decrease by refining the internal model for inferring the probability of the hidden state and its transition. More interestingly, an agent can decrease the free energy by refining the internal model for inferring the probability of the action, and this is the second type of process of minimizing free energy. From this viewpoint, it can be said that our action-making process is also the process of inference, as with perception. This process is called active inference. Agents based on the FEP perceive and act by minimizing variational free energy with these two types of minimization process.

In active inference, the desired distribution of actions at a given hidden state, $P(a_t|s_t)$, is expected to minimize free energy for a future state when an action is taken from the distribution. In previous works (Millidge, 2020), $P(a_t|s_t)$ was defined by

$$P(a_t|s_t) = \sigma(-\gamma G(s_t, a_t)). \quad (3)$$

where σ is a softmax function, γ is a precision weight, and $G(s_t, a_t)$ is the expected free energy. Expected free energy is the estimated free energy for future t. Eq. (3) means that the agent estimates free energy for a certain action, i.e., the agent makes a planning. If the agent is assumed to plan several steps ahead, the expected free energy is estimated for the sequence of actions, $\pi = \{a_t, a_{t+1}, \dots a_{t+T}\}$. In active inference literature, π is called a policy. From Eq. (2), expected free energy for a single time-step is expressed by

$$G(s_t, a_t) = -logP(o_t) + KL[Q(s_t)||P(s_t|o_t)]. \quad (4)$$

A neural network model is used to estimate expected free energy in (Millidge, 2020), and a Monte-Carlo tree search is used in (Fountas, 2020). In active inference literature, the first term $(-logP(o_t))$ is treated as preference of the agent to the observation, and the intention of the agent is encoded into this term as a reward signal (r),

$$G(s_t, a_t) = -r(o_t) + KL[Q(s_t)||Q(s_t|o_t)], \quad (5)$$

where $P(s_t|o_t)$ is approximated by $Q(s_t|o_t)$. The second term in Eq. (5) is called intrinsic value and expresses curiosity

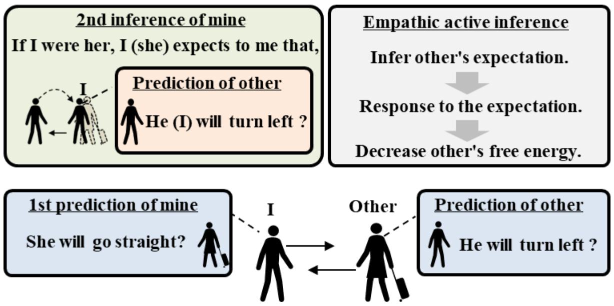

Figure. 2: Overview of empathic active inference

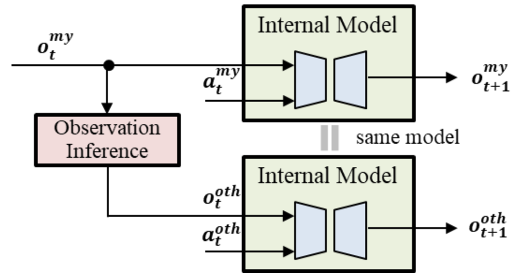

Figure 3: Processes of an inference of another's inference

to explore the environment. From Eq. (5), the agent acting on the basis of active inference takes into account both the intentional behavior expressed by the reward term and curiosity to explore the environment.

Simulation Theory and Mirror System

We extended active inference to generate social behavior. For this purpose, we embed the human capacity of empathy into active inference. There are mainly two types of human activity regarding empathy; (1) cognitive empathy and (2) emotional (or affective) empathy (Davis, 1983). Cognitive empathy is the ability of inferring another's mental state. Emotional empathy is the ability to feel what others are feeling as if it were your own feelings. In any types of empathy, understanding or sharing the experiences or feelings of others is thought to be related to our sociality (Eisenberg and Miller, 1987). Theory of mind and simulation theory are theoretical frameworks for understanding the mechanisms of these emotion. According to simulation theory, we can infer another's mental state by simulating what we would infer or feel if we were in the same situation as to other. Although there is still much room for discussion, it is suggested that mirror neurons and mirror systems are involved in such empathy (Keen, 2006; Gazzola et al., 2006). A Mirror neuron is a neuron that fires both when an animal acts and when the animal observes the same action performed by another. Mirror neurons were first observed in other primate species, and similar brain activities were suggested in humans. Neural systems related to mirror neurons are called mirror systems. Our proposed method enables an agent to empathize with others and is inspired by mirror system and simulation theory to virtually experience the experiences of others using its internal models as like simulating others' mental states with its own body.

Active Inference with Empathy Mechanism

Although there have been studies on applying active inference to control artificial agents, the application mainly focused on a case in which only a single controllable agent is assumed. We extend the situation to the multiple-agent case, especially, when multiple agents are us humans. In Figure 2, there are two types of agents, *I* and *others*, and both acting with active inference. The empathetic agent *I* always infers the future external environment and attempts to decrease the uncertainty of the inference. It is important that the *others* surrounding *I* also always infer the future external environment and attempt to decrease the uncertainty of inference. *I* is an uncertain factor for the *others*. Given these situations, there is another way to decrease free energy in addition to the ways described in the previous section about active inference. The way is to act as others' uncertainty decrease. Although it will not decrease *I's* free energy, it will decrease the total amount of free energy in the group of agents. Similarly, the free energy of *I* can decrease by the actions of the *others* if the *others* act to decrease the free energy of *I*. We can manipulate free energy not only through our actions but also through the actions of others by thinking of free energy collectively rather than individually. Actions based on this way can be said to be for others, i.e., social action. If each individual in a group behaves in this manner, it means that they behave as the others expect them to behave. We assume this state is a preferable social state for a group of agents, and we developed our method on the basis of this idea.

An empathetic agent first needs to infer another's inference to it. For this purpose, the empathetic agent uses an idea inspired by mirror systems and simulation theory. The core idea for inferring another's inference is that the other's inference is inferred using the same internal model used for inferring the empathetic agent's external environment. The process of inference of another's inference consists of two processes, as illustrated in Figure 3.

The first process is to estimate the observation of others (o_t^{oth}) from the observation of the empathic agent (o_t^{my}). If the observation is given by an image of vision, then the agent infers the observed image of the others. In this case, the method for novel view synthesis, such as Generative Query Network (GQN) (Eslami et al., 2018), can be used. If the observation is the set of positions of the others extracted from object detectors such as infrared sensors, the observation of others can be estimated simply by coordinate transformation. However, even in such a simple case, due to partial observability, it is not always possible to completely estimate the observation of others. For example, if there is an object that is not visible to the empathic agent due to occlusions or other factors, information about that object will be missing, even if the others see it. The second process generates prediction for the external environment of others (o_{t+1}^{oth}) by giving the estimated observation of others (o_t^{oth}) to the internal model that the empathic agent uses to predict its external environment. This method inspires a mirror system and simulation theory to infer the inferences of others by simulating the situation of if 'I were in the other's situation'. The action of the other (a_t^{oth}) is also required to infer the inference of the other. Simple methods can be applicable to generate the action of others such as randomly selected action. We assumed no-operation (NOP) action as the action of others. This is the idea of predicting 'when I stop, the other will act first'.

As explained in the previous section, an agent acting in accordance with active inference will take the action that will decrease the expected free energy the most. Intentional behavior is achieved by encoding the reward information corresponding to the behavioral intention as a preferable observation. We similarly extend Eq. (5) to a form that encodes intentional response to others' expectations to the agent as a reward,

$$\begin{aligned} G(s_t, a_t) &= -r(o_t) + KL[Q(s_t)||Q(s_t|o_t)] \\ &= -\{r_{my}(o_t) + \sum_i r_{oth}(o_t, e_t^i)\} \\ &\quad + KL[Q(s_t)||Q(s_t|o_t)]. \end{aligned} \tag{6}$$

The first term is the reward for the agent's goal (r_{my}), which is determined by the observed information, and is the same as the reward in Eq. (5). In Eq. (6), a reward term for the expectations of others (r_{oth}) are added, and they are determined by the observation (o_t) and expectation of others to the agent (e_t^i) which is estimated by the agent. The expectation from others is derived from the inference of others (o_t^{oth}). For example, if the positions of the others surrounding the others are estimated as the inference of others, the positions of the empathic agent inferred by the others are the expectations from the others. Similar to the flexibility in encoding rewards in active inference (Millidge, 2020), expectations from others can be encoded flexibly, such as by using probability distributions. For multiple others, the reward for each of the other (i) is summed. It is important that an inference of others to the empathic agent is interpreted as an expectation to the empathic agent from others. The proposed agent takes the action that will decrease the expected free energy described in Eq. 6 the most.

Evaluation

Simulation Setup

To evaluate the sociality of an empathetic agent, we ran multi-agent simulations in which multiple agents, including the empathetic agent, are walking from their initial positions to their destinations and need to avoid colliding with each other. The empathetic agent was controlled with the proposed method, and the others were controlled by the social force model (SFM) (Helbing and Molnar, 1995), which is a model for controlling pedestrians in a social space. Although there are a variety of extensions, the SFM basically models the motion of agents by the combination of a driving and repulsive forces. The driving force describes the motivation of agents to move toward the given goal at a certain desired velocity. The repulsive force represents the motivation of agents to avoid colliding with others or with obstacles such as walls.

To evaluate the sociality for a variety of scenarios, two types of conditions were changed depending on the scenario. The first condition was the situation of the scenario. Three types of situations were assumed for the simulation, as illustrated in Figure 4. In this section, the standard/empathetic agent is called the "player", and the other agents controlled with the SFM are simply called the "others" or "other". The player is the red dot in Figure 4. Situation (a) is the simplest in which two agents,

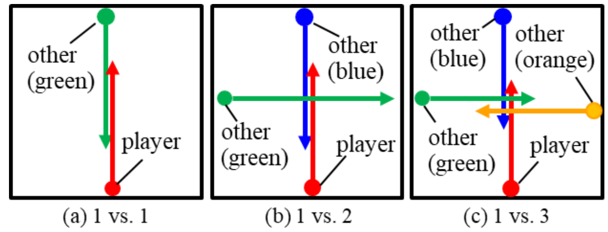

Figure 4: Types of simulation situations

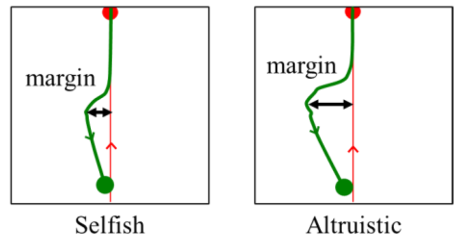

Figure 5: Two types of other

the player and one other, walk from their initial points to their destinations. Because their initial points and destinations are opposite each other, the player and other must take the non-shortest path (straight line between initial point and destination) to avoid colliding. Situation (b) is denser than (a) in which two others plus the player are walking and will cross each other at the center of the field. Situation (c) is the densest in which three others plus the player walk in a crossroad and will pass each other at the center of the field. There are no obstacles (i.e., walls) in all situations, and agents can walk in any area of the field.

The second condition is the type (characteristics) of other, i.e., selfish or altruistic. The type of other is controlled by the weight of the driving and repulsive forces in the SFM. When others are selfish, the weight of driving force toward the destination is set higher than that of the repulsive force with others. When others are altruistic, the weight of the repulsive force is set higher than that of the driving force toward the destination. Example trajectories of each type are illustrated in Figure 5. The player moves in a straight line toward the destination, and the other takes the non-shortest path to avoid colliding with the player in both cases. The difference between the types of other appears in the difference of the margin to avoid collision. When the other is altruistic, it walks with a larger margin with the player than that when it is selfish.

The observation of the player is the position of agents relative to the current position of the player. The observation is constructed for the two simulation steps, i.e., $[o_{t-1},\ o_t]$. As an ideal case of observation assumed in this simulation, there is no lack of data caused by occlusions, and there is no noise in the observation. The action space is defined as a discrete space. The player moves a constant distance in five directions $(-60°, -30°, 0°, 30°, 60°)$ toward the current direction. The constant distance of the player's movement is almost the same as that of others. In addition to these actions, the player can select NOP action, i.e., stop at the current position. A total of six actions is the action space for the player.

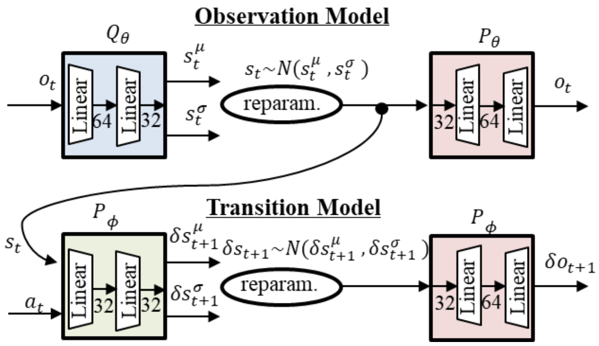

Figure 6: Models for empathic active inference

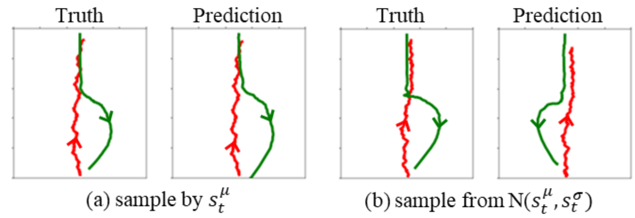

Figure 7: Examples of prediction results

Model and Training Setup

Two densities are assumed in the evaluation, and they are modeled using simple neural networks. The models are illustrated in Figure 6. The first model is the observation model for $Q_\theta(s_t|o_t)$ and $P_\theta(o_t|s_t)$ parameterized by θ. The second model is the transition model for $P_\phi(\delta s_{t+1}|s_t, a_t)$ and $P_\phi(\delta o_{t+1}|\delta s_{t+1})$ parameterized by ϕ. Both of the models consist of two fully connected layers (Liner), and the dimension of the latent space is set to 32. The observation and transition models use a re-parameterization trick to express probabilistic distribution with neural networks as like variational auto-encoder (VAE) (Kingma and Welling, 2013). The latent vectors in both the observation and transition models are assumed to be distributed in normal distribution. The distributions of the latent vectors of the observation model (Q_θ) and transition model (P_ϕ) are learned to get closer to each other. Here, the transition model and its learning are modified from the pure FEP in this paper. The transition model is constructed to model the difference in the observation (i.e., difference in positions) between the present and future, not the absolute position in the future. This is because there is continuity in this environment, namely, the position of agents continuously changes, and agents do not jump to far away places in a time. This feature can help to learn transition model. For this modification, each of the distributions of the latent vectors of the observation model (Q_θ) and the transition model (P_ϕ) are learned to become closer to the standard normal distribution, $N(0, I)$, respectively, and simultaneously, the distribution of the observation model (Q_θ) for s_{t+1} is learned to become closer to the sum of the normal distribution of the observation model (Q_θ) for s_t and the transition model (P_ϕ) for s_{t+1}. Namely, the KL divergence described in Eq. (7) also decreases in the learning process,

$$KL[N(s_{t+1}^\mu, s_{t+1}^\sigma)||N(s_t^\mu + \delta s_{t+1}^\mu, s_t^\sigma + \delta s_{t+1}^\sigma). \qquad (7)$$

The action is encoded into one-hot vectors.

In this evaluation, we do not assume the density of $Q(a_t|s_t)$ while it is assumed as a policy model in the previous works (Millidge, 2020; Fountas, 2020) because the action of the player is determined by the action probability $P(a_t|s_t)$ described in Eq. (3) with Eq. (6). The training of policy model is out of the scope in this study. The neural networks were trained in an offline manner since the aim was evaluating sociality of the empathetic agent's behavior, not validation of the feasibility of online training. Therefore, the training data for each scenario, i.e., three situations and two others' characteristics, are generated for a randomly acting player. During training-data generation, the player and one of the others have interchanged each other at random. This is for giving the player experiences of others' viewpoints, and it is necessary because the player must infer the other's inference with the other's viewpoint. When the player has no experience for the others' viewpoint, it cannot infer the inference of others. 1M samples were generated for each scenario. Adam was used for optimizing the neural network's parameters (Kingma and Ba, 2014), and its learning rate was constantly set to 1e-4. Training proceeded for 10,000 epochs. KL vanishment is a known difficulty in the learning of the VAE model, and KL annealing was proposed to tackle this problem (Bowman et al., 2016). In the learning process, the weight of KLD loss is doubled from 1e-4 at every 1000 epochs as KLD annealing.

Results

Examples of prediction from the trained observation and the transition models are shown in Figure 7 for situation (a) with the other's being altruistic. Predictions were cumulatively generated for a time horizon until reaching the destination. The player predicted the future for several steps by using only the first observation (o_1) and its action sequence ($a_1, a_2, \dots a_T$). The action sequence is randomly generated. The player predicted future observation (p_1) by using (o_1, a_1) then further predicted future observation (p_2) by using (p_1, a_1), and so on until a certain time. This procedure for predicting future by accumulating its prediction is necessary when the player decides the action on the basis of the expected free energy for several time steps ahead. In Figure 7, prediction for both (a) sampling by the center of distribution (s_t^μ) and (b) sampling from $N(s_t^\mu, s_t^\sigma)$ are shown. The prediction result matches the truth trajectory for (a). However, the prediction result does not match the truth trajectory for (b). This is because the player predicts with uncertainty, and the uncertainty is accumulated. This uncertainty intentionally fluctuates the prediction, and it makes the predictions diverse. The player decides the action that decreases expected free energy the most on the basis of these diverse predictions.

Next, we evaluated the sociality of the player. In this evaluation, sociality is discussed on the basis of the trajectory of the agents. If the player has less sociality, it goes straightly to the destination regardless of the others. On the other hand, the player actively takes a circuitous trajectory to avoid colliding with others when it has more sociality. The difference in sociality appears in the shape of a trajectory like that

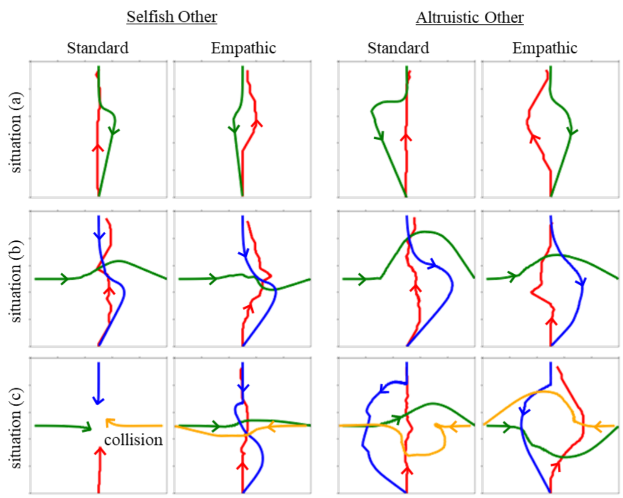

Figure 8: Trajectories of agents for each scenario

illustrated in Figure 5. We show the trajectory of the agent to intuitively understand its sociality of agent. Moreover, the total travel and minimum distances between the player and others during the simulation are also shown to quantitatively understand the sociality. The behaviors of the agents are shown in Figure 8. In this evaluation, the player estimated the expected free energy described in Eq. (6) by Monte-Carlo tree search (MCTS) with the learned models (Coulom, 2006; Fountas et al., 2020). The learned models were used for predicting the future state and value of each action (i.e., expected free energy) in MCTS. The maximum depth of the tree was set to 3 (i.e., three time steps are maximally estimated) and the search was run 3000 iterations for each time step. The action of the player is determined by Eq. (3) with the estimated expected free energy for each action. The precision weight (γ) in Eq. (3) is set to 1. The behavior of an agent controlled with standard active inference (hereafter, standard agent) was also evaluated for comparison. Although the standard agent is also based on the prediction about a future environment using internal models, it does not take into account the expectation from others. The reward (r_{my}) in Eq. (6) is defined by how close the agent is to the destination. If the agent moves closer to the goal, then positive reward is returned, and the expected free energy will be smaller than if it moves away from the destination. Moreover, the empathic player (i.e., empathic agent) estimates the reward of its future state for others' expectations, r_{oth} in Eq. (6). This reward is defined by the distance between the position of the player and others' expected positions.

From the result for situation (a), the standard agent almost walked straight to the destination. This is because it predicted that it is the highest reward (i.e., lowest free energy) when it walks straight to the destination and simultaneously predicted that another will pass without colliding with it even if it goes straight. On the other hand, the empathic agent, however, moved in a more circuitous trajectory. From Table 1, the total travel distance of the empathic player is larger than that of the standard player in situation (a) with regardless of the type of others (selfish or altruistic). On the other hand, the total travel distance of the others is smaller and minimum distance between

Table 1: Quantitative evaluation of behaviors

		Method	Travel distance (player)	Travel distance (Ave. of others)	Minimum distance
situation (a)	selfish	Standard	4.68	5.04	0.51
		Empathic	4.77	4.80	0.81
	alt.	Standard	4.65	5.93	1.20
		Empathic	5.03	5.12	1.51
situation (b)	selfish	Standard	5.35	5.17	0.49
		Empathic	5.45	5.38	0.42
	alt.	Standard	5.35	6.33	1.07
		Empathic	5.30	5.35	1.45
situation (c)	selfish	Standard	-	-	-
		Empathic	4.85	5.12	0.61
	alt.	Standard	4.85	6.45	0.87
		Empathic	5.40	5.74	1.39

the player and others is larger when the player is the empathic agent than those when the player is the standard agent. Comparing the empathic agent and the standard agent when the other is altruistic for situation (a), the travel distance of the empathic agent increased by 8.2% (4.65→5.03), while the travel distance of the other decreased by 13.7% (5.93→5.12) and the minimum distance between the player and the other increased by 25.8% (1.20→1.51). From these quantitative and qualitative results, the empathetic player behaved more socially than the standard player. In other words, the empathetic player behaved in a way that benefited the others, even if it was to its detriment. Because the difference between the empathetic and standard players was only the term of reward for the others (r_{oth}), the difference in behavior stems from this term. Motivation to respond to others' expectations changes the behavior of the player, in other words, the others surrounding the player leads to the player being social. The evaluation also showed that sociality of the player changes in according with the others around it. The total travel distance of the player when others were altruistic was larger by 5.5% (4.77→5.03) than that when others were selfish for situation (a). The difference in others reflected to the player's behavior.

The difference in others also changed the player's behavior for situations (b) and (c). Figure 9 shows the transitions in movement when others were selfish in situation (b). At first, the empathetic player (red) and blue-other started to turn to avoid colliding with the green-other (t1). The player and blue-other continued to turn (t2) and avoided each other by a small margin and passed each other (t3). Finally, the green-other passed the goal (t4). Meanwhile, Figure 10 shows the different transitions in movement when others were altruistic in situation (b). At first, the empathetic player (red) and blue-other started to turn to avoid colliding as similar to the case when others were selfish (t1). The empathic player turned to the left to avoid the blue-other unlike when the others are selfish, and the green-other also started to turn (t2). The all agents moved like in one

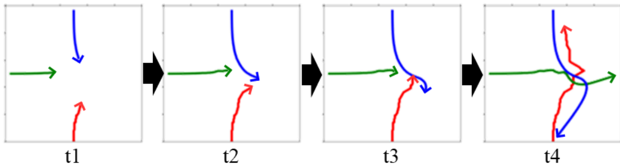

Figure 9: Transition of movements when others were selfish in situation (b).

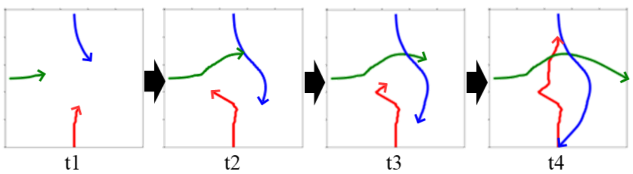

Figure 10: Transition of movements when others were altruistic in situation (b).

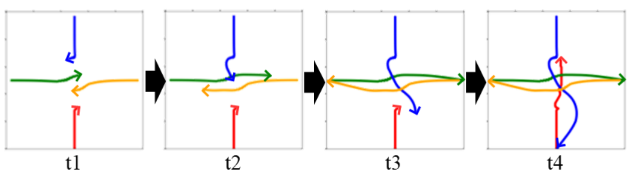

Figure 11: Transition of movements when others were selfish in situation (c).

circle, all together (t3). Finally, the all agents reached to the goals. The similar behaviors were also shown when the others are altruistic in situation (c). The all agents moved like in one circle as shown in Figure 8. On the other hand, the behaviors were more complex when the others are selfish in situation (c). In situation (c), when the player was controlled by the standard active inference and the others were selfish, the player and others collided. Meanwhile, Figure 11 shows the transitions in movement when the others were selfish and player was controlled by the empathic active inference in situation (c). In this situation, the green-other and orange-other first pass each other, while the blue-other and player waited (t1). The blue-other then passed the center point, while the player kept waiting (t2, t3). Finally, the player passed the center point (t4). The behavior of the player is similar to that of the blue-other.

From these results, the behavior of the empathetic agent was influenced by the surrounding others, i.e., it behaved in a similar manner to the surrounding others. This is because the empathetic agent responds to the expectations of others, and the expectation of others is predicted as 'how would I predict the observed other if I were in his situation by using the internal model learned by the other's behavior. Therefore, the empathetic agent behaves like others surrounding it. The empathetic agent can change its sociality for a given situation without manually changing the reward for the situation. Sociality is automatically adjusted to every scene from the behavior of others.

Conclusion and Future Work

Currently, artificial intelligence (AI)/robot ethics is recognized an important issue for integrating AI/robot technology into our society (Jobin et al., 2019). Although many concrete problems such as expandability, transparency, and safety, are being actively studied, one of the most important is to define what an ethical AI/robot is. The proposed method for socially behaved agent can be a solution for this problem in the ethics of AI /robots. From the viewpoint of this paper, an ethical AI/robot is considered to behave on the basis of not only its goal but also others' expectations of it and behaves like the others living around it in society. The mechanism behind the proposed method is empathy towards others. An agent controlled with the proposed method attempts to decrease not only its free energy but also that of others. It is as if human brains and agents are supposed to be shared. These sharing mechanisms generate social behavior. The others around artificial agents are its teachers for sociality. Namely, our behavior is reflected in the ethics of the artificial agent in the proposed method.

For future work, a more feasible training process is evaluated such as training of the policy network ($Q(a_t|s_t)$) and online-training. When the policy network is trained, the action of the empathic agent is directly determined by it. As another future work, we will improve the empathy mechanism. Currently, the empathy mechanism of the proposed method is mainly inspired by affective empathy and is evaluated for control tasks that require action decisions in a short time. However, there is another type of empathy, cognitive empathy, which is related to the ability to infer others' high-order mental states. This type of empathy is important for applying AI/robots to tasks that require more careful consideration such as planning strategies.

Related Work

Active inference has been evaluated for simple control systems (Pio-Lopez et al., 2016; Parr and Friston, 2017; Friston et al., 2018) and for more complex control systems by leveraging advances in deep learning (Ueltzhöffer, 2018; Millidge, 2020; Fountas, 2020; Tschantz, 2020; Catal, 2020; Catal, 2021). These studies mainly assumed a situation in which there is a single control target and no other agents around. Active inference is also discussed for two agents, and the synchronization process of the two agents has been examined (Friston and Frith, 2015). Recently, active inference is also applied to multi-agent cases for discussing emergences of collective intelligences from autonomous behaviors of individuals (Friedman, 2021; Kaufmann, 2021). Moreover, although it is not active inference, inference of others' mental states and sociality of agents for a simple discrete environment has been discussed (Yoshida et al., 2008).

In the field of multi-agent systems, there have been many studies on the emergence of cooperative behavior (Hernandez-Leal et al., 2019). In these multi-agent systems, it is mainly assumed that other agents are also machines, and the agents can explicitly communicate observations, model parameters, and prediction with each other (Foerster et al., 2016; Tampuu et al., 2017; Gupta et al., 2017). In this study, we assumed that the other agents are humans, and there is no explicit communication between an artificial agent and human agents. Similar to the situation discussed in this paper, social robotics assumes that a robot functioning around humans has sociality and required to infer mental states of humans from their behavior. For example, ProxEmo generates a trajectory for moving based on estimations of others' emotions from their gait behavior (Narayanan et al, 2020). Most of these studies modeled others as different agents from the agent itself, on the other hand, the proposed method models others as being the same as the agent, following the empathy model of the mirror system.

References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: active inference in the motor system. Brain Structure and Function, 218(3), 611−643.

Albarracin M, Demekas D, Ramstead MJD, Heins C. (2022). Epistemic communities under active inference. Entropy.; 24(4), 476.

Blei, D. M., Kucukelbir, A., & McAuliffe, J. D. (2017). Variational inference: A review for statisticians. Journal of the American statistical Association, 112(518), 859-877.

Bowman, S. R., Vilnis, L., Vinyals, O., Dai, A. M., Jozefowicz, R., & Bengio, S. (2016). Generating sentences from a continuous space. In International Conference on Computational Natural Language Learning (pp. 10−21). Association for Computational Linguistics.

Çatal, O., Verbelen, T., Nauta, J., De Boom, C., & Dhoedt, B. (2020). Learning perception and planning with deep active inference. In 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 3952−3956). IEEE.

Çatal, O., Verbelen, T., Van de Maele, T., Dhoedt, B., & Safron, A. (2021). Robot navigation as hierarchical active inference. Neural Networks, 142, 192−204.

Cattaneo, L., & Rizzolatti, G. (2009). The mirror neuron system. Archives of neurology, 66(5), 557−560.

Coulom, R. (2006). Efficient selectivity and backup operators in Monte-Carlo tree search. In International Conference on computers and games (pp. 72−83). Springer.

Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. Journal of personality and social psychology, 44(1), 113.

Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The helmholtz machine. Neural computation, 7(5), 889−904.

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. Psychological bulletin, 101(1), 91.

Eslami, S. A., Jimenez Rezende, D., Besse, F., Viola, F., Morcos, A. S., Garnelo, M., ... & Hassabis, D. (2018). Neural scene representation and rendering. Science, 360(6394), 1204-1210.

Feldman, H., & Friston, K. (2010). Attention, uncertainty, and free-energy. Frontiers in human neuroscience, 4, 215.

Friedman, D. A., Tschantz, A. D. D., Ramstead, M. J. D., Friston, K., & Constant, A. (2021). Active inferants: The basis for an active inference framework for ant colony behavior. Frontiers in Behavioral Neuroscience, 15, 126.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. Journal of physiology-Paris, 100(1−3), 70−87.

Friston, K. (2010a). The free-energy principle: a unified brain theory?. Nature reviews neuroscience, 11(2), 127−138.

Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010b). Action and behavior: a free-energy formulation. Biological cybernetics, 102(3), 227−260.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. Biological cybernetics, 104(1), 137−160.

Friston, K., & Frith, C. (2015). A duet for one. Consciousness and cognition, 36, 390−405.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2016). Active inference and learning. Neuroscience & Biobehavioral Reviews, 68, 862−879.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: a process theory. Neural computation, 29(1), 1−49.

Friston, K. J., Rosch, R., Parr, T., Price, C., & Bowman, H. (2018). Deep temporal models and active inference. Neuroscience & Biobehavioral Reviews, 90, 486−501.

Foerster, J., Assael, I. A., De Freitas, N., & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. Advances in neural information processing systems, 29.

Fountas, Z., Sajid, N., Mediano, P., & Friston, K. (2020). Deep active inference agents using Monte-Carlo methods. Advances in neural information processing systems, 33, 11662−11675.

Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. Current biology, 16(18), 1824−1829.

Gupta, J. K., Egorov, M., & Kochenderfer, M. (2017). Cooperative multi-agent control using deep reinforcement learning. In International conference on autonomous agents and multiagent systems (pp. 66−83). Springer.

Helbing, D., & Molnar, P. (1995). Social force model for pedestrian dynamics. Physical review E, 51(5), 4282.

Hernandez-Leal, P., Kartal, B., & Taylor, M. E. (2019). A survey and critique of multiagent deep reinforcement learning. Autonomous Agents and Multi-Agent Systems, 33(6), 750−797.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399.

Kaplan, R., & Friston, K. J. (2018). Planning and navigation as active inference. Biological cybernetics, 112(4), 323−343.

Kaufmann, R., Gupta, P., & Taylor, J. (2021). An active inference model of collective intelligence. *Entropy*, *23*(7), 830.

Keen, S. (2006). A theory of narrative empathy. Narrative, 14(3), 207−236.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. TRENDS in Neurosciences, 27(12), 712−719.

Millidge, B. (2020). Deep active inference as variational policy gradients. Journal of Mathematical Psychology, 96, 102348.

Narayanan, V., Manoghar, B. M., Dorbala, V. S., Manocha, D., & Bera, A. (2020). Proxemo: Gait-based emotion learning and multi-view proxemic fusion for socially-aware robot navigation. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 8200−8207). IEEE.

Parr, T., & Friston, K. J. (2017). Uncertainty, epistemics and active inference. Journal of the Royal Society Interface, 14(136), 20170376.

Pio-Lopez, L., Nizard, A., Friston, K., & Pezzulo, G. (2016). Active inference and robot control: a case study. Journal of The Royal Society Interface, 13(122), 20160616.

Quattrocki, E., & Friston, K. (2014). Autism, oxytocin and interoception. Neuroscience & Biobehavioral Reviews, 47, 410−430.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature neuroscience, 2(1), 79−87.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. Annu. Rev. Neurosci., 27, 169−192.

Tampuu, A., Matiisen, T., Kodelja, D., Kuzovkin, I., Korjus, K., Aru, J., ... & Vicente, R. (2017). Multiagent cooperation and competition with deep reinforcement learning. PloS one, 12(4).

Tschantz, A., Baltieri, M., Seth, A. K., & Buckley, C. L. (2020). Scaling active inference. In 2020 International joint Conference on neural networks (pp. 1−8). IEEE.

Ueltzhöffer, K. (2018). Deep active inference. Biological cybernetics, 112(6), 547−573.

Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. PLoS computational biology, 4(12).