Model of Tricking using Belief on Unobservable Other

Takashi Omori[1], Kazuhiro Okutani[2]

[1]School of Engineering, Tamagawa University
[2]Nippon Chemi-Con Corp.

Abstract: In a communication scene between human, we estimate the belief of other person on outer world, and change own action based on the estimation. Belief inference is indispensable for the human-human, and for the human-agent interaction. However, there are few models that represent dynamics of belief in the interaction scene. So, in this study, we use a chasing game that require the belief process for realizing higher level interaction, and evaluate a belief model we developed by a computer simulation. We also discuss on a possibility of tricking that makes use of the belief estimation process.

1 Introduction

Communication includes two processes, one that can be observed from outside like the speech, behavior and facial expression, and one that can't be observed directly like the intention or belief. The former is an information transmission process through a human perception. But the latter is realized through an inference on the mental state of other. The inference is triggered by the observation of former.

Most of conventional human interaction devices are designed based on evidences of observable human communication behavior. But the essence of communication exists in unobservable mental interaction because communication is just a way to change human behavior through a mutual understanding. To properly respond to other's intention in various communication situations, we must understand the unobservable intntion or belief of other people in addition to the superficial perception[1]. We call this psychological interaction as Mental Interaction.

There are some studies that use logic to describe the mental process [2]. But these studies were limited to a formal description of the mental process, and were not applied to a practical scene. However, intention and belief are important factors of communication, and can't be ignored. Yokoyama defined an intention as a hidden variable that affects on human action decision, and evaluated an effect of intention estimation through a computer simulation of collaborative game [3]. Abe applied the model to a robot-child play task and demonstrated effectiveness of the interest estimation in a collaborative interaction [4].

However, the model of belief still remains at a level of theory. There are few studyies on the model of belief in practical behavior task. So, in this study, we focus on a model of belief in the mental interaction process.

Belief in this paper is the facts on outer world that are recognized and believed by a behaving subject, an agent. Belief is acquired by observing environment, and updated by a recognition of change in the environment or an inference within the agent. An agent decides its action based on its belief.

In a mental interaction scene, we estimate belief of other agent by observing its behavior. The estimated belief enables more precise prediction of other's action. Prediction is a useful tool for realizing smooth interaction with other. That is, the estimation of other's belief is a fundamental problem of the mental interaction.

So, in this study, we try to clarify a model of other's belief estimation and action prediction by developing an action game that requires the belief estimation to play. Through a construction of action decision model for the game and its computer simulation, we discuss the principle of mental interaction in a social scene.

2 Modeling of Belief Based Action

2.1 Belief on Unobservable Other

Belief is independent between agents and is continuous within each agent. It is not updated when no observation nor inference is conducted. Thus far, there are few study on belief as a part of action decision or communication. As belief affects on the action decision of agent, we can estimate a belief, a recognition on a

situation, of an agent from its behavior at a specific situation.

Then what is the merit and demerit of the belief estimation from a viewpoint of action decision model? For the purpose, we discuss on a belief on other's action and its decision process by self in a simple chasing task, Tag Game. When a target to chase become unobservable behind an obstacle, we estimate its current position using knowledge about the target. By this reason, Chaser can keep chasing of unobservable other and the other, Runner, also keeps running away. Runner also has a belief on the behavior of Chaser. Though belief is not always correct, our action decision is based on the belief like this.

Sometimes we can make use of the belief. For example, Chaser can choose a strategy to wait for Runner by going reverse way when Runner is predicted to continue the same escaping strategy. In this case, Runner predicted Chaser to continue chasing behind it. But as Chaser took different way, the belief of Runner became uncorrect. We must discuss on an internal process of the strategic behavior chice.

In this stud, we apply a reinforcemnt learnig for the behavior decision in Tag Game. Runner and Chaser can take proper action while they can observe each other. But when they become unobservable each other by the obstacle, the agents can't decide their action in principle. For this problem, a human programmer used to implicitly embedd a mechanism of other's position estimation. However, this estimation process is a part of belief and we aim to discuss it in this paper. So, in this paper, we discuss models of a multiple strategy for other, an other's action prediction for the strategy, and a choice of self action decision strategy depending on a situation.

One problem is that the other agent also has a belief on itself and other. This mutual estination nature of interaction affects action decision of the other. Yokoyama discussed an action guiding model in a mutual action prediction situation [3]. But a level of estimation was shallow and they didn't consider a belief of other. The mutual belief model should be discussed to explain a prectical proble. But in this study, we don't discuss it to avoid complexty of the model. Fig.1 shows current model of action decision process with an intention and a belief process.

2.2 Consistency : Belief on Others Action Strategy

In our daily life, we can choose effective actions when we have knowledge on other, an other's model. If the other can use some strategies, we observe and learn proper action strategy for each of the strategies, choose one of them for other strategy of the moment, and keep applying it when the other become unobservable. Behind the choice to keep the strategy of the moment, there is a belief that the other agent also continues the current strategy after it became unobservable.

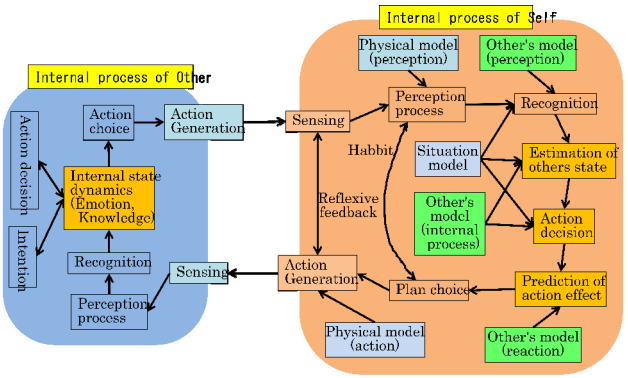

Fig. 1. A model of action decision process with belief.

To realize this nature of action decision, we consider a consistency of action decision as a parameter of action decision. The consistency represents a tendency of effective action choice. In Tag Game, the consistent action is a choice of shortest path toward a goal. A noisy action sequence is a sample of inconsistent action. We can estimate the degree of consistency by observing other's action sequence. A strong consistency will predict continuation of the same action strategy.

2.3 Tag Game and the Game World

For implementing Tag Game in this study, we set a simple 18-by-18 grid world with a shielding 11-by-11 square obstacle inside (Fig.2). There are two strategies for Chaser to capture Runner. One is Tracking Strategy in which Chaser tracks escaping Runner until it drives Runner to bay and capture it. Another is Waiting Strategy in which Chaser predicts Runner to continue current escaping action while it is behind the obstacle, go reverse way of the obstacle, and wait for Runner coming. Either strategy requires the agents having the belief to continue strategic actions while it can't be observed. And if Chaser has abilities to estimate an internal state of Runner and to predict other's action for longer time, Waiting Strategy becomes possible. In contrast, we just assumed Escaping Strategy for the runner that simply runs off from Chaser for simplification.

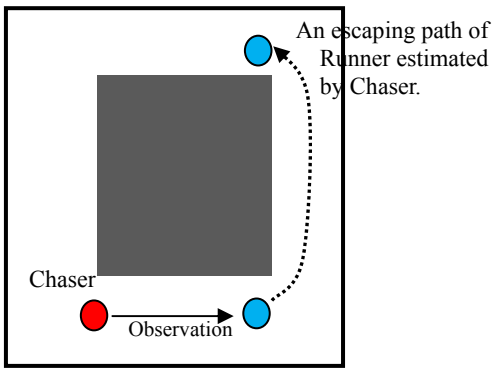

Fig.2 A game world assumed in Tag Game.

3 Model of Belief Estimation in Tag Game

3.1 Model of Self Action Decision

The basic strategy for Tag Game is Tracking Strategy for Chaser and Escaping Strategy for Runner. We used a reinforcement learning to realize Tracking and Escaping actions. Both agents learned relative position based state-action relation by TD learning in an obstacle free game world in advance. The belief of other's position is used for the action decision when the other is unobservable.

In TD learning, an action is decided in a probabilistic way using a temporal difference of expected reward. In the process, a temperature parameter T is used to introduce probability nature in the action sequence. When the temperature is high, the actions become noisy and more exploring, and noise free and more exploiting to acquire more reward in a low temperature condition.

The feature of temperature T agrees well with the features required for the consistency parameter K. When the consistency K is high, an agent should take action sequence with low noise, low temperature T. When the action sequence is noisy, K is low, it means high temperature T.

When Runner continues consistent Escape Strategy, Chaser can take Waitng Strategy. For its realization, Chaser must have abilities of long time span prediction of other and path planning to reach an predicted Runner position. When the time span of these abilities become long enough, Chaser becomes to be able to take the Waiting Strategy by knowing an Waiting way is more effective than usual Tracking.

3.2 Evaluation of Others Action Consistency

In Waiting Strategy, Runner is requested to keep the consistent Escaping action. In our simulation, the agents can predict other's behavior using the other's action model learned by the leinforcement learning. However, the action decision of other's model includes the consistency K, and a noise level of chasing and running action changes depending the K value. So, the consistency K of other must be estimated.

We tried the estimation of consistency from other's action sequence in a computer simulation. We used inner products, cosine, between the action direction vector of the agent and the direction of ideal chasing or escaping direction vector, and a leaky integrator for its temporal smoothing [5]. Fig.3 shows a computer simulation result in which the true K value is set 0.9 at first, changed to 0.1 at step 900 and again changed to 0.9 at step 1100. The estimated K value followed the true value within 100-200 steps that is permissible in a practical scene.

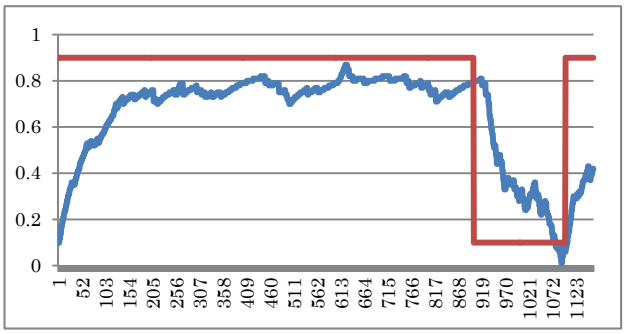

Fig.3 Estimation of other's consistency K through observation in a computer simulation.

3.3 Computer Simulation of Tricking

Fig.4 shows a case of Tricking in our simulation. The red blocks show the path plan of Chaser and the blue blocks shows the predicted path of Runner. Both pathes are created by Chaser at the moment that Runner hided behind the obstacle. The hight of blocks indicates pass time from start, now. The block is lowest for the first predicted position, and become higher along time. The red and bule larger blocks show some actual positions of Chaser and Runner. The green larger blocks with red and blue edge are the estimated position of Chaser and Runner by other agent.

In this case Chaser (red) predicted Runner (blue) to continue escaping counter clockwise, and Runner actually escaped as predicted (large blue). Runner expected the chaser to take Tracking strategy. But Chaser took Waiting Strategy actually as is shown by the action plan of Chaser. This is the tricking plan realized by Chaser agent.

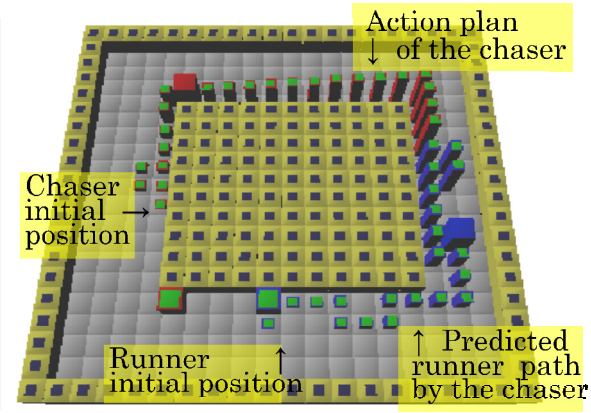

Fig.4 A case of Waiting Strategy planning by Chaser.

4 Discussion : Merit of Belief Model Use

The estimation of consistency from other's behavior corresponds to the reading of other's belief on its action continuation. The estimation lead us to more effective behavior prediction of other agent. Tricking is realized by leading other's internal state to a specific one and letting the other to take a desired strategy. Usually, the estimation of others internal state was thought as a way to improve prediction performance of the other's action. But here, we showed another possibility of making other to take desiable strategy by using the model of other.

In this paper, we embedded the process of self path plan search by the long range action prediction and the strategy change by hand. However, to realize automonous tricking strategies in various situations, we need some mechanism that the agent recognize a merit of taking other strategy and seek for a possible one. A traditional way of its realization is embedding the fuction by human hand. But to understand the belief and tricking ability of human, we should consider an automonous mechanism that creates new strategy by itself.

5 Conclusion

To reveal the mechanism of belief in action decision, we designed Tag Game that requires a belief for its solution, and proposed a simple agent model. The introduction of belief enabled the estimation of other's belief, and enabled more precise action decision of self. We also found the belief model is a foundation of tricking that is often seen in human sosial action. We don't think the tricking is not always necessary for a human assiting agent. But it will help development of human-agent interaction of higher level by considering a dymanics of human belief in real situations [6].

Acknowledgement

This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas "The study on the neural dynamics for understanding communication in terms of complex hetero systems (No.4103)" of The Ministry of Education, Culture, Sports, Science, and Technology, Japan

References

[1] Osawa: Computational Model of Intelligent Agent in Conversation, Computer Software,Vol.6, No.4, pp348-359, (1989) (In Japanese)

[2] Wooldridge M. : Reasoning About Rational Agents, MIT Press, (2000)

[3] Ayami Yokoyama, Takashi Omori : Modeling of human intention estimation process in social interaction scene, Proc. FUZZ-IEEE 2010,2306-2311,2010

[4] Kasumi Abe, Iwasaki Akiko, Tomoaki Nakamura, Takayuki Nagai, Ayami Yokoyama, Takayuki Shimotomai, Hiroyuki Okada, Takashi Omori, "Playmate Robots That Can Act According to Child's Mental State", IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.4660-4667, Portugal, 2012

[5] Kazuhiro Okutani, Takashi Omori : Development of chasing game for a study of other's belief estimation and action decision, , HAI2012, 2D-17, 2012 (In Japanese)

[6] Takahashi, H., Saito, C., Okada, H., Omori T., An investigation of social factors related to online mentalizing in a human-robot competitive game., Japanese Psychological Research, 55(2),144-153, 2013