Constructing a multi-agent mechanism for changing epistemic states

Kotaro Nishiyama (1210042)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 12, 2014

Keywords: multi-agent system, public announcement logic, angent annnouncement logic, Muddy Children Puzzle.

In recent years, a lot of communication with the text on the internet is developing rapidly, e.g., in social games and SNS. "Are You A Werewolf?" is a type of traditional table games but a talk game, and it has been played widely by using chat application and bulletin boards on the internet. The players of this game are 8-13 people, divided into human role and werewolf role, and they try to eradicate each other by performing a conversation. One of the features of this game is that the participants repeat conversations without knowing who plays the werewolf role, and try to find a liar by pointing out contradictions and instable opinion of players.

On the other hand, the study of modal logic in recent years has been developed in the context of epistemic logic and dynamic logic, under the name of Dynamic Epistemic Logic (DEL). DEL studies the notion of games, announcements, and communication based on belief change and update of shared knowledge in terms of modal logic. In DEL, actions such as private communication and public announcement can be regarded as events to change the epistemic state. For example, after a public announcement is performed, it is assumed that the agents will always know the content of the announcement.

However, if the knowledge obtained by the public announcement is inconsistent with knowledge that the agent had from the original, it could

be two cases. First, you may reject the knowledge gained from the public announcement and regard the agent making a false announcement. Second, we may update the present state of knowledge to accept the content of the announcement.

In addition, Agent announcement logic has an extension of the public announcement logic, where agents may execute bluffing announcements and lying announcements to other agents as well as truthful announcements. For example, if an agent knows that φ but announces of $\neg\varphi$ to the other agents, the agent could cheat on the other agents.

In this study, we treated as a model of multi-agent system a scene "Are You A Werewolf?" by using a modified form of the agent announcement logic. And we have constructed the system that can observe how to change the knowledge state when the public announcement has been performed. To be more specific, we propose a mechanism that can be confirmed that each agent accepts or reject the public announcement when the public announcement has been performed and the public announcement makes conflicts or correspond to original knowledge of each agents.

The purpose of this study is, to build a system in which knowledge of each agent is changed by announcement,with which agents try to achive the goal.

However, if some agents are incooperative make intentionally false statements to inhibit the achievement of objective, we can expect that the achievement difficult. In this study, where the incooperative agents and cooperative agents are mixed, each agent, needs to decide wheher to accept the other's announcement or to admit, according to his own knowledge.

In this paper, we apply DEL to the agent announcement logic with liars; the Muddy Children Puzzle is a classic logic puzzle as an application example of a multi-agent system, public announcement line the knowledge state of the agent we have proposed a mechanism that can be confirmed that when it is us, the knowledge state that each agent has, with the candidates knowledge against it or its equivalent, of whether to reject or accept its public announcement on each agent.

In this paper, we re-solved Muddy Children Puzzle that is a classic logic puzzle and as an application example of a multi-agent system by using the deformation of the agent announcement logic that is an extension of the

public announcement logic. And we proposed a mechanism can confirm that when the public announcement has been executed, and the knowledge state that each agent has, with the candidates knowledge against it or its equivalent, of whether to reject each agent, or accept its public announcement.

In addition, Muddy Children Puzzle was redefine the problem at two points that is assumed in order to approximate the conditions "Are You A Werewolf?" is carried out in stages the announcement, that the agent may telling lie. Further, if the knowledge obtained by the public announcement knowledge and the agent itself that received the announcement with conflict, the agent announcement logic redefined the semantics to reject the announcement. With this redefinition, the agents do not create a knowledge situation of "do not believe anything".

Experiments carried out in this paper, on the basis of the Muddy Children Puzzle, agent of three people be done in stages the public announcement, was confirmed that other agents that whether to accept or not the public announcement and knowledge from that public announcement. The announcement contents is only "I know if my face is muddy or not." or "I don't know if my face is muddy or not.".

It should be noted that, in this experiment, I have implemented a mechanism, the agent is possible to lie at any time, the agent can reject without accepting the announcement in case of conflict with the knowledge that the announcement has of his own agent who received the announcement.

As a result, we have shown that the following. In some cases, by lying announcement, it is possible to gain knowledge that could not have if agent does not tell a lie. If the agent gets a lying or bluffing announcement and the gained knowledge does not make contradiction to him original knowledge, the agent may believe the lying or bluffing announcement and accept the lie, and believe the other world different to real. In case of conflict with their own knowledge, the agent can then dismissed the announcement, to see through the lies.