JAIST Repository
https://dspace.jaist.ac.jp/

Title	不確実な状況における利己的な学習主体の相互協調
Author(s)	鳥居，拓馬
Citation	
Issue Date	2016-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/13511
Rights	
Description	Supervisor:橋本　敬，知識科学研究科，博士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Mutual cooperation between greedy learners under uncertainty

Student No: 1060005
Name: Takuma Torii

Cooperation is a common form of social interaction. The problem of cooperation is a conflict between individual rationality and collective rationality: selfish players maximizing their own profit obtain a smallest profit as a collective. Since in many situations in nature or society behaving selfishly is apparently a profitable option, how organisms solve the problem of cooperation is a long-lasting question. The problem was formulated as Prisoner's Dilemma in game theory, and various theoretical and empirical studies have shown possibilities of cooperation.

In game theory, a rational player takes a payoff-maximizing strategy (or action). A payoff to each player is a function of strategies of all players. A profile of strategies is called a Nash equilibrium if no one can improve his payoff by unilaterally changing his strategy. Prisoner's Dilemma is a game, in which each of two players has two options: Cooperate or Defect. It has been proved that the only Nash equilibrium in one-shot Prisoner's Dilemma is mutual defection. Iterated Prisoner's Dilemma (IPD) is its extensive-form variant. It has been proved that mutual cooperation can be a Nash equilibrium as well as mutual defection if both players consider sufficiently long-term future. One of the prominent findings in this area is the so-called tit-for-tat (TFT) or reciprocal strategy, whose behavioral rule is: if you cooperate (defect) I will cooperate (defect). Many theoretical studies has provided evidence supporting TFT for cooperation in some context, however, for what objective one acquires TFT-like behaviors is an unanswered question.

A key idea in recent game theory is bounded rationality: a decision maker has to choose an action based on limited information and restricted cognitive resources. Learning is a means to overcome uncertainty arising from incompleteness of information. Some psychological studies have shown that human cooperation is observed more frequency under uncertainty. Chapter 1 and 2 contains the background and reviews regarding the problem of cooperation in psychology and game theory.

This thesis aims at proposing a theoretical description for the problem of cooperation between two learning players under uncertainty. Concretely, the thesis aims at showing some conditions for mutual cooperation and what the learning players can maximize to establish mutual cooperation. Led by findings from psychology and game theory, in this thesis, the problem is formulated as Iterated Prisoner's Dilemma under uncertainty, where one can only get feedbacks to him in response to his actions. That is, no one can get information about the payoff matrix and related to the opponent. Reinforcement learning is a mechanism that can maximize its total profit only based on feedbacks (payoffs) in response to its actions. There are many evidence that reinforcement learning can solve various real world problems in uncertain environments.

In this thesis, we showed some conditions for mutual cooperation that can be established between selfish, reinforcement learners who attempt to maximize their own profit under uncertainty. Mutual cooperation is observed almost surely if both players make decisions based on sufficiently long-term experience. From a detailed analysis, TFT-like behaviors are observed during mutual cooperation between reinforcement learners. This finding gives TFT a position as a by-product of selfish/greedy learning. Chapter 3 includes the findings above.

Further conditions are investigated by approximating IPD of reinforcement learners. Using this approximated model, we formally studied several properties of the game, including payoff-related conditions for mutual cooperation. Based on the findings from the approximated model, we derived a payoff matrix that dramatically improve mutual cooperation between reinforcement learners. This can be interpreted as a practical mechanism for cooperation. The approximated model was studied in Chapter 4.

Chapter 5 contains a framework for future studies, which allows us to study reinforcement learning strategies within a more generalized class, although there are technical issues. In this thesis, we studied 1st-order strategy class, and discussed future directions.

Combining all the findings in this thesis and previous studies, a theoretical description for the problem of cooperation between learning players is argued in Chapter 6. It states that mutual cooperation can be established under uncertainty if both selfish players learn to maximize their own profit from their long-term trial-and-error experience.

The problem of cooperation is a special case of the free-rider problem and/or shared resource problem, which are more common in real social situations. For example, one reported that leaving the problem of cooperation unsolved declines the performance of team members. The problem of cooperation can be one of the common barriers that impede performance of organizational activity, such as organizational knowledge creation. The findings in this thesis will provide a clue to resolve conflict situations in real social interactions. Contrary to a common belief about uncertainty, our finding suggests the possibility that uncertainty regarding information, especially conflict relationship, might improve mutual cooperation, if participants can learn from their action-feedback experience.

Keywords: Cooperation, Reinforcement learning, Game theory, Uncertainty, Prisoner's dilemma