

## دستیابی به همکاری از طریق یادگیری تقویتی چند عاملی در معمای زندانی تکرارشونده

سمیرا فرزانه<sup>۱</sup>، فرشته زندی<sup>۲</sup>، جواد سلیمی سرتختی<sup>۳\*</sup>

<sup>۱</sup>دانشجوی کارشناسی ارشد مهندسی کامپیوتر نرم افزار، دانشکده مهندسی برق و کامپیوتر، دانشگاه کاشان، کاشان، ایران.

<sup>۲</sup>عضو هیئت علمی، استادیار، دانشکده مهندسی برق و کامپیوتر، دانشگاه کاشان، کاشان، ایران.

### چکیده

امروزه معمای زندانی یکی از مسائل اولیه و مهم در نظریه بازی ها است. در این معما نقطه تعادل نشی وجود دارد و چنانچه عامل ها منطقی رفتار کنند در آن نقطه بازی می کنند؛ بدین منظور عامل ها برای دستیابی به سود بیشتر از بین دو عمل همکاری و عدم همکاری، عدم همکاری را انتخاب می کنند. در حالیکه برای عامل ها نقطه بهتری نسبت به نقطه نش وجود دارد و آن هم این است که هر دو عامل همکاری را انتخاب کنند. بنابراین، در جهت افزایش میزان همکاری عامل ها معمای زندانی به صورت معمای زندانی تکرارشونده با یک رویکرد یادگیری تقویتی در نظر گرفته شده است. نتایج مقاله نشان دهنده این است که رویکرد مورد نظر سبب افزایش میزان همکاری عامل ها شده است و اگر عاملی همکاری را پیشه کند عامل دیگر نیز همکاری را انتخاب می کند و بالعکس.

**کلمات کلیدی:** عدم همکاری متقابل، معمای زندانی تکرارشونده<sup>۱</sup>، یادگیری تقویتی، همکاری متقابل، LSTM<sup>۲</sup>

## Achieving Cooperation Through Multi agent Reinforcement Learning In Iterated Prisoner's Dilemma

Samira Farzaneh<sup>1</sup>, Fereshteh Zandi<sup>2</sup>, Javad Salimi Sartakhti<sup>3\*</sup>

<sup>1,2</sup>Master Student of Computer Software Engineering, Faculty of Electrical and Computer Engineering, University of Kashan, Kashan, Iran.

<sup>2</sup>Faculty member, Assistant Professor, Faculty of Electrical and Computer Engineering, University of Kashan, Kashan, Iran.

### Abstract

Nowadays, the prisoner's dilemma is one of the primary and important issues in game theory. In this dilemma, there is a Nash Equilibrium, and if the agents behave rationally, they play at point; For this purpose, the agents choose defection between the two actions of cooperation and defection to achieve greater profit. However there is a better point for the agents than the Nash Equilibrium, it is that both agents choose the cooperation. However there is a better point for the agents than the Nash Equilibrium, it is that both agents choose the cooperation. Therefore, in order to increase the rate of cooperation of the agents, the prisoner's dilemma has been considered as iterated prisoner's dilemma with a reinforcement learning approach. The results of the article show that the desired approach let has increased the rate of cooperation of the agents, and if one agent choose the cooperation, the other agent also chooses cooperation and vice versa.

### تاریخچه مقاله:

تاریخ ارسال: ۱۳۹۹/۱۰/۲۹

تاریخ اصلاحات: ۱۳۹۹/۱۱/۲۲

تاریخ پذیرش: ۱۳۹۹/۱۱/۱۹

تاریخ انتشار: ۱۳۹۹/۱۲/۲۹

### Keywords:

Mutual Defection  
Iterated Prisoner's  
Dilemma  
Reinforcement learning  
Mutual Cooperation  
LSTM(Long Short Term  
Memory)

\*ایمیل نویسنده مسئول:

salimi@kashanu.ac.ir