

## Reinforcement learning in a prisoner's dilemma

*Thursday, June 20, 2024 1:30 PM (30 minutes)*

I characterize the outcomes of a class of model-free reinforcement learning algorithms, such as stateless Q-learning, in a prisoner's dilemma. The behavior is studied in the limit as players stop experimenting after sufficiently exploring their options. A closed form relationship between the learning rate and game payoffs reveals whether the players will learn to cooperate or defect. The findings have implications for algorithmic collusion and also apply to asymmetric learners with different experimentation rules.

**Primary author:** DOLGOPOLOV, Artur (Bielefeld University)

**Presenter:** DOLGOPOLOV, Artur (Bielefeld University)

**Session Classification:** Parallel session: Algorithmic collusion: Foundations for understanding the emergence of anticompetitive behaviour