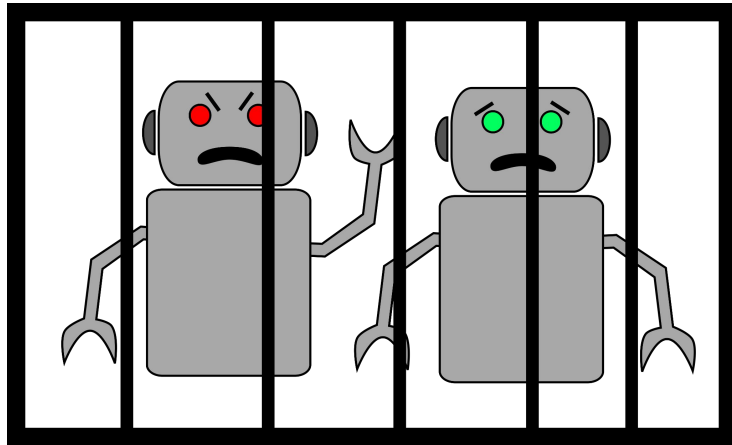


Oases of Cooperation:

An Empirical Evaluation of Reinforcement Learning in the Iterated Prisoner's Dilemma

Peter Barnett, John Burden

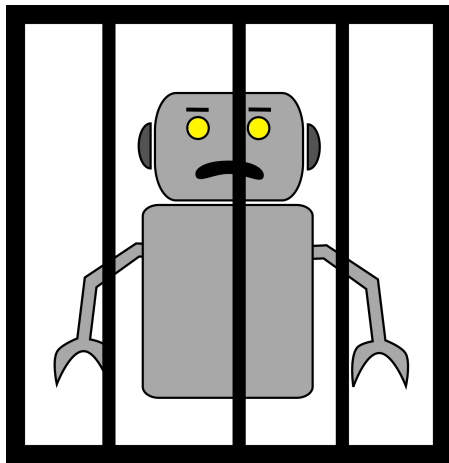
- Investigate cooperativity when players have incentives to not cooperate



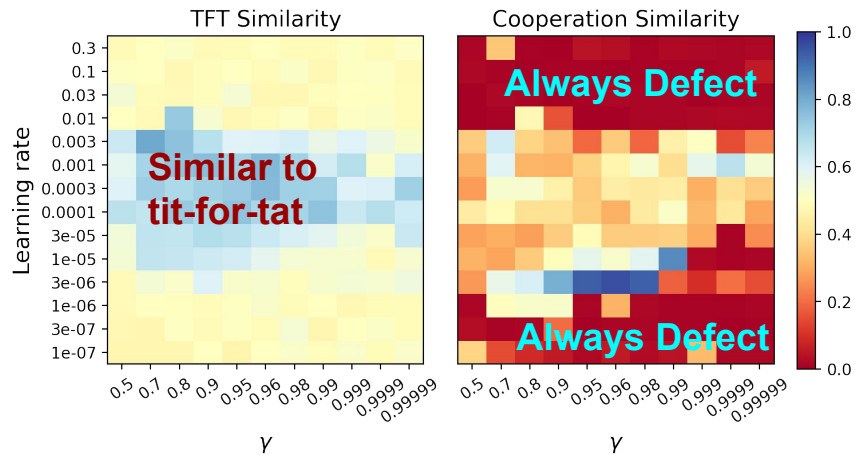
		Player 2	
		Cooperate	Defect
Player 1	Cooperate	3, 3	0, 5
	Defect	5, 0	1, 1

Single agent training

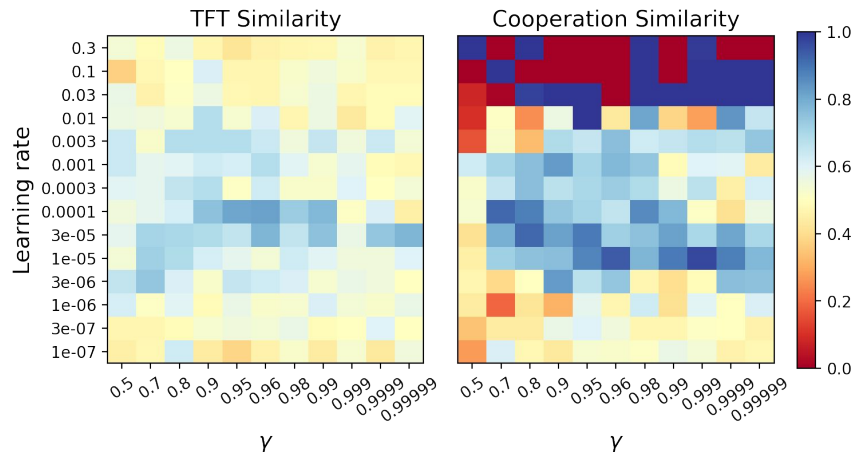
- Training RL agents against fixed policies
- Fixed policy:
 - Initially plays tit-for-tat
 - After a random number of turns defects forever



DQN



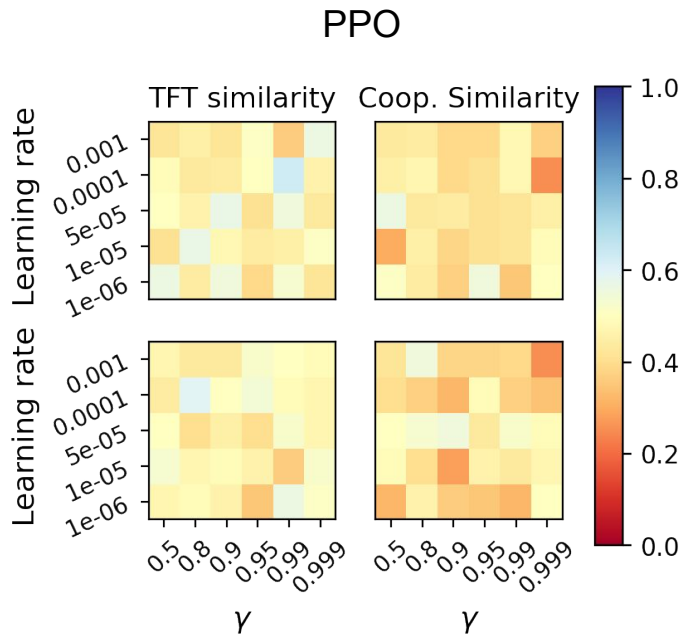
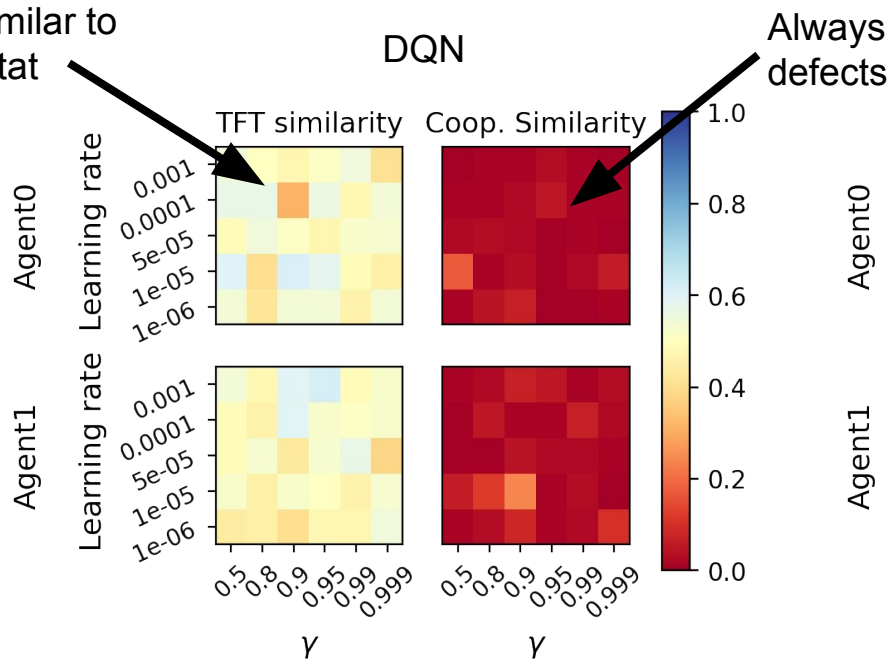
PPO



Multi-agent training

- Unable to find cooperative Nash Equilibria by default
 - For random game length there is a cooperative NE, but agents can't find it

Not similar to
tit-for-tat



Multi-agent pretraining

- Train agents with fixed policies to instil certain behaviour
- Influences behaviour in multi-agent training

