

Multi-Agent Learning



POLITECNICO
MILANO 1863

Setting

- We have multiple agents that simultaneously learn
 - Because some information is unknown (opponents' payoffs)
 - Because the opponents do not play optimally
 - Because an equilibrium cannot be computed exactly

A good MAL algorithm

- Exploitation of a sub-optimal player:
 - Given a player that is playing sub-optimal strategies, a good MAL algorithm should exploit such strategies and lead more utility than the equilibrium value to the learner
- Convergence to the equilibrium in self-play:
 - When all the players are using the learning algorithm, this algorithm should lead the players to play the equilibrium



Libratus (1.8 million USD won, 2017)
Against the top 4 international poker players