Reinforcement Learning

Final exam : Friday May 15th

- 7:15 PM - 10:15 PM
- Online
- Multiple choice
- Calculator needed
- Practice Gradiance Quiz will be posted
- One more Gradiance Quiz left (RL)

$\mathcal{A}$ - set of all actions

$A_t \in \mathcal{A}(S_t) \subset \mathcal{A}$

$R_t \in \mathbb{R}$

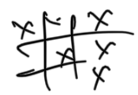
Policy π

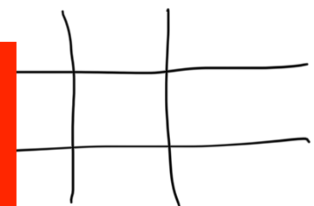
$S_1 \rightarrow A_{S_1}$

$S_2 \rightarrow A_{S_2}$

TicTacToe

Playing against an imperfect opponent.





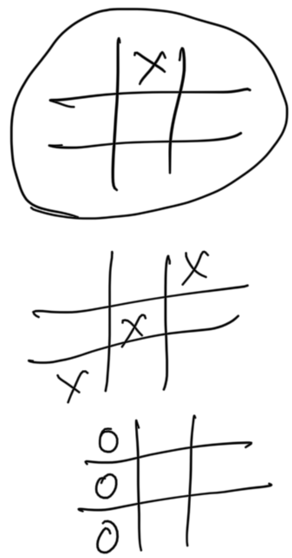
9^3 "possible" states

This should be 3^9 possible states

S

Value   Prob of winning from that state.

S is small fraction.



0.5

1

0

$A(S_0)$

S_1

$x - 10$

$0.5 + \alpha(0 - 0.5)$

$0.5 - 0.5\alpha$