CMPE 252 Artificial Intelligence
11/30/2021

Tic-Tac-Toe Reinforcement Learning

Alireza Mahin parvar, Casey Delaney , Jay Bharadva

Problem Description

- Using RL agents that learn to play the game of Tic-Tac-Toe: one follows the SARSA algorithm and the other follows Q-learning

- Initially we are applying SARSA Learning on Tic-Tac-Toe game.

- Analyze the results of RL agents using plots and showing the differences in them

Q-Learning

Q-learning (off-policy TD control) for estimating $\pi \approx \pi_*$

Algorithm parameters: step size $\alpha \in (0, 1]$, small $\varepsilon > 0$
Initialize $Q(s, a)$, for all $s \in \mathcal{S}^+, a \in \mathcal{A}(s)$, arbitrarily except that $Q(\mathit{terminal}, \cdot) = 0$

Loop for each episode:
  Initialize S
  Loop for each step of episode:
    Choose A from S using policy derived from Q (e.g., ε-greedy)
    Take action A, observe R, S'
    $Q(S, A) \leftarrow Q(S, A) + \alpha\big[R + \gamma \max_a Q(S', a) - Q(S, A)\big]$
    $S \leftarrow S'$
  until S is terminal

Task assignments among team members

- Casey
  - GUI development of the Tic-Tac-Toe board
  - Game mechanics

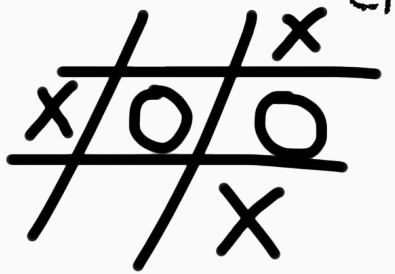
- Alireza & Jay
  - Designing Training Agent in RL section for Q Learning and Sarsa Algorithm
  - Creating training agent in RL section
  - Generating analysis plots

- All team members worked on slides, report and game implementation

Current Progress Demonstration

CMPE 252

Tic Tac Toe



Created by
Casey
Alireza
Jay

Loading . . .

Future Steps

- GUI Bug Fixes
- Finish implementing SARSA algorithm
- Finish implementing Q-Learning algorithm
- Project Completion!


```
SystemExit
```

```
/usr/local/lib/python3.7/site-packages/IPython/core/interactiveshell.py
:3445: UserWarning: To exit: use 'exit', 'quit', or Ctrl-D.
  warn("To exit: use 'exit', 'quit', or Ctrl-D.", stacklevel=1)
```