Reinforcement Learning for Games

Timothy Lillicrap & Blake Richards

Tutorial #4
Train a policy network

Imitate the expert!

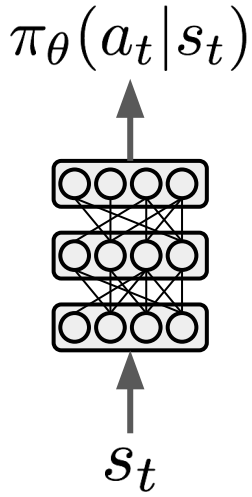
Goal: Learn how to train a policy network via supervised learning / behavioural cloning.

Train a policy network from expert games

- Our policy function will produce a categorical distribution over all possible discrete actions.
- We train on the moves provided in the expert dataset and the output is the softmax function.t

- We train to maximize the likelihood of the actions:

$$\mathcal{L}(\theta) = \mathbb{E}_{\tau \sim D_{\pi^*}} \left[\sum_{t=0}^{T} \log \pi_\theta(a_t|s_t) \right]$$



Exercise

- Train a network to predict the next move in an expert dataset
- Do so by minimizing the negative log likelihood of the next action given the current board position.

Tutorial #5
Play games using a policy network

Beat the value function player!

Goal: Learn how to use a policy network to play games.

Creating an agent using a policy function

Now that we have a policy that has been trained via imitation learning on expert data, we can use it to play games.
For the current state, we can sample actions according to:

$$\pi_\theta(a_t|s_t)$$

1. The action with the highest probability (the "greedy" action).
2. An action according to the categorical distribution determined by the softmax output:

$$p(a^i|s) = \frac{\exp(-z_i/T)}{\sum_{j=1}^{K}\exp(-z_j/T)}$$

Exercise

- Use the policy network to determine the next move for a player.
- Build a player that takes the move given the maximum probability by the network.
- Compare this to another player that samples moves according to the probability distribution output by the network.

Tutorial #10
Unbeatable Opponents

Will we still play games?

Goal: Engage with the social and ethical ramifications of having unstoppable computer players for games.

Exercise

Read a recent article on technology based cheating in chess:

Chess's cheating crisis: 'paranoia has become the culture'

Exercise

Reflect on related questions:

- Why do humans play games?
- What kinds of games can AI be applied to to win / what kinds of games remain out of reach and why?
- What happens if top human play hinges increasingly on access to compute?
- What will happen to fairness for other intellectual pursuits?
    - https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3