

RLSP2021 最終課題

立体N目並べに対する DQNの適用

加藤 愛斗¹, 久保谷善記², 竹味 和樹³

1. 豊田工業高等専門学校 2. 早稲田大学 3. 名古屋工業大学

立体N目並べとは



Connect4 (2D)

- (別名)重力つき平面四目並べ
- 交互に駒を置き、縦・横・斜めに連続して4つ以上並べたら勝利



(重力つき)立体N目並べ

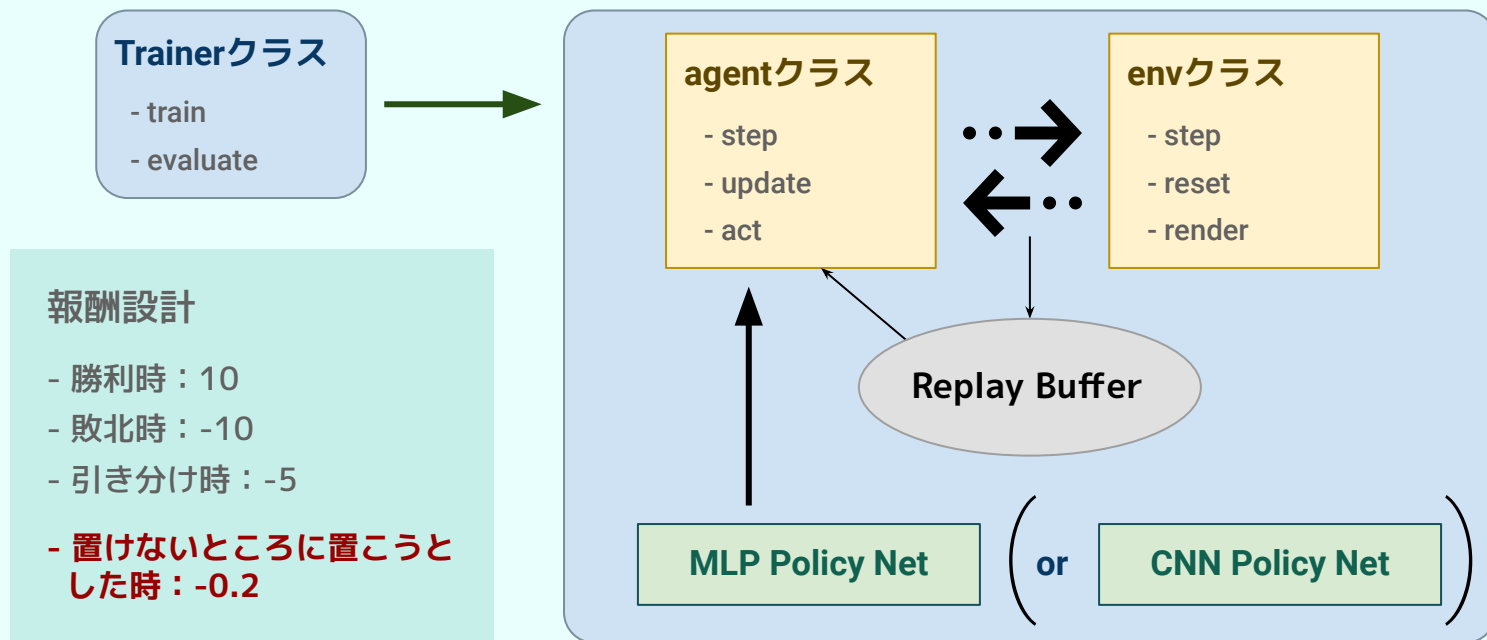
- 交互に駒を置き、縦・横・斜め・**対角**に連続して**N個**以上並べたら勝利

重力付き

「最下段」か「真下のマスが埋まっているマス」にしか置けないという制約

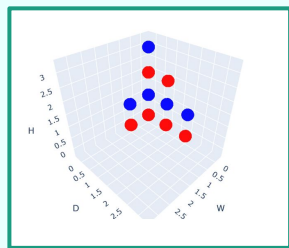
実装の概要

・ アルゴリズムフロー

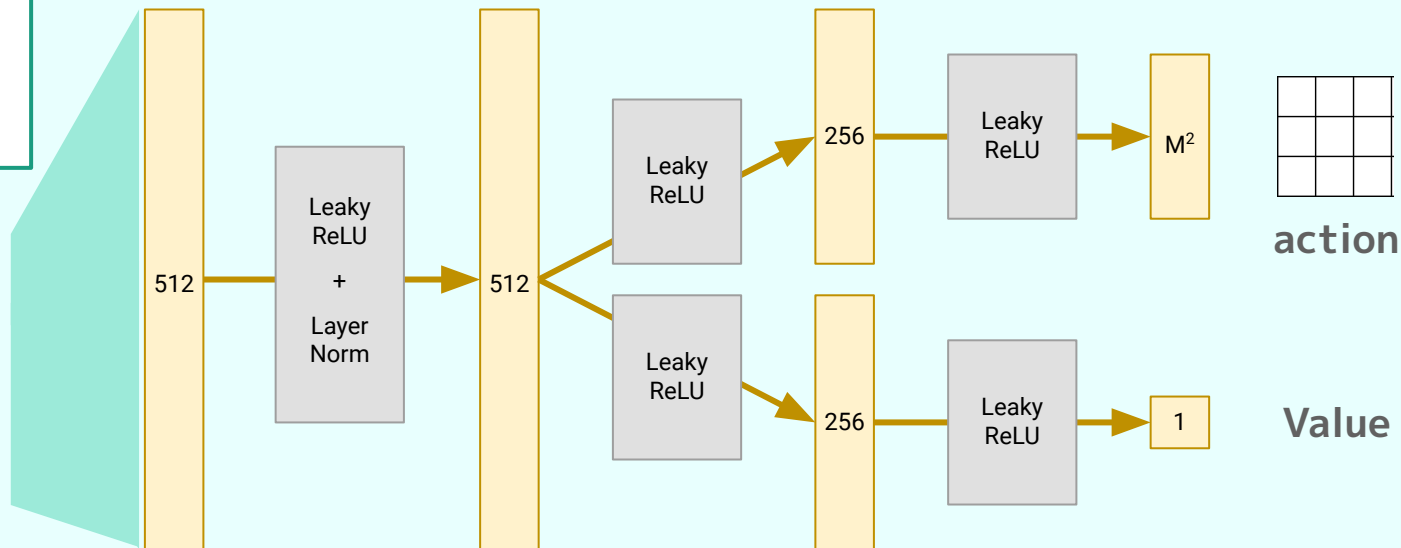


方策ネットの構造

・ MLPモデル



[[[1., 1., 1., 0.,
[1., 0., 0., 0.,
[0., 0., 0., 0.,
[0., 0., 0., 0.],
[[-1., -1., -1., 0.,
[-1., 0., 0., 0.,
[0., 0., 0., 0.,
[0., 0., 0., 0.],
.
.
.



実装上の工夫

1. バッチサイズ

バッチサイズを小さくし、一つ一つの経験データにフォーカス

2. 正規化レイヤーの追加

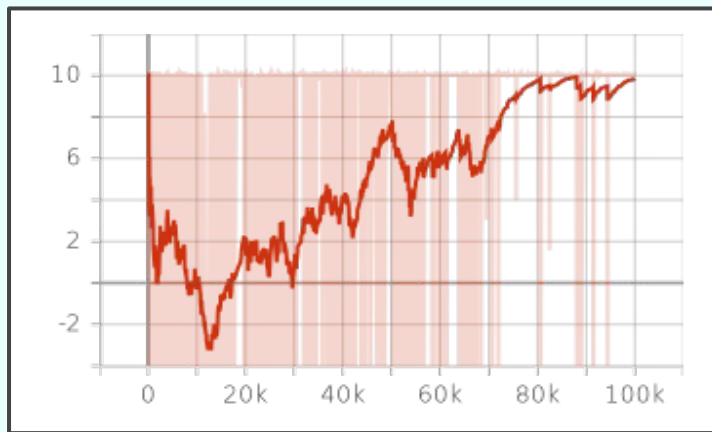
勾配爆発を抑制

3. validation

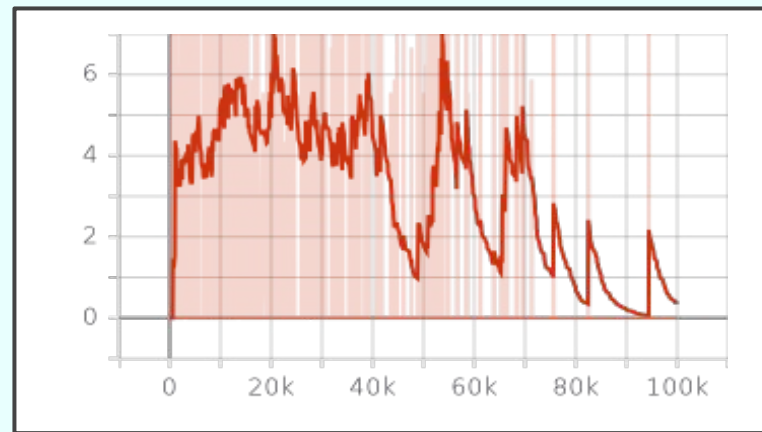
一定エピソードごとに方策評価を行い、学習の進み具合を把握

実験結果

- ・ $4 \times 4 \times 4$ 盤面内, 4 目並べ (MLP)
 - 学習終了時には安定して敵プレイヤーに勝利している
 - Random エージェントに勝率85.67%



↑Reward(episode)



↑置けない場所に置こうとした回数

参考文献

<https://www.slideshare.net/tomo3141592653/ai-119864507>

<https://github.com/s-horiguchi/Yonmoku3D>

<https://www.bodoge-intl.com/strategy/rittai4ruleststrategy/>

<https://www.gurobi.com/resource/3d-tic-tac-toe-jupyter-notebook/>

<https://web.stanford.edu/class/psych209/Readings/MnihEtAlHassibis15NatureControlDeepRL.pdf>

<https://masamunetogetoge.com/gradient-vanish>