

ぶよぶよの連鎖構成法のための 機械学習と人間からの知識抽出に関する 基礎的研究

電気通信大学 情報理工学部 総合情報学科
橋山研究室

柴澤弘樹

背景：ゲームAIの現状

- AIがプロの人間プレイヤーに勝つ

- 機械学習AI

- 強い、事前知識不要
- × 学習結果の解釈が難しい

- ルールベースAI

- 処理の解釈、改良が容易
- × 知識のルール化が難しい



2つの手法を

「ふよふよ」に適用し

比較・検討

人の知識を適用したAI

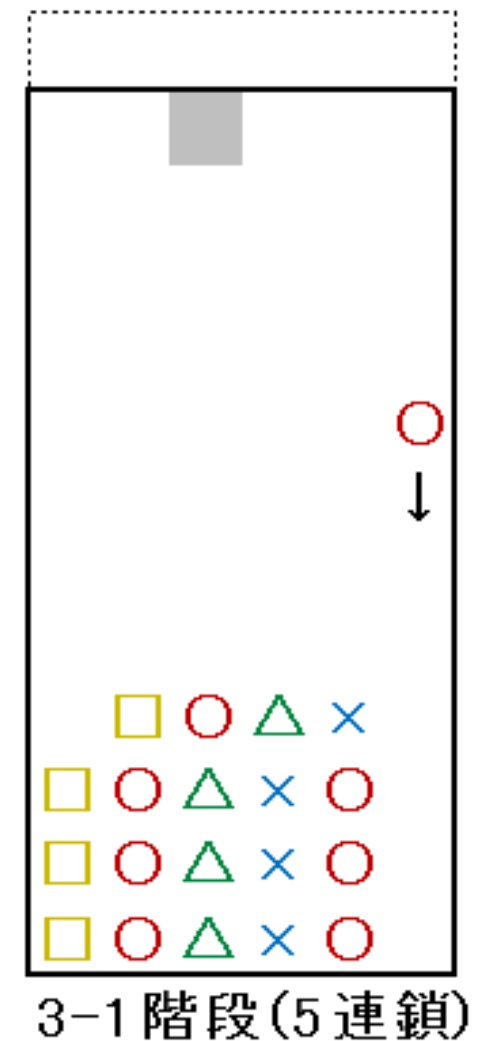
従来手法

ポテンシャル最大化法^{[1][2]}

- 見えている手のみ（3手分）を全探索
- 2,3手目のスコアを最大化

提案

3-1階段の構築ルールを適用
if-thenルールを書き下し



[1] 富沢大介, 池田心, シモンビエノ. 落下型パズルゲームの定石形配置法とぷよぷよへの適用. 情報処理学会論文誌, Vol. 53, No. 11, pp. 2560–2570, nov 2012.

[2] 大月龍, 前田新一, 石井信. 不完全情報ゲームに対する階層化したモンテカルロ探索とそのぷよぷよへの適用. 電子情報通信学会技術研究報告. NC, ニューロコンピューティング, Vol. 113, No. 500, pp. 275–280, mar 2014.

実験：DQN vs 人の知識AI

DQN^[3]

- 実装
DQN-Chainer^[4]
RLE^[5]
- 学習
ゲーム内AIと対戦
50000ステップ
× 100回

人の知識を適用したAI

- 画像認識による
入出力を実装

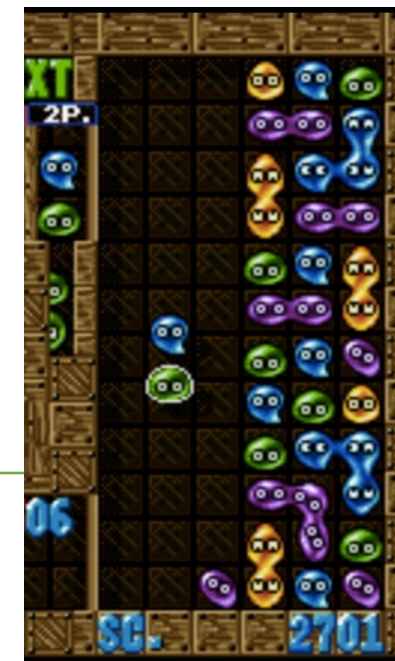


0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
5	4	0	0	0	0
5	4	1	0	0	0
5	4	1	0	0	0

- 連鎖発動のスコア
閾値2100点

ゲーム内AI

- 「のほほ」
- まぐれによる連鎖
- 時として5連鎖以上



[3] Volodymyr Mnih et al. Human-level control through deep reinforcement learning.
Nature, Vol. 518, No. 7540, pp. 529–533, 2015.

[4] <https://github.com/ugo-nama-kun/DQN-chainer.git>, Last Visited 2017/2/13.

[5] <https://github.com/nadavbh12/Retro-Learning-Environment.git>, Last visited 2017/2/13.