

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

柴澤弘樹 橋山智訓 田野俊一

背景・目的

ゲームAIの発展: 人に勝てるまでに

・機械学習AI

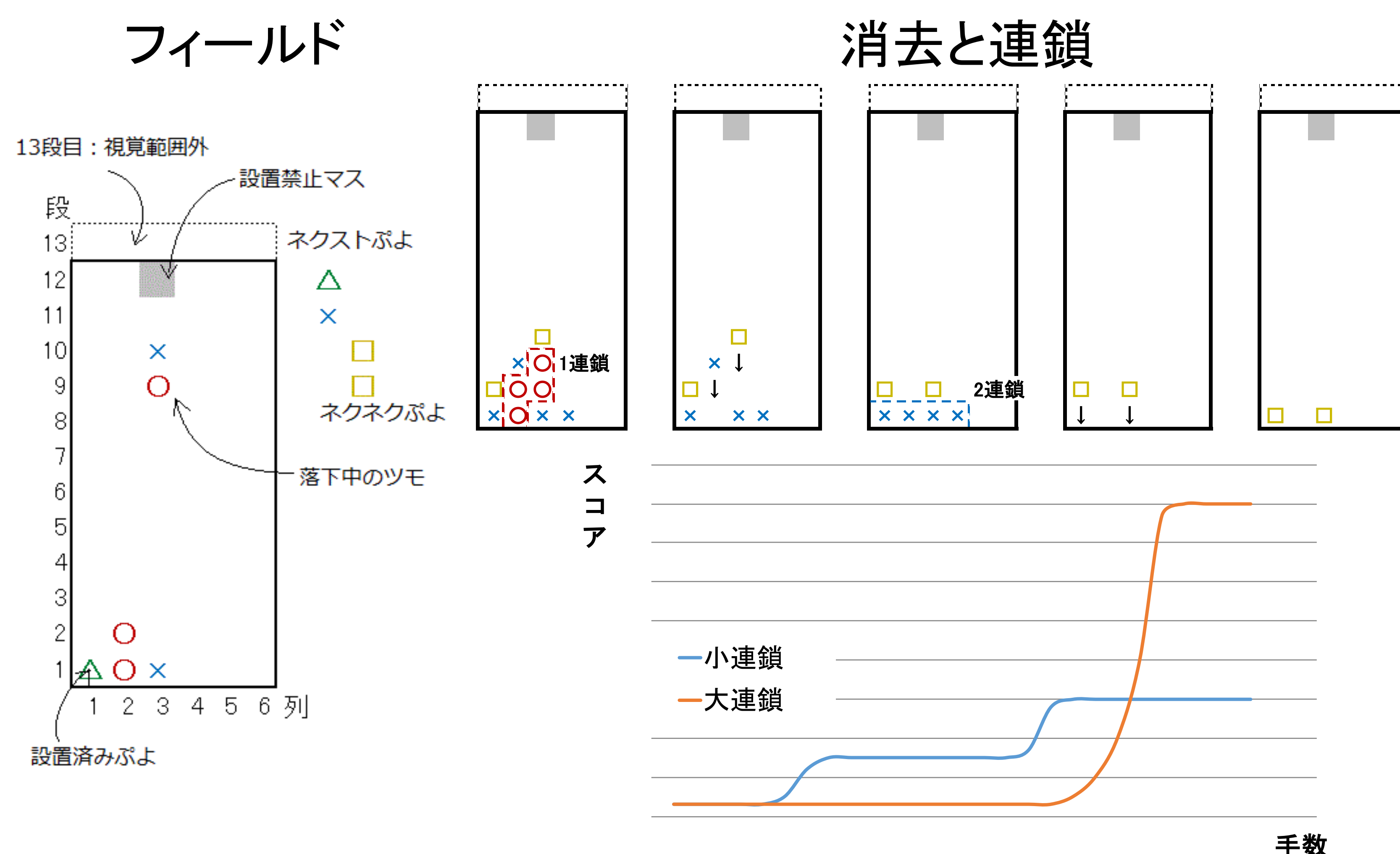
- 強い、事前知識不要
- × 学習結果の解釈が難しい

・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「ぶよぶよ」に適用し比較・検討

ぶよぶよのルール



限られた情報から連鎖構築する必要

人の知識を適用したAI

・ポテンシャル最大化法

見えている手のみ(3手分)を探索
1手目で消去する手を除外



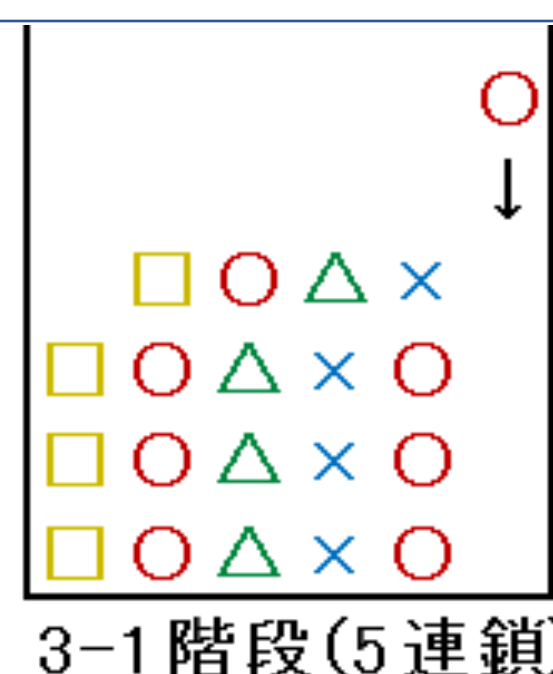
・全幅探索法

3手目を全ぶよに関して探索
1手目の消去を許可



・人の知識適用

3-1階段の構築ルールを適用
if-thenルールを書き下し



3-1 階段(5連鎖)

ツモがソロである:
土台が1つでツモが同色の列が存在 or
土台が0個で、隣列に同色の土台がない列が存在:
その列に縦置き
土台が3つ以下でツモが同色の列が存在し、かつ
右隣の列の土台が3つ以上で仕掛けが無い:
その列と右隣の列に横置き
その他:
左隣の列に立てて置く
ツモがソロでない:
ツモの一つと同色の土台が存在する:
ツモと同色の土台が隣り合って存在し、かつ
両土台が3つ以下:
色が揃うように横置き
左端列の土台が2つでツモに土台の色が含まれる:
左端列に縦置き
ツモと同色の土台が隣り合って存在し、かつ
右列の土台が2つ:
土台と仕掛けを完成させる縦置き
ツモの1つと同色の土台が存在し、かつ
その右の列の土台が3つ以上であって仕掛けが無い:
さらにその右の列の土台が(ツモの1つと同色or空):
仕掛けの列とその右の列に横置き
その他:
仕掛けの列に縦置き
ツモの1つと同色の土台が存在し、その右列が空列:
土台を1つ増やす横置き
その他:
隣り合う2列が空列:
その左の列の土台にツモの色が含まれる:
左の土台と繋げないように、空列に横置き
ツモがABBBである or
ツモがABACである or
ツモがABCCBBである:
空列に対し左がB、右がAとなるよう横置き
その他:
空列に対し左がB、右がAとなるよう横置き
その他:
左端の列に縦置き

DQN: Deep Q-Learning

ディープラーニング+Q学習

画面状態と報酬から、行動の評価であるQ値を学習

ゲーム内AIとの対戦を学習

- ・ 状態: ゲーム画像 4フレーム
- ・ 報酬: (自スコア-相手スコア)の変化
- ・ 50000ステップ × 100回

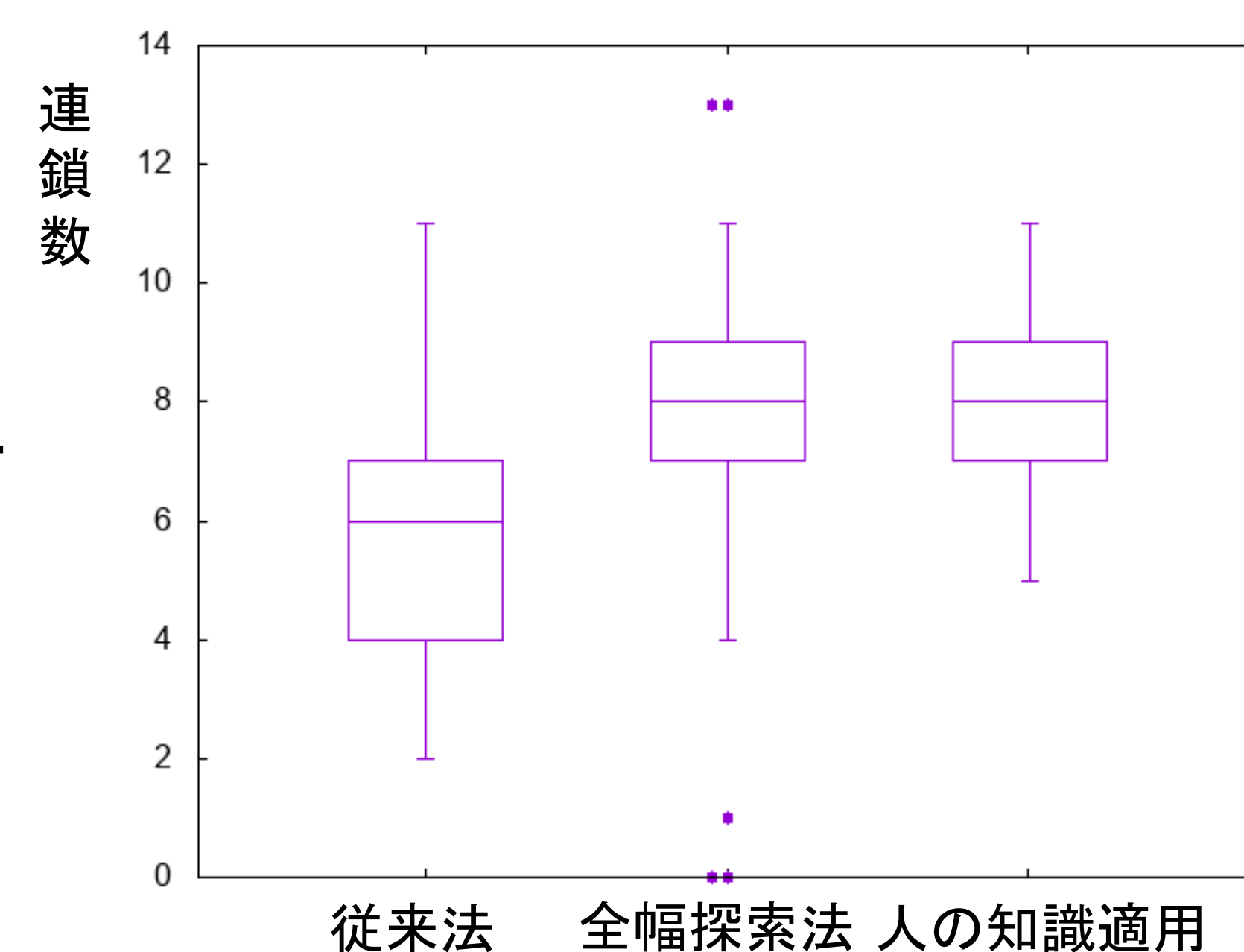


実験1: 連鎖構築シミュレーション

- ・ シミュレータを実装
- ・ 32手での構築連鎖
- ・ 同配石でそれぞれ50試行



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



全幅探索で連鎖数**向上**
人の知識適用で連鎖数が**安定**

実験2: DQN vs 人の知識を適用したAI

ゲーム内AI「のほほ」とそれぞれ対戦

勝利数		1試合あたり平均スコア	
	実装AI - のほほ	実装AI	のほほ
DQN	2 - 48	743.30	1790.12
人の知識適用AI	24 - 26	4762.52	2703.74

人の知識適用AIが強い

まとめ

・機械学習AI(DQN)

- 事前知識不要
- × 弱い、ルールの解釈が困難、改善方針が不明

・人の知識を適用したAI

- 強い、ルールの解釈が容易、結果の再利用が可能
- × 知識のルール化に手間

今後

機械学習と知識の**組み合わせ**を検討