

模倣学習エージェントを用いたゲーム戦略学習環境の構想

The Plan of Game Strategy Learning using Imitation Learning Agents

植野 雅之^{*1}, 和田 慎二郎^{*2}, 高見 友幸^{*1}

UENO Masayuki^{*1}, WADA Shinjiro^{*2}, TAKAMI Tomoyuki^{*1}

^{*1} 大阪電気通信大学 総合情報学部

^{*1} Osaka Electro Communication University, Faculty of Informatics

^{*2} プール学院大学 短期大学部

^{*2} Poole Gakuin Collage

Email: ueno@osakac.ac.jp

あらまし: ディープラーニング技術を用いることで人間の学習者がある状況のとき, どのような振るまいをするかを機械に学習させることで, その振る舞いを模倣するエージェントとして利用することが可能である. このような模倣学習エージェントは, 様々な教育的な利用が考えられる. 本発表ではこの模倣エージェントを論理的ボードゲームにおける戦略の学習のための教育環境として用いる構想について述べる.

キーワード: 模倣学習エージェント, 論理的ボードゲーム, ゲーム戦略学習, 模倣学習

1. 序

知識情報処理などの人工知能技術を教育システムに生かそうとする試みは, 知的教育システム (Intelligent Tutoring Systems: ITS) 研究として, 様々なシステムが提案・開発されてきた. しかし, 現時点の技術水準では, 人間の教師をシミュレートする人工知能を作ることは難しいと考えられるため, 機械知と人間知の非対称性を越えて刺激を与えることができる知的教育システムが求められており, 教育エージェント (Pedagogical Agent) として研究が進められている. これは, 一定の役割を担うことのできるエージェントにより, 教育の支援をおこなうという発想である.

本研究において提案するのは, 機械学習を用いて, 人間の学習者の振る舞いを模倣する能力を持つ教育エージェント「模倣学習エージェント」である (図1). このエージェントに学習者が持つ様々なスキルだけでなく, 学習者が持っている思い込みなど含めて反映させることができれば, 学習者にとって自分を振り返る教育的な意味での一種の「鏡」として機能させることができ, 自分が持っている戦略の問題点や見落とししていた視点, 思い込みなどに気付くことができるようになると考えられる. 特に相手が必要な学習, 例えば, 討論, 外国語会話, スポーツなどの対戦, ボードゲームなどの対局, 教師教育などの学習においては, このような教育エージェントを用意することによって体験的に学習することができるようになると考えられる.

2. ゲーム戦略の学習

本研究の対象領域としては, 「スポーツ」「討論」「外国語会話」等の領域への適用には技術的問題が多いと思われるので, 学校教科的な内容から外れることとなるが, 将棋・リバーシなどの論理的ボードゲームの戦術・戦略的スキルの学習を扱うこととし

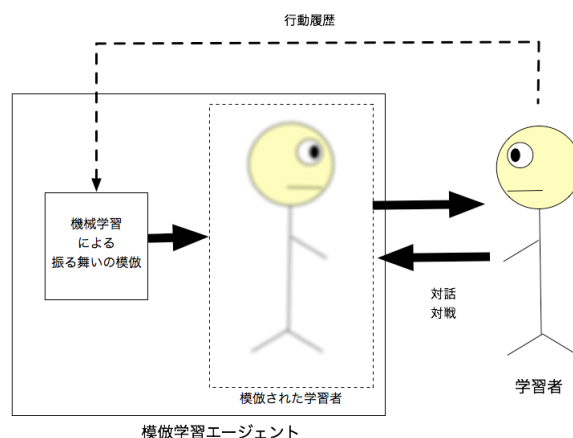


図1 模倣学習エージェント

たい.

ここで「論理的ボードゲーム」とは学術的には「二人零和有限確定完全情報ゲーム」と呼ばれるボードゲームであり, 両者の駒の位置などの勝敗に関わる情報が相互にオープンで, 有限の時間で勝敗が決着するタイプのゲームである. このようなゲームでは, 偶然の要素がないため, 論理的かつ空間的な思考力を鍛える知育として用いられることもある. このような論理的ボードゲームでは, 問題解決空間やそこでの行為が明確であるため, 振り返りなどもおこなやすい. また, 一般的な人間の能力を越えるレベルの AI プレイヤーが既に作られており, 技術的にも利用しやすいというメリットがある. さらにボードゲームなどの戦略的スキルを学習することを目的とする教育システムはあまり例が見当たらず, 新たな領域を切り開くものとして期待することができる.

3. 模倣学習の実現

将棋における模倣学習については, (2)(3)などにあ

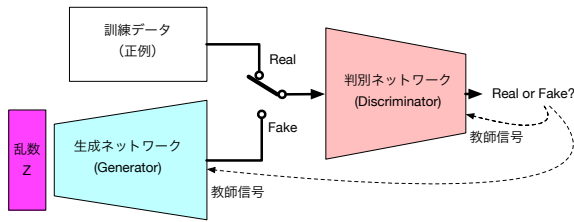


図 3 敵対的生成ネットワーク (GAN)

るようにプロ棋士の振る舞いを再現するレベルの模倣学習が実現されている。技術的には大量のデータを必要とする深層学習において、敵対的生成ネットワーク (Generative Adversarial Network : GAN) と呼ばれる技術を用いて、サンプルに類似する偽のデータを生成する生成ネットワークとその真偽を判定する判別ネットワークを競合させて学習をおこなうことで少数のサンプルからの学習を可能にしている (図 2)。(2)(3)では、棋譜の各手と生成ネットワークにより偽の手を生成したものを判別ネットワークによりその真偽を見分ける競合状態で両者のネットワークの学習を繰り返すことで実現している。

4. 模倣エージェントを用いたゲーム戦略学習のための教育環境

模倣学習エージェントをゲーム戦略を学習するための対戦環境として用いる方法としては、次の2段階が考えられる。

- [1]単独での利用
- [2]マルチユーザーでの利用

[1]単独での利用では、仮想的な自己対戦をおこなうことができる (図 1)。①仮想的な自己対戦により、自己が持つ戦略を立場を変えて見ることができ、様々な問題点を認識することができる可能性がある。

[2]マルチユーザーで利用する場合には、仮想的な自己対戦以外に②他人の模倣エージェントとの対戦 (図 3)、③模倣エージェント間の仮想対戦 (図 4)をおこなうことができる。人間同士の対戦では、単にスケジュールを合わせる以外に強いストレスを与える可能性があり、簡単におこなうことが難しいが、模倣エージェントとの仮想対戦ではこれらの負担を無くすことができる。さらに模倣エージェント間の仮想対戦により、大体の力量の差などを推し量ることが可能となり、仮想的なランキングを作るなど、人間間の対戦をおこなう際の参考にすることができる。

模倣エージェントの導入により、このような多様な形での対戦が可能となり、一般的な人間同士の対戦以外の様々な体験を得ることができる学習環境、教育環境として機能させることが可能になると考えられる。

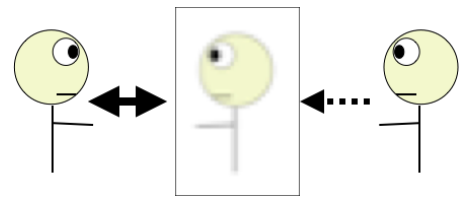


図 2 他者の模倣エージェントとの対戦

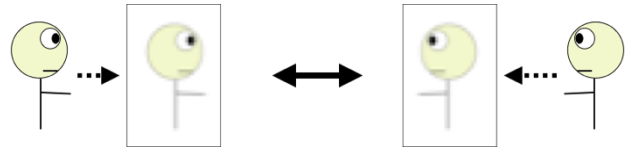


図 4 模倣エージェント同士の対戦

5. 結論と展望

技術的には3章に述べた方法により、特定プレイヤーの棋譜を元にそのプレイヤーの戦略的スキルを模倣する仮想プレイヤーを作り出すことができると考えられる。プロではないプレイヤーの棋譜に対して、この技術を適用する場合に考えられる問題点としては、(1)プロではないプレイヤーの打ち方はプロほど洗練されておらず、安定していないと考えられるため、機械学習がうまくいかない可能性がある点、(2)棋譜がそれほど多くないため、有効な機械学習が成立しない可能性がある点などがある。

一方で棋譜からそのプレイヤー自身と同程度の仮想プレイヤーを作り出すことができれば、4章で述べたようなゲーム戦略を学習する闘技的な学習環境を構築することができる。さらに記号処理的な AI 技術を用いて、模倣エージェントに新たな戦略を容易に付け加えて強化することができれば、戦略に関する試行錯誤をかなり効率化できるのではないかと考えている。また、機械学習されている戦略を何らかの形で抽出することができれば、初学者に対する明示的な指導に利用することも可能となると考えられる。今後の研究にて、実装・検証していきたい。

謝辞

本研究は JSPS 科研費 JP19K12260 の助成を受けたものです。

参考文献

- (1) エージェント技術の教育応用, 吉川, 高橋, 教育システム情報学会誌, Vol.35, No.1, 2018, pp5-12
- (2) Chidambaram, M. and Qi, Y.: Style Transfer Generative Adversarial Networks: Learning to Play Chess Differently, arXiv preprint arXiv:1702.06762 (2017).
- (3) Shanchuan Wan and Tomoyuki Kaneko. Style Transfer in Playing Shogi Based on Generative Adversarial Networks. GPW 2017, pp. 138-143(2017)
- (4) 人間プレイヤーの振る舞いを模倣する対戦エージェントの可能性, ゲーム学会第 17 回全国大会論文集, 2018, p.21-22