

完全情報下における戦略に基づいた ルールベースな HANABI エージェントの有効性調査

Investigation on rule-based agent for a card game, HANABI, under complete information

市来 正裕(Ichiki Masahiro)¹ 佟 菲(Tong Fei)² 中里 研一³

名古屋大学大学院情報学研究科¹ 京都大学大学院情報学研究科² ボッシュ株式会社中央研究所³

Abstract—人工知能と人間の協力による課題の解決には、「他者と協力する」という社会的知能が必要不可欠である。人工知能に社会的知能を獲得させるための研究に、HANABI という協力型カードゲームを用いたものがある。本研究ではこのゲームのいかなる局面においても、観測可能な情報から最善手を判断できるという仮説の下、予め考案した戦略に従ってゲームをプレイする、ルールベースなエージェントを開発した。ゲームに参加する全エージェントが同じ戦略を持つという状況を設定して実験を行い、得点及びゲーム終了までの過程を調査した。結果として現在の戦略による高得点の達成は困難であるとわかった。しかし戦略次第では得点が向上する可能性が示唆された。

1. はじめに

現在、ゲームをプレイする人工知能エージェントの開発が盛んに行われている。対戦ゲームの分野における、人間に勝てるエージェントを作るという目標は、アルファ碁などの登場により達成されつつある[1]。その一方で、協力ゲームの分野における、他者と協力して課題を達成するエージェントの開発はいまだ途上である。このようなエージェントの開発は、「他者の思考を読む」という人間らしさの構成要素の一つである社会的知能の獲得や、人間らしさそれ自体の解明に貢献すると考えられる。

本研究では、協力ゲームの題材として HANABI というカードゲームを採用し、高得点をとるエージェントの開発を目的とした。HANABI を選んだ理由としては、このゲームが不完全情報ゲームであることや、自然言語処理による発言の意味理解が不要なことなどが挙げられる。不完全情報ゲームでは、各プレイヤーの確認可能な情報が異なっており、完全情報ゲームの将棋や囲碁などとは性質が異なる。また、このゲームではプレイヤー間の意思疎通がトークンによって厳しく制限されており、チープトークなどから

プレイヤーの意図を推測する必要がない。

本稿では HANABI をプレイするエージェントをルールベースで実装し、その性能の評価及び高得点獲得のための作戦の検討を行った。

2. 先行研究

大澤は HANABI の 2 人プレイを想定した戦略を考案し、その戦略に基づいた行動をとるエージェントを実装した[2]。この戦略では、お互いのエージェントが自分と同じ戦略に基づいて行動するという設定の下で自分の状態推定を行い、その有効性を示した。佐藤らは[2]の結果を踏まえ、2 人のプレイヤーの内、一方を人間に変更し、さらにエージェントの戦略の指標として思考時間を利用した実験を行った。人間は行動に自信があるときに思考時間が短く、自信がないときに思考時間が長いという仮定の下で思考時間の閾値を決定し、エージェントはその時間に応じて戦略を選択した。思考時間に関する仮説は支持されたが、思考時間閾値による性能向上は確認できなかった。

本研究では、[2]の考え方にに基づきさらに得点調査及び得点向上方法の検討を行った。

3. ゲーム戦略

本研究でエージェントに実装したゲーム戦略を示す。手札に色と数字の両方が判明しているカード(判明カード)がある場合とない場合で行動を定めた。判明カードを所持しており、それを場にプレイできる場合はそのカードをプレイ、そうでないならカードの破棄もしくはヒントの提示を行う。判明カードを所持していない場合は、ヒントの提示もしくはカードの破棄を行う。

3.1 ヒントの提示

ヒントは、誰に、どの色、どの数字、の 3 要素からなる。ここでは、場に出せるもの、さらに高得点を狙えるものを優先することとした。ただし、実装においては簡単のため、カードの

数字を 1 とそれ以外に分類して次のようなルールを定めた。まず、場にプレイできるカードを所持するプレイヤーを手番の早い順に見つける。見つけたプレイヤーのカードから、場にプレイできるもの、さらに 1 以外のカードを優先して選択する。これは、1 のカードが最も多くゲームに含まれているためである。その後、選んだカードの数字もしくは色のうち、まだ提示されていないものをヒントとして提示する。どちらも提示されていない場合はランダムに決定する。

3.2 カードの破棄

カードを破棄する場合、まず判明カードから捨ててよいカードを選ぶ。捨ててよいカードとは、既に場にプレイされたカードと同種類、もしくは複数枚の同じカードの存在が確認できるものである。判明カードを所持していない場合はランダムに選択し、破棄する。

3.3 カードのプレイ

1 以外のカードを優先するが、プレイできるカードの数字が 1 のみである場合、他のプレイヤーの手札に最も多く登場する色を優先する。

4. 実験

4.1 完全情報ゲームとしての HANABI

比較対象として自分の手札を確認できるルールに変更した場合の実験から行った。この条件下ではヒントが必要ないため、ヒントの提示は手番のスキップとした。1000 回のゲーム試行結果を図 1 に示す。平均点は 24.1 点であり、本戦略の有効性が確認できた。

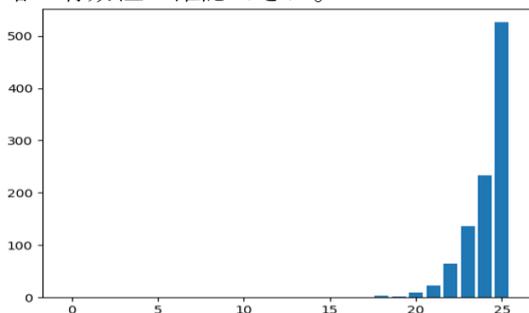


図 1: 完全情報下における得点分布 (横軸が得点, 縦軸が出現回数)

4.2 不完全情報ゲームとしての HANABI

HANABI 本来のルールである、自分の手札を隠した状態での実験を行った。1000 回のゲーム試行結果を図 2 に示す。平均点は 11.6 点である。完全情報ゲームの場合と比べ、大幅な得点低下が確認された。

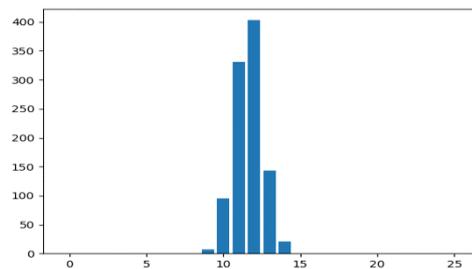


図 2: 不完全情報下における得点分布 (縦軸が得点, 横軸が出現回数)

4.3 青コイン枚数の影響調査

次にヒントの効率性を調べるためにヒント回数を制限する青コインを、 $16 \cdot 24 \cdot 32$ 枚に変更して実験を行った。結果の平均点及び、ゲームの全行動ターンに対してランダムなカード破棄の回数の割合の変化を図 3 に示す。青コイン数と得点、ランダムなカード破棄の回数割合に相関が確認できた。つまり、ヒントの出し方及びカード破棄方法が改善されればさらなる性能向上が見込まれる。

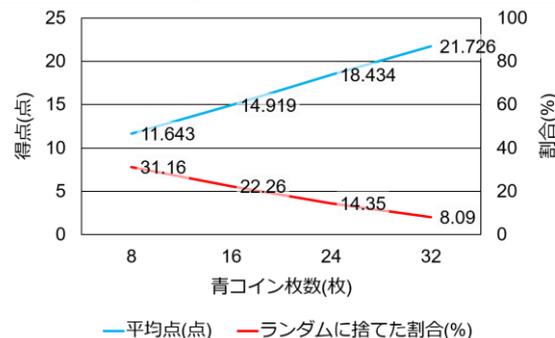


図 3: 平均点とランダムカード破棄割合の推移

5. 結論

本研究では、協力ゲーム HANABI をプレイするエージェントをルールベースで実装し、その得点力の調査と性能向上のための検討を行った。現状のルールでの高得点は難しいものの、カード破棄やヒント提示などに改善の余地がある。

今後の課題としては、色か数字片方だけの情報からのカード推測、ヒント対象となるカードを増やすことが挙げられる。また、人間を混ぜた場合の得点力変化や、人間によるエージェント評価等を調査する予定である。

- [1] David Silver et al., “Mastering the game of Go without human knowledge,” nature, 2017.
- [2] 大澤 博隆, “協力ゲーム Hanabi を用いた他者行動からの自己状態推定手法の検討,” Human-Agent Interaction Symposium, 2014.