

他者視点の推定を用いたエージェント評価の向上: カードゲーム Hanabi の AI との対戦分析から

Increasing Agent's Impression with Estimation of Other's Viewpoint:
From Plays of Card Game Hanabi with AI

加藤拓也¹ 大澤博隆¹

Takuya Kato¹, Hirotaka Osawa¹

¹ 筑波大学

¹University of Tsukuba

Abstract: AI research on incomplete information games has aspects that imitate human intelligence. The author chose a card game called Hanabi as an evaluation of imitating human intelligence with AI. Hanabi is a cooperative and incomplete information card game, and this game has a unique feature that a player cannot observe his / her own cards. Player selects whether to build a set of cards or discard the card while providing hints of the card with the cooperator. In the previous research, it was shown that the imitation of human behavior that corrects incomplete information increases the score in Hanabi play between agents. The author evaluated agent's function to modify incomplete information based on behavior of cooperator in the game with human and agent. I experimented with human and two kinds of agents with a difference whether to imitate modify incomplete information as a cooperator of experiment participants and I analyzed the game results and impression evaluation of participants. The result is that while the imitation of reflex intelligence is not effective for increasing the score, the participant has a good impression on the agent when the agent modifies the incomplete information properly, and the difference with the case of human collaborator was not seen.

1. はじめに

ゲームの人工知能研究の分野では、強いアルゴリズムの開発の他に、人間を楽しませる、人間に似たふるまいをするエージェントも研究対象となっている。多くの盤上ゲームで人間に勝てるアルゴリズムの開発が達成されつつある現在、人間らしいエージェントの追及が今後の焦点になるとされている[1]。

プレイヤーがゲームの全ての情報を知ることができる完全情報ゲームでは、人間を楽しませるエージェントは人間に合わせて強さを調整する手法を取っている[2]。一方で不完全情報ゲームはプレイヤーが知ることができない情報が存在するようなゲームである。不完全情報ゲームでは協調した意思決定[3]、ノンバーバル情報の提示[4]、人間の手からの学習が有効である[5]。

こうした中、本研究では他者視点の推定がエージェント評価の向上にどのように影響するか評価した。題材として Hanabi というカードゲームを課題とした。Hanabi の特徴として、自身の手札を見ることができない点と協力型のカードゲームであるという点が挙げられる。自身の手札の情報を持たないカード

は相手の情報提示から、相手の視点を再現し、手札を推定し得点を得ることになる。他者視点をシミュレートし不完備情報を推測することで、プログラム同士での得点が向上することは我々の前の研究で分かっているが[6]、本研究では、こうした他者視点推定がゲームを行った人間にどのような印象を与えるかを検証する。評価実験ではゲームの相手として他者の観測と行動のシミュレーションの機能の有無に違いがある 2 種類のエージェントと人間の 3 条件を設定し、実験参加者が抱く印象とゲーム結果への影響を分析した。本研究ではエージェントの相手は人間のため、相手のモデル化の誤りが生まれる。そこで、他者の観測と行動のシミュレートが高い確率で成功したゲームに限定した分析を行った。

2. Hanabi のモデル

Hanabi は 2 人から 5 人のプレイヤーによって行われる協力型のカードゲームである。本研究ではプレイヤーの人数は 2 人とした。50 枚のカードと情報公開のためのトークン 8 つを用いてゲームを行う。カードには 5 種類の色(白・赤・青・黄・緑)及び 1 から 5 の数(色や数をカードの要素と呼ぶ)が割り当てられ、

各色で数が1のカードは3枚、2から4のカードは2枚、5のカードは1枚存在する。色ごとに数字の昇順にカードを並べていき(並べた山を花火という)、全色で数字の5まで花火を完成させることがゲームの目標となる。プレイヤは5枚のカードを手札として保有する。各プレイヤは自身の手札のカード要素を見ることはできず、相手の手札のカード要素を見ることができる。残りの40枚は山札に積まれる。花火は全色で0から、情報トークンは8枚の状態からゲーム開始となる。プレイヤには交互にターンが回り、各ターンにおいて、プレイヤは情報提示・破棄・プレイの3種類のアクションのうち1つを選択する。

情報提示では相手手札にあるカード要素を1つ選択し、該当のカードを相手に教えることができる。相手の手札に存在するカード情報のみを教えることができ、また該当するカードは漏れなく教えなければならない。従って情報を教えられなかったカードはその情報に該当しないことが求まる。例えば、手札の1,3番目が白であると情報提示された場合、2,4,5番目のカードは白でないことがわかる。また、情報提示では情報トークンを1枚消費する必要がある。従って情報トークンが0枚のターンではプレイヤは情報提示を行うことができない。

破棄では、プレイヤは自身の手札からカードを1枚選択し捨て、捨てたカードの代わりに山札から1枚カードを補充し手札に加える。その際、情報トークンを1枚増やす。情報トークンが8枚のターンでは、カードを破棄しても情報トークンの数は変化しない。破棄されたカードの要素は全プレイヤに公開される。

プレイでは、プレイヤは自身の手札からカードを1枚選択し花火に並べることを試みる。プレイしたカードの数字が現在の同色の花火の数字より1だけ大きい場合、プレイは成功となる。この時、プレイしたカードは該当色の花火に並べられ、花火の数が1増加する。プレイに失敗した場合、プレイしたカードは破棄される。プレイの成功・失敗に関わらず、プレイしたカードは全プレイヤに公開される。

ゲームの終了条件には以下の3種類がある。

- 3回プレイに失敗した場合
 - 山札のカードが0枚になった後、全プレイヤがアクションを1回ずつ行った場合
 - 全ての色の花火が完成した場合
- ゲームが終了した時点での全色の花火の数の総和が得点となり、25点が最大となる。

3. アルゴリズム

本研究のエージェントの行動は以下の7種の順番

で表される。

1. プレイ可能カードの情報の提示

相手がプレイ可能カードを持つことが分かり、そのカードの情報が不完備な場合、情報提示を行う。数も色も情報が不完備な場合、教える情報をランダムに選択する。

2. 破棄可能カードの情報の提示

相手が破棄可能カードを持つことが分かり、そのカードの色か数を提示することで破棄可能カードと断定できる場合、情報提示を行う。

3. 破棄可能カードの破棄

破棄可能を断定できるカードが自身の手札に存在する場合、そのカードを破棄する。

4. ランダムなカードの破棄

カードの情報が最も少ないカードを破棄する。カードの情報が最も少ないカードが複数枚の場合、その中からランダムにカードを選択し破棄する。

5. プレイ可能カードのプレイ

H_{pi} と D を比較してプレイ可能カードが自身の手札に存在する場合、そのカードをプレイする。

6. 数の確定情報からプレイ可能カードをプレイ

H_{pi} で色が不明だが、 I_{pi} からプレイ可能カードであることが分かるカードをプレイする。

7. 自己推定に基づいてプレイ可能カードをプレイ

自身の手札を推定した場合、 H_{pi} , I_{pi} は異なる情報を表す(例えば、 H_{pi} において色が確定しているが

I_{pi} においては複数の色の可能性が残る)。プレイされるカードのうち、情報の相違が見られるカードをプレイする。

シミュレーションと自己推定を適用したエージェントは基本的には確定的な情報のみに基づきアクションを行うが、自己推定を行う際にのみ、確率論的戦略を用いる。推定戦略の比較として、3.1のエージェントから推定機能を除いたエージェントを用いた。このエージェントは確定的な情報のみに基づきアクションを行う。

4. 他者視点推定の影響評価

本研究では相手による参加者の行動と印象の影響を調べるため、ゲームの記録の他に参加者の様子を記録する動画撮影と印象評価アンケートを各参加者に対して行った。アンケートでは以下のような質問項目を用いた。

- 相手についての印象

Q1. 意思を持っていると感じたか

- Q2. ゲームに慣れていたか
 Q3. 親しみやすかったか
 Q4. 自分に配慮していたか
 Q5. 賢かったか
 Q6. 相手は自分の思い通りに行動していたか
 Q7. 相手の行動の意図が分かったか
 ● ゲームについての印象
 Q8. 得点を上げようと心掛けたか
 Q9. ゲームはやりやすかったか
 Q10. 失敗した行動があったか
 Q11. 主体的にゲームを進めたか
 Q12. ゲームの結果に満足しているか
 Q13. 自由記述欄

Q1 から Q5 は Bartneck らの The Godspeed Questionnaire Series[以下 GQS]を参考にした[7]。ゲーム相手について、Q1, 2 は擬人化、Q3, 4 は可愛さ、Q5 は知性を測定する項目である。Q6, 7 はゲーム相手との意思の疎通が図れたかを問う項目である。Q8 から Q12 は異なる相手とゲームを行った際のゲームの印象についての項目も用意した。全ての質問は Likert scale に従って 7 段階で評価した(そう思わない 1-7 そう思う)。各ゲームが終了した毎にその時のゲームについて回答してもらった。また、実験の最後にインタビューを行った。

4.2. 参加者と実験手順

実験には Hanabi のゲーム経験がない 12 名が参加した。12 名はいずれも 20 代男性の大学生または大学院生である。参加者は 2 人 1 組で実験を実施した。参加者 2 人は Hanabi のルールとインタフェースの操作方法について説明を受けた後、各々別の部屋で実験を行った。参加者はゲームを 6 回を行い、最初の 3 回は自己推定戦略に基づくエージェントとのゲーム(以下 R 条件)、決定論的戦略に基づくエージェントとのゲーム(以下 C 条件)、参加者同士でのゲーム(以下 I 条件)の 3 条件のゲームを 1 回ずつ行った。次に、前半 3 回の条件と同じ順番で 1 回ずつゲームを行った。実験に用いる山札は 6 種類用意し、1 人の参加者が 6 種類の山札で 6 回ゲームを行うようにした。また、参加者のゲームの習熟効果を考慮し、条件の順序を全参加者で均等に割り振ることでカウンターバランスを取った。1 回のゲームが終わる毎にアンケートを行い、6 回目のアンケートの回答とインタビューを終え次第、実験を終了した。

本実験では以下を仮説とした。

1. 自己推定戦略に基づくエージェントは決定論的戦略に基づくエージェントに対して Q1, 2 の評価が高くなること。

2. 自己推定の成功率が高い場合、Q3 から Q7 の印象評価が高くなること、及び得点が上昇すること。

5. 結果と考察

本実験では、R 条件でのゲームを 24 回行ったが、カードの推定は 1 ゲームで平均 5.6 回(SD 3.4)行われ、そのうちカードを正しく推定した割合は 1 ゲームで平均 63%(SD: 24)であった。また推定したカードのプレイは 1 ゲームで平均 3.7 回(SD: 2.4)行われ、そのうちプレイに成功した割合は 1 ゲームで平均 63%(SD: 38)であった。R 条件のゲームにおいて、情報を推定したカードをプレイした回数が 0 回のゲームが 1 件見られたため、検定から除外した。

また、本研究では自己推定戦略における相手のモデル化及び推定が正しく行われた場合を想定するため、R 条件のうちエージェントの他者視点推定の成功率が高い条件(以下 R(S)条件)、C 条件、I 条件と印象評価、得点、思考時間の影響を調べるため、多重分析を行った。R 条件のうち、自己推定で要素を推定したカードをプレイした際のプレイ成功率が 0.5 以上のゲームを相手のモデル化及び推定が正しく行われたゲームとし R(S)条件に分類した。実験の結果、24 回の R 条件のゲームのうち、17 回のゲームが R(S)条件に分類された。

3 条件の各評価項目の差について Bonferroni 法による多重比較検定を行った結果、全ての評価項目は R 条件と C 条件の間で有意差は確認できなかった。続いて、自己推定戦略に基づくエージェントを成功率の高いエージェントに限定した場合の得点とし、検定を行った。結果、印象評価項目の Q3,4,5,6 は R(S)条件と C 条件の間で有意差が確認された。また、R(S)条件と I 条件の間においても有意差が確認できたが、C 条件と I 条件の間では有意差がなかった。

表 1 R(S)条件、C 条件、I 条件間の検定結果

	平均値の差		
	R(S)条件 -C 条件	R(S)条件 -I 条件	C 条件 -I 条件
Q3	0.941*	-0.088	-1.029*
Q4	0.912*	-0.353	-1.265*
Q5	1.059*	-0.235	-1.294**
Q6	1.118**	-0.088	-1.206*

**: $p < 0.01$, *: $p < 0.05$

評価実験の結果、印象評価項目の Q3 から Q6 については R(S)条件と C 条件の間、C 条件と I 条件の間に有意差または有意傾向が見られた一方で、R(S)条件と I 条件の間には有意差は見られなかった。

GQSにおいて、Q3, 4はエージェントの可愛さ、Q5はエージェントの知性に関するアンケート項目である。このことから、エージェントが他者視点の推定を正確に行うことでプレイヤはエージェントに人間の相手と同程度に親近感や知性を感じることが判明した。また、他者視点の推定を正確に行う場合、エージェントはプレイヤの希望に人間の相手と同程度に応えることができるといえる。

印象評価項目と推定を行ったカードのプレイ成功率の関係を調べるため、R条件について相関分析を行った結果、Q1からQ7はプレイ成功率との間に正の相関が見られた。自己推定戦略において他者視点の推定をエージェントが正確に行うことでプレイヤはエージェントに人間らしさ、親近感、知性を感じ、また意思疎通が上手く行われたと感じる。

自己推定戦略の決定論的戦略に対しての得点の有意性は確認できなかった。そこで、得点と他の評価項目の関係を調べるため、R条件について相関分析を行った。その結果、得点と推定したカードのプレイ成功率に正の相関が見られた($r=0.662$, $p<0.01$)。このことから、自己推定戦略において他者視点の推定をエージェントが正確に行いことで、得点が有意に上昇することがわかった。

参加者の平均思考時間は自己推定戦略の決定論的戦略との間に得点の有意な差は見られなかった。そこで、平均思考時間と他の評価項目の関係を調べるため、R条件とC条件各々について相関分析を行った。分析の結果、自己推定戦略に基づくエージェントとのゲームでは、平均思考時間と印象評価項目Q11(主体的にゲームを進めたか)の間に正の相関が見られた($r=0.533$, $p<0.01$)。また、決定論的戦略に基づくエージェントとのゲームでは、平均思考時間と印象評価項目Q10(失敗した行動があったか)の間に負の相関が見られた($r=-0.412$, $p<0.05$)。このことから、参加者はエージェントとのゲームにおいてアクションを考える時間が長いほど、自分のアクションに自信を持つこと、ゲームに貢献したと感じることがいえる。

6. 結論

本研究では、相手の観測情報のシミュレーションから自身の不完備な情報を推定するという認知機能を模倣することでエージェントが人間に与える印象を検証した。協力ゲームHanabiを用いた調査の結果、エージェントの自己の情報推定が正確である程、相手はエージェントに対して人間らしさ、親近感、知性、意思疎通能力を感じ、高いゲーム得点を得ることが分かった。特に、エージェントの自己の情報推定が高確率で成功した場合、相手がエージェントに

感じる親近感、知性、意思疎通能力は人間同士のゲームで相手に感じる印象と同程度に高いことが確認できた。しかし、相手の行動のモデルと人間の行動に相違がありエージェントが誤った推定を行う場合、人間がエージェントに対して人間らしさを感じなくなるという結果も得た。

今後の課題として、エージェントと人間の行動に相違がある場合においても、人間がエージェントに対して人間らしさを感じさせることが必要になる。相手の行動のモデルを人間がゲームを行う際の行動に近づけ、エージェントの情報推定の正確性を高めることでこの課題は達成できると期待される。

謝辞

本研究はJSPS 科研費 JP26118006 の助成を受けたものです。

参考文献

- [1] 松原仁, “豊かな体験をつくるエンタテインメントコンピューティング技術: 2. エンタテインメントと人工知能 -思考ゲームを例として-,” 情報処理, vol. 58, no. 1, pp. 14–17, Dec. 2016.
- [2] 仲道隆史, 伊藤毅志, “プレイヤの技能に動的に合わせるシステムの提案と評価,” 情報処理学会論文誌, vol. 57, no. 11, pp. 2426–2435, 2016.
- [3] 大曾根圭輔, 鬼沢武久, “ユーザに親近感を持たせるポーカーパートナーエージェント,” 知能と情報, vol. 21, no. 6, pp. 1127–1142, 2009.
- [4] 小林優, 大澤博隆, 稲葉通将, 篠田孝祐, 鳥海不二夫, 松原仁, 片上大輔, “擬人化エージェント用いた議論の雰囲気を実現する人狼対戦システムの開発,” in HCGシンポジウム2014, 2014, p. A-5-3.
- [5] 柘刈哲彦, 伊藤毅志, “手の流れを考慮して自然な手を選ぶ将棋AIの試作,” 研究報告ゲーム情報学 (GI), vol. 2015-NaN-33, no. 12, pp. 1–8, Feb. 2015.
- [6] H. Osawa, “Solving Hanabi: Estimating Hands by Opponent’s Actions in Cooperative Game with Incomplete Information,” AAAI Work. Comput. Poker Imperfect Inf., pp. 37–43, 2015.
- [7] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, “Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots,” Int. J. Soc. Robot., vol. 1, no. 1, pp. 71–81, 2008.