

協力ゲーム Hanabi を用いた協調行動のモデル化 — 共通経験の蓄積に関する検討 —

Effect of accumulating successful experiences and using common examples in Cooperative Game Hanabi

桑原 涼香¹, 長島 一真¹, 森田 純哉¹, 川越 敦², 大澤 博隆³

Ryoka Kuwabara, Kazuma Nagashima, Junya Morita, Atsushi Kawagoe, Hirotaka Osawa

¹ 静岡大学, ² 筑波大学, ³ 慶応義塾大学

Shizuoka University, Tsukuba University, Keiogijuku University

kuwabara.ryoka.19@shizuoka.ac.jp, nagashima.kazuma.16@shizuoka.ac.jp, j-morita@inf.shizuoka.ac.jp,

kawagoe.atsushi.tj@alumni.tsukuba.ac.jp, hiro.osa@gmail.com

概要

本研究では、協調行動の分析に適したカードゲーム「Hanabi」を題材としたシミュレーションを行った。ゲーム中にコミュニケーションが成功した際、自他の行為を事例として蓄積し、それを利用することで他者の行為の意図を推定する認知モデルを構築した。2体のモデルが対戦するシミュレーションの結果、試行の進行に伴う事例の蓄積・利用、およびゲーム得点の上昇が確認された。

キーワード：認知モデル, 協調課題, Hanabi

1. はじめに

人間は、直接的な指示がなくともメッセージに内包される意図を推測し、他者との協調を成功させることができる。こういった成功の背後には、他者の意図を心的にモデル化する能力が必要である。認知科学の広い分野において、人間同士のインタラクションを支えるこの種の社会的知能は「心の理論」[11]や「他者モデル」[12]などの概念によって説明されてきた。これらの原理/機構を理解することで、協調作業の支援やコミュニケーション障害に関する臨床的な応用が導かれると考える。

認知科学において、人間のコミュニケーションに関する研究は、様々なアプローチによって行われている。そのなかでも、計算機モデルを構築することによるシミュレーションは、心の理論や他者モデルなどの言語で説明される現象に対して、分解された処理レベルでの表現を導くという点で有用である。

特に本研究では、意図伝達の背後にある処理として、共通事例の蓄積と利用を想定する。他者とのコミュニケーションのプロセスを履歴として保持し、その履歴を参照することで他者からのメッセージを解釈し、他者へメッセージを送信する。送信者と受信者が想定す

る事例が共通したときに、コミュニケーションは成功し、2者間での協調が成功することになる。

本研究では、上記の処理を課題レベルで具体化するため、協調行動の分析に適した題材である協力ゲーム「Hanabi」を扱う。そのうえで、「共通の経験を蓄積することによる協調の成功」に関わる計算機モデルを構築する。本研究で構築するモデルは、シンボリックな事例の利用とサブシンボリックな事例の運用を含むものであり、既存の認知アーキテクチャ [1] を用いて構築される。構築されたモデルによって、事例の蓄積・利用により、協調に関わるゲーム得点の上昇が見込めることを確認する。

2. 関連研究

人間による複雑なコミュニケーションを捉えるために、抽象化された状況での行動を計測する研究が行われている。その1つの手法が、日常で利用されるコミュニケーションの手段が制限された実験室において、コミュニケーションの生起を観察する実験記号論である [3]。

実験記号論に立つ研究のなかで、金野 [5] は、図形を組み合わせたメッセージを送信しあうことで、グリッドワールドにおける共通の位置で2者が待ち合わせをする課題（メッセージ付き協調ゲーム）を考案している。この課題では、待ち合わせの試行を繰り返す中で、意味を有さなかったメッセージに意図がのせられ、協調課題に成功するようになる。この課題の成功の背後にあるメカニズムを探るための脳波計測 [6, 7]、あるいは個人特性との関連に関する検討 [6, 7] が行われている。さらに、それらの行動の背後に存在する認知プロセスを検討する計算機モデルが構築されている [8]。

森田ら [8] の構築したモデルは、メッセージ付き協

調ゲームにおけるコミュニケーションの成功事例を蓄積し、蓄積された成功事例を現在の状況に適用する。現在の自分の位置、あるいは他者から送信された図形を手がかりとして過去の事例を検索し、検索された事例と同様のメッセージを送信し、グリッドワールドにおいて移動する。この手続きを繰り返すことで、モデルは相手との協調（待ち合わせ）に成功するようになる。なお、このようなインタラクション場面における事例の利用は、他の研究においても用いられている [2, 4]。これらの研究において、様々なインタラクション課題における事例の検索や利用の仕方が検討されている。

しかし、現実社会に適用可能な知見を得るためには、上記の検討に比べ、より複雑な状況におけるモデルを構築する必要がある。意図の伝達に一定の困難さが存在する状況での検討により、コミュニケーションに内在する処理に関するより高い粒度での知見が得られる。この件に関して、一般に市販されるカードゲームを用いることは、上記のような実験室実験用の課題に比べ、有用なテストベッドになりうる。カードゲームには、明確なルールが存在しつつ、プレイヤーが楽しみ競うための一定程度の複雑さが存在するからである。

協調をプレイヤーに要求するカードゲームとして、Hanabi を用いた研究が存在する。大澤による Hanabi のモデルは、他者行動からの自己推定戦略のような高度な心の理論を示している [10]。また、大澤によるモデルは、人と対戦するエージェントに実装され、エージェントの振る舞いのバリエーション（リスク志向の個人差、反応時間など）が、対戦する人間が受け取る印象の変化を検討している [9]。しかし、Hanabi を題材とした研究において、協調に至るプロセスの検討は未だなされていない。そこで、本研究では、上記で述べた共通経験の蓄積を利用することで、Hanabi における協調の成立プロセスをモデル化する。

3. 課題

本節では本研究において採用する課題である Hanabi について説明する。この説明を行うなかで、Hanabi におけるコミュニケーションの特質を具体化する。

3.1 Hanabi ルール

3.1.1 構成要素

協力型ゲーム Hanabi は通常 2 から 5 人のプレイヤーで行われ、本研究では 2 人のプレイヤーとし、2

体のモデルによるシミュレーションを行う。構成要素は 50 枚のカード、8 個の青トークン、3 つの赤トークンである。カードは 5 色（白、赤、青、黄、緑）あり、各色について 1 が 3 枚、2 が 2 枚、3 が 2 枚、4 が 2 枚、5 が 1 枚ある。はじめに各プレイヤーへ 5 枚の手札が配布され、残りの手札は山札となる。青トークンは、情報提供を行うことのできる最大数が 8 回に制限されている事を示すカウンタであり、全プレイヤーが共有する。赤トークンは、誤った色あるいは数字を場に出す事で追加されるエラーカウンタである。

本ゲームのゴールは、場に同色のカードを 1 から 5 まで昇順に重ねて異なる 5 色の花火を完成させる事である。ゲーム終了までに可能な限り得点を高く積み上げる事を目指す。Hanabi のゲーム終了条件は以下の 3 つである。

- 場に誤ったカードを 3 回出した（赤トークンが 3 つ追加された）場合
- 山札が 0 になってから 1 ターンが経過した場合
- 全ての色について 5 まで並べた場合

3.1.2 プレイヤーの行動

各プレイヤーは、各ターンにおいて次の 3 つの行動のうちいずれかを選択する。

1. 相手へのヒント提供 (情報提供)

この行動は青トークンを 1 つ消費し、プレイヤーは相手の持つ手札の色か数字の情報を教えることができる。教える色あるいは数字が相手の手札に複数ある場合、その色または数字の手札全ての場所を教えなければならない。例えば、相手プレイヤーが左から赤 1、白 2、白 4、緑 3、黄 2 という手札を持ち、2 の情報を教える場合について考える。白である 2 枚目と 5 枚目両方の位置を教えなければならない、2 枚目のみ、5 枚目のみという教え方はできない。青トークンが 0 のときこの行動は選択できない。

2. 手札のプレイ

この行動は自分の手札 1 枚を場に出すことができる。既存の同色の花火より 1 つ大きい数字をそこに重ねた場合は成功である。但し、完成させることができるのは 1 つの色につき 1 つの花火のみであり、既に 1 から 5 まで積みあがっている色の手札を出すことはできない。場の花火に繋がらない異なる色あるいは 1 つ大きい数字以外の数字を出した場合は失敗となり、赤トークンが 1 つ追加される。例えば、場に白の花火が 1 から 2 ま

での数字で構成されている場合、白3を出せば成功、白4や青3をだせば失敗となる。失敗時に出したカードは破棄され、ゲーム終了まで使用されない。プレイ後は山札から1枚引いて手札に加える。

3. 手札の破棄

この行動は自分の手札から1枚を選んで破棄し、青トークンを1つ回復する事ができる。最大数を超える青トークンを戻す事はできない。

3.2 Hanabi におけるコミュニケーション

本研究では、Hanabi におけるコミュニケーションとして、ヒントとプレイの一連のやり取りを想定する。Hanabi におけるヒントは、色や数字の情報を明確にするだけでなく、確実に出すことができるとは限らないものの、出すことを示唆するために利用することも可能である。例えば、プレイヤー1がプレイヤー2に「2枚目は青」というヒントを与え、プレイヤー2が2枚目の手札の数字が分かっていない状態でプレイして、成功するなどのケースである。このような不確実な状況下でのヒントがプレイの成功を導いた場合、その経験は将来の類似の場面でも再利用され、ゲームの得点を上昇させる一因となることが想定される。

4. モデル

事例の蓄積と利用が協調の成功に貢献するかを検討するため、Hanabi を遂行するモデルを一般的な認知アーキテクチャである ACT-R (Adaptive Control of Thought-Rational [1]) を用いて構築した。ACT-R は2節に述べた事例の利用に関わる認知モデル [4, 2, 8] で利用されており、先行研究に即したモデルを構築するために有用と考えた。

本研究で構築したモデルの処理の流れを図1に示す。このフローチャートは Hanabi に関する過去のモデル [10] をベースとしつつ、事例の蓄積と利用に関わる処理を含めたものとなっている。本節の残りの部分にて、モデルによって蓄積される事例およびその利用を説明する。

4.1 事例の蓄積

フローチャート (図1) に示す通り、自分あるいは相手がプレイを成功させて得点を取得した場合に事例は蓄積される。フローチャートの上部では、相手の直前の行動が得点に結びついたか否かを判別する。ま

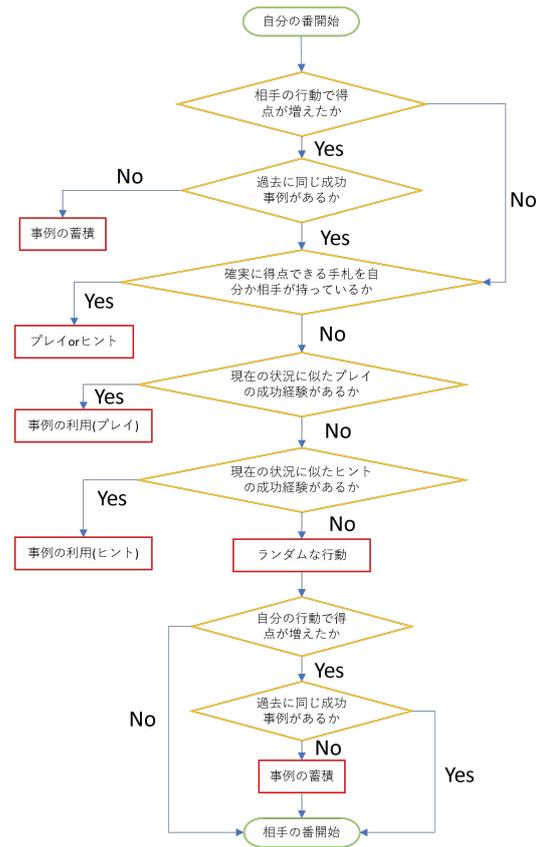


図1 各プレイヤーのターンにおけるフローチャート

た、フローチャートの下部では、自分の行動が得点に結びついたか否かを判別する。これらの判別に当てはまり、かつそのときの状況と同一の状況での成功がそれ以前になかった場合に、新たな成功経験を両プレイヤーは記憶する。

いずれの状況においても、成功事例に含められる情報は、

- 場の状態
- 自分の手札の情報
- 相手の手札の情報
- 直前にプレイした手札
- そのプレイを導いたヒント

で構成される。これら例を、図2、図3、図4、図5に示した。

場の状態は、白、赤、緑、青、黄の5色に対して、それぞれの花火がどの数字まで場に出ているかを示す。図2では緑が場に3まで出されており、他の色はまだ場に出されていないことを示している。

自分の手札の情報は、現在自分が認識している自分の手札の色と数字の情報を示す。図2において、自分が認識していない色や数字は黒くマスクされている。相手の手札は、自身から見える相手の手札の情報と相

W	R	G	B	Y
0	0	3	0	0

図2 場の情報の例

1	2	3	4	5
		G	Y	R
2			2	

図3 自分の手札の例

手が認識している相手の手札の情報から構成される。図4の上段2つは自身の見ているカードの状態であり、下段2段は相手が認識している宛位のカードの状態である。下段では、上段2段のうち自分がヒントを与えていないカードの状態がマスクされている。

直前にプレイした手札の情報は、成功時にプレイした手札の色と数字の組み合わせにより示される。プレイを導いたヒントは、ヒントの対象となった手札によって明らかになった色、もしくはヒントにより明らかになった数字を意味する。図5は、図2、図3、図4に示される場や手札の状態において、自分が得点した場合と相手がプレイした場合のそれぞれにおけるプレイした手札とヒントの例を示している。

自分が得点した場合に関しては、図2より「場に出ている青が出ていない」状態であることが既知である。その際に、相手からのヒントにより「1枚目と2枚目が青」であるということが既知となる（図5における受けたヒント）。ここから、自分は2枚目の数字が明らかでなくても、青を場に出してほしいという意図だと推測してプレイし、実際に青1が場に出て成功となったことを示している。

1	2	3	4	5
Y	R	G	Y	R
1	2	1	3	4
	R	G		
			3	4

図4 相手の手札の例

自分が得点した場合

プレイした手札	受けたヒント
2	1 2
B	B B
1	2

相手が得点した場合

プレイした手札	与えたヒント
1	1 3
Y	G
1	1 1

図5 成功時のプレイとヒントの例

相手がプレイして得点した場合は、相手がプレイする直前に自分が与えたヒントが成功経験として記憶される。図4より、「3枚目が緑」であること、図2より、「場に出ている黄が出ていない」状態であることが既知である。そして、自分が「1枚目と3枚目が1」であるというヒントを教えることで、図5のように相手の1枚目と3枚目の数字が1となる。この状態のときに、相手は1枚目の色が明らかでなくても、緑以外の1の手札を場に出してほしいという意図だと推測してプレイし、実際に黄1が場に出て成功となったことを示している。

4.2 事例の利用

図1に示す通り、自分あるいは相手が確実に得点できる手札を持っていなかった場合に事例は利用される。まず、現在の状況で自分がプレイして得点した事例（プレイ経験）を検索し、検索に成功すればその事例に即したプレイを行う。失敗すれば、自分がヒントを与えて相手が得点した事例（ヒント経験）を検索する。検索に成功すればその事例に即したヒントを与える。

事例を検索する際に参照する情報（検索クエリ）は、自分のプレイで成功した場合と、自分のヒントによって相手がプレイで成功した場合で異なる。自分のプレイで得点した事例を検索する場合は、場の状態、自分の手札の情報、直近の相手からのヒントの情報が必要となる。一方、相手のプレイで得点した事例を検索する場合は、場の状態、相手の手札の情報が必要となる。

検索クエリと蓄積された事例とのマッチングを考える際に、本研究では完全一致によらず、類似に基づく手法を採用した。場の状態および手札の状態を組み合わせた事例のパターンは膨大であり、完全一致の事例を利用することが極めて困難だからである。

現在	W	R	G	B	Y
	0	0	3	0	0
事例	W	R	G	B	Y
	2	2	2	0	0

2/5

図6 場の類似

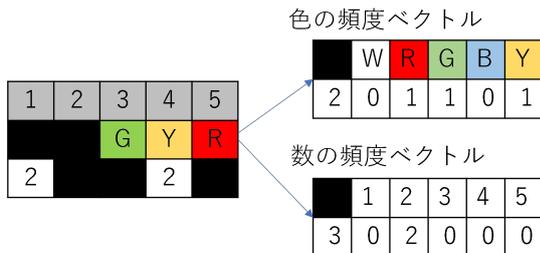


図7 手札の類似に利用する頻度ベクトル

類似の計算は、場および手札で異なる手法を採用した。図6は場の状態の類似を説明するものである。蓄積された成功事例の場の状態（白2、赤2、緑2、青0、黄0）と、現在の場の状態（白0、赤0、緑3、青0、黄0）で数字のマッチする色の数をカウントする。手札の類似では、図7に示すように、類似の計算に先立ち、手札の属性（色と数字）ごとに頻度ベクトルを計算する。色と数字それぞれの頻度ベクトルからコサイン類似度を計算する。これらの類似をACT-Rの機能を利用することで合計し、検索クエリ全体（プレイ経験利用の場合は相手のヒントとの一致を含む）との適合を、蓄積されたすべての事例に対して計算する。その結果、最も高い適合となった事例でなされた行動をそのターンにて実行する。

5. シミュレーション

3節で示したモデル2体でHanabiのシミュレーションを行った。このシミュレーションでは、ゲーム終了までを1試行とし、連続した100試行を20回行った。シミュレーション結果から、1から100の各試行における20回分の平均得点を計算し、試行を重ねることによる学習過程を観察した。

図8に結果を示す。図の横軸は試行回数、縦軸は平均得点である。図より、試行を経ることに得点の増加を認めることができる。はじめの試行では得点が5.5点程度であったものが、中盤や後半の試行では7.8点を超える箇所も存在する。なお、100試行をまとめた平均は6.76点となる。また、ゲームの試行による得

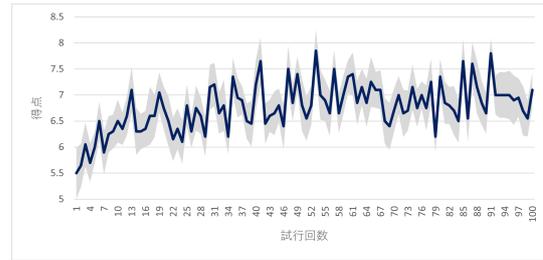


図8 シミュレーション結果

点の上昇の傾向を確かめるために試行と得点の相関を計算したところ、ピアソンの積率相関係数は0.518となった。

以上より、共通経験の蓄積・利用が導いた結果であるとは断定できないにせよ、事例の蓄積・利用が協調成功に貢献している可能性が示唆された。

6. まとめ

本研究では、意図伝達の背後にある処理として、事例の蓄積・利用を想定し、2プレイヤーによるHanabiのモデルを構築した。構築したモデルはHanabiにおけるヒントとプレイの関係をコミュニケーションとして扱い、成功した事例における場や手札の情報を事例として蓄えた。さらに、複雑な状況での事例の利用を可能にするために、事例の類似系計算を実装した。このような工夫を行なうことにより、モデルは試行を重ねることによるゲーム得点の増加を達成した。

本研究で構築したモデルは、事例の蓄積によるコミュニケーションの成立に関する過去の研究[8]に比べ、複雑な状況を扱っている。また、Hanabiに関する過去のモデル[10]とも異なり、協調の成功に至るまでのプロセスを扱っている。これらの先行研究をベースとしつつ、それらを融合する試みの中で、場や手札に関する類似度の計算などの新たな仕組みの開発につながった。

しかし、本研究で示したシミュレーションは予備的なものであり、事例の蓄積・利用による協調効果をより明確にするために検討すべき課題がいくつか挙げられる。第一に、今回の結果に影響を及ぼした要因が複数考えられることである。例えば、今回のモデルには、ACT-Rに実装される他の学習機能（ユーティリティ学習、プロダクションコンパイル）との切り分けがなされていない。したがって、これらの学習と事例の蓄積・利用とを切り分けた検討を行い、結果に反映されている要因の特定が必要である。

第二に、現状のモデルの平均点は人間が行った場合の成績（10点以上）に届いていないことである。現状

のモデルに対して、より構造的な類似の考案、破棄した手札の情報を取り入れる、色や手札の枚数を減少させる、等の改良を加えることでより有意な結果が見込めるかもしれない。また、人間同士の得点データを取得し、モデルシミュレーションのデータと比較する統計的な分析も必要であると考えられる。

参考文献

- [1] J. R. Anderson. *How Can the Human Mind Occur in the Physical Universe*. Oxford Press, 2007.
- [2] Y. Du, P. Aggarwal, K. Singh, and C. Gonzalez. Modeling of multi-defender collaboration in a cybersecurity scenario. In *International Conference on Cognitive Modeling 2022*, 2022.
- [3] Bruno Galantucci. Experimental semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science*, Vol. 1, No. 2, pp. 393–410, 2009.
- [4] Atsushi Hosokawa and Junya Morita. Cognitive models leading to behavior-reading in a card game. In *HAI '21: Proceedings of the 9th International Conference on Human-Agent Interaction*, pp. 337–341, 2021.
- [5] 金野武司, 森田純哉, 橋本敬. 言語的コミュニケーションシステムの創発に関する実験的アプローチ. 計測と制御, Vol. 53, No. 9, pp. 801–807, 2016.
- [6] Guanhong Li, Takashi Hashimoto, Takeshi Konno, Jiro Okuda, Kazuyuki Samejima, Masayuki Fujiwara, and Junya Morita. The mirroring of symbols: An eeg study on the role of mirroring in the formation of symbolic communication systems. *Letters on Evolutionary Behavioral Science*, Vol. 10, No. 2, pp. 7–10, 2019.
- [7] 森田純哉, 小嶋暁, 金野武司, 橋本敬. 新規な記号コミュニケーションシステムの形成に及ぼす自閉傾向の影響. 認知科学, Vol. 29, No. 3, 2022.
- [8] 森田純哉, 金野武司, 奥田次郎, 鮫島和行, 李冠宏, 藤原正幸, 橋本敬. 協調的コミュニケーションを成立させる認知的要因-認知アーキテクチャによるシミュレーション. ヒューマンインタフェース学会論文誌, Vol. 20, No. 4, pp. 435–446, 2018.
- [9] Hirotaka Osawa, Atsushi Kawagoe, Eisuke Sato, and Takuya Kato. Emergence of cooperative impression with self-estimation, thinking time, and concordance of risk sensitivity in playing hanabi. *Frontiers in Robotics and AI*, Vol. 8, p. 658348, 2021.
- [10] 大澤博隆. 協力ゲーム hanabi におけるエージェント間の協調行動の分析. 人工知能学会全国大会論文集 第29回 (2015), pp. 1F23–1F23. 一般社団法人人工知能学会, 2015.
- [11] David Premack and Guy Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, Vol. 1, No. 4, pp. 515–526, 1978.
- [12] 植田一博. 『認知的インタラクションデザイン学』の展望: 時間的な要素を組み込んだインタラクション・モデルの構築を目指して. 認知科学, Vol. 24, No. 2, pp. 220–233, 2016.