

協調的なエージェント間インタラクションにおける感情の役割— 認知シミュレーションから見えるリスクと制御

The Role of Emotion in Cooperative Agent Interactions: Insights into Risk and Control from Cognitive Simulations

河路 塁生 森田 純哉
Ruiki Kawaji Junya Morita

静岡大学
Shizuoka University

Abstract: 協力ゲーム Hanabi の認知シミュレーションにより、人工的な感情の形成が協働の成果に及ぼす影響を検討した。結果、ポジティブ感情により、(1) リスク志向的行動が生起すること、(2) 難易度の高い環境において成果を低下させること、(3) 協調の制度化がそれらを緩和させることが示唆された。そのため人間との生産的な関係を実現する AI の設計には、協調のプロトコルによる感情の制御が重要だと考えられる。

1 はじめに

これまでの Human-Agent Interaction (HAI) および Affective Computing 研究の進展により、人工システムによる感情の表出と誘発、それによる人間の感情の操作が進展している [1]。例えば、チャットボットをはじめとする対話エージェントにおける感情的表現は、人間の印象形成や行動選択に影響を与えることが示されている [2]。また、広告などにおいても、古くから閲覧者の感情状態の喚起による購買意欲を増大することが狙われている [3]。このように、人工的に設計された感情は人間の意思決定を左右する重要な要因として広く用いられている。

一方で、こうした人工的な感情の設計が、人間を含むエージェント間の協調的インタラクションにどのような影響を及ぼすのかについては、未だ十分に体系的な理解がなされていない。過去感情についての研究が行われてきた心理学の分野では、協調と感情の関係についていくつかの重要な知見が蓄積されている。感情が協調を促進するのに有効 [4] であり、協調を安定化させ成功に導くには制度化やプロトコルの形成が重要であること [5] が知られている、また、感情は記憶と結びついており [6]、共有された経験が共同体の形成に寄与すること [7] が示されている。感情の生起には予測と実際の出来事とのズレが起因しており [8]、共同体の一致した予測が集団的な感情となり集団的行動を引き起こしている [9]。

これらの知見の多くは、自然発生的な感情を前提としているため、感情が人為的に設計・操作された状況に

おいても同様に成立するののかの検討が必要である。特に、人工的な感情がどのように協調に影響を及ぼすかを明らかにすることは、協調エージェントを社会実装にあたって重要な示唆を与えられられる。本研究ではこの問題に対して、認知モデルを用いたマルチエージェントシミュレーションを実施する。心理学的知見に基づく感情モデルを認知アーキテクチャ上に構築し、エージェントの感情を操作したシミュレーションをすることで、人工的に設計された感情が協調にもたらす影響を検討する。

2 シミュレーション課題

本研究ではプレイヤー間の協力が必要なゲーム、Hanabi を課題とする。協調 AI のベンチマークとして多くの研究が行われており [10, 11]、過去、このゲームにおける協調や意図推定に関する研究が行われ [12, 13]、その計算機モデル [14, 15] が構築されている。

Hanabi は 2 人から 5 人用のゲームであるが、本研究においては簡略化のために 2 人でのプレイのみを扱う。ゲームには 5 色 (白・赤・青・黄・緑) 50 枚のカード (各色ごと 1 のカードが 3 枚, 2 から 4 のカードが 2 枚, 5 のカードが 1 枚) と赤と青の 2 種類のトークンが含まれている。ゲームの目標は、協力して同じ色のカードを数字の昇順に重ねていき、より多くのカードを場に出すことである。

ゲーム開始時、カードをシャッフルしそれぞれのプレイヤーに 5 枚ずつ配り、そのプレイヤーの手札とす

る。この手札の内容は所持しているプレイヤー自身は確認できず、他のプレイヤーのみがその内容を確認できる。手札とならなかったカードは山札とする。全プレイヤーで共有のものとして青トークンを8枚、赤トークンを0枚所持した状態で最初のプレイヤーの手番が開始される。

これらの Hanabi のルールは、計算論的には以下のように記述することが可能である [16]。ゲームの状態 S は以下の集合である。

- 手札 H_i : プレイヤー P_i の保持する順序付けられた通常5枚の手札のカードのセット
- 山札 D : 順序を含む残りの山札のカードのセット
- 捨て札 W : 破棄、またはプレイ失敗により使用不可能になったカードのセット
- 場の花火 $F = \{f_{white}, \dots, f_{green}\}$: f_j は色 j において場に出された最大の数であり、初期値は0である
- 青トークン (情報トークン) I : ヒントを行う際に必要なリソース (初期値=最大値 8)
- 赤トークン (ライフトークン) L : プレイ失敗の許容回数 (初期値 0, 最大値 3)

プレイヤー P_i の観測 O_i は、 S から自分の手札の情報 H_i と山札 D を除き、ヒント行動によって公開情報となったプレイヤーの手札 H の情報を加えたものである。

プレイヤーには順番に自分の手番が与えられ、以下の三つの行動 a のいずれかを行い次のプレイヤーの手番に移る。

- ヒント: プレイヤー P_i が他プレイヤー P_j に対し、特定の色または数値いずれか一つについての情報を教える。前提条件として青トークンが1つ以上存在しなければヒントを出すことはできない ($I > 0$)。ヒントが行われると青トークン1枚が消費される ($I \leftarrow I - 1$)。ヒントの効果として、プレイヤー P_j の手札 H_j に存在する指定された属性を持つカードの位置が公開情報となる。
- 破棄: プレイヤー P_i が手札の x 番目のカード $card$ を破棄する (H_i から削除する)。カードは捨て札に送られ ($W \leftarrow W \cup \{card\}$)、青トークンが最大値でなければ1枚回復する ($I \leftarrow I + 1$)。その後カードをデッキから一枚手札に加える (D の一番上を H_i に移動する)。破棄されたカードは自分も含め全員が確認することができ、ゲーム中は使用できなくなる。
- プレイ: プレイヤー P_i が、手札の x 番目のカード $card = (color, value)$ を出す (H_i から削除する)。 $value = f_{color} + 1$ の場合成功となり、 $f_{color} \leftarrow value$ に遷移する。さらに $value = 5$ であれば $I \leftarrow I + 1$ に遷移する (I が最大値でない場合)。その他の場合は失敗と判定され、 $L \leftarrow L + 1$ および $W \leftarrow W \cup \{card\}$ の遷移が発生する。その後

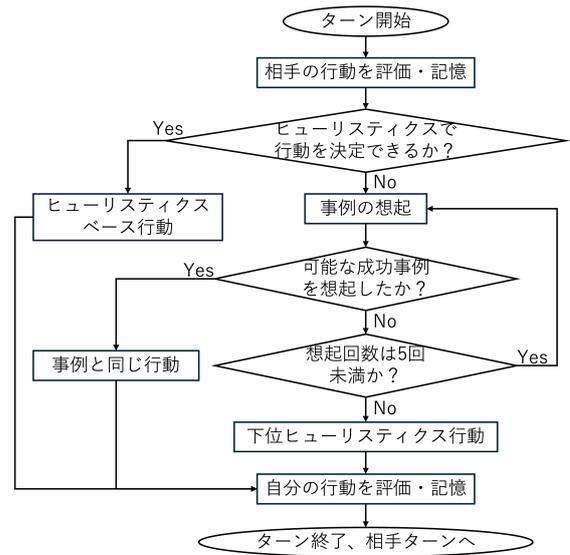


図 1: モデルの行動決定フローチャート

成功失敗問わずにカードをデッキから一枚手札に加える (D の一番上を H_i に移動する)。

ゲームの終了条件は三つ存在する。一つ目は赤トークンを3枚獲得すること。二つ目は山札が0枚になったのちに全てのプレイヤーが1度ずつ行動すること。三つ目は5色の花火全てが5枚重ねられ完成することである。終了条件を満たしたならば得点を計算する。得点は場に重ねられたカードの枚数の合計となり、最大値は5色 \times 5枚の25点である。

3 モデル

本研究のモデルはゲームの状態 S を元に行動 A を決定する方策 $\pi(a|s)$ を認知アーキテクチャ ACT-R [17] を用いて表現したものである。先行研究 [15] を基に本研究のモデルを開発した。モデルは事例の利用に基づく行動決定を、感情に該当するパラメータを介して実行する。さらに協調の熟達を表現するために、ヒューリスティクスに基づく行動決定のルールを実装した。

モデルはゲームの状態 S から得られる観測 O_i に基づいて行動を決定する。先に述べた通り、この観測はプレイヤーごとに異なっている。図1はモデルのターンが開始されてから終了するまでの行動を表したフローチャートである。

3.1 行動決定のヒューリスティクス

行動の決定に際し、モデルは手続的知識として獲得されたヒューリスティクスに優先的に従う。これらの

ヒューリスティクスは2者間の共通経験によって制度化されたプロトコルを積み込み、ルールとしたものである。これらはルールに該当する状況で行動の決定を行うため、モデル全体の方策 π の一部分である部分方策 $(\pi_1, \pi_2, \dots, \pi_n)$ に相当する。これらヒューリスティクスの部分方策および後に示す想起による行動決定 π_{rec} の集合が全体の方策である ($\pi = \{\pi_1, \pi_2, \dots, \pi_n\} \cup \pi_{rec}$)。これらのヒューリスティクスは、先行研究 [18] においての人間のプレイヤーへの調査を基に、期待効用に基づくユーティリティ値によって以下に示す順に発火の優先度を設定した。

1. 確定カードのプレイ：色と数字の両方が判明している自分の手札のカードでプレイして成功するものがあればそれをプレイする。
2. プレイ可能カードへの確定化ヒント：ヒントが出せる状況で、色と数字の片方のみが判明している相手の手札のカードにプレイして成功するものがあれば、そのカードのまだ判明していない情報のヒントを出す。
3. 単独カードへのヒント：ヒントが出せる状況で、相手の手札のカードにプレイして成功するものがあり、相手の手札に同じ色または数字のカードが存在しない場合、その単独のカードの情報のヒントを出す。
4. 確定カードの破棄：色と数字の両方が判明している自分の手札のカードで既に同じ色・同じ数字のカードがプレイ成功しているものがあればそのカードを破棄する。または両方が判明していて同じ色・同じ数字のカードを2枚以上持っている場合そのカードを破棄する。
5. プレイ可能カードへのヒント：ヒントが出せる状況で、相手の手札のカードにプレイして成功するものがあれば、そのカードのまだ判明していない情報のヒントを出す。
6. 破棄可能カードへのヒント：ヒントが出せる状況で、相手の手札に既に同じ色・同じ数字のカードがプレイ成功しているものがあれば、そのカードのまだ判明していない情報のヒントを出す。

3.2 事例に基づく行動決定

上記の手続き化されたヒューリスティクスのいずれも適用できない場合、記憶された事例に基づいた π_{rec} による行動の決定を試みる。記憶検索に使用される事例は、自分もしくは相手が行動を実行し、その結果の成否が判明した時に記憶される。事例に含まれる情報は観測 O に加えて以下に示すその状況で行った行動 A と行動の結果発生した観測 O' の一部分から構成される。

- 行動の種別：ヒント、破棄、プレイのいずれか。

- 行動の対象となったカードの情報：行動の種別がプレイもしくは破棄であった場合はそのカードの色と数、行動の種別がヒントの場合は指定された色または数。色や数が未判明のまま行動した場合には、未判明として記憶。
- 行動の成否：その行動が成功したか否か。プレイの場合は得点の加算、ヒントの場合は直後に相手が得点、破棄の場合は捨てたカードが既に場に出て成功したカードと同じ色・同じ数字であれば成功と判定。

記憶された事例は、図1に示す想起の手続きで利用される。初めに現在の状況を基に過去の事例の検索を行う。この際、ACT-Rの部分一致機能を用いることで、完全に同一でなくとも最も類似したものが想起される。事例の検索が為され、成功した事例であった場合はその事例通りに行動を行おうとする。想起された事例が失敗のものであった場合や、何も想起できなかった場合には、既にこの行動決定で想起した事例以外から再度想起を行う。想起の繰り返しは最大5回まで行われる。繰り返しの度に、部分一致のペナルティを半減させ、ランダム性を高める。5回の繰り返しを行ったのちにも事例が検索されなかった場合は下位ヒューリスティクスにて行動を決定する。下位ヒューリスティクスは行動のコストを最小化するものであり、

1. ランダムなヒント
2. ヒントを出されてないカードのランダムな破棄
3. ランダムな破棄

の3つのヒューリスティクスを順に適用を試みる。

想起の部分一致メカニズムでは、現在の状況に対する記憶項目 i の重要度を示すパラメータとして、

$$Activ_i = B_i + \sum_l PM_{li} + \varepsilon \quad (1)$$

を求める。 B_i は記憶の利用頻度や利用されてからの時間経過によって計算されるベースレベルであり、その事例が一般的にどの程度役立つかを表す。 M_{li} は事例 i に含まれるそれぞれの要素 l の現在の状況との一致度であり、この記憶検索にどれだけ関連しているかを表す。 P は部分一致の不一致ペナルティ係数 (mp) である。 ε はロジスティック分布に従うノイズであり、人間の想起の揺らぎを表す。

類似度である M_{li} は、完全一致を0、完全不一致を-1とし、事例に含まれる各情報ごとに計算される。場に出ているカード及び残りのカードについては、(一致した要素の数/要素数)-1と計算される。要素数は場のカードであるならば色の数である5、残りのカードの情報であるならば5色×5数字の25である。手札の情報の場合は、色・数それぞれの頻度を要素とするベクト

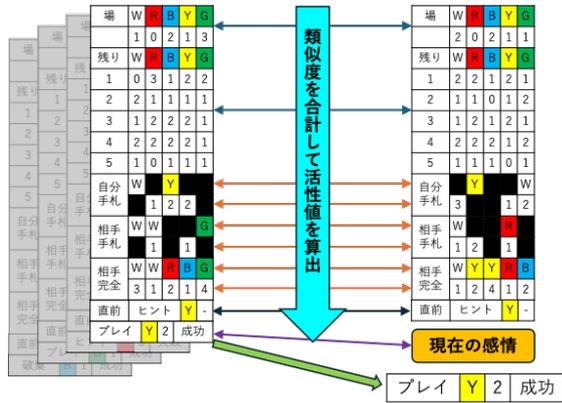


図 2: 想起の例

ルのコサイン類似度 -1 が用いられる。直前の行動の種別、及び対象の情報は類似を持たず、一致していなければ完全不一致とした。行動の成否に関しては、次節に示される感情の影響に従った類似度が計算される。これらの類似度を合計し、最も類似している事例が想起される。想起の例を図 2 に示す。事例が想起され、想起された行動が成功していた場合は事例と同じ行動を試みる。

なお、式 1 はベイズの定理の自然対数をとったものと一致し、 $P(i|O) = P(i) \times P(O|i)$ のそれぞれの項と式 1 の ε を除いた各項が対応しており、 ε のノイズにより確率的な決定が行われる。

3.3 感情の影響

モデルの感情は最小 -1 、最大 1 である二つのパラメータ、Valence(感情価)と Arousal(覚醒度)として表され、その値により事例の想起が変化する。感情価については、類似度の計算として事例の想起に影響する。

$Activ_i$ における感情価の影響は、類似度 M_i に対し、

$$M_{emotion,i} = -|V - V_i|/2 \quad (2)$$

を含めることにより計算される。 V は現在の感情価 (Valence), V_i は対象の事例の成否である。この時 V_i は成功事例であれば完全にポジティブであるとして 1 、失敗事例であれば完全にネガティブであるとして -1 となる。

この式は感情価の想起への影響として気分一致効果 [6] を認知モデル的に表現したものである。記憶した事例が成功したものであればポジティブ、失敗であればネガティブであるとラベルづけをすることにより現在の感情価との距離を測り、近いものが想起されやすくなる。

覚醒度は、式 1 における B_i に影響する。このパラメータのオフセット値を 0.5 倍から 1.5 倍に変動させ

表 1: 行動と感情の変化

成功評価	事例利用	感情価	覚醒度
成功	有り	増加 大	増加 大
成功	無し	増加 小	減少 小
失敗	有り	減少 大	増加 大
失敗	無し	減少 小	減少 小

る。これにより覚醒度が高い場合には想起が成功しやすく、低い場合には失敗しやすくなる。

感情の更新には予測誤差の考えを基とし、予測誤差が解決に向かうことが快、離れることが不快とする Joffily によるモデルを参考とした [19]。これを計算機モデルに実装するためにヒューリスティックスペース行動、事例利用行動、のどちらかを基に行動を行ったかによって変動の仕方に違いを持たせた。上位のヒューリスティックスペース行動の場合は、ほぼ必ず成功し予測と結果との差異が少ないため感情は大きく変動しない。下位ヒューリスティックスペース行動では、不確定なままに行動しているために予測・経験の精度が共に低いために大きく変動しない。事例利用行動では予測を行い、成功を期待しているために成功すれば予測誤差の解決に向かうため大きく快に変動、失敗すれば解決から遠ざかるために大きく不快に変動する。また覚醒度に関しても大きな変動が起きる事例利用行動を行った場合に増加、大きな変動が起きないその他の場合には時間経過による微量の減少を適用した。表 1 に感情パラメータの増減をまとめた。

感情の更新には以下の式 3 を用いた。

$$V_{t+1} = V_t + \alpha(r - V_t) \quad (3)$$

感情価の変動の場合 r は成功時は 1 、失敗時は -1 であり、事例利用時は $\alpha = 0.2$ 、非利用時は $\alpha = 0.02$ とした。覚醒度の変動の場合事例利用時は $r = 1$ 、 $\alpha = 0.2$ 、非利用時は $r = -1$ 、 $\alpha = 0.01$ とした。

4 シミュレーション

4.1 狙いと設定

事例に基づく行動決定の過程に、前節で示した感情がどのような影響を及ぼすのかを観察するため、モデル 2 体が参加するシミュレーションを実施する。状況に応じた感情の影響を検討するために、3.1 におけるヒューリスティックスの獲得状況および課題の難易度を以下のように操作する。

- 難易度：感情の影響を排除したモデル同士による事前シミュレーションにより、開始時のデッキ配

列によってゲームのスコア平均が大きく変化することが確認されている。無作為に作成された10種類の初期配列を用いてそれぞれ100事例シミュレーションしたところ、平均スコアが最も高くなったデッキでは16.0点、最も低くなったデッキでは11.3点というスコアが得られた。この最も平均スコアが高くなったデッキ配列を「低難易度条件」、最もスコアが低くなったデッキ配列を「高難易度条件」とし、難易度ごとに各試行で同一の初期配列を用いた。

- 制度化度: 3.1に示した行動決定のヒューリスティクスは、人間にとって所与のものではなく、事例に基づく行動決定の繰り返しにより集団でのプロトコルが制度化されることにより獲得されたと考える。この考えに基づき、3.1のヒューリスティクスを操作することで、未制度化条件と制度化済条件を設定した。未制度化条件は、相手の意図を用いず、直接確認できる情報のみによるヒューリスティクスから行動決定を行う。具体的には、「1. 確定カードのプレイ」、「2. プレイ可能カードへの確定化ヒント」、「4. 確定カードの破棄」のみを利用する。それに対して制度化済条件では、1から6の全てのヒューリスティクスを利用した行動決定を行う。

4.2 シミュレーション1: 固定感情

本研究の焦点である感情についての段階的な検討を行うため、感情パラメータを固定したシミュレーションを行った。前節に示した式3を含まないモデルは、初期に設定した値に固定された感情パラメータに従った事例の想起を行う。

4.2.1 シミュレーション1 手続き

一回の実行において、モデルは、ゲーム開始から終了までを1試行とする連続した2試行を実施する。連続した試行の間で、モデルは事例の記憶と利用を行う。「ネガティブ条件(覚醒度0,感情価-1)」、「ニュートラル条件(覚醒度0,感情価0)」、「ポジティブ条件(覚醒度0,感情価1)」を設定し、それぞれにおいて難易度と制度化度の各条件ごとに100実行ずつ繰り返す。

感情条件が異なることにより、行動決定に用いられる経験の想起が変化する。ネガティブ条件の場合、失敗事例との類似度が高いため失敗事例が想起されやすくなり、経験を用いた行動決定はされづらい。ポジティブ条件の場合、成功事例との類似度が高く想起されやすいため、経験を用いた行動決定を多く行う。ニュートラル条件では失敗事例と成功事例との間で同じ類似

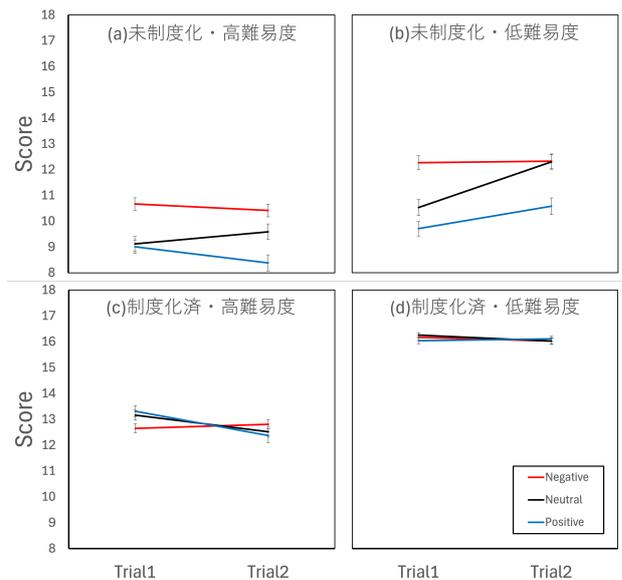


図 3: シミュレーション1: 固定感情のスコア

度を持つため、経験の利用頻度は中間になる。感情条件によって事例に基づく行動決定の頻度が変化し、協調に適した感情パラメータの確認を行う。

4.2.2 シミュレーション1 結果

結果を図3に示す。それぞれのグラフの縦軸は100実行のスコア平均値とその標準誤差である。未制度化条件と制度化済条件で、スコアに及ぼす感情の影響が大きく異なった。未制度化条件は、ネガティブ条件で最もスコアが高く、ポジティブ条件で最もスコアが低い結果となった。また、試行間でのスコアの変動はネガティブ条件では認められず、ニュートラル及びポジティブ条件において認められた。未制度化・高難易度におけるポジティブ条件では学習(事例の蓄積)によりスコアが低下し、未制度化・高難易度のニュートラル条件、未制度化・低難易度のニュートラル及びポジティブ条件においてはスコアが上昇した。一方、制度化済条件では、試行間の差異や感情による影響がほとんど認められなかった。

上記の結果のうち、制度化済条件において学習効果が観察されなかった理由は、行動がヒューリスティクスによって大きく固定されているためと考えられる。制度化済モデルの行動決定は、協力が成功するように最適化されており、経験の利用により、協力が向上する余地は残されていないと考えられる。

未制度化・高難易度条件において、事例利用が負の効果を示すことは、リスク容認的な行動の観点から説明できる。ポジティブ条件は、成功事例の想起における部分一致のペナルティを小さくし、記憶事例と多少

の差があった場合でも事例に基づく行動を行う。そもそも行動の成功率の小さい高難易度条件において、リスク容認的な振る舞いは、点数の低下を導くと考えられる。その一方、低難易度条件においては、高難易度条件と比べてリスク容認的な行動が成功しやすい環境であったと考えられる。

また、4つの条件全てにおいてネガティブ条件ではスコアの変動が見られなかった。これはポジティブ条件とは逆にリスク否定的な行動決定によるものと説明できる。部分一致のペナルティが増加し、記憶事例との差が少なくても事例に基づく行動決定は発生しづらい。そのため想起による行動決定は減少し、リスクを取らずに学習前の試行と同様の行動決定をすることでスコアの変動も発生しなかったと考えられる。

4.3 シミュレーション2: 変動感情

シミュレーション1で観察された感情の負の効果が、感情変動の仕組みを取り入れることで調整されるのかを検討した。式3をモデルに含むことにより、行動の成否に応じた感情パラメータの変動が発生する。変動した感情を用いた想起により状況に適応した想起を行うかを確認する。

4.3.1 シミュレーション2 手続き

シミュレーション1と同様に、一回の実行において、モデルはゲーム開始から終了までを1試行とする連続した2試行を実施する。連続した試行の間で、モデルは事例の記憶と利用を行う。覚醒度0、感情価0を初期値として、式3に従い感情パラメータが変動する「感情変動条件」を設定し、難易度と制度化度の各条件ごとに100実行ずつ繰り返す。

感情が変動することにより、行動決定に用いられる経験の想起が変化する。感情価の変化により事例の成否との類似度が変化するため、感情価の低下により経験を利用した行動は少なくなり、感情価が上昇すると多くなる。覚醒度の変化は成否に関わらず全ての事例の想起されやすさにつながり、覚醒度が低い場合は事例が想起されない可能性がある。そのため覚醒度の低下は経験利用行動の減少に、覚醒度の上昇は増加につながる。シミュレーション1において観察された、リスクの観点によって環境に応じた適切な感情価が存在することから、感情価が調整されることによりスコアは上昇すると考えられる。

4.3.2 シミュレーション2 結果

感情変動条件とニュートラル条件の結果を比較した。図4に条件別のスコアの変化を示す。

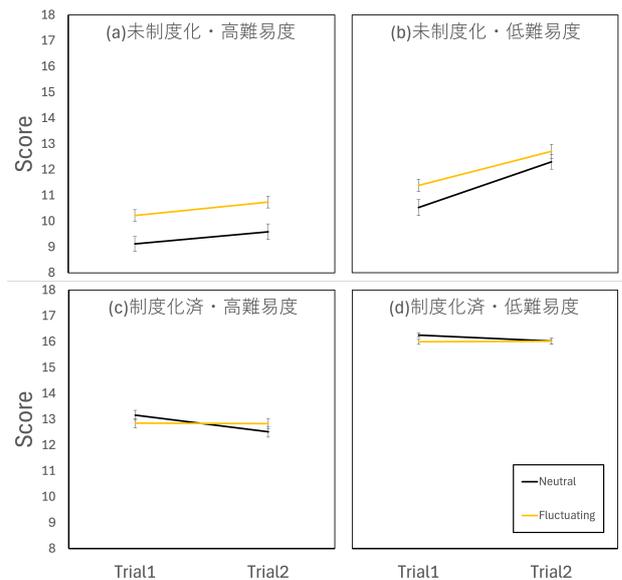


図4: シミュレーション2: 変動感情のスコア

未制度化条件では、高難易度、低難易度ともに感情変動条件がニュートラル条件のスコアを上回った。学習の効果は、ニュートラル条件と同様、スコアの上昇として観察された。一方、制度化済条件では感情条件によるスコアの顕著な差は見られなかった。特筆すべき点として、熟達者・高難易度条件において認められたニュートラル条件におけるスコアの低下が抑制された。

初心者モデルにおいて感情変動条件のスコアがニュートラル条件を上回ったことは、状況に応じた感情パラメータの調節が有効に機能したことを示している。感情パラメータが適切な値になることによりリスクが調節され、スコア自体と試行間の正のスコア変動として観察された。このような感情調整の効果は、熟達者条件においても、経験の利用による負の影響の抑制という形で観察された。

5 総合考察

本研究では、人工的な感情の設計が協調に及ぼす影響を検討するために、感情に基づく記憶想起を組み込んだ Hanabi エージェントモデルを構築した。協調の制度化（集団内プロトコル形成）の水準をモデル側の要因として操作し、さらに環境側の要因としてゲーム難易度を設定した。これらの条件下で、(1) 外部から感情状態を操作するシミュレーション1と、(2) フィードバックによって感情が自然に変動するシミュレーション2を行い、感情操作が協調過程に与える影響を比較した。

シミュレーション1の結果、制度化が未整備で利用可能なヒューリスティクスが少ないモデルでは、感情

を想起に利用することにより行動決定に関わる記憶内容が偏り、協調の成否が大きく変化した。特に、環境に適合しないポジティブな感情状態では、リスクを不相応に容認する意思決定が生じ、結果として協調の失敗に繋がった。一方で、同じくリスク受容の増加は、状況によっては協調形成を促進する側面も示し、感情がリスク制御を介して協調を「阻害も促進もする」両義的な働きを持つことが確認された。これに対して制度化の進展したモデルでは、未制度化モデルに比べて協調成功率が高く、感情操作による影響は相対的に小さかった。このことは、集団内プロトコルの形成が、感情に伴う意思決定の変動を吸収し、協調の頑健性を高める可能性を示唆する。

続くシミュレーション2では、フィードバックによって感情が変動することで、環境に応じた感情状態の調節が生じることが観察された。未制度化モデルにおいても、感情が適切に調整される条件下では、リスクが過大にも過小にもならない形で管理され、協調の成功と協調形成の促進が両立した。一方、制度化済モデルでは、感情が変動する状況を導入しても、協調成功の顕著な増加には繋がらなかった。これは、制度化済モデルでは、協調の成否が主としてプロトコルにより規定され、感情の寄与が限定されるためと解釈できる。

以上より、本研究のシミュレーションは、人工的な感情がリスク制御を通じて協調に影響しうること、そしてその影響が制度化（集団内プロトコル形成）の水準に依存することを示した。感情によるリスク制御は協調形成に寄与する一方で、環境に不適合な感情が導入されると、過剰なリスク選好が誘発され協調の破綻を招く。したがって、人工的感情を協調AIへ組み込む際には、感情の利得を活かしつつ負の側面を抑制する制度設計が不可欠である。

ただし本研究は、感情が協調に与える影響を主としてリスク制御の側面から検討しており、その他のメカニズムについては十分に扱えていない。例えば感情表出は、コミュニケーション信号として相手の推論や関係形成を方向づける可能性を持つ。Hanabiではゲーム内制約により明示的なコミュニケーションが限定される一方、人間プレイヤーを想定しゲーム外の要素を介して情報伝達を捉える研究も存在する[20]。今後は、リスク制御に加えて、感情の信号機能や社会的調整機能を含む枠組みに拡張する必要がある。

また、本研究では感情を想起過程に影響するパラメータとして実装したが、感情は想起以外の行動決定にも作用しうる。本研究の意思決定では期待効用に基づきヒューリスティクスのユーティリティ値を決定したが、感情によってこの期待効用評価自体が変調される可能性や、想起の実行に対するユーティリティが変動する可能性もある。これらを導入することで、感情の昂りによって規範的ルールが排除される状況など、より現

実的な失敗様式を表現できると考えられる。

本研究で得られた知見は、人工的な感情を持つ協調AIの設計に対し、制度と感情の両輪の重要性を示すものである。すなわち、感情は協調形成に資する一方で、安定的な協調の実現には、感情の負の影響を吸収・抑制する集団内プロトコルの形成が不可欠である。本研究は、インタラクションにおける感情と制度の相補性を示し、人間社会に即した協調エージェント開発に向けた設計論への寄与が期待される。

参考文献

- [1] Rosalind W Picard, et al. *Affective computing*, 1997.
- [2] Martin Adam, Michael Wessel, and Alexander Benlian. Ai-based chatbots in customer service and their effects on user compliance. *Electronic markets*, Vol. 31, No. 2, pp. 427–445, 2021.
- [3] Richard P Bagozzi, Mahesh Gopinath, and Prashanth U Nyer. The role of emotions in marketing. *Journal of the academy of marketing science*, Vol. 27, No. 2, pp. 184–206, 1999.
- [4] Robert H Frank. *Passions within reason: The strategic role of the emotions*. WW Norton & Co, 1988.
- [5] Elinor Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge university press, 1990.
- [6] Gordon H Bower. Mood and memory. *American Psychologist*, Vol. 36, No. 2, p. 129, 1981.
- [7] Michael Tomasello. *Why we cooperate*. MIT press, 2009.
- [8] Charles S Carver and Michael F Scheier. Origins and functions of positive and negative affect: a control-process view. *Psychological review*, Vol. 97, No. 1, p. 19, 1990.
- [9] Diane M Mackie, Thierry Devos, and Eliot R Smith. Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of personality and social psychology*, Vol. 79, No. 4, p. 602, 2000.
- [10] Nolan Bard, Jakob N Foerster, Sarath Chandar, Neil Burch, Marc Lanctot, H Francis Song, Emilio Parisotto, Vincent Dumoulin, Subhodeep

- Moitra, Edward Hughes, et al. The hanabi challenge: A new frontier for ai research. *Artificial Intelligence*, Vol. 280, p. 103216, 2020.
- [11] Jakob Foerster, Francis Song, Edward Hughes, Neil Burch, Iain Dunning, Shimon Whiteson, Matthew Botvinick, and Michael Bowling. Bayesian action decoder for deep multi-agent reinforcement learning. In *International Conference on Machine Learning*, pp. 1942–1951. PMLR, 2019.
- [12] 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, 2015.
- [13] Kansei Miyata and Hirotaka Osawa. Enhancing cooperative behavior through subtle communication cues: The impact of hint order and hesitation in hanabi. In *Proceedings of the 12th International Conference on Human-Agent Interaction*, pp. 444–446, 2024.
- [14] 桑原涼香, 長島一真, 森田純哉, 宮田貫誠, 川越敦, 大澤博隆. 協力ゲーム hanabi を用いた察するコミュニケーションのモデル構築. may 2023.
- [15] 河路墨生, 森田純哉, 大澤博隆. 協力型ゲームのシミュレーションを通じた情動とその伝染による協調行動への影響の理解. 第 2024-ICS-214 巻, pp. 1–8, 2024.
- [16] Joseph Walton-Rivers, Piers R Williams, Richard Bartle, Diego Perez-Liebana, and Simon M Lucas. Evaluating and modelling hanabi-playing agents. In *2017 IEEE congress on evolutionary computation (CEC)*, pp. 1382–1389. IEEE, 2017.
- [17] J. R. Anderson. *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press, 2007.
- [18] 川越敦, 大澤博隆. Hanabi ゲームにおける認知モデルを用いた協力行動分析の提案. 人工知能学会全国大会論文集 第 36 回 (2022), pp. 4I3OS26b04–4I3OS26b04. 一般社団法人 人工知能学会, 2022.
- [19] Mateus Joffily and Giorgio Coricelli. Emotional valence and the free-energy principle. *PLoS Computational Biology*, Vol. 9, No. 6, p. e1003094, 2013.
- [20] 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.