

利害関係にある人間集団の集団規範に適応を試みる エージェントが示す公平さ

Fairness Exhibited by Group Norm-Aware Agent in Ultimatum Game

布施 陽太郎^{1*} 芦田 美那² アイエドゥン エマヌエル² 徳丸 正孝²
Yotaro Fuse¹, Biina Ashida², Emmanuel Ayedoun², Masataka Tokumaru²

¹ 富山県立大学

¹ Toyama Prefectural University

² 関西大学

² Kansai University

Abstract: 集団規範を考慮しながら意思決定するエージェントが利害関係下にある人間集団の一員として振舞うことの影響を調査する。エージェントが自律した人間社会の一員として受容されるために、我々はエージェントによる集団内の暗黙の了解(暗黙的集団規範)への適応に着目している。本研究では、エージェントの適応がもたらす影響を、利害関係下の集団ゲームシナリオの基で調査する。実験の結果より集団規範へ適応的なエージェントは公平な振舞いを表出することが示された。

1 はじめに

今日、アシスタントやパートナーとしてエージェントやロボットが人間と共生する社会の到来が予想される中、人間とロボットの間でのやり取りに関する様々な研究が取り組まれている [1, 2]。人間とロボットの共生のためには、ロボットに対する人間の受容が不可欠であり、ロボットの感情表出やジェスチャ、視線などの社会的な振舞いについて研究されている。

社会性とは集団を作って生活しようとする性質である。特に、人間集団において人間は直接的なやり取りや指示無しに、その場の状況に応じて集団の一員である他者に特定の振舞いを期待することがある [3]。そのようなとき、その集団では集団規範が共有されているといえる。人間の社会ではその場で求められる振舞いは明文化されず、しばしば暗黙的に共有される。その規範に適応できるかどうかは、ある集団メンバーが集団の一員としてふさわしいか否かの判断のためのそのほかの集団メンバーにとっての基準のひとつになりえる。

我々は直接的なやり取りが無いシチュエーションでも、暗黙的に形成される集団規範に適応するエージェントやロボットの研究に取り組んできた [4]。暗黙的に期待される振舞いの推察が可能なロボットやエージェントの実現を目的にしている。そのようなロボットは従来

に比べてロボットへの親しみやすさや人間らしさ、ロボットの行動の予測しやすさを向上させ、人間とロボットの共生社会の発展に将来貢献することが期待される。ロボットに対する人間の受容の観点からも集団規範への適応について考慮し探求すべきである。

我々の先行研究では比較的単純な実験シナリオにおいてエージェントやロボットを評価しており、集団内の利害や自身の行動へのフィードバックがある状況が考慮されていなかった。人間が集団を形成するとき、各集団メンバーの利益が他者の損になる状況が発生し得る。集団内に利害関係がある場合、人間は他者に公平であることを望み、他者と自分との利益の差があることを好まない [5]。加えて、人間は集団内のメンバーの行動に対して直接的な返答やフィードバックを実施することがある。先行研究では、以上のようなメンバー間の損得や直接的な返答のない実験環境で提案モデルを検証した [4]。しかし、より現実的な人間集団に即した評価実験を実施するためには、そのような要素のある実験シナリオでの検証が必要である。

利害関係にある人間は所属する集団内の他者が公平に振舞うことを望む。加えて、自分の得を好み、自分の損を好まない。このことから以下の2点のことが言える。

- 自分の極端な利他的振舞いは自身の利益を損ない、公平さの観点から自身が許容しない。
- 自分の極端な利己的振舞いは他者の利益を損な

*連絡先：富山県立大学工学部知能ロボット工学科
〒939-0398 富山県射水市黒河 5180
E-mail: fuse@pu-toyama.ac.jp

い、公平さの観点から他者が許容しない。

したがって、利他と利己を織り交ぜた意思決定が人間らしく理にかなった集団内での振舞いとして求められる。ここでは、ある程度他者の利益に配慮を見せることで利他的に振舞いつつ、自らの利益を守るよう利己的に振舞うことが重要である。これは従来研究のシナリオに比べて、より複雑なシナリオにおいてロボットが集団メンバとして振舞う必要があることを意味する。以上のような振舞いが求められる実験シナリオ環境下において、先行研究 [6] では提案モデル搭載エージェントの振舞いを観察しただけであった。本研究ではそのエージェントを評価し、そのエージェントが集団全体に与える影響を調査する。

本研究では、集団内で利害関係と行動へのフィードバックを考慮して振舞う必要がある実験シナリオとして集団最後通牒ゲームを提案した。集団型最後通牒ゲームは最後通牒ゲーム [7] を基にしている。ゲームの構造上、各ゲームプレイヤは他者のすべての振舞いを観察できない。しかし、我々が提案した集団規範アウェアな意思決定モデル [4] は集団内のすべての振舞いを観察可能であることを前提にしている。したがって、本研究はエージェントの意思決定が集団規範に適応的であることの影響を調査することを目的とする。人間プレイヤとは異なり、エージェントは集団内のすべての振舞いを観察した上で意思決定する。

2 実験シナリオ

2.1 最後通牒ゲーム

最後通牒ゲームでは、2名のゲームのプレイヤが“提示者 (proposer)”と“受け手 (responder)”に任命され、金銭の分配と受領に関する意思決定に従事する。初めに、提示者は実験の主催者から金銭を与えられ、それを受け手と自分との間でどのように分割するかを決定し、分割した金額を受け手に提案する。受け手が提案を受諾する場合 (accept), 主催者から初めに与えられた総額は提示者の提案通りに分割される。一方で、受け手が提案を拒否する場合 (reject), 両者とも獲得金額がゼロになる。文化の違い考慮したうえで、提示者は平均して得た金額の 40% を受け手に提示する傾向があることが報告されている [8]。

2.2 集団型最後通牒ゲーム

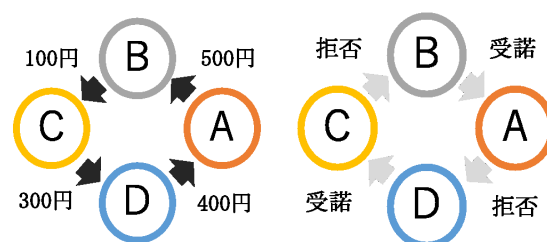
本項では、本研究で用いられた集団型最後通牒ゲームについて説明する。ゲームの基本的なルールとして、お金を他者に分配し、その金額について受領ないしは拒否するという骨組みは同様である。集団型最後通牒

ゲームにおいて、各被験者は金額を提示する提示者と提示された金額に対して受領か拒否かを返答する受け手という 2 つの役割を担う。

1 ラウンドにおけるお金の取引の流れを図 1 に示す。図中の A, B, C, D はそれぞれゲームのプレイヤを表している。図 1(a) は、プレイヤが金額を提示した際の構図である。ここで、金額を提示する操作を全員が一斉に行う。すると、提示された金額に対する受諾又は拒否は図 1(b) のように図 1(a) の逆となる。この受諾または拒否の返答も全員が一斉に行う。

まずラウンドの初めに各被験者は主催者から 1,000 円を貰い、自分以外の被験者のうちある 1 人の被験者と分割する金額を決める。この際、被験者は自分が誰に提案し誰から提案されるかは知らされない。本実験では、提案する相手は実験終了まで固定である。その後、全員が「提示者」になり一斉に決められた相手に分配する金額を提案する。すると、全ての被験者が「受け手」になるので、提示された金額に対し受諾するか拒否するかを決める。以上の過程を 1 回のラウンドと定義する。実験ではラウンドを 20 回繰り返してからゲームが終了する。

加えて、各プレイヤはゲーム内のすべての情報を得られない。プレイヤは自身が行った取引についての情報と 1 ラウンドが終了する時にどのプレイヤが受諾または拒否を表明したのかの公開情報のみを得る。しかしながら、本実験においては利害関係下にある集団における集団規範に適応することの影響の調査を目的とするため、提案モデルはすべてのプレイヤの取引情報を得て、その情報を基に意思決定した。



(a) プレイヤが金額を提示する際の一例 (b) 提示された金額に対する判断の一例

図 1: プレイヤ A, B, C, D が参加する集団型最後通牒ゲームの取引例。

2.3 ゲームインターフェース

人間のゲームプレイヤはディスプレイとゲーム用コントローラをインターフェースとして、集団型最後通牒ゲームに取組んだ。図 2 はゲーム用仮想空間の全体を示す。中心の灰色の円形の机を囲っている 4 体の口

ポットがゲームに取組む。各プレイヤーはそれらのロボットのいずれかの視点に立ち、そのロボットを操作することでゲームをプレイする。図3から図5はあるプレイヤーが見るディスプレイに表示されるインターフェースの例を示す。プレイヤーはコントローラを用いて自身のゲーム内の振舞いを決定する。図5における1ラウンドの取引の結果を示す画面は、各色のロボットの返答のみを示す。

実験主催者と各被験者は同じ実験環境空間の中でゲームに取組む。各被験者は他のプレイヤーのゲーム画面を見ることはできない。加えて、実験主催者がゲームの進行を常に確認していることを被験者は知らされる。

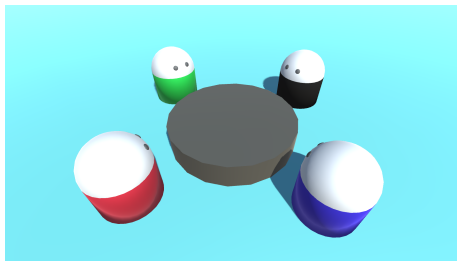


図 2: ゲーム環境全体。



図 3: ある特定のプレイヤーの提示金額入力画面。

3 方法

3.1 概要

本実験では、利害関係を形成する人間集団内の集団型最後通牒ゲームに計算機エージェント1体を参加させた。ゲームは4名のプレイヤーが参加し、そのうちの3名は人間被験者で、1体はエージェントであった。2種類のエージェントを用意した。1体のエージェントは提案された意思決定モデル[4]に基づいて集団規範に適応を試みながら意思決定した。一方で、もう1体のエージェントは比較のための意思決定モデルに基づいて意思決定した。提案された意思決定モデルでは適応的な振舞いを表出するために、強化学習的なフレームワークと価値関数の更新方法を基にしている。

ゲームにおいて、プレイヤー p が提示した金額を m_p 、受諾か拒否かの返答を r_p と定義する。プレイヤー p は集

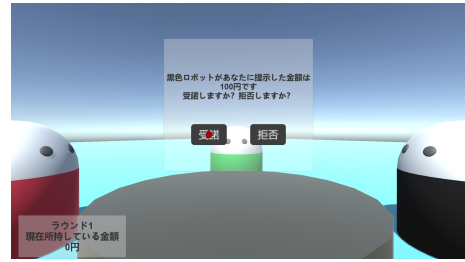


図 4: ある特定のプレイヤーの返答 (受諾か拒否) の入力画面。

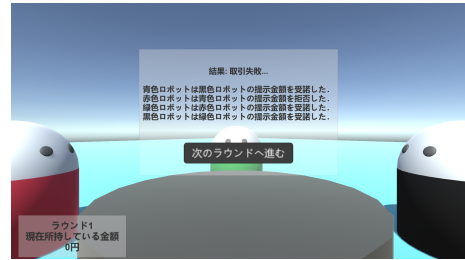


図 5: 1ラウンド目における全てのプレイヤーの判断の表示。全てのプレイヤーが同じ情報を得る。この後に第2ラウンドに移行する。

合 $\{1, 2, 3, a\}$ の要素であり、 a はエージェント、 $1, 2, 3$ は人間被験者を意味する。式(1)に示す通り、返答 r_p の値は受諾 (accept) の場合は0、拒否 (reject) の場合は1である。

$$r_p = \begin{cases} 0 & \text{if ACCEPT} \\ 1 & \text{if REJECT} \end{cases} \quad (1)$$

エージェントは他のプレイヤー p の m_p, r_p を入力とし、自身の意思決定である $m_{p=a}, r_{p=a}$ をゲームのプレイとして出力する。ただし、本実験においては人間のプレイヤーは自分の取引に関連するプレイヤーの情報のみに対して、エージェントは集団型最後通牒ゲームにおいて本来得られないプレイヤー全員の情報を得ている。

3.2 提案モデル搭載エージェント

3.2.1 環境とエージェント

図6はラウンド t における提案モデルに基づく意思決定の概要を示す。1回のラウンドにおいてエージェントは以下の順番に処理を実施する。

1. ゲームにおける行動 (Action) の表出
2. ゲームにおける他者行動の観察 (Observation)
3. 内部状態の更新 (Renew)

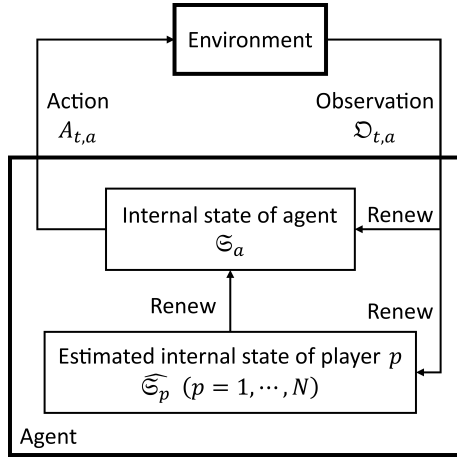


図 6: ラウンド t における集団規範形成モデルに基づく意思決定方法の概念図.

図 6 において、環境 (Environment) はエージェント (Agent) の外部であり、仮想ゲーム環境や他の人間プレイヤーを指す。エージェントは行動 ($A_{t,a}$: Action) を環境に出力し、環境の観察情報 ($\mathcal{D}_{t,a}$: Observation) を得る。式 (2) と式 (3) とに示す通り、行動 $A_{t,a}$ は自らの行動の出力 $m_{t,p=a}, r_{t,p=a}$ であり、観察情報 $\mathcal{D}_{t,a}$ は自分を含む全てのプレイヤーの情報である。ただし、ゲーム内の他者の人数は 3 名であるので、 $N = 3$ となる。

$$A_{t,a} = \{m_{t,a}, r_{t,a}\} \quad (2)$$

$$\mathcal{D}_{t,a} = \{A_{t,p}\} \quad (3)$$

$$(p = 1, \dots, N, a)$$

観察情報 $\mathcal{D}_{t,a}$ を基にして、エージェント自身の内部状態 (\mathcal{S}_a : Internal state of agent) とエージェントが推定するプレイヤー p の内部状態 ($\widehat{\mathcal{S}}_p$: Estimated internal state of player p) を更新 (Renew) する。加えて、エージェントによる他者推定結果である $\widehat{\mathcal{S}}_p$ に基づいて、エージェント自身の内部状態 \mathcal{S}_a がさらに更新される。以上で 1 ラウンド内での処理は完了する。次のラウンド $t+1$ においては更新された内部状態 \mathcal{S}_a に基づいて、行動 $A_{t+1,a}$ が環境に出力される。

3.2.2 行動と価値関数

ラウンド t でのエージェントの行動 $A_{t,a}$ は $m_{t,a}, r_{t,a}$ で構成され、それぞれ以下の式 (4) と式 (5) に示されるように、価値関数 V_{offr} と V_{diff} に基づいて導出される。式 (5) にある $m_{t,\rightarrow a}$ はエージェントに提示される金額のことを示す。 $m_{t,\rightarrow a} - m_{t,a}$ は自身の提示金額と自身に提示された金額の差分を意味する。その差分を許容

するか否かは提示された金額を受諾するか拒否するかの返答に基づく。

$$m_{t,a} = \arg \max V_{offr} \quad (4)$$

$$r_{t,a} = \begin{cases} 0 & \text{if ACCEPT} \\ & (\arg \max V_{diff} \leq m_{t,\rightarrow a} - m_{t,a}) \\ 1 & \text{if REJECT} \\ & (\arg \max V_{diff} > m_{t,\rightarrow a} - m_{t,a}) \end{cases} \quad (5)$$

価値関数 V は意思決定として実行や選択が可能な行動の集合の要素を引数とし、その行動を特定の集団内で実行したり選択したりする価値を数値で出力する関数である。価値関数の出力の値が高いほど、その出力を出す引数に当たる行動は集団内の集団規範への適応するための振舞いとして適しているものとエージェントは判断する。加えて、価値関数は他者の振舞いの観察情報を基に常に更新される。

本実験において、価値関数 $V_{offr}(m_o)$ はある提示金額 m_o を実際に表出することの集団内での価値を出力する。また、価値関数 $V_{diff}(m_d)$ は自身の提示金額 m_a と自身に提示された金額 $m_{\rightarrow a}$ の差を許容することの集団内での価値を出力する。 m_o と m_d はそれぞれ式 (6) と式 (7) に示される集合 M_o 、集合 M_d の要素である。

$$m_o \in M_o = \{0, 10, 20, \dots, 990, 1000\} \quad (6)$$

$$m_d \in M_d = \{-1000, -990, \dots, 0, \dots, 990, 1000\} \quad (7)$$

3.2.3 内部状態の構成

図 6 に示される内部状態 $\mathcal{S}_a, \widehat{\mathcal{S}}_p$ は以下の式 (8) と式 (9) のように構成される。

$$\mathcal{S}_a = \{V_{offr}(m_o), V_{diff}(m_d)\} \quad (8)$$

$$\widehat{\mathcal{S}}_p = \{\widehat{V}_{offr}^p(m_o), \widehat{V}_{diff}^p(m_d)\} \quad (9)$$

ゲーム内の他者である他のプレイヤーもエージェントと同様の意思決定の構造を持つと仮定して、エージェントが推定するプレイヤー p の内部状態 $\widehat{\mathcal{S}}_p$ は推定された価値関数 \widehat{V}_{offr}^p と \widehat{V}_{diff}^p によって構成される。これらの推定された価値関数は他者の振舞いの観察情報を基に常に更新される。

集団内でのインタラクションが開始される前の内部状態 $\mathcal{S}_a, \widehat{\mathcal{S}}_p$ は初期価値関数によって構成される。本実験では、特定の値を出力するように初期の価値関数を設定した上で実験を実施した。

3.2.4 自身の内部状態の更新

図 7 は集団型最後通牒ゲームのプレイヤー P_p の立場から見える周辺の取引の概念図を示す。集団メンバ P_p

の取引相手となるのは P_l と P_r であり、それぞれのプレイヤーから提示金額 $m_{\rightarrow p} = m_{P_r}$ や返答 $r_{\rightarrow p} = r_{P_l}$ がプレイヤー P_p に入力されている。したがって、プレイヤー P_p が直接かかわる取引の集合 T_p を式 (10) で表現する。

$$T_p = \{m_p, r_p, m_{\rightarrow p}, r_{\rightarrow p}\} \quad (10)$$

提示金額の価値関数 $V_{offr}(m)$ の更新式を式 (11)、価値関数 $V_{offr}(m)$ の更新に用いられる報酬関数 $R_{offr}(m)$ を式 (12) のように定義する。また、 $m \in M_o$ であり、プレイヤー番号の集合は $P = \{1, 2, \dots, N\}$ である。

$$V_{offr}(m) \leftarrow (1 - \alpha)V_{offr}(m) + \alpha (R_{offr}(m) + \gamma \max V_{offr}(m)) \quad (11)$$

$$R_{offr}(m) = \sum_{p \in P} \left\{ (1 - r_{\rightarrow p}) \exp \left(-\frac{(m_p - m)^2}{kurtosis} \right) \right\} + \sum_{p \in P} \left\{ r_{\rightarrow p} \exp \left(-\frac{(m_p + \Delta - m)^2}{kurtosis} \right) \right\} \quad (12)$$

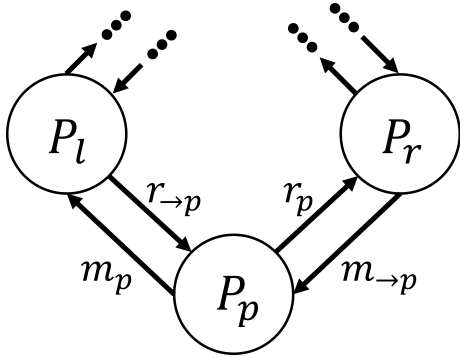


図 7: 集団型最後通牒ゲームのプレイヤー P_p の立場から見た取引の概念図。プレイヤー P_r の提示金額 m_{P_r} は $m_{\rightarrow p}$ 、プレイヤー P_l の返答 r_{P_l} は $r_{\rightarrow p}$ のように代替して表現している。

また、差額の基準の価値関数 V_{diff} と報酬関数 R_{diff} を式 (13) と式 (15) に示す。差額の基準の価値関数から、プレイヤーが持つ拒否の傾向で最も価値が高いものが導出される。差額 $m \in M_d$ である。各プレイヤーの提示金額 m_p に対する返答 $r_{\rightarrow p}$ を基にして報酬が算出される。

$$V_{diff}(m) \leftarrow (1 - \alpha)V_{diff}(m) + \alpha (R_{diff}(m) + \gamma \max V_{diff}(n)) \quad (13)$$

$$R_{diff}(m) = \sum_{p \in P} \{(1 - r_{\rightarrow p})E(m_{\rightarrow p}, m_p, m)\} + 0.5 \sum_{p \in P} \{r_{\rightarrow p}E(m_{\rightarrow p}, m_p, m)\} \quad (14)$$

$$E(m_{\rightarrow p}, m_p, m) = \exp \left(-\frac{(N(m_{\rightarrow p} - m_p) - m)^2}{kurtosis} \right) \quad (15)$$

$$N(m) = \frac{m - |m|}{2} = \begin{cases} 0 & (m \geq 0) \\ -m & (m < 0) \end{cases} \quad (16)$$

γ は割引率、 α は学習率を表している。 $kurtosis$ は、ガウス関数のパラメータであり、価値関数を更新するための報酬関数の形状を決定する。以上のパラメータは、エージェントの内部状態を構成する定数であり、内部状態の更新に関するエージェントの個性が反映されているといえる。

3.2.5 他者の推定内部状態の更新

エージェントが全てのプレイヤーの取引情報を得るのに対して、人間のプレイヤーは自身の周囲の取引情報のみを得る。エージェントと同様の方法で人間プレイヤー p が意思決定すると仮定して、その人間プレイヤーの内部状態 $\widehat{\Theta}_p$ を推定する。そこから導き出される推定される意思決定を基にして、エージェント自身の内部状態 Θ_a を再度更新する。

エージェントはプレイヤー p の内部状態を構成する価値関数 $\widehat{V}_{offr}^p(m_o)$ 、 $\widehat{V}_{diff}^p(m_d)$ を観察情報を基にして報酬関数 $\widehat{R}_{offr}^p(m_o)$ 、 $\widehat{R}_{diff}^p(m_d)$ を用いて更新する。価値関数の更新式と報酬関数の構造は式 (11) から式 (15) に示されるものと同様である。

プレイヤー p は式 (10) に示される取引情報のみを知るため、プレイヤー p の取引相手の集合 $\{p_p, p_r\}$ が式 (12) と式 (15) 内のプレイヤー集合 P に代入される。この時、 $\alpha, \gamma, kurtosis$ の値はエージェントのパラメータのものと同じとした。

3.2.6 他者推定に基づく自身の内部状態の更新

t ラウンド目における推定された他者 p の内部状態 $\widehat{\Theta}_p$ を構成する価値関数を基にしてエージェント自身の価値関数を更新する。エージェントが利益を得つつ、エージェントが提示相手から拒否を受けないために価値関数を調整することを目的とする。

式 (17) に示される方法でエージェント自身の価値関数 V_{offr} を更新することにより、エージェントが損をせず得をしない金額を提示するような価値関数の形状の形成を促す。式 (17) の不等式はエージェントの直接の取引相手 p_l, p_r の提示金額や差額の許容範囲の下限の推定を基にして、エージェントが提案すべき金額 m の範囲が示されている。差額の許容範囲の下限とは自身

が提示した金額と自身に提示された金額を差額としたときに、その差額を許容する限度のことを指す。例えば、差額が -100 円よりも大きいのであればその差額を許容するという考えを持つプレイヤーが居たときに、そのプレイヤーの差額の許容範囲の下限 $infAD$ は -100 である。

図 7 におけるプレイヤー p がエージェントであるとす。エージェントはプレイヤー P_r の差額の許容範囲の下限 $infAD_r$ を満たし、プレイヤー P_l の提示金額 \widehat{m}_l を推定して自身の差額の許容範囲の下限 $infAD$ を満たすような金額 m_a の提示をプレイヤー P_r に提示することが望ましい。式 (17) はその 2 つを両立可能な提示金額の範囲があれば、その範囲の金額を提示する価値を相対的に高めるように価値関数を更新する。

$$V_{offr}(m) \leftarrow c_v \times V_{offr}(m) \quad (17)$$

(not if $infAD_l + \widehat{m}_l \leq m \leq \widehat{m}_r - infAD$)

$$\widehat{infAD}_l = \arg \max \widehat{V}_{diff}^l \quad (18)$$

$$\widehat{m}_l = \arg \max \widehat{V}_{offr}^l \quad (19)$$

$$\widehat{m}_r = \arg \max \widehat{V}_{offr}^r \quad (20)$$

$$infAD = \arg \max V_{diff} \quad (21)$$

3.3 比較用エージェント

本実験では、集団規範への適応を試みる提案モデルの比較対象として、適応的でない振舞いを表出するエージェントを用意した。具体的には、提示金額 m_a と返答 r を常に 100 円の提示と受諾 ($A_a = \{m_a = 100, r_a = 0\}$) として出力し続けるエージェントが本研究における比較対象である。

初期の価値関数は両エージェントとも同じでラウンド 1 における行動は $m_a = 100, r_a = 0$ を出力し、ラウンド 2 以降の行動は異なる。提案モデル搭載エージェントは意思決定と観察の後に内部状態を更新するが、比較用エージェントは内部状態を更新しない。これによって、集団規範に適応的なエージェントと適応的ではないエージェントを比較する。

4 実験

4.1 実験の概要

図 8 に集団内の被験者のゲーム内のポジションを示す。 P はプレイヤーを意味し、 P_1, P_2, P_3 は人間被験者であり、 P_A は計算機エージェントである。実験内の 3 名の被験者はプレイヤー P_1, P_2, P_3 のうちのいずれかにラ

ンダムに決まり、各被験者は取引相手が現実空間の誰なのかを知ることはできない。つまり、取引相手が人間またはエージェントのどちらであるかも被験者は分からない。

本研究では 3 人の被験者と 1 体のエージェントが仮想空間内の集団として集団型最後通牒ゲームに取り組む。集団内の 3 名の人間プレイヤーと実験主催者 1 名は同じ空間で実験に取り組んだ。15 名の被験者はそれぞれ 5 個の集団 (A, B, C, D, E) のいずれかの集団に所属し、各集団においてゲームに取り組んだ。被験者は全員男性で、平均年齢は $22.53 (SD = 0.8844)$ であった。加えて、エージェントが提案モデルに基づいて意思決定する場合のパラメータを表 1 に示す。

集団内のプレイヤーは集団型最後通牒ゲームを 2 度だけプレイする。1 度目のゲームから 2 度目のゲームに移行する際に被験者のポジションは変わる。したがって、金額提示もしくは受諾か拒否かの判断の提示を 2 回のゲームで同じプレイヤーと行うことはない。

本実験終了後に被験者はゲーム内で得られた金額に比例する形で成功報酬を得る。具体的には、500 円から 1000 円の Amazon ギフトカードを配布した。被験者は 2 回のゲームの内のより多くのスコアを得たゲームに基づいて報酬を得る。被験者は集団型最後通牒ゲームのルール説明の際に以上の報酬についての説明を受けた。

本実験は富山県立大学「人を対象とする研究」倫理審査部会の承認を得て実施された。

表 1: 提案モデルのパラメータ

| | |
|------------------------------|-------|
| 学習率 α | 0.2 |
| 割引率 γ | 0.8 |
| 報酬関数パラメータ Δ | 50 |
| <i>kurtosis</i> | 10000 |
| initial $\arg \max V_{offr}$ | 100 |
| initial $\arg \max V_{diff}$ | -100 |

4.2 結果

表 2 は各集団における 2 回分のゲームの結果を示している。各集団のプレイヤーが各ゲームにおいて最終的に得た金額、集団内プレイヤーの得た金額の合計、エージェントが得た相対的な利益の絶対値 (集団内プレイヤーが得た金額の平均値とエージェントが得た金額の偏差の絶対値) が示されている。集団 (Group) A と B には 1 回目のゲームに提案モデル搭載エージェントが参加し、2 回目のゲームに比較用モデル搭載エージェントが参加した。集団 (Group) C, D, E には逆の順番でエージェントが参加した。

表 2: 各集団のプレイヤーが得た金額と計算機システムの相対的な利益の絶対値.

| Group | Model | Total amount of player | | | | Sum of total amount | Absolute value of agent's relative gain |
|-------|----------------|------------------------|--------|--------|----------|---------------------|---|
| | | 1 | 2 | 3 | 4(agent) | | |
| A | proposed | 9,060 | 8,720 | 9,080 | 9,140 | 36,000 | 140 |
| B | proposed | 9,870 | 10,850 | 12,150 | 11,130 | 44,000 | 130 |
| C | proposed | 9,800 | 9,990 | 10,138 | 10,072 | 40,000 | 72 |
| D | proposed | 10,810 | 8,950 | 10,050 | 10,190 | 40,000 | 190 |
| E | proposed | 10,060 | 12,130 | 11,350 | 10,460 | 44,000 | 540 |
| A | for comparison | 7,320 | 7,480 | 8,000 | 9,200 | 32,000 | 1,200 |
| B | for comparison | 9,950 | 7,730 | 9,330 | 12,990 | 40,000 | 2,990 |
| C | for comparison | 3,700 | 2,600 | 4,581 | 5,119 | 16,000 | 1,119 |
| D | for comparison | 8,500 | 8,650 | 10,342 | 8,508 | 36,000 | 492 |
| E | for comparison | 6,000 | 7,000 | 7,150 | 7,850 | 28,000 | 850 |

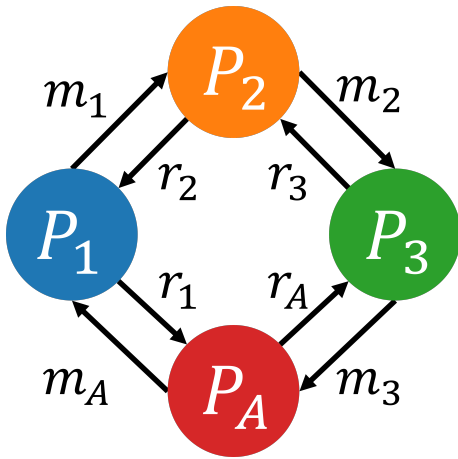


図 8: 人間プレイヤーとエージェントの取引 (プレイヤーの提示する金額 M と提示金額に対する判断 R) の流れ. P_1, P_2, P_3 は人間プレイヤーで, P_A はエージェントのプレイヤーであることを意味する.

図 9 はエージェントが得た最終的な金額とエージェントが所属する集団内の他プレイヤーが得た最終的な金額の差の絶対値を示す. 集団規範への適応性を持つモデルと適応性を持たない比較用モデルの側面から比較している. 比較用モデルに基づくエージェントと比較して, 提案モデルに基づくエージェントの方が集団内の他のプレイヤーに近い金額を得る傾向があることがわかる. Brunner–Munzel 検定 [9] によって両群の間には有意差が見られた ($W_N^{BF} = 3.476$, $p\text{-value} = 0.002367$).

図 10 は被験者が得た最終的な金額の比較を, 提案モデルに基づくエージェントが参加する集団と, 比較用モデルに基づくエージェントが参加する集団で比較している. 提案モデルに基づくエージェントが参加する集団の人間メンバは比較的高い利益を得ていることがわかる. Brunner–Munzel 検定 [9] によって両群の間には有意差があった ($W_N^{BF} = -8.391$, $p\text{-value} < 0.001$).

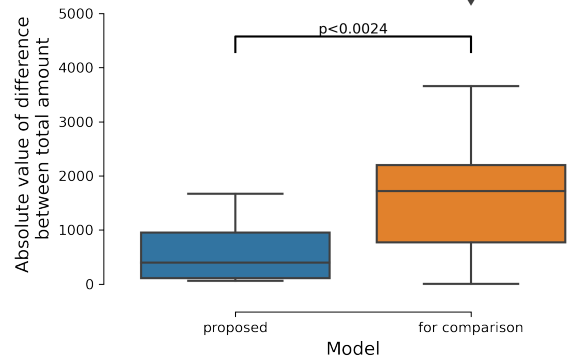


図 9: エージェントが得た最終的な金額とエージェントが所属する集団内の他プレイヤーが得た最終的な金額の差の絶対値. エージェントが集団内の他プレイヤーと比較してどの程度の利益を得たのかを示す指標.

4.3 考察

図 9 から比較用モデルに基づくエージェントと比較して, 提案モデルに基づくエージェントの方が集団内の他のプレイヤーに近い金額を得る傾向があることがわかる. 提案モデルに基づくエージェントは集団に集団規範が発生すると想定し, 集団規範に適応的な振舞いを表出した. 一方で, 比較用モデルは集団規範に適応的な意思決定を行わず, 全てのラウンドで同じ金額 100 円を提示し, エージェントに提示された金額を必ず受諾した. その結果, 提案モデルに基づくエージェントの集団内他者に対する相対的な利益は低く, 比較用モデルに基づくエージェントの利益はほとんどの場合相対的に高い利益を得た. このことは, ゲームでの各集団メンバの利益に偏りが無いことを公平であると定義するなら, 比較用モデルと比較して提案モデルは利害関係下で公平な意思決定を実現したことを示している.

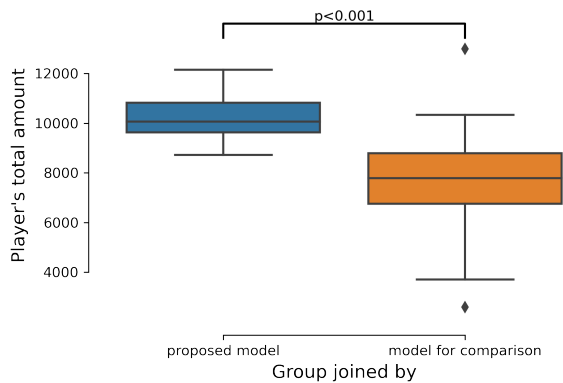


図 10: 提案モデルに基づくエージェントが参加する集団と比較用モデルに基づくエージェントが参加する集団にそれぞれ参加した被験者が得た最終的な金額の比較。どちらの種類のエージェントと共にゲームに参加する方が利益を得られるのかを示す指標。

図 10 は集団に参加した被験者が得た利益の群間での比較である。表 2 が示すように、比較用モデルに基づくエージェントは集団内の他のプレイヤーと比べて、相対的に高い利益を得た。しかしながら、集団内のプレイヤーが得られる利益は、提案モデルに基づくエージェントが参加するプレイヤーの利益と比べて、相対的に低いものとなった。このことから、集団規範に適応的な振舞いを表出するエージェントが公平さを示すことで、結果的に集団全体の利益を押し上げる、集団内プレイヤー間の互恵的な振舞いを実現した可能性がある。

5 おわりに

本研究では利害関係のある人間集団シナリオに参加したエージェントが示した振舞いやその人間への影響を調査した。特に、人間集団が形成する暗黙的集団規範に着目し、集団規範に適応的なエージェントと適応的でないエージェントの間での比較によって、集団内のエージェント自身や他者にどのような影響を与えるのかを調査した。

実験結果より、集団規範に適応的な振舞いを示すエージェントは、利害関係下でもエージェント自身が極端に得をしたり損したりしないよう意思決定できたこと、その適応的な意思決定が集団内プレイヤーへの公平な利益の分配につながったことが明らかになった。加えて、集団規範に適応的な振舞いを示すエージェントが参加する集団のプレイヤーは比較的多くの利益を得ることより、互恵的な振舞いが表出される傾向がみられた。

将来的には、提案モデルの他者推定モジュールの開発、コミュニケーションロボットを用いた人間とロボットの集団での実験の実施、提案モデルに基づくエージェ

ントやロボットに対する人々の主観的な印象の調査に取り組む。

謝辞

本研究は JSPS 科研費 JP21K12099 の助成を受けたものである。

参考文献

- [1] Leite, I. et al.: Social robots for long-term interaction: A survey, *International Journal of Social Robotics*, Vol. 5, No. 2, pp. 291–308 (2013)
- [2] Sheridan, T. B.: A review of recent research in social robotics, *Current Opinion in Psychology*, Vol. 36, pp. 7–12 (2020)
- [3] Feldman, D. C.: The development and enforcement of group norms, *Academy of Management Review*, Vol. 9, No. 1, pp. 47–53 (1984)
- [4] Fuse, Y., Tokumaru M.: Social Influence of Group Norms Developed by Human–Robot Groups, *IEEE Access*, Vol. 8, pp. 56081–56091 (2020)
- [5] Fehr, E., Schmidt, K. M.: A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics*, Vol. 114, No. 3, pp. 817–868 (1999)
- [6] Fuse, Y. et al.: Decision-Making Model for Robots that Considers Group Norms and Interests, *Human–Computer Interaction. Technological Innovation: Thematic Area, HCI 2022, Held as Part of the 24th HCI International Conference*, pp. 475–485 (2022)
- [7] Thaler, R. H.: Anomalies: The ultimatum game, *Journal of economic perspectives*, Vol. 2, No. 4, pp. 195–206 (1988)
- [8] Oosterbeek, H. et al.: Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis, *Experimental Economics*, Vol. 7, pp. 171–188 (2004)
- [9] Neubert, K., Brunner, E.: A studentized permutation test for the non-parametric Behrens–Fisher problem, *Computational Statistics and Data Analysis*, Vol. 51, pp. 5192–5204 (2007)