

ローカルクエリが人-AI間の交渉の Win-Win 解に与える影響

佐藤幹晃^{1*} 寺田和憲¹
Motoaki Sato¹ Kazunori Terada¹

¹ 岐阜大学
¹ Gifu University

Abstract: チームでの Win-Win の合意形成には相手がなにを重要視するか (選好) の理解が重要であるが, その認知過程は不明であった。本研究は, ヒューリスティックな選好探索戦略「ローカルクエリ」の学習が Win-Win となる意思決定に寄与するか, 参加者実験で検証する。参加者 (n=76) はローカルクエリ戦略学習有無の条件で, AI との 4 論点の最後通牒 (交渉) ゲームに参加する。実験の結果, ローカルクエリ学習を行うことで, 交渉の提案の Win-Win の程度と選好理解の正確さが向上することが明らかとなった。このことは, ローカルクエリ戦略学習が Win-Win の合意形成に寄与することを示唆する。

1 はじめに

交渉は多目的最適化問題として定式化でき, 理論的には数的手法を用いて解を求めることが可能である [Wierzbicki 92]. しかし, これは情報が完全に開示され, 意思決定が合理的に行われるシナリオに限定される [Fisher 81]. 現実には, 交渉における競争的インセンティブと情報の非対称性から, 交渉者は選好やリミットに関する情報を隠蔽したり, 相手を欺いたり疑ったりすることが多い [Schweitzer 05, Fulmer 09, Gaspar 12]. また, 交渉者は必ずしも完全に合理的ではなく [Bazerman 09], 認知バイアス [Caputo 13], 例えば相手が自分と同じ選好であるというバイアス (固定バイアス) [Bazerman 83] や他方が得をしたらもう片方は必ず損をするというバイアス (ゼロサムバイアス) [Meegan 10, Johnson 22, Davidai 23] の影響を受けやすい。さらに, Win-Win の計算に必要な最適化計算の難しさ [Schweitzer 05] や交渉者間の信頼の欠如 [Lewicki 10] が問題解決を妨げている。これらの相互に関連する認知計算上の課題により, Win-Win 解の達成が困難になっている。本研究では, これらの課題を計算論的な観点から分解することで, 特に非言語情報のみがやりとりされる状況において, Win-Win 解の達成に貢献できるかどうかを明らかにする。

交渉における主要な認知計算上の課題の多くは, 情報の非対称性を軽減するための効果的な方法として知られている相手モデリングによって対処できる [Billings 98]. 相手モデリングとは, 相手の行動を理解し予測するた

めに, 相手の目標, 選好, リミットを理解することである [Herik 05, Wu 17, Ong 15]. 交渉におけるリミットとは, 交渉者が受け入れることができる最低の条件であり, それを超えると合意を受け入れるよりも交渉から撤退することを選択する。このリミットは通常, 交渉者の交渉合意に代わる最良の選択肢 (Best Alternative to Negotiated Agreement: BATNA) の価値によって決定される [Fisher 81]. 選好を推論することは, それぞれの当事者が価値を置くものの違いを明らかにすることで, 両者が同じものを望んでいるというバイアス (固定バイアス) の軽減に役立つ [Dreu 00a]. これにより, 交渉者は, 一方の当事者の利得が必ずしも他方の損失を意味するわけではなく, 交渉で生み出される全体的な価値を拡大できる可能性があると認識できる。

相手の選好, リミットを推論し相手モデリングをする上で, 感情表現は重要な情報であるが, これらの表現に基づく相手モデリングには大きな認知計算上の課題がある [Sharma 20, Kleef 18]. これらの課題は, 主に感情表現に内在する曖昧さに起因する [Barrett 11, Wu 17]. 特定の感情表現には複数の解釈が可能であり, 様々な文脈 [Gelder 06, Barrett 11] や心的状態 [Schmid 10] が同様の外見的表現を生み出す可能性がある [Barrett 11]. リミットと選好を感情が符号化している場合, ローカルクエリにより選好のみの推論に焦点を当てることができる。ローカルクエリは, AI アルゴリズムにおける選好学習の手法の一つであり, 交渉において一つの論点の配分を変化させながら, 他の論点の配分を固定することで, 対象の論点に対する選好を効率的に学習するものである [Braziunas 08].

交渉における感情シグナルのデコーディングの重要性を示した多くの研究がある一方で [Dehghani 14, Lewicki 10,

*連絡先: 岐阜大学工学研究科工学専攻
〒501-1193 岐阜県岐阜市柳戸 1-1
E-mail: motoaki.sato@ai.info.gifu-u.ac.jp

Hoegen 22, Martinovski 10], このプロセスに伴う認知計算上の課題, 特に選好とリミットの同時推論の文脈での課題に取り組むことに焦点を当てた研究は比較的少ない. 本研究では, 相手モデリングの複雑なタスクを, 選好学習, リミット推定, Win-Win 計算というより管理しやすい構成要素に分解する. これにより, 固定バイバイアス [Bazerman 83, Dreu 00a] とゼロサムバイバイアス [Meegan 10, Johnson 22, Davidai 23] を軽減し, 最終的により多くの Win-Win の結果につながると仮説を立てている.

1.1 関連研究

交渉研究では, 相手の選好を理解し, リミットを推定し, 互いに有益な解を見つけることの重要性が長い間認識されてきた. しかし, これらの認知的タスクは, 特に非言語的な文脈において, 異なるステップではなく相互に関連するプロセスとして研究されることが多かった. 本節では, 交渉における感情表現の研究の発展を概観する. 初期の研究では感情が結果に与える影響を検討していたが [Pillutla 96, Morris 00], 最近の研究では, 選好を伝える手段として感情フィードバックを取り入れている [Pietroni 21, Lelieveld 22]. 交渉プロセスの明確かつ重要な最初のステップとして, 特に感情表現が重要な役割を果たす非言語的交渉において, 選好学習に焦点を当てることの重要性を示している.

Pietroni らは, エージェントの選好を反映した感情フィードバック (テキストまたは表情) を提示する 2 論点 4 ラウンドの交渉 (売り手と買い手での付帯利益と福利厚生との交渉) を行った [Pietroni 08]. 彼らは, 参加者の固定バイバイアスと要求を測定した. その結果, 感情フィードバックの存在が固定バイバイアスを減らし, より統合的な解につながる事が確認され, 参加者が感情から相手の選好を読み取ることができたことが示された. しかし, テキストによる感情のフィードバックは, 「モニターの配分には満足しているが, ハードディスクの配分には納得していない」のように, 各論点に対する評価が独立して与えられ, 選好を直接読み取ることが可能であった. さらに, 表情は, 提案自体に対する評価的フィードバックではなく, 各論点に対して独立して提供された. 認知タスクには, ローカルクエリにより表情から選好を復元することは含まれていなかった.

これらの感情のみの交換に焦点を当てた研究に加えて, 感情を補足的な要因として扱い, 選好やリミットに関するテキスト情報の交換と組み合わせた研究も多数ある. 例えば, IAGO (Interactive Arbitration Guide Online) プラットフォームを使用した研究では, 感情的手がかりをテキストコミュニケーションと統合する

さまざまなアプローチが実証されている [Gratch 15, Gratch 16, DeVault 15, Mell 17, Johnson 20]. IAGO では, 半構造化されたインタラクションを通じて, 交渉者は絵文字を使って自分の選好を表現したり, 直接相手の選好を尋ねたり, トレードテーブルを使ってオファーをしたりすることができる. このインタフェースでは 12 の定義済みメッセージと 4 つのクリック可能なアクションを使ってやり取りし, 実験のセットアップに応じて通常 7 分から 15 分の指定された時間内に必要な回数だけメッセージを交換することが許された. 交渉者はコミュニケーションにおいて戦略的あるいは曖昧になることを選択するかもしれない, 真の選好とリミットを確かめるタスクを複雑にしている.

佐藤らは, 選好とリミットの両方を推測する必要がある非言語的交渉タスクにおいて, 参加者がある程度の選好学習を行うことができることを示した [Sato 23]. しかし, Win-Win 解に到達する代わりに, 参加者は自分の取り分を減らす傾向があった. これらの結果は, 交渉中の選好とリミットのデコーディングという二重の認知タスクの難しさ, 特に感情表現の読み取りを通じたそれが, Win-Win の割り当ての追求を妨げた可能性があることを示唆している. このゲームは, 2 人のプレイヤー (人間の提案者とエージェントの応答者) が複数論点のアイテムを分割する一方向の一回勝負の交渉であった. 交渉中に相手の選好を復元するために, 参加者はローカルクエリを行う必要があった. 参加者はスライダーバーを操作してオファーのセットを作成することができた. 各オファーのセットに対して, エージェントは総効用に基づいて表情を提示した. これは, Pietroni らの個々の論点に対する評価が提供されたタスクとは異なり, 総効用を構成する重みを計算するタスクであった [Pietroni 08]. しかし, エージェントの表情にはリミットもエンコードされていた. 参加者はローカルクエリを通じて相手の選好を復元することができたかもしれない. 選好学習のパフォーマンスは真の選好との順位相関係数で約 0.4 であり, 参加者自身の効用の減少は, 交渉中に 3 つの認知タスク (Win-Win の選好とリミットの推定, Win-Win 計算) を同時に行うことの難しさを示唆している. 特に, 選好学習はローカルクエリを通じて行う必要があったため, そのような戦略に不慣れな参加者は, 相手の選好を十分に復元できなかった可能性がある.

1.2 本研究

本研究の目的は, 非言語的交渉における選好推定の重要性と, 特にローカルクエリの役割を明らかにすることである. 先行研究では, リミットと選好の推論という認知的タスクが混在した形で提示されることが多

く、交渉の難しさを認知計算的な観点から十分に説明することが難しかった。本研究では、非言語的交渉における事前のローカルクエリ戦略を学習することの重要性と、交渉中のローカルクエリの役割に焦点を当てる。以下の仮説を検証する。

H1 交渉前にローカルクエリ戦略を学習することで、Win-Win 解に到達することに貢献する

H2 交渉前にローカルクエリ戦略を行うことで、固定バイバイアスを解消することに貢献する

これらの仮説を検証するために、選好とリミットについて数学的に定義された評価モデルを持つエージェントと人間の参加者との交渉を行った。参加者は、感情表現を通してエージェントの選好とリミットをデコードするタスクを与えられた。具体的には、独立した選好学習（ローカルクエリ戦略学習）の有無と、交渉中のローカルクエリの使用が、客観的に測定された交渉パフォーマンスと固定バイバイアスに与える影響を統計的に調査した。本実験の設計には以下の重要な要素が含まれている。

- 参加者が感情表現を通してエージェントの選好とリミットをデコードできるように、選好とリミットについて数学的に定義された評価モデルを備えたエージェント
- 選好学習の機会を独立して与えられた参加者のサブセット。この機会において、参加者は交渉の各論点に対する好みの度合いを反映したエージェントの感情反応を観察できた。
- 参加者がエージェントにオファーのセットを提示し、エージェントがオファーのセットから計算された効用値に基づいて感情表現を表示する複数論点最後通牒ゲーム。この過程で、参加者は交渉中にローカルクエリを使用して、エージェントの選好を推定することができる。

このアプローチにより、リミットの推定とローカルクエリの効果を分離し、非言語的交渉における Win-Win 解の達成と固定バイバイアスの低減に与える影響を検証することができる。

2 実験方法

2.1 実験参加者

実験参加者は Prolific で募集した。適切なサンプルサイズを決定するために、Cohen らが提案した方法で、G*Power を用いて検定力分析を行った。2水準（ローカルクエリ戦略学習：有 vs. 無）の参加者間計画で、効

果量 $f = .25$, $\alpha = .05$, 望ましい統計的検定力 .95 の場合、分析では総サンプルサイズ 76 人が推奨された。ID の不一致とアテンションチェック該当者を取り除き、19 歳から 70 歳までの男性 37 人、女性 38 人、その他 3 人を分析に用いた ($M_{age} = 33.34$, $SD_{age} = 12.09$)。実験計画は 2(ローカルクエリ戦略学習：有/無) × 2(複数論点最後通牒ゲーム：1 回目/2 回目) の参加者間要因配置で行った。

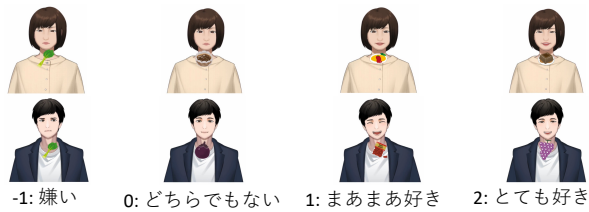
2.2 複数論点最後通牒ゲーム

標準的な単一論点の最後通牒ゲームを多論点に拡張した。このゲームでは、人間の提案者と AI エージェントの応答者の 2 人のプレイヤーが、一方向の一発勝負のやり取りで複数のアイテムについて交渉する。合意に達した場合、各プレイヤーは交渉で得た取り分を受け取る。合意に達しない場合、両者ともに何も受け取れない。

佐藤らが提案した多論点インターフェースを使用した [Sato 23]。このインターフェースでは、人間のプレイヤーが各論点のスライダーバーを使ってアイテムをエージェントに割り当てることができる (図 1e を参照)。AI エージェントは、オファーの評価に対応する表情を表示することでこれらの割り当てに回答する。標準の最後通牒ゲームはゼロサムであるが、本研究の多論点版では、プレイヤーの選好が論点によって異なる場合、非ゼロサムの結果が可能になる。この設計により、Win-Win となる解決策の可能性を探ることができる。

2.3 ローカルクエリ戦略学習

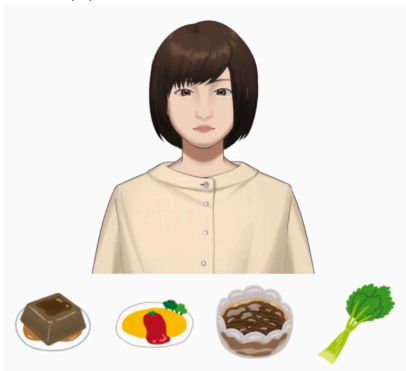
参加者は、エージェントに拒否されることを避けながら、自分のポイントを最大化することだった。これを達成するためには、エージェントの選好 ($w^{\text{エージェント}}$) とエージェントのリミットの両方を推論する必要があった。図 1e に示すソフトウェアを使用し、選好とリミットの同時推論は理論的には可能であるが、佐藤らによる先行研究では、この同時タスクの難しさが示されている [Sato 23]。この困難に対処するため本研究では、ローカルクエリ戦略学習条件有りの参加者は図 1d に示すインタフェースで、エージェントの表情にエンコードされた選好を観察することで、選好を学習した。エージェントの選好はに設定した。-1 はエージェントが嫌いな争点に、2 はエージェントが非常に好きな争点に割り当てられた。 w^{Agent} と表情の関係を図 1a に示す。



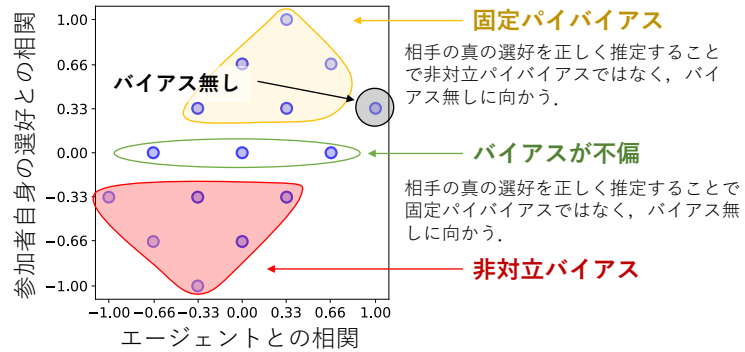
(a) アイテムごとに表出するエージェントの表情

アイテム (個数)	A (7)	B (5)	C (5)	D (5)
$w_{\text{参加者}}$	2	1	0	-1
$w_{\text{エージェント}}$	2	0	-1	1

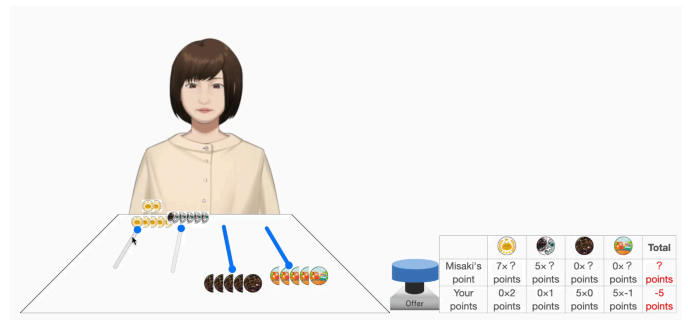
(b) アイテムの重み



(d) ローカルクエリ戦略学習



(c) 自分と相手との選好の順位相関の可能な組み合わせ。



(e) 複数論点最後通牒ゲーム

図 1: 方法: 選好学習のインターフェース

2.4 手順

参加者は Prolific から Qualtrics で作成したアンケートページに進み、最初にインフォームドコンセントを提供し、人口統計学的質問に回答した。次に、参加者は複数論点最後通牒ゲームタスクの詳細な説明を読み、クイズを完了した。次に、参加者は 16 の食品アイテムを 1 から 4 の順位 (-1: 嫌い, 0: どちらでもない, 1: 好き, 2: 非常に好き) で順位付けするよう求められた。この順位は $w_{\text{参加者}}$ の順序を決定するために使用された。次に、参加者は、タスクを開始する前にエージェントの選好をどのように推測したかを評価するために、エージェントが最も好きな論点を 1 から 4 の順位で並べ替えるよう求められた。

次に、ローカルクエリ戦略学習あり条件の参加者のみが、ウェブインターフェース上でローカルクエリ戦略学習の説明を読み、ローカルクエリ戦略学習を開始した。図 1d は、学習フェーズにおける食品の提示する例を示している。最初に、無表情のエージェントと 4 つの食品の画像が表示される。参加者が画像をクリックすると、クリックされた画像が拡大されてエージェントの胸に表示される。その後、条件に応じて、エージェントが無表情、喜び、または怒りの表情を示す。感情表現を示した後、エージェントは無表情に戻り、拡

大された画像が消える。その後、画面は最初の画面に戻る。参加者は 4 つの食品の画像を任意の順序で 1 つずつクリックし、エージェントに提示した。参加者は、エージェントの感情表現の変化に基づいて、エージェントの食品の選好を推定するよう求められた。ローカルクエリ戦略学習の後、参加者はエージェントが最も好きな論点を 1 から 4 の順位で並べるよう求められた。一方、ローカルクエリ戦略学習なし条件では、参加者はこの学習フェーズを行わなかった。そして全ての参加者が、これらの食品を配分する 4 論点最後通牒ゲーム (1 回目) を行い、エージェントが最も好きな論点を 1 から 4 の順位で並べ、エージェントのリミットを推定するよう求められた。その後、食品からシールに配分するアイテムを変えて、4 論点最後通牒ゲーム (2 回目) を行った。シールについては、全ての参加者は、食品とは異なり、ローカルクエリ戦略学習は行わなかった。1, 2 回目の最後通牒ゲームにおいてエージェントのアイテムの重み (選好) はどちらも表 1b に基づいて決定された。

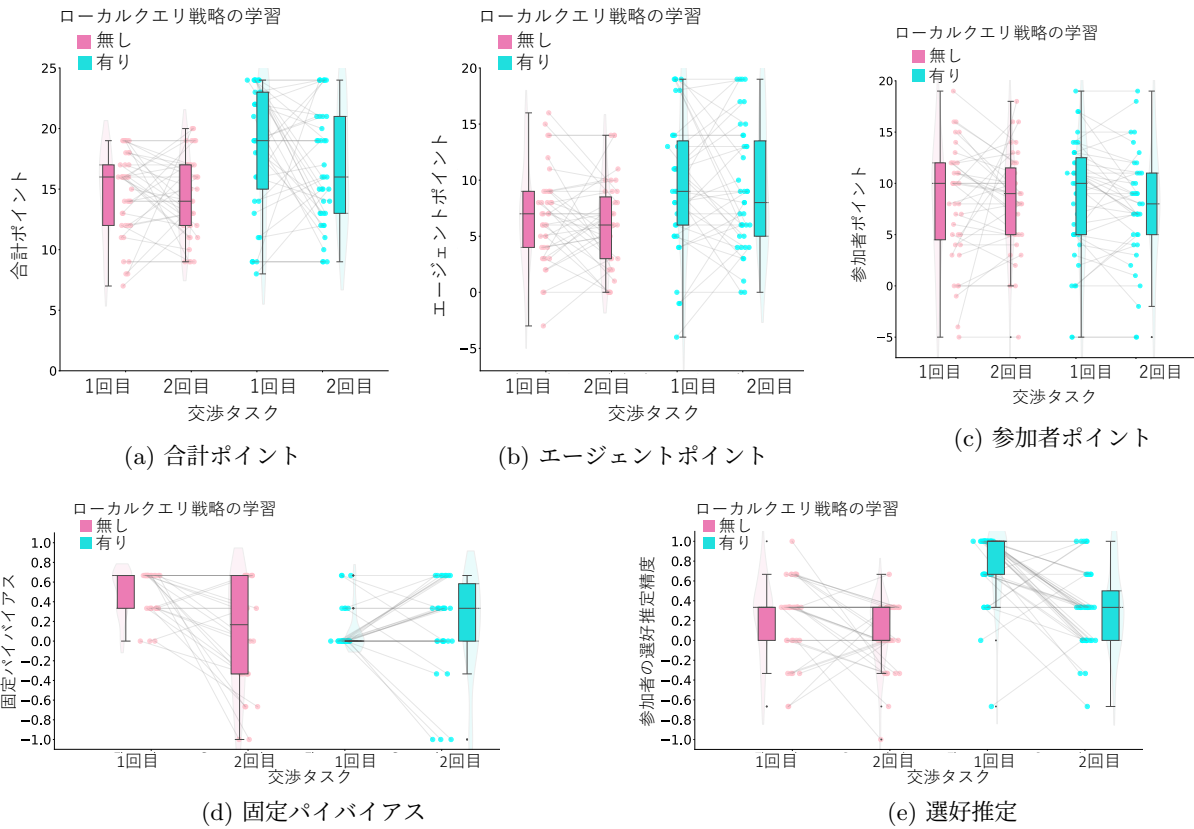


図 2: 実験結果

2.5 報酬

参加者には、参加費として5.5ドルを支払った。さらに、参加者が自分自身の得点を高くする動機づけのために、複数論点最後通牒ゲームで参加者自身が獲得した得点に応じて追加の報酬を得る機会を提供した。本研究は、岐阜大学大学院医学系研究科の医学審査委員会の承認を得ている。参加者には、データセットに参加者の情報が使用されることについて、インフォームドコンセントを行った。

2.6 測定と分析

複数論点最後通牒ゲームにおいて、エージェントのポイント、参加者のポイント、両者の合計ポイントを測定し、交渉のパフォーマンスを評価した。先行研究 [Johnson 20, Sato 23] に倣い、ケンドールの順位相関係数を用いて、相手の選好を推定する際のバイアスを測定した。具体的には、参加者が推測したエージェントの選好の順序と、参加者自身の選好およびエージェントの実際の選好との順位相関係数を計算した。これにより、固定バイバイアスと非対立バイアス（参加者が自分の選好はエージェントの選好と対立していないという思い込み）の程度を評価することができた。

図 1c は、自分と相手の選好の順位相関係数の可能な組み合わせを示している。相関係数は -1 （完全に逆の

順位）から $+1$ （完全に正しい順位）の範囲を取る。まず、横軸に注目すると、エージェントの選好との順位類似度が 1 である右端の点 $(1, 0.33)$ は、完全に正確な推定、すなわちバイアスがないことを表している。逆に、 -1 の左端の点は、参加者がエージェントの真の選好とは完全に逆の選好を推定したことを示している。この点では、参加者自身の選好との順位類似度が 0.33 であることに注意が必要である。しかし、これは固定バイバイアスを示すものではなく、実験設計における選好的部分的な一致を反映している（論点 A は完全に分配的、論点 B と C は部分的に分配的、論点 D は互換的。表 1b を参照）。このポイントの選好設計により、参加者がエージェントの選好を完全な精度で推定した場合でも、参加者自身の選好との類似度は 0.33 となる。しかし、これはバイアスとして解釈すべきではなく、実験設定に固有の特性である。むしろ、この点は、参加者が自分の選好とエージェントの選好の類似点と相違点の両方を正確に認識している理想的で偏りのない状態を表している。次に、縦軸に注目する。参加者自身の選好との類似度は、2種類のバイアスを区別する。固定バイバイアス（参加者がエージェントの選好を自分の選好と類似していると考え）は参加者自身の選好との類似度の正の領域で表され、非対立バイアスは参加者自身の選好との類似度の負の領域で表される。相関係数が 0 である縦軸上の3点、すなわち $(-0.66, 0)$ 、 $(0, 0)$ 、 $(0.66, 0)$ は、偏りのない不正確な推定を表して

いる。

交渉フェーズ中のスライダバーの移動回数（ローカルクエリ）を測定し、固定バイバイアスを低減し、バイバイアス無しの状態になるために必要な要因を検証した。

2.7 分析

エージェントが獲得したポイント、参加者が獲得したポイント、両者が獲得したポイント（合計ポイント）について、2水準（ローカルクエリ戦略学習：有/無）の繰り返しありの分散分析を行なった。参加者は、複数論点最後通牒ゲームの前と後の2回、エージェントの選好を推定したので、固定バイバイアスと選好推定について、1回目・2回目の4アイテムの複数論点最後通牒ゲームのインタラクション要因（1回目/2回目水準：参加者内）×ローカルクエリ戦略学習の有無の要因（ローカルクエリ戦略学習：有/無）の2要因混合計画で行った。

3 実験結果

合計ポイントについて、繰り返しありの二要因分散分析を行った結果、外見要因が有/無の主効果が観測された ($F(1, 76) = 15.28, p < .001, \eta_p^2 = .167$) が交互作用はなかった ($F(1, 76) = 1.00, p = .320, \eta_p^2 = .013$)。Bonferroniの方法による多重比較を行った結果、1回目交渉における合計ポイントはローカルクエリ戦略学習有条件 ($M = 18.26, 95\% CI [16.85, 19.67]$) の方が、無条件 ($M = 14.64, 95\% CI [13.23, 16.05]$) よりも有意に高く ($M_{diff} = 3.62, p < .001$)、2回目もローカルクエリ戦略学習有条件 ($M = 16.82, 95\% CI [15.51, 18.14]$) の方が、無条件 ($M = 14.36, 95\% CI [13.05, 15.67]$) よりも有意に高いことがわかった ($M_{diff} = 2.46, p = .010$) (図 2a)。この結果は、ローカルクエリ戦略学習が、エージェントと参加者にとって Win-Win な分配をすることに寄与することを示す。

エージェントポイントについて、繰り返しありの二要因分散分析を行った結果、ローカルクエリ戦略学習の主効果が観測された ($F(1, 76) = 9.89, p = .002, \eta_p^2 = .115$) が交互作用はなかった ($F(1, 76) = 0.00, p = 1.000, \eta_p^2 = .000$)。Bonferroniの方法による多重比較を行った結果、1回目の交渉ではローカルクエリ戦略学習有条件 ($M = 9.72, 95\% CI [8.07, 11.36]$) の方が、無条件 ($M = 6.64, 95\% CI [5.00, 8.29]$) よりも有意に高く ($M_{diff} = 3.08, p < .010$)、2回目もローカルクエリ戦略学習有条件 ($M = 9.18, 95\% CI [7.65, 10.71]$) の方が、無条件 ($M = 6.10, 95\% CI [4.57, 7.63]$) よりも有意に高いことがわかった ($M_{diff} = 3.08, p = .006$) (図 2b)。この結果は、ローカ

ルクエリ戦略学習により、参加者がエージェントに譲歩するようになったことを示している。

参加者ポイントについて、繰り返しありの二要因分散分析を行った結果、ローカルクエリ戦略学習の主効果がなく交互作用もなかった ($F(1, 76) = 1.30, p = .257, \eta_p^2 = .017$) (図 2c)。エージェントポイントの結果も踏まえるとこの結果は、エージェントの選好を正しく推定することで、参加者は自身のポイントを下げることなく、エージェントのポイントを上げることを示している。

固定バイバイアスについて、全参加者のうち、選好学習なし条件では 39 人中 28 人、選好学習あり条件では 39 人中 30 人が固定バイバイアスを持っていた。繰り返しありの二要因分散分析を行った結果、交渉の回数と選好学習の交互作用が有意であった ($F(1, 56) = 8.370, p = .005, \eta_p^2 = .130$)。Bonferroniの方法による多重比較を行った結果、1回目はローカルクエリ戦略学習有条件 ($M = .444, 95\% CI [.362, .527]$) の方が、無条件 ($M = .821, 95\% CI [.736, .907]$) よりも有意に低く ($M_{diff} = .377, p < .001$)、2回目はローカルクエリ戦略学習有条件 ($M = .489, 95\% CI [.307, .671]$) と、無条件 ($M = .476, 95\% CI [.288, .664]$) で差がなかった ($M_{diff} = .013, p = .923$) (図 2d を参照)。このことから、ローカルクエリ戦略を学習することで固定バイバイアスを低減でき、2回交渉をすることによっても固定バイバイアスは低減できるとわかる。

選好推定について、繰り返しありの二要因分散分析を行った結果、繰り返しありの二要因分散分析を行った結果、交渉の回数と選好学習の交互作用が有意であった ($F(1, 56) = 5.656, p = .020, \eta_p^2 = .069$)。Bonferroniの方法による多重比較を行った結果、1回目はローカルクエリ戦略学習有条件 ($M = .752, 95\% CI [.634, .870]$) の方が、無条件 ($M = .239, 95\% CI [.122, .357]$) よりも有意に低く ($M_{diff} = .513, p < .001$)、2回目もローカルクエリ戦略学習有条件 ($M = .308, 95\% CI [.191, .424]$) の方が、無条件 ($M = .068, 95\% CI [-.048, .185]$) よりも優位に高かった ($M_{diff} = .240, p = .058$) (図 2e)。このことは、ローカルクエリ戦略学習により、参加者がエージェントの選好をより良く推定できるようになったことを示している。

4 議論

本研究では、参加者がローカルクエリ戦略学習をすることで、交渉において感情表現から相手の心的状態を推測する能力が向上し、Win-Win な解決策を導くことができるかを検討した。その結果、ローカルクエリ戦略学習により Win-Win 解の改善 (H1) と固定バイバイアスの低減 (H2) に貢献するという仮説を支持してい

る。これは、交渉の早期段階での相手の選好学習の重要性を指摘する一連の研究 [Thompson 90a?, Dreu 00b] と一致するものであり、選好学習を交渉プロセスから完全に分離することの効果を示すものである。

また、実験結果は、人がローカルクエリ戦略を学習すると、エージェントの選好をより正確に推定できるようになることを示している。この結果は、ローカルクエリ戦略学習が、相手理解を促進し、より受け入れやすい提案を提示することにつながる可能性を示唆している。これは、交渉者がエージェントの感情を読み取ることを動機づけ、相手の真の選好や優先順位に関する情報をより効果的に処理し、固定パイバイアスを克服できるという de Dreu らの研究結果と一致している [Dreu 00a]。さらにこれらの結果は、ローカルクエリ戦略学習の効果が、複数回の交渉を通じて持続する可能性を示唆している。ローカルクエリ戦略学習を経験した参加者は、1 回目だけでなく 2 回目の交渉においても、高い合計ポイントを達成し、エージェントにより多くの譲歩を行った。この結果は、ローカルクエリ戦略学習が、長期的な交渉成果の改善につながる可能性を示している。

本研究の結果は、交渉者が相手の選好を効果的に学習することの重要性を強調している。ローカルクエリ戦略学習を通じて、参加者は相手の選好に関する情報を体系的に処理する機会を得た。この経験が、交渉で提案することへの動機づけを高め、固定パイバイアスの克服につながったと考えられる。これは、情報交換が固定パイバイアスを減らし、より創造的な合意につながることを示した Thompson らや Pinkley らの研究結果と一致している [Thompson 90b, Pinkley 95]。また、本研究の結果は、相手モデリングの観点からも解釈できる。ローカルクエリ戦略学習は、参加者が観察された表情から選好を推論するプロセスを構造化し、明示的に行う機会を提供した。この経験が、参加者の相手モデリングを促進し、より包括的な状況認識につながったと考えられる。これは、Billings らが示した、交渉者が変化する状況を随時分析し、相手との交流から新しい情報を得ることで相手モデリングするプロセスと関連している [Billings 98]。

本研究の限界として、複数論点最後通牒ゲームという特定の交渉設定に焦点を当てたことが挙げられる。今後の研究では、より多様な交渉の文脈で、ローカルクエリ戦略学習の有効性を検討する必要がある。

謝辞

本研究は JSPS 科研費 JP24H00718 および JST, CREST (JPMJCR21D4), 未来社会創造事業 (JPMJMI22J3) の支援を受けたものである。

参考文献

- [Barrett 11] Barrett, L. F., Mesquita, B., and Gendron, M.: Context in Emotion Perception, *Current Directions in Psychological Science*, Vol. 20, No. 5, pp. 286–290 (2011)
- [Bazerman 83] Bazerman, M. and Neale, M.: Heuristics in negotiation: Limitations to disputeresolution effectiveness, in Bazerman, M. H. and Lewicki, R. J. eds., *Negotiations in organizations*, pp. 51–67, Sage Publications (1983)
- [Bazerman 09] Bazerman, M. and Moore, D.: *Judgment in Managerial Decision Making*, Wiley (2009)
- [Billings 98] Billings, D., Papp, D., Schaeffer, J., and Szafron, D.: Opponent Modeling in Poker, in *AAAI/IAAI* (1998)
- [Braziunas 08] Braziunas, D. and Boutilier, C.: Elicitation of Factored Utilities, *AI Magazine*, Vol. 29, No. 4, pp. 79–92 (2008)
- [Caputo 13] Caputo, A.: A literature review of cognitive biases in negotiation processes, *International Journal of Conflict Management*, Vol. 24, No. 4, pp. 374–398 (2013)
- [Davidai 23] Davidai, S. and Tepper, S. J.: The psychology of zero-sum beliefs, *Nature Reviews Psychology*, Vol. 2, No. 8, pp. 472–482 (2023)
- [Dehghani 14] Dehghani, M., Carnevale, P., and Gratch, J.: Interpersonal effects of expressed anger and sorrow in morally charged negotiation, *Judgment and Decision Making*, Vol. 9, No. 2, pp. 104–113 (2014)
- [DeVault 15] DeVault, D., Mell, J., and Gratch, J.: Toward Natural Turn-Taking in a Virtual Human Negotiation Agent, in *AAAI Spring Symposium 2015* (2015)
- [Dreu 00a] Dreu, de C. K. W., Koole, S. L., and Steinel, W.: Unfixing the fixed pie: A motivated information-processing approach to integrative negotiation., *Journal of Personality and Social Psychology*, Vol. 79, No. 6, pp. 975–987 (2000)
- [Dreu 00b] Dreu, C. K. W. D., Weingart, L. R., and Kwon, S.: Influence of social motives on integrative negotiation: A meta-analytic review and test of two theories., *Journal of Personality and Social Psychology*, Vol. 78, No. 5, pp. 889–905 (2000)

- [Fisher 81] Fisher, R., Ury, W., and Patton, B.: Getting to Yes: Negotiating Agreement Without Giving in (1981)
- [Fulmer 09] Fulmer, I. S., Barry, B., and Long, D. A.: Lying and Smiling: Informational and Emotional Deception in Negotiation, *Journal of Business Ethics*, Vol. 88, pp. 691–709 (2009)
- [Gaspar 12] Gaspar, J. and Schweitzer, M.: The Emotion Deception Model: A Review of Deception in Negotiation and the Role of Emotion in Deception, *Negotiation and Conflict Management Research*, Vol. 6, (2012)
- [Gelder 06] Gelder, B., Meeren, H., Righart, R., Stock, Van den J., Riet, W., and Tamietto, M.: Beyond the face: Exploring rapid influences of context on face processing, *Progress in brain research*, Vol. 155, pp. 37–48 (2006)
- [Gratch 15] Gratch, J., DeVault, D., Lucas, G. M., and Marsella, S.: Negotiation as a Challenge Problem for Virtual Humans, in *Intelligent Virtual Agents*, pp. 201–215, Springer International Publishing (2015)
- [Gratch 16] Gratch, J., DeVault, D., and Lucas, G.: The Benefits of Virtual Humans for Teaching Negotiation, in *Intelligent Virtual Agents*, pp. 283–294, Springer International Publishing (2016)
- [Herik 05] Herik, H., Donkers, H., and Spronck, P.: Opponent modelling and commercial games, *IEEE 2005 Symposium on Computational Intelligence and Games, CIG'05* (2005)
- [Hoegen 22] Hoegen, J., DeVault, D., and Gratch, J.: Exploring the Function of Expressions in Negotiation: the DyNego-WOZ Corpus, *IEEE Transactions on Affective Computing*, pp. 1–12 (2022)
- [Johnson 20] Johnson, E. and Gratch, J.: The Impact of Implicit Information Exchange in Human-agent Negotiations, in *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, ACM (2020)
- [Johnson 22] Johnson, S. G. B., Zhang, J., and Keil, F. C.: Win-win denial: The psychological underpinnings of zero-sum thinking., *Journal of Experimental Psychology: General*, Vol. 151, No. 2, pp. 455–474 (2022)
- [Kleef 18] Kleef, van G. A. and Côté, S.: Emotional Dynamics in Conflict and Negotiation: Individual, Dyadic, and Group Processes, *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 5, No. 1, pp. 437–464 (2018)
- [Lelieveld 22] Lelieveld, G.-J. and Dijk, van E.: *Emotions in Bargaining*, pp. 433–454, Springer International Publishing (2022)
- [Lewicki 10] Lewicki, R., Saunders, D., and Barry, B.: *Negotiation: Readings, Exercises, and Cases*, Higher Education, McGraw-Hill Irwin (2010)
- [Martinovski 10] Martinovski, B.: *Emotion in Negotiation*, pp. 65–86 (2010)
- [Meegan 10] Meegan, D. V.: Zero-Sum Bias: Perceived Competition Despite Unlimited Resources, *Frontiers in Psychology*, Vol. 1, (2010)
- [Mell 17] Mell, J. and Gratch, J.: Grumpy & Pinocchio: Answering Human-Agent Negotiation Questions through Realistic Agent Design, in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS '17*, p. 401–409, Richland, SC (2017), International Foundation for Autonomous Agents and Multiagent Systems
- [Morris 00] Morris, M. W. and Keltner, D.: How emotions work: The social functions of emotional expression in negotiations, *Research in Organizational Behavior*, Vol. 22, No. 0, pp. 1–50 (2000)
- [Ong 15] Ong, D. C., Zaki, J., and Goodman, N. D.: Affective cognition: Exploring lay theories of emotion, *Cognition*, Vol. 143, pp. 141–162 (2015)
- [Pietroni 08] Pietroni, D., Kleef, G. A. V., Dreu, C. K. D., and Pagliaro, S.: Emotions as strategic information: Effects of other’s emotional expressions on fixed-pie perception, demands, and integrative behavior in negotiation, *Journal of Experimental Social Psychology*, Vol. 44, No. 6, pp. 1444–1454 (2008)
- [Pietroni 21] Pietroni, D., Verdi, S. H., Giuliani, F., Rosa, A., Missier, F. D., and Palumbo, R.: The interpersonal effects of emotion on rejection of severely unfair ultimatum proposal, *International Journal of Conflict Management*, Vol. 33, No. 1, pp. 1–21 (2021)

- [Pillutla 96] Pillutla, M. M. and Murnighan, J.: Unfairness, Anger, and Spite: Emotional Rejections of Ultimatum Offers, *Organizational Behavior and Human Decision Processes*, Vol. 68, No. 3, pp. 208–224 (1996)
- [Pinkley 95] Pinkley, R. L., Griffith, T. L., and Northcraft, G. B.: “Fixed Pie” a la Mode: Information Availability, Information Processing, and the Negotiation of Suboptimal Agreements, *Organizational Behavior and Human Decision Processes*, Vol. 62, No. 1, pp. 101–112 (1995)
- [Sato 23] Sato, M., Terada, K., and Gratch, J.: Teaching Reverse Appraisal to Improve Negotiation Skills, pp. 872–884 (2023)
- [Schmid 10] Schmid, P. and Mast, M.: Mood effects on emotion recognition, *Motivation and Emotion*, Vol. 34, pp. 288–292 (2010)
- [Schweitzer 05] Schweitzer, M., Dechurch, L., and Gibson, D.: Conflict Frames and the Use of Deception: Are Competitive Negotiators Less Ethical?, *Journal of Applied Social Psychology*, Vol. 35, pp. 2123 – 2149 (2005)
- [Sharma 20] Sharma, S., Elfenbein, H. A., Sinha, R., and Bottom, W. P.: The Effects of Emotional Expressions in Negotiation: A Meta-Analysis and Future Directions for Research, *Human Performance*, Vol. 33, No. 4, pp. 331–353 (2020)
- [Thompson 90a] Thompson, L.: *Social perception in negotiation*, PhD thesis, Northwestern University (1990)
- [Thompson 90b] Thompson, L. and Hastie, R.: Social perception in negotiation, *Organizational Behavior and Human Decision Processes*, Vol. 47, No. 1, pp. 98–123 (1990)
- [Wierzbicki 92] Wierzbicki, A. and Makowski, M.: Multi-Objective Optimization in Negotiation Support. (1992)
- [Wu 17] Wu, Y., Baker, C. L., Tenenbaum, J. B., and Schulz, L. E.: Rational Inference of Beliefs and Desires From Emotional Expressions, *Cognitive Science* (2017)