

令和 6 年 6 月 22 日現在

機関番号：13701

研究種目：基盤研究(B)（一般）

研究期間：2021～2023

課題番号：21H03782

研究課題名（和文）利害対立を解消する他者心的状態推定のための感情情報処理

研究課題名（英文）Affective computing for mental states inference to solve social dilemma

研究代表者

寺田 和憲 (Terada, Kazunori)

岐阜大学・工学部・教授

研究者番号：30345798

交付決定額（研究期間全体）：（直接経費） 13,200,000円

研究成果の概要（和文）：本研究では、人が、心的状態と感情表現に関する生成モデル（Appraisal Model）を尤度関数としてベイズ推論によって、他者の心的状態を推定し、利害対立を解消していることを検証した。検証では、囚人のジレンマ、複数論点最後通牒ゲーム、集団的リスクジレンマなどのゲーム理論的課題を用い、コンピュータシミュレーションおよび人を対象とした実験を実施した。その結果、社会的価値志向性（SVO）理論、個人の嗜好と限界を表現する多次元効用関数、複数エージェントの感情に基づく生成モデルの妥当性が明らかになった。また、これらのモデルを用いた心的状態推論によって、人が実際に利害対立を解消していることが確認された。

研究成果の学術的意義や社会的意義

この研究成果の学術的意義は、感情表出と心的状態の関係を定量化し、ベイズ推論による意図推定プロセスを解明したことで、人の協力的行動および利害対立解消を下支えする認知、感情メカニズムの計算論的理解とそれに基づく複雑な社会的相互作用のモデル化が可能となったことにある。社会的意義は、トレーニングプログラムによる対人スキルの向上、集団間の対立や社会的ジレンマの解決に対する人認知を考慮した新たなアプローチの設計、より自然で協力的な人-AI相互作用の設計に応用できる基礎的な知見を提供し、社会に実在する利害対立を実際に減らすことができる可能性を提示したことにある。

研究成果の概要（英文）：This study verified that humans estimate the mental states of others and resolve conflicts of interest through Bayesian inference, using generative models (appraisal models) of mental states and emotional expressions as likelihood functions. This verification employed game-theoretical tasks such as the Prisoner's Dilemma, the Multi-Issue Ultimatum Game, and the Collective Risk Dilemma, utilizing both computer simulations and human experiments. The results demonstrated the validity of generative models based on Social Value Orientation (SVO) theory, multidimensional utility functions representing individual preferences and limits, and models based on emotions of multiple agents. Furthermore, we confirmed that humans actually resolve conflicts of interest through mental state inference using these generative models.

研究分野：人工知能

キーワード：認知科学 アフェクティブコンピューティング 感情 心の理論 利害対立解消 協力的行動

## 様式 C - 19, F - 19 - 1 (共通)

### 1. 研究開始当初の背景

人間社会において、利害の対立は避けられない現象である。日常生活では、外食時の行き先決定や中古車の売買交渉など、法的・数理的に解決が困難な利害の対立がしばしば発生する。これらの対立を解消する上で、感情表出が効果的であることが多くの研究によって明らかにされてきた。例えば、怒りの表出は相手から譲歩を引き出し、喜びの表出は協力を促し、後悔の表出は関係修復を可能にする。このプロセスにおいては、表出された感情から相手の意図、信念、選好といった心的状態を推定する必要がある。これは感情の逆評価と呼ばれる過程であり、人間の社会的相互作用において重要な役割を果たしている。しかしながら、利害対立の状況下で人々がどのように相手の心的状態を推定し、また、その相互心的状態推定を通じた意思決定によってどのように対立が解消されるかについては、まだ十分に解明されていない。感情の社会的機能に関する研究は、進化心理学的な観点からも重要性が指摘されている。感情表出は、生存に関わる問題への適応反応として進化してきたと考えられており、社会的な協調や競争において重要な役割を果たしてきた。この観点から、これまで、研究代表者の寺田は、ロボットやバーチャルエージェントを用いて感情の社会的機能に関する多くの研究を行ってきた。近年、感情情報処理研究に大きな転換をもたらしたのが、ベイズ推論によって心の理論を計算モデル化する、ベイジアン心の理論である。ベイズ推論は、過去の経験や知識（事前確率）と新たな観察（尤度）を組み合わせ、最も確からしい推論（事後確率）を導き出す方法である。この手法は、人が他者の心的状態を推論する認知過程をモデル化する上で非常に有効であることが示されている。本研究では、ベイズ推論を用いた精緻なモデル化が人の感情を用いた社会的意思決定をより良く説明できるのではないかとこの着想に至った。

### 2. 研究の目的

本研究の目的は以下であった。

利害対立状況における、心的状態から表情や身体運動の感情表現を出力する精緻な生成モデル（尤度関数、Appraisal Model）を獲得すること。

人が生成モデルを用いたベイズ推論によって相互意図（選好）推定を行い利害対立を解消していることを検証すること。

### 3. 研究の方法

上記の目的を達成するために、多数の方法を用いたが、ここでは、囚人のジレンマ、複数論点最後通牒、集団的リスクジレンマを用いた以下の、国際的にハイインパクトな3つの研究に絞って報告する。

#### 3 - 1. 囚人のジレンマによる検証

Ito, R., de Melo, C. M., Gratch, J. & Terada, K. (2024, September 17). Emotional Expression Help Regulate the Appropriate Level of Cooperation with Agents. *The 12th International Conference on Affective Computing and Intelligent Interaction (ACII '24)*.

この研究では、囚人のジレンマゲームを用いて、人が社会的価値志向（Social Value Orientation: SVO）理論に基づく生成モデルを用いたベイズ推論によって、相手のSVOを推論し、協力が搾取の意志決定を行っているかをコンピュータシミュレーションと人を対象とした実験によって検証した。

SVOは、相互依存性理論（Interdependence Theory）に基づき、個人が社会的な状況や対人関係において、自己と他者の利益をどのように重視するかを表す概念であり、個人が社会的ジレンマ状況でどのように行動するかを予測するのに役立つ。主なものとして、自己と他者の利益を同程度に重視する協力（Cooperative）、自己の利益を最大化することを重視する個人主義（Individualistic）、自己と他者の利益の差を最大化することを重視する競争（Competitive）、他者の利益を最大化することを重視する利他（Altruistic）が知られているが、本研究では、殉教者-マゾヒスト（Martyrdom-Masochism）と個人主義-競争主義（Individualism-Competition）を用いた。

SVOは $0^\circ$ から $360^\circ$ までの角度 $\theta_{SVO}$ で与えられ、SVOが決定すると、 $w_{self} = \cos(\theta_{SVO})$ 、 $w_{other} = \sin(\theta_{SVO})$ のように、角度に対する余弦と正弦によって自他に与える重み $w_{self}$ と $w_{other}$ が決定する。自己と他者が得る報酬 $r_{self}$ 、 $r_{other}$ が与えられると、効用は $u = [w_{self}, w_{other}] \cdot [r_{self}, r_{other}]$ として決定する。報酬は囚人のジレンマにおける利得であり、自他が協力（C）を選択するか裏切り（D）を選択するかの組み合わせによって異なる。 $u$ は意思決定に影響を与えるために、意思決定に関する尤度関数 $p(\text{Decision} | \text{SVO}, \text{Matrix})$ を与える。すなわち、協力的な人は相互協力（CC）に対する効用が高いためCを選択しやすく、競争的な人は相手を裏切り（CD）ことに対する効用が高いためDを選択しやすい。さらに $u$ は状況に対する評価（appraisal）であるために、表情出力 $EE$ を決定する尤度関数 $p(EE | \text{SVO}, \text{Outcome})$ を与える。 $u$ がポジティブな場合には喜び、 $u$ がネガティブな場合には怒り、もしくは悲しみを表出する。表情表出パターンはどのようなSVOを持つかによって異なる。殉教者-マゾヒストでは相互協力（CC）に対する効用が低く、ネガティブな表情である悲しみを表出し、相手に裏切り（DC）に対する効用が高く、ポジティブな表情である喜びを表出する。一方、個人主義-競争主義では相互協力（CC）に対して効用が高く、喜び、相手の裏切りに対して効用が低く怒りを出さる。2つの尤度は意思決定と表情を観察した場合に以下の式のベイズ推論による相手のSVO推定を可能にする。

$$p(SVO | Matrix, EE, Decision) \propto p(Decision | SVO, Matrix)p(EE | SVO, Outcome)p(SVO)$$

本研究では、常にC(協力)を選択するAIエージェントが殉教者-マゾヒスト、個人主義-競争主義、ニュートラル(感情表出なし)のいずれかの表情パターンを表出する条件で、人が20ラウンドの囚人のジレンマゲームを行った。これにより、人がAIエージェントの感情表出からそのSVOを推定し、自身の意思決定を変えるかどうかを検証した。特に、殉教者-マゾヒストの表情パターンを示すAIエージェントに対してより搾取(非協力的)になり、個人主義-競争主義の表情パターンを示すAIエージェントに対してより協力的になるかどうか注目した。また、エージェントの外見(男性、女性、ロボット)による影響も検討した。

Yahoo! Japan Crowdsourcingから379名が実験に参加した。実験の結果、AIエージェントの感情表現は、人の協力行動に影響を与えることが明らかになった。特に、殉教者-マゾヒスト的な感情表現をするAIに対して、人はより非協力的(搾取的)な行動を取る傾向が見られた。一方、競争主義-個人主義的な感情表現をするAIに対しては、より協力的な行動を取る傾向が見られた。感情表現の影響は、ゲームの後半(12ラウンド目以降)でより顕著になった。これは、人がAIの感情表現パターンを学習し、それに基づいて行動を調整していく過程を示唆している。人は、AIの感情表現からそのSVOを推論していることが示された。これは、人が感情表現を手がかりにしてAIの意図や選好を解釈していることを示している。AIエージェントの外見は、人の行動や判断に有意な影響を与えなかった。これは、感情表現のパターンが、外見よりも重要な要因であることを示唆する。

この研究の目的、に対する貢献は以下の通りである。

目的 に対する貢献

- ✓ SVOに基づく生成モデルの提案：社会的価値志向性(SVO)理論を基礎とした感情表現の生成モデルを提案した。このモデルは、心的状態(SVO)と状況(ゲームの結果)から感情表現を生成する過程を数学的に定式化している。
- ✓ 異なるSVOに対応する感情表現パターンの設計：殉教者-マゾヒストと個人主義-競争主義という異なるSVOに対応する具体的な感情表現パターンを設計した。これは、心的状態の違いが感情表現にどのように反映されるかを示す実践的な例となっている。
- ✓ 感情表現の社会的機能の考慮：感情表現が単なる内部状態の表出ではなく、相手の行動を調整する社会的機能を持つことを考慮している。これは、生成モデルがより現実的で機能的なものになることを示唆する。

目的 に対する貢献

- ✓ 意図推定と行動調整の関連性：人の協力行動がAIの感情表現パターンによって変化したことは、推定された意図(SVO)に基づいて人が自身の行動を調整していることを示唆しています。これは、意図推定が利害対立の解消につながる可能性を実証的に示している。
- ✓ 推論プロセスの具体化：感情表現から意図を推論する逆評価(reverse appraisal)の概念が、ベイズ推論のプロセスをより具体的に説明できることを示した。

### 3 - 2 . 複数論点最後通牒ゲームによる検証

Sato, M., Terada, K. & Gratch, J. (2023). Teaching Reverse Appraisal to Improve Negotiation Skills. *IEEE Transactions on Affective Computing*, 1-14. <https://doi.org/10.1109/taffc.2023.3285931>

この研究では、複数論点最後通牒ゲームを用いて、人が生成モデルを用いたベイズ推論によって相手の選好や意図を推定し、それを基に利害対立を解消するプロセスを検証した。特に、感情表現を通じて相手の心的状態や選好をどのように解釈し、交渉においてどのように活用できるかを明らかにすることを目的とした。

一般的な複数論点交渉ゲームでは、参加者が論点について繰り返し提案を行い、相互に影響を与え合いながら合意形成を目指す。一方、複数論点最後通牒ゲームは、最後通牒ゲームと同じく、提案者が一度だけ提案を行い、応答者が受け入れるか拒否する非対称なワンショット交渉ゲームである。最後通牒ゲームとの違いは、論点が複数であるために、提案者は複数の選好を考慮に入れながら戦略的に相手に拒否されない提案を構築する必要がある点である。この複雑さは、応答者が受け入れるか拒否する限界にも影響を与え、単一論点の最後通牒ゲームよりも戦略的な考慮が増える。そのため、交渉の結果はより予測が難しく、合意に至るプロセスも一層複雑になる。また、交渉ゲームはゲームマトリックスが未知であるという点で、囚人のジレンマと異なる。囚人のジレンマにおいては、ゲームマトリックスは通常、プレイヤーにとって既知であるが、交渉ゲームにおいては論点(選択)の価値が個々に異なり明かされていないために未知である。交渉者は相手の意志決定を予測するために、限られた情報から応答者の選好および限界を推測する必要がある。

この研究では、人が相手の効用関数を反映した表情生成モデルから逆に相手の選好と限界を推論し、交渉に役立てられるかを検証した。我々は、AIエージェントの効用関数に基づいて表情を出力する生成モデルを設計した。エージェントの効用は $u = w \cdot x$ で与えた。ここで $w$ は各論点に対する重み(エージェントの選好を示す)、 $w$ は各論点のレベル(提案内容)である。生成モデルは、エージェントの効用を怒り、中立、7段階の喜びの表情に変換した。参加者は、このAIエージェントの表情表出を観察し、それを手がかりに逆推論(reverse appraisal)を行う必要があった。つまり、表出された感情から、エージェントの

内部状態である選好と限界を推測しなければならなかった。参加者は、この推論結果を基に交渉戦略を立て、提案を行った。理想的には、相手の選好をよく理解することで、双方にとって有利な提案（統合的解決）ができるはずである。

Yahoo! Japan Crowdsourcing から 148 名が実験に参加した。実験では、一部の参加者にこの逆推論のトレーニングを行わせ、トレーニングの効果を検証した。具体的な検証のポイントは、参加者が正確に相手の選好と限界を推論できるか、その推論が実際の交渉結果（joint points や個別の得点）にどう影響するかであった。実験結果の要約は次の通りである。逆推論トレーニングを受けた参加者は、AI エージェントの選好をより正確に推定できるようになり、交渉時により積極的に相手の選好を探索する行動を示した。また、エージェントの提案受諾率が高くなり、自身の得点を減少させてエージェントの得点を増加させる傾向が見られた。しかし、joint points の有意な増加は見られず、固定バイバイアス（交渉者が、交渉対象となる資源や利益の総量が固定されていると誤って認識する傾向。相手の利益と自分の利益が完全に相反すると考えること。）も改善されなかった。選好推定の正確さと joint points の間には正の相関が見られたが、これは 2 回目の交渉でのみ観察された。これらの結果は、選好推定能力の向上が必ずしも統合的な解決につながらないという、交渉スキル向上と実際の交渉結果改善の間の複雑な関係を示唆している。

この研究の目的、に対する貢献は以下の通りである。

目的 に対する貢献

- ✓ モデルの具体化：効用関数に基づいた具体的な Appraisal Model を提案した。このモデルは、エージェントの内部状態（選好と限界点）を感情表現（怒り、中立、7 段階の喜び）に変換する方法を明確に定義している。
- ✓ 交渉状況への適用：複数の問題を扱う交渉状況に適用可能なモデルを提示した。これは、より複雑で現実的な状況におけるモデルの有効性を検討する上で重要である。

目的 に対する貢献

- ✓ 計算論的アプローチの有効性：感情表出から相手の選好と限界の心的状態を推論するプロセスを計算論的に表現し検証することで、従来の心理学的アプローチでは捉えきれなかった側面を明らかにすることができた。
- ✓ 交渉スキルトレーニングへの応用：ベイズ推論に基づく意図推定のメカニズムを明らかにすることで、より効果的な交渉スキルトレーニング方法の開発につながる可能性がある。
- ✓ バイバイアスの影響の定量化：固定バイバイアスなどの認知バイバイアスが、ベイズ推論による意図推定プロセスにどのように影響するかを定量的に評価することができる。

### 3 - 3 . 集団的リスクジレンマによる検証

de Melo, C. M., Santos, F. C. & Terada, K. (2023). Emotion Expression and Cooperation under Collective Risks. *iScience*, 26(11), 108063. <https://doi.org/10.1016/j.isci.2023.108063>

この研究では、集団的リスクジレンマ(Collective Risk Dilemma: CRD)という、気候変動対策など、長期的で不確実な利益のために短期的な犠牲が必要な状況をモデル化した、個人と集団の利害対立状況において、感情表出が利害の解消（協力行動）に影響を与えることを実証した。

CRD は、個人の利益と集団の利益が対立する社会的ジレンマの一種である。参加者は共通の目標（閾値）を達成するために個人的な資源を投資するかどうかを決定する。具体的な構造は以下である。まず、参加者に初期の資金（トークン）を与えられる。参加者は複数のラウンドで共通プールに投資するかどうかを決める。全参加者の総投資額が設定された閾値に達しない場合、全員が災害のリスク（全資金喪失）に直面する。閾値に達した場合は災害のリスクが回避され、各参加者は、投資せずに手元に残した資金を保持することができる。閾値に到達することで集団全体が利益を得る（災害を回避する）一方で、個人レベルでは投資を控えた参加者がより多くの個人的利益を得る可能性がある。すなわち、個人にとっては投資しないことが短期的に有利だが、それによって集団全体が失敗するリスクが高まる。

実験 1 では、Amazon Mechanical Turk から 93 名が実験に参加した。3 人グループで 10 ラウンドの CRD を行う設計とした。参加者は初期資金として 40 トークンを与えられた。各ラウンドで、参加者は 0 から 4 トークンの範囲で投資を決定できた。グループ全体で 60 トークンの投資が必要とされた。閾値未達成の場合、50%の確率で全員が残りの資金を失うリスクがあった。参加者は実験参加に対して \$2.50 を受け取り、さらに保持したトークンは \$30 の抽選券に変換された。低要求条件と高要求条件の 2 条件を設定した。低要求条件では 4 トークン、高要求条件では 8 トークンを閾値とした。仮想の顔アバターを用いて感情を表現した。投資が要求水準を下回ると怒りを、上回ると喜びを表出した。実験 2 には Prolific から 283 名が参加した。基本的な設計は実験 1 と同様だが、実験 2 では、3 条件（低要求、無感情、高要求）を設定した。また、各ラウンド後に参加者の感情状態を自己報告させた。また、実験後にグループに対する印象を評価させた。

実験 1 の結果、閾値達成については、高要求条件のグループは、低要求条件のグループよりも投資閾値を達成する確率が高かった。つまり、怒りをより多く表出するグループの方が、集団的目標を達成しやすかったといえる。投資額については、高要求条件の参加者は、低要求条件の参加者よりも多くのトークンを投資する傾向があった。これは、怒りの表出が参加者の協力行動を促進したことを示唆している。実験 2 の結果は次の通りである。計的に明確な差

は見られなかったものの、高要求条件のグループが最も高い確率で閾値を達成する傾向が見られた。これは実験 1 の結果と一致する傾向である。実験 1 と同様に、高要求条件の参加者は低要求条件の参加者よりも多くのトークンを投資する傾向があった。これにより、怒りの表出が協力的行動を促進するという結果が再現された。低要求条件の参加者は他の条件よりも多く喜びを経験し、高要求条件の参加者は低要求条件よりも多く怒りを経験した。これは、グループの感情表出が参加者自身の感情経験に影響を与えることを示している。高要求条件のグループは最も要求が高いと認識され、低要求条件のグループは最もポジティブだと評価された。しかし、グループの公平さに関しては条件間で差が見られなかった。これらの結果から、怒りの表出など高要求を示す感情表現が、集団的リスクジレンマにおける協力的行動を促進する可能性が示唆された。一方で、そのような感情表出はグループの印象にネガティブな影響を与える可能性もあることが分かった。

この研究の目的、に対する貢献は以下の通りである。

目的 に対する貢献

- ✓ 感情表出の文脈依存性の実証：この研究は、集団的リスクジレンマという特定の利害対立状況において、怒りと喜びの表出が異なる効果を持つことを示した。これは、感情表出が状況に応じて異なる機能を果たすことを実証しており、精緻な生成モデルにはこの文脈依存性を組み込む必要があることを示唆している。
- ✓ 要求水準と感情表出の関連性：実験では、要求水準（低要求・高要求）と感情表出（喜び・怒り）を明確に関連付けている。これは、心的状態（要求水準）から感情表現への変換過程をモデル化する上で重要な知見となる。
- ✓ 複数エージェントの感情表出モデルへの示唆：グループ全体の感情表出が個人の行動に影響を与えるという結果は、複数エージェントの感情表出をモデル化する際の考慮点を提供している。

目的 に対する貢献

- ✓ 不確実性下での意思決定：集団的リスクジレンマという社会的状況で、参加者が他者の感情表出を手がかりに意思決定を行っていることが示された。これは、ベイズ推論のような確率的推論プロセスが働いている可能性を示唆している。
- ✓ 集団レベルでの推論：個々の参加者が全体の投資額を考慮しながら意思決定を行っていることは、集団レベルでの意図推定が行われている可能性を示唆している。これは、複数エージェントのベイズ推論モデルの適用可能性を示している。

なお、この研究では直接的にベイズ推論モデルを適用または検証しているわけではない。むしろ、人間の行動がベイズ推論モデルで説明可能である可能性を示唆するデータを提供している。

#### 4. 研究成果

本研究では、人が、心的状態と感情表現に関する生成モデル（Appraisal Model）を尤度関数としてベイズ推論によって、他者の心的状態を推定し、利害対立を解消していることを検証した。検証では、囚人のジレンマ、複数論点最後通牒ゲーム、集団的リスクジレンマなどのゲーム理論的課題を用い、コンピュータシミュレーションおよび人を対象とした実験を実施した。目的 に対する具体的な成果は以下の通りである。

目的 に対する成果

1. 社会的価値志向性（SVO）理論に基づく感情表現の生成モデルを提案し、心的状態と感情表現の関係を定量的に表現することに成功した。
2. 効用関数に基づいた具体的な Appraisal Model を提案し、エージェントの内部状態を感情表現に変換する方法を明確に定義した。
3. 感情表出の文脈依存性を実証し、状況に応じて感情表出が異なる機能を果たすことを示した。
4. 要求水準と感情表出の関連性を明らかにし、心的状態から感情表現への変換過程のモデル化に重要な知見を提供した。

目的 に対する成果

1. 人間が相手の感情表出から SVO や選好を推論し、それに基づいて行動を調整することを実証的に示した。これは、ベイズ推論による相互意図推定の可能性を支持する証拠となる。
2. 複数論点最後通牒ゲームにおいて、逆推論（reverse appraisal）のトレーニングが相手の選好推定の正確性を向上させることを示した。
3. 集団的リスクジレンマにおいて、参加者が他者の感情表出を手がかりに不確実性下での意思決定を行うことを示し、ベイズ推論のような確率的推論プロセスの存在を示唆した。
4. 感情表出が単なる反応ではなく、情報伝達的手段として機能することを示し、交渉における感情の役割を新たな視点から捉え直した。

この研究成果の学術的意義は、感情表出と心的状態の関係を定量化し、ベイズ推論による意図推定プロセスを解明したことで、人の協力的行動および利害対立解消を下支えする認知、感情メカニズムの計算論的理解とそれに基づく複雑な社会的相互作用のモデル化が可能となったことにある。社会的意義は、トレーニングプログラムによる対人スキルの向上、集団間の対立や社会的ジレンマの解決に対する人認知を考慮した新たなアプローチの設計、より自然で協力的な人-AI 相互作用の設計に応用できる基礎的な知見を提供し、社会に実在する利害対立を実際に減らすことができる可能性を提示したことにある。

## 5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 4件/うち国際共著 2件/うちオープンアクセス 3件）

1. 著者名 Takagi Hisashi、Terada Kazunori	4. 巻 11
2. 論文標題 The effect of anime character 's facial expressions and eye blinking on donation behavior	5. 発行年 2021年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 9146
掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s41598-021-87827-2	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 de Melo Celso M.、Santos Francisco C.、Terada Kazunori	4. 巻 26
2. 論文標題 Emotion expression and cooperation under collective risks	5. 発行年 2023年
3. 雑誌名 iScience	6. 最初と最後の頁 108063 ~ 108063
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.isci.2023.108063	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Sato Motoaki、Terada Kazunori、Gratch Jonathan	4. 巻 Early Access
2. 論文標題 Teaching Reverse Appraisal to Improve Negotiation Skills	5. 発行年 2024年
3. 雑誌名 IEEE Transactions on Affective Computing	6. 最初と最後の頁 1 ~ 14
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/taffc.2023.3285931	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 寺田 和憲	4. 巻 39
2. 論文標題 人は人の理解をやめるか？ AI を用いた詳細価値観モデリングと精密行動予測の可能性？	5. 発行年 2024年
3. 雑誌名 人工知能	6. 最初と最後の頁 230 ~ 239
掲載論文のDOI（デジタルオブジェクト識別子） 10.11517/jjsai.39.2_230	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 NAKAMURA Yuya、KATO Kunihito、TERADA Kazunori、TAKAHASHI Toshikazu、SHIOZAWA Yasuo、AIZAWA Hiroaki	4. 巻 89
2. 論文標題 Sensory Inspection of Wood Grain by Traversing the GAN Latent Space	5. 発行年 2023年
3. 雑誌名 Journal of the Japan Society for Precision Engineering	6. 最初と最後の頁 213 ~ 220
掲載論文のDOI (デジタルオブジェクト識別子) 10.2493/jjspe.89.213	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計45件 (うち招待講演 0件 / うち国際学会 17件)

1. 発表者名 松田 治真, 寺田 和憲
2. 発表標題 作り笑顔あふれる社会 ポライトネス社会規範と語用論的笑顔の共進化
3. 学会等名 HAIシンポジウム2024
4. 発表年 2024年

1. 発表者名 伊藤 諒哉, 寺田 和憲, Celso M. de Melo
2. 発表標題 自己犠牲エージェント一人を搾取的にさせるエージェントの表情表出
3. 学会等名 HAIシンポジウム2024
4. 発表年 2024年

1. 発表者名 佐藤 幹晃, 寺田 和憲, Jonathan Gratch
2. 発表標題 非言語交渉における認知課題の分割が統合的解決能力の向上に与える影響
3. 学会等名 HAIシンポジウム2024
4. 発表年 2024年

1. 発表者名 Kazunori Terada
2. 発表標題 Towards an AI Agent System Promoting the Computational Development of Social Emotional Skills
3. 学会等名 The 11th Congress of The Asian Society for Child and Adolescent Psychiatry and Allied Professions (ASCAPAP2023) (国際学会)
4. 発表年 2023年

1. 発表者名 Kazunori Terada
2. 発表標題 The ambiguity of the generative model of rational agents in relation to mind-reading deficits
3. 学会等名 The 11th Congress of The Asian Society for Child and Adolescent Psychiatry and Allied Professions (ASCAPAP2023) (国際学会)
4. 発表年 2023年

1. 発表者名 藤井 俊輔, 滝 之弥, 吉田 温登, 西川 菜月, 小森 政嗣, 植田 一博, 加藤 邦人, 原 武史, 熊崎 博一, 寺田 和憲
2. 発表標題 目玉焼きの質感を決定づける潜在空間上の効用関数を用いた自閉スペクトラム症者と定型発達者の感覚処理特性の比較
3. 学会等名 2024年度人工知能学会全国大会(第38回)
4. 発表年 2024年

1. 発表者名 中村 真士, 岩松 秀俊, 伏屋 貴文, 服部 雅典, 速水 悟, 寺田 和憲
2. 発表標題 大規模言語モデルを用いた労働災害の原因推論
3. 学会等名 2024年度人工知能学会全国大会(第38回)
4. 発表年 2024年



1. 発表者名 高木 寿, 李 楊, 寺田 和憲, 小森 政嗣
2. 発表標題 アルゴリズム嫌悪と裏切り嫌悪がAIに対する信頼に与える影響
3. 学会等名 2024年度人工知能学会全国大会(第38回)
4. 発表年 2024年

1. 発表者名 松田 治真, 寺田 和憲
2. 発表標題 笑顔の多義性ーポライトネス社会規範と語用論的笑顔の共進化ー
3. 学会等名 2024年度人工知能学会全国大会(第38回)
4. 発表年 2024年

1. 発表者名 佐藤 幹晃, 内田 貴久, 吉川 雄一郎, グラッチ ジョナサン, ドゥメル セルス, 寺田 和憲
2. 発表標題 人らしいバーチャルエージェントと不気味なアンドロイドの感情表現が交渉での協力に与える影響
3. 学会等名 2024年度人工知能学会全国大会(第38回)
4. 発表年 2024年

1. 発表者名 伊藤 諒哉, セルス ドゥメル, ジョナサン グラッチ, 寺田 和憲
2. 発表標題 人を搾取的にさせるエージェントの表情表出
3. 学会等名 2024年度人工知能学会全国大会(第38回)
4. 発表年 2024年

1. 発表者名 Manato Nakamura, Kazunori Terada, Satoru Hayamizu, Hattori Masanori, Takafumi Fuseya, Hidetoshi Iwamatsu
2. 発表標題 Few-Shot Multi-Label Annotation of Causes for Incident Texts Using Large Language Models
3. 学会等名 The 28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES '24) (国際学会)
4. 発表年 2024年

1. 発表者名 Hisashi Takagi, Yang Li, Masashi Komori, Kazunori Terada
2. 発表標題 Measuring Algorithm Aversion and Betrayal Aversion to Humans and AI using Trust Games
3. 学会等名 The 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN2024) (国際学会)
4. 発表年 2024年

1. 発表者名 Motoaki Sato, Takahisa Uchida, Yuichiro Yoshikawa, Celso M. de Melo, Jonathan Gratch, Kazunori Terada
2. 発表標題 People Negotiate Better with Emotional Human-Like Virtual Agents than Android Robots
3. 学会等名 The 12th International Conference on Affective Computing and Intelligent Interaction (ACII '24) (国際学会)
4. 発表年 2024年

1. 発表者名 Ryoya Ito, Celso M. de Melo, Jonathan Gratch, Kazunori Terada
2. 発表標題 Emotional Expression Help Regulate the Appropriatelevel of Cooperation with Agents
3. 学会等名 The 12th International Conference on Affective Computing and Intelligent Interaction (ACII '24) (国際学会)
4. 発表年 2024年

1. 発表者名 M. Sato, K. Terada, and J. Gratch
2. 発表標題 Preference learning from emotional expressions contributes integrative solutions between human-AI negotiation
3. 学会等名 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (国際学会)
4. 発表年 2023年

1. 発表者名 K. Terada, Y. Noma, and M. Hattori
2. 発表標題 Personalized Information Presentation based on Detailed Individual Values for Persuasive Agents
3. 学会等名 International Conference on Human-Agent Interaction, in HAI '23 (国際学会)
4. 発表年 2023年

1. 発表者名 佐藤 幹晃, 寺田 和憲, ジョナサン グラッチ
2. 発表標題 Win-Win関係構築のための感情認知計算
3. 学会等名 2023年度日本認知科学会第40回大会
4. 発表年 2023年

1. 発表者名 寺田 和憲, 長谷川 智大, セルス ドゥメル, フランシスコ サントス
2. 発表標題 善人と悪人を識別する認知計算
3. 学会等名 2023年度日本認知科学会第40回大会
4. 発表年 2023年

1. 発表者名 野田 和弥, 寺田 和憲, セルス ドゥメル
2. 発表標題 産業用ロボットの表情と説明が人の意思決定に与える影響
3. 学会等名 人工知能学会全国大会(第37回)
4. 発表年 2023年

1. 発表者名 佐藤 幹晃, 寺田 和憲, ジョナサン グラッチ
2. 発表標題 感情的プリプレイコミュニケーションはAIと人の交渉におけるWin-Win関係に寄与する
3. 学会等名 人工知能学会全国大会(第37回)
4. 発表年 2023年

1. 発表者名 高木 寿, Yang Li, 小森 政嗣, 寺田 和憲
2. 発表標題 読影性能が胸部 CT 読影検査における AI の信頼性に与える影響
3. 学会等名 人工知能学会全国大会(第37回)
4. 発表年 2023年

1. 発表者名 天野 燎, 長谷川 智大, 伊藤 諒哉, 寺田 和憲, セルス ドゥメル
2. 発表標題 人に搾取的意思決定を促す自己犠牲エージェント
3. 学会等名 人工知能学会全国大会(第37回)
4. 発表年 2023年

1. 発表者名 山元 佑京, 大本 義正, 熊崎 博一, 岩永 竜一郎, 今村 明, 清水 日智, 寺田 和憲
2. 発表標題 ASDスクリーニングのための図形描画における筆跡特徴の検討
3. 学会等名 人工知能学会全国大会(第37回)
4. 発表年 2023年

1. 発表者名 寺地 哲平, 寺田 和憲, 原 武史, 阿見 翼, 小森 政嗣
2. 発表標題 人間・AIの画像診断ベンチマークテストを目指した二肢判断課題にもとづく適応テストの構築
3. 学会等名 人工知能学会全国大会(第37回)
4. 発表年 2023年

1. 発表者名 長谷川 智大, 寺田 和憲, フランシスコ サントス, セルス ドゥメル
2. 発表標題 ベイジアン心の理論による社会的ジレンマ状況における読心の計算モデル
3. 学会等名 人工知能学会全国大会(第37回)
4. 発表年 2023年

1. 発表者名 Kazunori Terada
2. 発表標題 The ambiguity of the generative model of rational agents in relation to mind-reading deficits
3. 学会等名 The 11th Congress of The Asian Society for Child and Adolescent Psychiatry and Allied Professions (ASCAPAP2023) (国際学会)
4. 発表年 2023年

1. 発表者名 Kazunori Terada
2. 発表標題 Towards an AI Agent System Promoting the Computational Development of Social Emotional Skills
3. 学会等名 The 11th Congress of The Asian Society for Child and Adolescent Psychiatry and Allied Professions (ASCAPAP2023) (国際学会)
4. 発表年 2023年

1. 発表者名 Junya Yamada, Souichiro Yoshikawa, Hirokazu Kumazaki, Hideki Kozima, Kazunori Terada
2. 発表標題 People learn other's preferences on a latent feature space using emotion expressions as labels
3. 学会等名 The 44th Annual Meeting of the Cognitive Science Society (国際学会)
4. 発表年 2022年

1. 発表者名 Celso de Melo, Kazunori Terada, Francisco Santos
2. 発表標題 Human Social Norms and Reputations are Shaped by Emotion Expressions
3. 学会等名 Conference of the International Society for Research on Emotion (ISRE2022) (国際学会)
4. 発表年 2022年

1. 発表者名 Maria Angelika-Nikita, Celso de Melo, Kazunori Terada, Gale Lucas, Jonathan Gratch
2. 発表標題 Impact of Partner Expressions on Felt Emotion in the Iterated Prisoner's Dilemma: An Event-level Analysis
3. 学会等名 Conference of the International Society for Research on Emotion (ISRE2022) (国際学会)
4. 発表年 2022年

1. 発表者名 寺田和憲, Li Yang
2. 発表標題 AIに対する信頼と意図帰属
3. 学会等名 2022 年度人工知能学会全国大会 (第36 回)
4. 発表年 2022年

1. 発表者名 佐藤 幹晃, 寺田 和憲
2. 発表標題 心的状態の可視化が複数論点最後通牒ゲームにおける人のWin-Win計算能力獲得に与える影響
3. 学会等名 2022 年度人工知能学会全国大会 (第36 回)
4. 発表年 2022年

1. 発表者名 臼井公希, 寺田和憲, Celso M. de Melo
2. 発表標題 産業用ロボットの感情表出が人の協調的意思決定に与える影響
3. 学会等名 2022 年度人工知能学会全国大会 (第36 回)
4. 発表年 2022年

1. 発表者名 長谷川智大, 寺田和憲, Celso M. de Melo, Francisco C. Santos
2. 発表標題 他者モデルの推定による行動の予測と一貫性のある協力行動の実現
3. 学会等名 2022 年度人工知能学会全国大会 (第36 回)
4. 発表年 2022年

1. 発表者名 臼井 公希, 寺田 和憲, セルス ドゥメル
2. 発表標題 産業用ロボットの感情表出が人の協調的意思決定に与える影響
3. 学会等名 2022年度人工知能学会全国大会(第36回)論文集
4. 発表年 2022年

1. 発表者名 佐藤 幹晃, 寺田 和憲
2. 発表標題 心的状態の可視化が複数論点最後通牒ゲームにおける人のWin-Win計算能力獲得に与える影響
3. 学会等名 2022年度人工知能学会全国大会(第36回)論文集
4. 発表年 2022年

1. 発表者名 長谷川智大, 寺田和憲, Celso M. de Melo and Francisco C. Santos
2. 発表標題 他者モデルの推定による行動の予測と一貫性のある協力行動の実現
3. 学会等名 2022年度人工知能学会全国大会(第36回)論文集
4. 発表年 2022年

1. 発表者名 Motoaki Sato, Kazunori Terada and Jonathan Gratch
2. 発表標題 Visualization of social emotional appraisal process of an agent
3. 学会等名 2021 9th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACII) (国際学会)
4. 発表年 2021年



1. 発表者名 Jerome Dinet, Hanna Verdel, Hirokazu Kumazaki, Kazunori Terada, Sylvie Vitel, Eloise Zehnder
2. 発表標題 A Hybrid Robotic System for Children with Autism: A longitudinal Study of Social Interaction in Classroom
3. 学会等名 ERGO-IA' (国際学会)
4. 発表年 2021年

1. 発表者名 Maria Angelika-Nikita and Celso de Melo and Kazunori Terada and Gale Lucas, Jonathan Gratch
2. 発表標題 The Impact of Partner Expressions on Felt Emotion in the Iterated Prisoner's Dilemma: An Event-level Analysis
3. 学会等名 Advances in Cognitive Systems (ACS) (国際学会)
4. 発表年 2021年

1. 発表者名 佐藤 幹晃, 寺田 和憲
2. 発表標題 エージェントの感情と社会的評価プロセスの可視化
3. 学会等名 インタラクション2022
4. 発表年 2022年

1. 発表者名 野間 靖央, 佐藤 幹晃, 高木 寿, 寺田 和憲, ジョナサン グラッチ
2. 発表標題 人はエージェントの表情から選好を推定できるか
3. 学会等名 ヒューマンインタフェースシンポジウム
4. 発表年 2021年

1. 発表者名 Koki Usui, Kazunori Terada, Celso M. de Melo
2. 発表標題 The influence of emotional expressions of an industrial robot on human collaborative decision-making
3. 学会等名 The 10th International Conference on Affective Computing and Intelligent Interaction (ACII '22)
4. 発表年 2022年

1. 発表者名 Kazunori Terada, Yasuo Noma, Masanori Hattori
2. 発表標題 Persuasion by Shaping Beliefs about Multidimensional Features of a Thing
3. 学会等名 The 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS '24) (国際学会)
4. 発表年 2024年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	山田 誠二  (Yamada Seiji)  (50220380)	国立情報学研究所・コンテンツ科学研究系・教授    (62615)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------