

行動変容システムの誤動作による不信感に対する 寛容性醸成のための擬人化に関する研究

胡 伊端†, 辻 愛里‡, 藤波 香織‡

† 東京農工大学大学院 生物システム応用科学府 生物機能システム科学専攻

‡ 東京農工大学大学院 工学研究院 先端情報科学部門

1. はじめに

近年のテクノロジーの進歩が、行動変容技術の発展と普及を促進している。人々は自分の生活習慣や健康を管理するため、個々の行動パターンや目標に合わせたパーソナライズされたサポートを利用する。しかし計測や判定にエラーがあると、システムに不信感を抱き、やがて使用されなくなってしまうことから、行動改善の目的を達成できないことが懸念される。

システムに対するユーザの信頼性を高めることを目的とした様々な研究手法において、システムの擬人化 [1] は、一般的なコンピュータよりもユーザから信頼されることが確認された。しかし、表現力だけでは信頼性を得るのに十分ではないことを報告した研究 [2] によって、機械に十分な信頼度を得るためには、脆弱性などロボットの他の特性と組み合わせる必要があることが示されている。「弱いロボット」 [3] は、人間と不完全なロボットやシステムとの間に調和的な共生関係を作り出すことを目的として導入された。脆弱表現を模倣することで自己の状態をユーザに示し、それによって共感を呼び起こす。さらに、その脆弱表現が、ユーザの信頼性にプラスの影響を与えるという報告もある [4]。

これらの知見を元に、本研究ではシステムのエラーは避けられないという前提のもとで、利用者がエラーに寛容になるようにすることで途中の離脱を回避する手法の開発に取り組んでいる。本稿では 2 章で提案する擬人化と脆弱性の表現方法について説明し、3 章で行動変容向けの提案表現方法を導入した実験システムについて説明する。4 章で不具合があるシステムにおける提案手法を評価し、5 章ではそれを踏まえて考察する。そして 6 章では、提案手法の問題点を元に、システムの計測や判定に対する「システムの確信度」を定量化し、システムに対する信頼性の低下を防止する

ための改善策を議論する。

2. 提案方法

本研究では、システムの擬人化と脆弱性表現を組み合わせたシステムエラー対応手法を提案する。ユーザの信頼を高める擬人化アプローチに加え、エラー発生時のシステムの脆弱表現を加えることで、システムへの寛容感を醸成するための共感を喚起し、システムエラーにおけるユーザの信頼低下を回避することを目指す。擬人化とその脆弱性の組み合わせが、長期的な条件下で利用者の信頼や共感に与える影響を検証するために、本研究ではバーチャルエージェントの感情表現を用いて表現した。物理的なロボットの方が説得力が高いという研究結果もあるが、そのパフォーマンスを長期間にわたってコントロールすることは難しい [5]。そこで、実際の状況におけるユーザの日常的な活動を励ますために、仮想的なエージェントを検討した。

エージェントの外見については、デフォルメキャラクターの特徴である表情や体の動きの可愛らしさを活かした。また、システムがエラーを起こしていないと推測される場合、Wong-Baker FACES Pain Rating Scale [6] を用いて、図 1 に示したエージェントの基本表現を提示した。



図 1 擬人化エージェントの基本表現

擬人化エージェントは、エラーがあると推定されるとき、ユーザの共感を呼び起こすために脆弱な感情を用いる。システムのエラーの大きさに対応するために Marartelaro によって提案された 2 段階の脆弱表現 [4] を採用し、図 2 のように弱い脆弱表現と強い脆弱表現を使い分ける。



図 2 擬人化エージェントの脆弱表現

A study on anthropomorphism to foster tolerance for distrust due to error in behavior change systems

Yiduan HU†, Airi TSUJII‡, Kaori FUJINAMI†‡

†Department of Bio-Functions and Systems Science, Graduate School of Bio-Applications and Systems Engineering, Tokyo University of Agriculture and Technology

‡Division of Advanced Information Technology and Computer Science, Institute of Engineering, Tokyo University of Agriculture and Technology

3. 実験システム

3.1. システム概要

プラスチックごみの過剰投棄による環境破壊という社会的課題を行動変容によって解決することを目的とし、本実験はゴミ捨て行動を対象としている。システムは、行動変容コンテンツと、システムの誤動作に対する寛容性醸成コンテンツで構成される。ゴミ検出部で検出されたゴミの量を、木とエージェントの状態に関連付けてユーザにフィードバックすることで、ゴミ削減行動を長期的に動機づける。図3のシステムフローに従い、以下で各部を説明する。

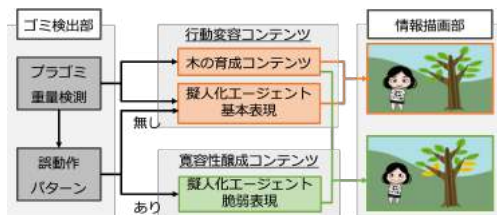


図3 実験システムフロー

3.2. ゴミ検出部

ゴミ検出部として、捨てられたプラゴミから重量検出を行い行動変容コンテンツとエージェントの振る舞いに反映する。ここで高精度な計測が可能な重量に着目し、計測結果に意図的に誤差を埋め込むことで誤動作パターンを生成し、エージェントに対する寛容性醸成評価に使用する。誤動作パターンをゴミの量における増加エラー（ゴミの量を多く誤る）と減少エラー（ゴミの量を少なく誤る）に分け、それぞれをさらに重大なエラーと軽微なエラーに分けた。詳細な設定については、表1に示す。

表1 誤動作パターン

	軽微なエラー	重大なエラー
増加エラー	重さの 200%	重さの 500%
減少エラー	重さの 50%	重さの 20%

3.3. 行動変容コンテンツ

行動変容コンテンツは、木の成長状態の変化として設計されたものと、エラーがない時に擬人化エージェントの表情を変化させることで、より効果的な行動変容を促すように設計されたものの2種類で構成されている。日中のゴミの重さによって図4に示すように木の成長状態が変化し、1日経つと前日の状況として図5に示す木の状態に反映される。システムは、その日のゴミの量に対して、最大限に茂っている状態から完全に枯れている状態まで5段階で表示する。



図4 当日の木の状態



図5 前日の木の状態

基本表現では、行動変容の追加コンテンツとして擬人化エージェントを採用している。表情は図1で例示するように、その日の廃棄ゴミ量が目標値以下であることを「とても嬉しい」「嬉しい」「無表情」の3段階と、目標を超えてしまったことを悲しむ様子を表す3段階がある。このため、ユーザはゴミの量をより具体的に認識することができる。

3.4. 寛容性醸成コンテンツ

ユーザのシステムの誤動作に対する寛容性を醸成するため、ゴミ検出に対するエラーがある場合に、基本表現に代わり図2に示す脆弱表現を用いる。3.2節で述べたように意図的に埋め込み誤動作を発生させ、それを弱い脆弱表現と強い脆弱表現で提示する。

4. 評価実験と結果

4.1. 実験概要

行動変容システムにエージェントとその脆弱表現が及ぼす影響を明らかにするための印象評価実験を実施した。日常空間内における行動を対象とするため、20代の大学生および大学院生9名を被験者として研究室内のゴミ箱の代わりに設置した実験システムを利用させる。実験システムを用いて、条件1) エージェント無し・誤動作無し、条件2) エージェント無し・誤動作有り、条件3) エージェント有り・誤動作無し、条件4) エージェント有り・誤動作有りの4条件を10日づつ合計40日間実施した。各実験条件で使用するコンテンツと実験システムを図6に示す。システムの誤動作パターンについては、実験条件2と4の誤動作有り条件では、各時間内にエラーなし、軽微なエラー、重大なエラーの順に誤動作が起きよう設計した。この場合、エラーなし、増加エラーと減少エラーパターンが異なる実験日に実行された。

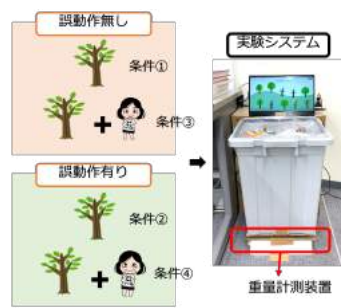


図6 実験条件のコンテンツと実験シナリオ

評価項目は、システムの利用評価とエージェント追加前後の信頼性の変化を検証すること、エージェントの表情状態に対してユーザが抱く印象と共感性を明らかにすることで構成される。異なるシステムにおけるユーザの利用体験を比較するため、システムユーザビリティスケール (SUS) を評価に用いた。信頼性評価には信頼性スケール [7] を利用した。さらに、システムユーザビリティの結果を合わせて、ユーザの信頼度とシステムに対する満足度の関係も評価した。実験条件 3 と 4 において、エージェントに抱く印象評価に GodSpeed[8] を用い、多次元共感性尺度 (MES)[9] から他者志向的反応に関するものを共感性の調査として、アンケートを実施した。さらに、各実験条件の終わりに、参加者にインタビューを行い、実験中の使用感や意見を調査した。

4.2. 実験結果

4 つの実験条件において、10 日間に 1 日の目標通りに行動した期間は、それぞれ条件 1 が 7 日、2 が 4 日、3 が 8 日、4 が 8 日であった。これは、参加者がエラーのないシステムを使用した場合、多くの日数、適切な廃棄処理を行ったことを示している。誤動作ありの実験条件 2 と 4 の比較結果から、擬人化とその脆弱な表現が行動改善を促す可能性があることが示された。図 7 に示す SUS スコアの結果から、システムの使いやすさに対する被験者の評価はエラーのないの方がエラーのあるシステムよりも高い。しかし、提案手法を用いることで、システムの使いやすさが向上するという結果は得られなかった。また、図 8 に示す信頼スコアの多重比較の結果、エージェントを用いない条件 1 と 2 の比較では、エラーがユーザのシステムに対する信頼度を低下させることが確認された。条件 1 と 3 に有意差が存在することから、本デザインの擬人化エージェントは、エラーのないシステムに対するユーザの信頼度を向上させるものではなかったが、条件 2 と 4 にわずかな差が存在することから、擬人化エージェントとその脆弱な表現は、システムエラーによる信頼度低下を改善する可能性があると考えられる。

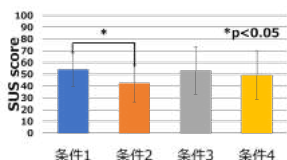


図 7 SUS スコア

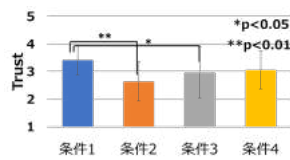


図 8 信頼スコア

GodSpeed アンケートの評価結果から脆弱表現の有無はユーザが擬人化エージェントに抱く印象に影響しないことが確認された。しかし、擬人化エージェントを対象とした MES の結果、脆弱表現を持つエージェントがより高い共感性を獲得した結果が確認された。インタビュー調査において、情報の提示に関して「エラーがあった時とゴミが多かった時の表情を区別するのが少し難しかったから、エージェ

ントがない時より少し使いにくかった」というコメントを確認した。また、エージェントの表現に関する「エージェントは少量のゴミには反応しない」、「もう少しエージェントの表情が変われば、関心を引くことができるかもしれないと思った」とようなコメントを得られた。

5. 考察

検証実験で得られた知見については、提案手法の不足について 2 つの観点からまとめる。

5.1. 情報提示の認知負荷

擬人化されたエージェントはより高いレベルの信頼を得られなかったという実験結果に対してインタビュー結果から参加者は、エージェントが重さをフィードバックすることで、情報を理解する難易度が上がると感じていたことが明らかとなった。この理由として、木の状態でエージェントの状態が同じ情報を伝えるのに用いられ、冗長であったことが考えられる。また、他の研究でも複数の情報フィードバックの対象が重なることは、ユーザの認知負荷につながる可能性があることが示唆されている [10]。このように、異なる要素に対して異なる種類の情報を提示し、擬人化されたエージェントが情報のフィードバックのメディアとしてではなく、社会的な役割としてのみユーザと相互作用することは、ユーザが情報を理解するのに役立つ可能性がある。

5.2. エージェントの表現力

エージェントの感情表現は、先行研究で提案された感情尺度 [4, 6] に沿って設計した。しかし、インタビューの調査結果によると、参加者は、エージェントがより豊かな表情を持つことで、ユーザの現在の行動に合わせて適切に反応し、自然なインタラクションが生まれることを期待していた。豊かな表情を持たせて行動に応じた反応が行われることで、ユーザはエージェントとのコミュニケーションがより深化し、よりリアルで満足度の高い体験を得ることができると考えられる。これらの知見に基づいて、今後は、特に擬人化されたキャラクターの表情をより豊かにし、表情やボディランゲージを充実させ、ユーザの現在の行動とのインタラクションを増やす必要があることが示唆された。

6. 今後の方針

検証実験は、意図的に制御条件下で行った。実用に即したルールでのシステム運用ではエラーが発生し、実際のシステムとは異なる。また、WoZ(Wizard-of-Oz) 法を用いることで、特定のシナリオを模倣する際、本物と異なる場合があるため、ユーザーのリアクションや知見は、実際のシステムでは発生しない問題や状況に基づいている可能性がある

[11]. そのため、実際にゴミの種類を検知するシステムを開発し、その結果を擬人化されたシステムに反映させることで、実利用での評価につなげる必要がある。実際のシステムでユーザの信頼度の変化をリアルタイムに観察し、エージェントの表現力を向上させる取り組みにより、擬人化表現技術と脆弱性表現技術の組み合わせによって、ユーザの信頼度が維持され、長期的に利用されるシステムになる可能性がある。今回の検討結果を踏まえ、実際に利用可能なゴミ種別認識システムを開発し、ユーザの信頼度を観測する機能を追加することで、既存の問題点を解決する。詳細は以下で述べる。

6.1. ゴミ種別認識システムの実装

提案手法の実使用における有効性を検証するために、まず現実の行動変容システムを開発する必要がある。そこで、図9に示すような画像認識に確信度推定を追加したゴミ種別認識システムを開発する。軽量かつ効率的な MobileNetV2 モデルと TensorFlow などのフレームワークを使用して、ゴミの分類に適した TrashNet データセットで事前学習させ、Raspberry Pi 上でゴミの種類をリアルタイムな検知を行う。このシステムを用いた検証実験としては、ユーザのゴミ分別行動を行動改善目標とし、捨てたゴミの種類を判定する。実験システムと同様に、木の成長状態をフィードバック内容として利用し、利用者のゴミ分別行動変容への動機付けを行う。

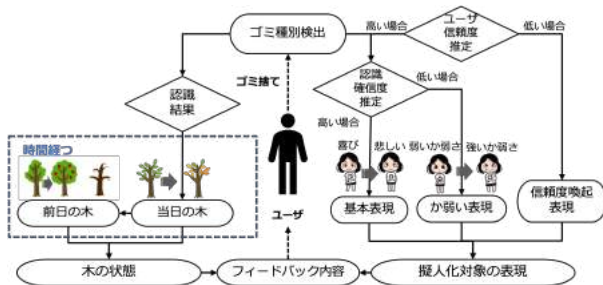


図9 システム概要

5.1 節で述べた情報提示の認知負荷の問題に対して、今後はユーザの分類行動の正誤は木の状態を用いてフィードバックし、ゴミ認識の確信度はエージェントの感情状態を用いてシステム自身の「確信度」として表現する。システムの確信度が低い場合、すなわちゴミ種別の認識精度が高くない場合、エージェントはユーザから寛容さを得るために脆弱表現を提示する。しかし、システムが高い確信を持ったときに、エージェントの基本表現をどのようにフィードバックすべきかは、検討中である。

6.2. ユーザの信頼度観測

システムに対するユーザの信頼度をリアルタイムに把握するため、信頼度測定機能を追加する。また、5.2 節で示し

たエージェントの表現力不足に対処するため、ユーザの信頼度が低下すると、エージェントは脆弱性とは異なる感情状態を表現し、ユーザの信頼度を回復させるために、ユーザとのインタラクションや親近感を増加させる。図9に示すように、信頼度判定の結果が低下した場合、エージェントは元の基本的な表情や脆弱な表情の代わりに、追加の動作や感情表現を表示する。なお本システムの利用シナリオは公共空間であるため、利用者の個人情報の侵害を低減する必要がある。従って、利用者の信頼度の観測手法としては、利用者の位置情報、すなわち、本システム周辺における利用者の滞在時間を判定内容とする予定である。

7. おわりに

本研究では、擬人化機能に対するアプローチを組み合わせた不具合のある行動変容システムとの長期的なインタラクションを通じて、擬人化システムの脆弱な表現がユーザの信頼や共感性に与える影響を調査した。誤動作を組み込んだゴミ箱システムとの長期的なインタラクションにより、擬人化システムの脆弱表現がユーザの信頼や共感性に与える影響を調査し、検証実験の結果により、提案アプローチが有効であることが示された。情報の冗長な提示やエージェントの表現力不足などの課題も明らかになった。

今後の研究では、擬人化手法の改良によって、検証実験で生じた問題点を解決し、提案手法の実利用における有効性を検証することに取り組む。さらに、ゴミ問題に限らず汎用的な場面での有効性についても検討する。日常空間におけるオブジェクトの擬人化が果たす役割を実験的に調べることで、長期的な条件下で信頼関係の構築に与える影響や効果を検証することが期待される。

参考文献

- [1] Kulms, P. et al.: More human-likeness, more trust? The effect of anthropomorphism on self-reported and behavioral trust in continued and interdependent human-agent cooperation, *Proc. of MUC'19*, pp. 31–42 (2019).
- [2] Dehn, D. et al.: The impact of animated interface agents: a review of empirical research, *J. Hum. Comput. Stud.*, pp. 1–22 (2000).
- [3] Okada, M.: Weak robots, *Proc. of JSAP*, p. 220409 (2022).
- [4] Martelaro, N. et al.: Tell me more designing HRI to encourage more trust, disclosure, and companionship, *Proc. of HRI '16*, pp. 181–188 (2016).
- [5] Li, J.: The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents, *Proc. of IJHCS*, Vol. 77, pp. 23–37 (2015).
- [6] Wong, D. et al.: Pain in children: comparison of assessment scales, *Pediatr Nurs*, Vol. 14, No. 1, pp. 9–17 (1988).
- [7] Madsen, M. et al.: Measuring human-computer trust, *Proc. of ACIS*, Vol. 53, pp. 6–8 (2000).
- [8] Nomura, T.: Humans' Subjective Evaluation in Human-Agent Interaction (HAI)(<Special Issue> Benchmarks in AI Research: Standard Problems, Data Sets, Evaluation Methods), *Proc. of JSAI, 31(2)*, pp. 224–229 (2016).
- [9] Suzuki, Y. et al.: Development of the Multidimensional Empathy Scale, *Proc. of Japanese J. Educ. Psychol.*, pp. 487–497 (2008).
- [10] Osawa, H. et al.: Using attachable humanoid parts for realizing imaginary intention and body image, *Proc. of Int. J. Soc. Robot.*, Vol. 1, No. 1, pp. 109–123 (2009).
- [11] Syrdal, D. et al.: Views from within a narrative: Evaluating long-term human-robot interaction in a naturalistic environment using open-ended scenarios, *Cognitive computation*, pp. 741–759 (2014).