

令和 6 年 6 月 25 日現在

機関番号：92707

研究種目：若手研究

研究期間：2019～2023

課題番号：19K20391

研究課題名（和文）高次な感情に対する理由付きコーパスの構築と、ユーモア認識および生成への応用

研究課題名（英文）Construction of a higher emotion causality corpus, and its application to language humor recognition and generation

研究代表者

山根 宏彰 (Yamane, Hiroaki)

富士通株式会社（富士通研究所）・その他部局等・研究員

研究者番号：70825562

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究では、人工知能がユーモアとモラルを理解し、これらの感情を含むコンテンツを生成する技術の基盤構築を行った。具体的には、ユーモアに対する道徳的判断を考慮したジョーク生成システムを構築し、N-gramコーパスとリカレント・ニューラル・ネットワーク（RNN）を用いてモデルがユーモアとモラルを考慮したジョーク生成が可能であることを示した。このシステムにより、モラルを考慮しながらユーモアのあるコンテンツを自動生成する手法を提案し、ヒューマンフレンドリーなAIの実現に向けた新たなアプローチを確立した。

研究成果の学術的意義や社会的意義

本研究は、人工知能が高度な感情を理解し、それを基にコンテンツを生成する技術の基盤を築いた。学術的には、ユーモアとモラルの融合に関する新しい知見を提供し、AIの感情理解能力を大きく向上させる可能性がある。社会的には、道徳的判断を考慮したユーモア生成技術は、安全で倫理的なAIコンテンツの生成に貢献し、AIが人々の日常生活により良く寄り添うための基盤となる。この成果は、教育、エンターテインメント、医療など幅広い分野での応用が期待される。

研究成果の概要（英文）：This study established a foundation for technology that aims to enable artificial intelligence to understand humor and morality, and generate content that incorporates these emotions. Specifically, a joke generation system was constructed that considers moral judgments on humor. Using Recurrent Neural Networks (RNN) and N-gram corpora, the research demonstrated the model's ability to generate jokes while considering both humor and morality. This system proposes a method for automatically generating humorous content while considering moral standards, establishing a new approach towards realizing human-friendly AI.

研究分野：自然言語処理

キーワード：ユーモア モラル

## 様式 C - 19、F - 19 - 1 (共通)

### 1. 研究開始当初の背景

本研究課題は、「高次な感情、特にユーモアやモラルを人工知能に理解させ、人間にその感情を感じさせるコンテンツを生成する」という挑戦的なテーマに取り組むものである。これは、人間の感情の複雑な側面を AI に理解させることで、ヒューマンフレンドリーな人工知能の実現を目指すものであり、感情コンテンツの生成においても大きな意義を持つ。

「ユーモアから誘発されるおかしみの感情」や「モラルの基準に照らした正義感」を人工知能 (AI) に理解させること、また逆に、人間にそのような感情を覚えさせる文章等のコンテンツを自動で生成させることは、人間と親和性が高い AI 構築への大きな一歩である。ユーモアは人間の生活を豊かにするものであるが、中には道徳性が欠如しているために人々を楽しませることができないジョークもある。容姿をブタ扱いする「ユーモア」が社会の注目を集める現代社会において、SNS の浸透を背景にしたモラルに反する投稿を原因として、自殺やうつなどに陥る深刻な事件・事故が後を絶たない。モラルを担保し、それでいて面白いユーモアに対する潜在的需要があるといえる。

### 2. 研究の目的

本研究の目的は、人工知能が高次な感情を理解し、人間にその感情を喚起させるコンテンツを生成する基盤を構築することである。具体的には、ユーモアやモラルの判断を含むコーパスを活用し、ユーモア生成時にモラルを考慮する AI モデルを開発することを目指す。

### 3. 研究の方法

#### ・データ収集：

ユーモアに関する国際会議 ISHS への参加と文献調査を通じて、認知科学のおよび AI の知見を吸収した。ユーモアとモラルに関する知見をもとにし、ジョーク関連のデータセット (Headline のジョークを集めたデータセットなど) や Twitter の Tweet に対する善悪の評価に関するデータを収集。

#### ・モデル開発：

ジョーク生成システムの提案と開発を行った。情報量的なアプローチを用いて  $N$ -gram コーパスからジョーク生成をするモデルを拡張し、道徳面でのラベルが付与されているコーパスで訓練されたりリカレント・ニューラル・ネットワーク (RNN) を組み合わせることにより、道徳的判断を含めたジョーク生成が可能であることを確認した。

より具体的には、「モラルのあるジョーク生成」のため、道徳的な基準に基づいてユーモアを選択可能なメカニズムを提案した。まず、候補となるジョークの生成を行う。具体的には、膨大な Web 上のテキストから構築された  $N$ -gram コーパスを用い、ジョークになりうる統計量をベースにした、様々なテンプレートパターンを用いてジョーク候補の生成を行った。ここで、他人が何を考え、何を感じているかを推論し、その推論に基づいて特定の状況下で相手が何を行うかを類推する「心の理論」を考慮したテンプレートパターンも用いている。

次に、生成されたジョークの選択を行った。具体的には、リカレント・ニューラル・ネットワークの一つである Long-Short Term Memory (LSTM) に対して、モラル価が付与された 35,000 個のツイッターから構成される Moral Foundations Twitter Corpus を学習させることにより、ジョークの選択に利用した。

Best-Worst Scaling から得られた実験結果は、この方式が道徳的なカテゴリーラベルを持つジョークを生成できることを示した。我々の研究で「モラル的に良い」とされている忠誠 (Loyalty) と権威 (Authority) に分類された分類器に関するジョークは、公正 (Fairness)、清純 (Purity)、損害 (Harm)、不正 (Cheating)、墮落 (Degradation) に関するジョークよりも優位に面白いことが確認された。

#### ・その他：

また、従来の RNN 手法に代わり、大規模学習済み言語モデルを用いた新たなモデルのトライアルも行った。機能的磁気共鳴画像法 (fMRI) を活用した神経科学的アプローチの試行もした。マルチモーダルなミームに関する共同研究を実施し、視覚情報を含むユーモアの評価データセッ

トを構築、評価を行った。

#### 4 . 研究成果

プロジェクト初期においては、ユーモアに関する国際会議 International Society for Humor Studies (ISHS) への参加と文献調査を通じた基礎知見の吸収を行った。

また、適切なデータセットの探索をすることで、ユーモアとモラルを融合させるのに適切なデータセットを見出すことができた。

その後は、RNN を訓練することでモラルを考慮しフィルタリングを行うことが可能なジョーク生成システムの提案を行った。クラウドソーシングを用い、モラル判断を含めたジョーク生成の可能性を確認した。成果は国際英文誌「Information Processing & Management」(Impact Factor: 8.6) に採択されている。

その他のトライアルとして、fMRI を活用した神経科学的アプローチを試行した。また、マルチモーダルなミームに関する共同研究で視覚情報を含むデータセットを構築し、ベースラインの評価し、国際会議ワークショップで発表した。さらに、大規模言語モデルを用い、ユーモアとモラルの融合を探求しており、ユーモアに対するモラル制約を行う新しいアプローチを試行しており、今後の研究に繋げる基盤を築いたと考えている。

今後は、大規模言語モデルの本格的な利用と心理学的知見の援用を通じて、より高性能なユーモア生成システムの構築及びモラル理解の深化を目指す予定である。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 Yamane Hiroaki, Mori Yusuke, Harada Tatsuya	4. 巻 58
2. 論文標題 Humor meets morality: Joke generation based on moral judgement	5. 発行年 2021年
3. 雑誌名 Information Processing & Management	6. 最初と最後の頁 1-15
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.ipm.2021.102520	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 -

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 Kohtaro Tanaka, Hiroaki Yamane, Yusuke Mori, Yusuke Mukuta, Tatsuya Harada
2. 発表標題 Learning to Evaluate Humor in Memes Based on the Incongruity Theory
3. 学会等名 Proceedings of the Second Workshop on When Creative AI Meets Conversational AI（国際学会）
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

2022年発表論文概要 <a href="https://www.m.i.t.u-tokyo.ac.jp/papers/papers2022">https://www.m.i.t.u-tokyo.ac.jp/papers/papers2022</a>
--

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
---------------------------	-----------------------	----

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
---------	---------