

意味的類似性と単語親密度に基づく意外性と納得感を考慮したなぞかけの生成

光畑 陽登

人間のコミュニケーションには、言語的なユーモアが存在する。そのため、ユーモアは人間の知的活動の一種ともいえることから、近年、機械が人間のユーモアを表現、理解することへの研究が活発になっている。本研究では、日本語のユーモアの一種として「なぞかけ」に着目した。なぞかけは、『「真夏日」とかけて「犯罪」ととく。そのころは、「帽子/防止」が必要でしょう』のように表現され、お題「真夏日」、解「犯罪」、心「帽子」、「防止」による定型的パターンが存在する。面白さがあるなぞかけには、意外性と納得感が求められる。そこで本研究では、Word2vec と同音異義語辞書を用いて、入力であるお題から解と心の候補を生成し、大規模言語モデルの一種である Sentence-BERT を利用して意味的類似性を算出し、その結果から意外性を判定し、単語親密度情報を利用して難しい表現を除外することにより納得感のあるなぞかけを生成する手法を提案した。

まず、意外性を推定できるようにするために、人によって考えられたなぞかけについて、その意外性を人手で分類し、なぞかけのお題と1つ目の心からなる文と、解と2つ目の心からなる文の2文間の意味的類似性と意外性の関係を分析した。意味的類似性と意外性の関係性を分析した結果、意味的類似性が低くなるに連れて意外性が高くなる結果が得られた。また、なぞかけの文においては、普段は使わないが知識として持っている語彙も用いられると考えられることから単語親密度データベースの“知っている”評定値を活用して難しい表現を除外することを試みた。

以上の提案手法で生成されたなぞかけと、既存研究での手法で生成されたなぞかけと人が考えたなぞかけに対して同一のお題で評価実験を実施し、評価結果に対して t 検定 (有意水準 5%, 両側検定) で有意差の検定を行った。その結果、既存手法に対して意外性においては 52% 有意に向上することを確認し、意味的類似性に基づいて意外性を推定することの有効性を示した。また、単語親密度情報を使用してなぞかけを生成することで、人手で作成されたなぞかけと比肩できる納得感を得られるなぞかけの生成可能性を示した。そのため、意外性と納得感については本研究の提案手法の有効性を示すことができた。また、面白さの指標においては、提案手法、既存手法いずれにおいても人手手法と比較して有意差が見られる結果となった。評価者から意外性が高く、納得感を得られるなぞかけではあるが、文末表現において違和感のあるなぞかけであるために、面白さを感じられなかったとの意見が得られた。しかし、有意差は検出されなかったものの、提案手法が既存手法を上回る結果が得られたことから、提案手法が面白さ指標の改善に寄与した可能性が示唆される。

最後に、納得感、意外性といった要素以外に面白さに寄与する因子があるのかを調査した。その結果、同音異義語辞書の拡充や、Word2vec の使用方法を工夫することにより、より面白いなぞかけを生成することができる可能性を示した。

(指導教員 関 洋平)