

福岡工業大学 学術機関リポジトリ

俳句を利用した語の感情属性の取得

メタデータ	言語: Japanese 出版者: 福岡工業大学総合研究機構 公開日: 2023-12-06 キーワード (Ja): キーワード (En): sentiment analysis, haiku, latent semantic indexing, dialogue system 作成者: 馬場 謙介, 新 伊織, 井崎 茉奈 メールアドレス: 所属: 情報工学科, 情報工学科, 情報工学科
URL	http://hdl.handle.net/11478/0002000063

俳句を利用した語の感情属性の取得

馬場 謙介 (情報工学部情報工学科)

新 伊織 (情報工学部情報工学科)

井崎 茉奈 (情報工学部情報工学科)

Extracting Emotional Attributes of Words Using Haiku

BABA Kensuke (Department of Computer Science and Engineering, Faculty of Information Engineering)

SHIN Iori (Department of Computer Science and Engineering, Faculty of Information Engineering)

IZAKI Mana (Department of Computer Science and Engineering, Faculty of Information Engineering)

Abstract

There is a need to support the independence of the older people, especially to assist them in leading spiritually fulfilling lives. To meet this need, we are developing a dialogue system that can estimate the user's emotional state. To realize that dialogue system, in this study, we implemented a function for estimating scenes put in haiku. We performed latent semantic indexing on a large amount of haiku data. The semantic can be regarded as scenes in haiku, and the vector representation of words obtained here represents the relationship between a word and multiple scenes. If the scene in a haiku is related to the emotion of the author, the vector representation of the words is an attribute about the emotion. In other words, by estimating the scene of the input haiku, we can identify the user's emotional state.

Keywords : sentiment analysis, haiku, latent semantic indexing, dialogue system

1. はじめに

高齢者の自立支援, 特に, 精神的に充実した生活を送るための補助が求められている. 日本の総人口に対する 65 歳以上の人口の割合は 2030 年には 30%を超えることが予測されており⁽⁶⁾, 現役世代の負担を減らすために介護予防をすすむことを確立することが急務である. 日々の会話や文芸創作などの精神的な活動を補助することができれば, 介護の負担を軽減することに加え, 高齢者の生活の質の向上が期待できる.

著者らは, 利用者の感情を推定できる対話システムの開発を行っている. 推定された感情によって発話を変化させることによって, より複雑で人間的な対話の実現が期待できる. Peng ら⁽²⁾による, テキストデータからの機械学習を用いた感情分析についてのサーベイによると, ソーシャルネットワークサービス等の短いテキストからの感情の理解は挑戦的な分野である. 本研究で開発を目指すシステムは, 対話の手がかりとして俳句を利用することが特徴であり, 俳句での伝統的なルールを利用して, 短いテキストから感情推定を行おうとするものである. 鈴木ら⁽⁷⁾は, 俳句は形式の極端な制限にもかかわらず, 作品投稿や鑑賞での共感を形成しやすい点を指摘している. 本システムは, 俳句創作のための歳時記的な検索機能に加え, 俳句に対して情景を

推定し, 感想を述べる機能を有する. これらの機能により, 高齢者の俳句創作を支援することに加えて, 話し相手としてより興味を持てる対話システムを提供することができる.

上記対話システムの実現を目標として, 本研究では, 俳句に詠まれた情景を推定する機能を実装した. 近年, 機械学習ベースの文書生成手法を俳句に適用する研究が多く行われており^(3-5, 8-10), 横山ら⁽¹⁰⁾は生成の手がかりとして情景の重要性を指摘しているが, 情景の直接的な推定やその感情推定への適用は行っていない. 本研究では, 大量の俳句データについて潜在意味解析⁽¹⁾を行った. ここで得られた意味には専門家により俳句でよく詠まれる情景との対応を付けた. よって, 得られた語のベクトル表現は複数の情景との関連を表している. 俳句に詠まれた情景は詠み手の感情に関係があるとすると, この語のベクトル表現は感情についての属性である. つまり, 入力された俳句について情景を推定することで, 利用者の感情に関する状態を特定することができる.

本稿の構成は以下の通りである. 第2章では, 俳句に詠まれた情景の推定機能について説明する. 第3章では, 開発を行った対話システムについて説明する. 第4章では, 今後の方針について述べる.

2. 俳句に詠まれた情景の推定

俳句の情景の推定は、俳句の情景が使用される語によって決まるという仮定のもとに、語の複数の情景との関連を明らかにすることによって実現する。著者らはこれまで俳句の解析を行ってきて、俳句の典型的な情景が、使用された語の組み合わせによって説明できるのではないかとの見解を持った(図1)。これは一般的な文書解析における「トピック⁽¹⁾」として扱うことができると考えた。具体的に、俳句データについて潜在意味解析を行うことによって、語をベクトル値で表現することができ、このベクトルの各要素は意味に対応している。

ぶらんこをこいでこいで火星まで
ぶらんこの地球とび出すまで濃きぬ
ぶらんこの軌道を放れ星の国

ブランコに腰かけながら見る夕日
本音出てブランコゆれて陽が沈む
ブランコの揺れ衰えて夕日去る

図1: 情景の近い俳句の例(俳句コンテスト入賞句より)

本研究では、実際に、約80万句の既発表句について潜在意味解析を行い、語のベクトル表現を得た。具体的には、各句を語に分解し、句と出現語の統計情報を表す行列に対して特異値分解を行った。ただし、一般的な文書から機械学習で得た表記揺れや類義語を同一視した。また、専門家の監修により、一部の季語と句について同一視できる集合を得た。情景の推定機能では、入力された俳句を語に分割し、各語のベクトル表現を用いて情景を推定する。出力される情景は、潜在意味解析で得られた語のベクトル表現の次元に対応しており、次元と人間が理解可能な言葉との対応付けでは俳人の助言を得た。

3. 俳句を用いた対話システムの開発

3.1 プロトタイプの作成

俳句の情景推定機能を持つ対話システムのプロトタイプとして、既存のチャットシステムをインターフェイスとした実装を行った。情景推定機能では、俳句を入力として、情景を出力する。キーボードによりテキストでの入力を行い、出力もチャットシステム上にテキストとして表示される。エンジン部分はサーバ上に実装しており、汎用的なPC等からインターネット経由で利用が可能である。図2に本システムの完成のイメージを示す。

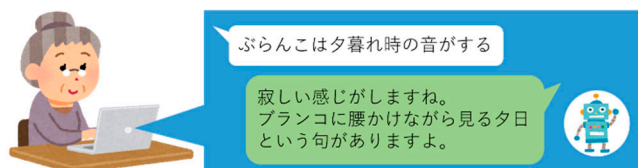


図2: 対話システムのイメージ

3.2 付加的な機能

開発を目指す対話システムでは、俳句の情景推定機能に加えて、俳句創作の補助機能を提供する。本研究において、前述の情景推定機能の開発の際に、副次的に以下の機能を開発した。

1. 入力キーワードを含む既存句の検索
2. 入力キーワードに対する併用語の推薦
3. 入力句に類似する既存句の検索

機能1については、大量の既存句について語への分割を行い、高速な検索用にインデックスを作成している。また、本研究中に得た語のベクトル表現により、厳密一致だけでなく、意味的に近い語での検索も可能である。機能2については、語の共起に関する統計情報に基づいて、入力キーワードと一緒に用いる語句の推薦を行う。これは、俳句創作において「とりあわせ」と呼ばれる語句の組み合わせの技法を、機械的に実現する試みとみなすことができる。機能3については、語句の集合間の基本的な類似度に、語のベクトル表現を導入したものをを用いており、従来よりも細かい類似度を、従来と同等の計算コストで求めることができる。

最終的な対話システムとしては、俳句の情景推定機能を利用者の心情の推定機能として実装するほか、機能1~3を俳句創作支援の機能として実装する。これらの付加的な機能についても、高齢者を被験者とする実証実験において有用性の評価を行う予定である。

4. おわりに

本研究では、俳句に詠まれた情景を推定する機能を開発した。俳句データに対し潜在意味インデックス付けを行うことによって、俳句の潜在意味として情景を推定する機能を実現した。また、この機能を、チャットボットとして実装し、対話システムのプロトタイプを開発した。

今後の方針として、情景推定機能と付随する機能を対話システムとして実装し、実証実験による評価を行う。対話システムを複数の被験者に利用してもらい、入出力の採取とアンケート調査を行う。被験者が実際に入力した文章に対する出力の妥当性を実測データによって評価するとともに、被験者の実際の心情に対する出力の妥当性をアンケートにより評価する。特に、高齢者を対象とした実証実験として、本学の近隣自治体との包括連携により、高齢者コミュニティの協力を得る予定である。

謝辞

本研究は、福岡工業大学総合研究機構「研究スタートアップ支援制度」によるものである。

文 献

- (1) C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval, Cambridge University Press (2008).

- (2) S. Peng, L. Cao, Y. Zhou, Z. Ouyang, A. Yang, X. Li, W. Jia, and S. Yu, A survey on deep learning for textual emotion analysis in social networks, *Digital Communications and Networks*, vol. 8, no. 5, pp. 745-762 (2022).
- (3) 宇野健太, 上垣外英剛, 高村大也, 奥村学, 古典単語と年代情報を活用した俳句の言い換え, *言語処理学会第 26 回年次大会発表論文集*, pp. 585-588 (2020).
- (4) 加藤智一, 竹田晃人, 季語情報付与による俳句自動生成器の改良と評価, *言語処理学会第 28 回年次大会発表論文集*, pp. 927-930 (2022).
- (5) 佐藤汰亮, 菊地正弥, 大竹孝樹, 今野颯人, Ana Brassard, 大内啓樹, 乾健太郎, プロンプトに基づく俳句生成モデルの生成内容評価, *言語処理学会第 27 回年次大会発表論文集*, pp. 1470-1474 (2021).
- (6) 内閣府, 令和 4 年版高齢社会白書, 令和 3 年度高齢化の状況及び高齢社会対策の実施状況(2022).
https://www8.cao.go.jp/kourei/whitepaper/w-2022/zenbun/04pdf_index.html (accessed Jun. 2023)
- (7) 鈴木雅実, 片桐恭弘, 萩田紀博, コミュニケーション・メディアとしての俳句を通じた共感形成に関する一検討, *IPSJ SIG technical reports 2002* (107), pp. 1-7 (2002).
- (8) 和田周, 寺井あすか, 季節を考慮した俳句の自動生成システムにおける評価機構, *日本認知科学会論文集*, vol. 38, pp. 469-472 (2021).
- (9) 廣田敦士, 岡夏樹, 荒木雅弘, 田中一品, 学習データセットを分けた seqGAN による俳句生成, *言語処理学会第 24 回年次大会発表論文集*, pp. 1292-1295 (2018).
- (10) 横山想一郎, 山下倫央, 川村秀憲, 深層学習を用いた俳句の生成と選句, *人工知能*, vol. 34, no. 4, pp. 467-474 (2019).