

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 An evaluation mechanism in an automatic HAIKU generation system considering seasons

和田 周, 寺井 あすか
Shu Wada, Asuka Terai

公立はこだて未来大学
Future University Hakodate
g2120054@fun.ac.jp

概要

計算機により生成された俳句に対する季語を中心とした単語関連度に基づく評価機構について検討した。まず、現代俳句データベースを用いて SeqGAN の学習を行い、季節情報を付与された俳句生成が可能なシステムを開発した。生成された俳句に対し、季語及び構成語の関連度を推定し、評価実験で得られた俳句の良さに関する評定値との関係性を分析した。その結果、意味の理解しやすい句において季語と対立語の関連度が低いほど良い俳句であると評価される傾向が示唆された。

キーワード：俳句生成, 評価機構, 単語関連度, SeqGAN

1. はじめに

近年、人工知能技術の応用として、音楽、小説、動画などの自動生成が盛んに行われている。その中で、限られたことばの制約の中で多様な表現を求められる俳句が注目されている。

俳句は 17 文字で構成された定型詩であり、季語を含んでいる、切れ字を 1 つ以下含むといったルールが定められている。この制約の中で作者の心情や情景を表現し、読み手に伝える必要がある。俳句を詠むという行為において、俳句の生成と同様に生成した俳句を評価し選ぶ「選句」のプロセスも重要である。しかし、俳句の自動生成では、生成した俳句の評価が難しく、俳句の評価機構の開発が行われている(横山ら 2019)。

一方、俳句の良さについては、俳句を読解する際季語を起点にイメージや連想を広げ、句を理解し評価をしていると指摘されている(皆川 2005)。したがって、俳句の評価において、季語及び季節感が重要だと考えられる。そこで本研究では心理実験により得られた知見を比喻生成システムの評価機構に応用することを目的とし、自動生成された句の評価における季語や季節感に関する指標の有用性を検討する。具体的には、

SeqGAN を用いることで季節を考慮した柔軟な俳句の生成を行うとともに、生成された俳句を構成する季語を中心とした単語関連度に基づく評価機構の可能性について考察する。

2. 関連研究

俳句読解に関して、皆川(2005)は評価が高い俳句と評価が低い俳句を刺激材料とし、俳句を構成する単語関連度に関する評定実験をおこなった。その結果、評価の高い俳句のほうが季語と他の語との結びつきが強いことが示唆された。また、吉田(2005)は俳句を構成する季語と対立語(季語と対立するかたちで用いられている語句)の関連性を基にアイデア探索の広さと評価の関連性について分析を行った。その結果、季語と対立語の関連性が低く、探索空間が広いと良い俳句になることが示唆された。

一方、俳句の自動生成に関して、Wu ら(2017)は 4 種類のモデルを用いて俳句生成システムの検討を行っている。しかし、この研究では生成された俳句を Perplexity を用いて評価しており、季語の有無を含め、季語を考慮した評価は行っていない。また、横山ら(2019)は現代俳句と評定値を教師データとする俳句の生成器と評価器を LSTM で開発した。しかし、この研究も同様に、季語を考慮した評価は行っていない。

3. 俳句自動生成

3.1 俳句生成システム

本研究では GAN を系列データに拡張した SeqGAN(Yu et al, 2017)を用いて、柔軟な俳句の生成を行う。学習データとして、現代俳句データベース(<http://www.haiku-data.jp/index.php>)からスクレイピングした俳句・季節・季語のデータを用いた。現代俳句データベースは、現代俳句協会により運営され、総句数 39,292 句の俳句(春の俳句が 9327 句、夏の俳句が 10242 句、秋の俳句が 8114 句、冬の俳句が 8404

句, 新年の俳句が 1014 句, 無季・欠損・その他の俳句が 2191 句) が登録されている。季節を考慮した俳句生成を行うため, 春・夏・秋・冬の俳句を用い, 俳句の文頭に季節情報を追加して学習を行った。例えば, 秋の俳句「赤蜻蛉 筑波に雲も なかりけり」という句に対し, 文頭に季節情報を付与し「<秋> 赤蜻蛉 筑波に雲も なかりけり」として学習を行った。SeqGAN による文章生成では, 俳句が満たすべき制約 (5/7/5 の拍数, 切れ字の個数) を設けていない為, 生成された文に対しフィルタリングを行った。フィルタリングでは, 切れ字 (「かな」, 「や」, 「けり」) の出現が 1 つ以内であるか, 拍数が 5/7/5 であるかの 2 種類の条件を設け, これらの条件を満たす文を生成された俳句として抽出する。一般的に, これらの条件に加えて季語に関する条件を設けるが, 本研究では生成された句と季語の関連性を分析する事が目的である為, フィルタリングに季語に関する条件を含めなかった。

3.2 生成結果

SeqGAN による俳句生成に対する季節情報の有無が与える影響について分析を行った。季節情報を付与した (または付与しない) 学習データに基づく SeqGAN を用いて生成された文にフィルタリングを行い, 季節情報を付与した SeqGAN では 2528 句, 季節情報を付与しない SeqGAN では 2675 句の俳句が抽出された。まず, 俳句データベースにおける季語リストに基づき, 1 句に出現する季語の個数について分析を行った。その結果, 俳句データベースに含まれる俳句においては 0 個が 22%, 1 個が 66%, 2 個以上が 12% であった。季節情報を付与しない SeqGAN で生成された俳句では 0 個が 25%, 1 個が 55%, 2 個以上が 20%, 季節の情報を付与した SeqGAN では 0 個が 24%, 1 個が 57%, 2 個以上が 19% であった。この結果から, 学習に用いられたデータベースと比較し, 季節情報の有無に関係なく, SeqGAN では 2 個以上の季語を含む俳句がより多く生成される傾向がみられた。

次に季語を 2 語以上含んでいる句を対象に, 季語が分類されている季節の不一致率に関する分析を行った。季節の不一致率は, 俳句データベースでは 9%, 季節情報を付与しない SeqGAN では 27%, 季節情報を付与した SeqGAN では 12% であった。この結果から季節情報を付与することで, 異なる季節の季語を含む俳句の生成を抑制できる事が示唆された。

4. 評価指標

4.1 評価指標に用いる季語

生成された句における季語は, 現代俳句コーパスの季語リストに含まれ, かつ, 生成された句に付与された季節タグと一致する季節に分類された語を季語とみなした。季語とみなされる語が複数存在する場合, 季節タグと一致する季節の意味ベクトルと最も関連度の高い語を季語とした。生成された句において, 現代俳句コーパスの季語リストに含まれ, かつ, 生成された句に付与された季節タグと一致する季節に分類された語が存在しない場合, 季語リストに含まれ, かつ, 季節タグと一致する季節の意味ベクトルと最も関連度の高い語を季語とした。

4.2 意味ベクトル

単語の意味ベクトルは白ヤギコーポレーションが配布している日本語 wikipedia コーパスで学習された word2vec 学習済みモデルを用いた。しかし, 学習済みモデルを, 俳句に対して用いた場合未知語が多く, 意味ベクトルが推定できない単語が多数存在した。そこで, 学習済みモデルに対し俳句データベースの俳句を用いて追加学習を行い, 単語の意味ベクトル推定を行った。季節 (春, 夏, 秋, 冬) の意味ベクトルは, 各季節に分類されている季語の意味ベクトルのセントロイドを用いて推定した。意味ベクトル間の関連度は推定された意味ベクトルの \cos 類似度を用いた。

4.3 評価指標の定義

評価指標について, 先行研究 (吉田 2005; 皆川 2005) に基づく俳句の良さに影響を与えると想定される 3 指標に基づき, 以下のように定義した。

季語と対立語の関連度 (季語:対立語)

句を構成する各語に関し, 季語に対して最も関連度が低い語を対立語とみなし, 季語と対立語の関連度

季語と他語の関連度と句全体の関連度の差分

(季語:他語-句全体)

季語を除く句を構成する各単語と季語の平均関連度と句を構成する各語同士の平均関連度の差分

季節と句全体の関連度 (季節:句全体)

俳句に付与された季節の意味ベクトルと句を構成する各語の意味ベクトルの平均関連度

表 1 俳句データベースの句に対する意味の理解しやすさ毎の良さと評価指標との相関係数

| | 季語:対立語 | 季語:他語-句全体 | 季節:句全体 |
|------------------------|--------------------|-----------|--------|
| 意味の理解しやすさ >=1(N = 108) | -0.028 | 0.115 | 0.038 |
| 意味の理解しやすさ >=2(N = 107) | -0.039 | 0.109 | 0.061 |
| 意味の理解しやすさ >2(N = 103) | -0.031 | 0.125 | 0.046 |
| 意味の理解しやすさ >=3(N = 86) | -0.091 | 0.037 | -0.031 |
| 意味の理解しやすさ >3(N = 73) | -0.099 | -0.006 | 0.020 |
| 意味の理解しやすさ >=4(N = 44) | -0.20 [†] | 0.019 | -0.045 |

[†]p<0.1

表 2 SeqGAN により生成された句に対する意味の理解しやすさ毎の良さと評価指標との相関係数

| | 季語:対立語 | 季語:他語-句全体 | 季節:句全体 |
|------------------------|--------|-----------|--------|
| 意味の理解しやすさ >=1(N = 108) | 0.018 | 0.098 | 0.044 |
| 意味の理解しやすさ >=2(N = 106) | -0.044 | 0.082 | -0.002 |
| 意味の理解しやすさ >2(N = 102) | -0.079 | 0.070 | 0.044 |
| 意味の理解しやすさ >=3(N = 74) | -0.024 | 0.107 | 0.135 |
| 意味の理解しやすさ >3(N = 56) | -0.167 | 0.093 | -0.006 |
| 意味の理解しやすさ >=4(N = 22) | -0.47* | 0.033 | -0.108 |

*p<0.05

5. 評価実験

5.1 実験方法

季語を中心とした単語の関連度と俳句に対する評価の関係性を調査することを目的として、季節情報を付与した SeqGAN により生成された句・データベースに登録されている句を対象とした評価実験を実施した。評定者は大学生 3 名。俳句を前述の 3 評価指標（季語と対立語の関連度、季語と他語の関連度と句全体の関連度の差分、季節と句全体の関連度）の高・中・低 3 群の組み合わせについて分類し、各分類ごとに各季節情報を付与された句を 1 つ無作為に抽出し、SeqGAN で生成された 108 句、データベース内の 108 句を用いた。評定者は各句に対して俳句の良さと意味の理解しやすさについて 5 段階評定を行った。

5.2 結果

俳句データベースに含まれる句と SeqGAN により生成された句の良さに関し、各句に関する評定者 3 名の平均評定値を用いて比較したところ、俳句データベースの句が 3.45(SD=0.46) に対し生成した句が 3.21(SD=0.30) であり、俳句データベースの句は SeqGAN により生成された句に比べて有意に良い俳句であると評価された ($t(210)=3.38, p<.01$)。また、俳句の良さと理解しやすさに関する評定者 3

名の平均評定値を用い、俳句の良さと理解しやすさの相関係数を求めた。その結果、俳句データベースの句の相関は 0.71、SeqGAN により生成された句の相関は 0.7 とどちらも、俳句の良さと理解しやすさに関して正の強い相関 ($p<.01$) が得られた。

同様に、俳句の良さと季語と対立語の関連度、季語と他語の関連度と句全体の関連度の差分、季節と句全体の関連度の相関係数を求めた。その結果、俳句データベースに登録された句・SeqGAN により生成された句共に有意な相関は見られなかった。そこで、理解しやすさの違いを考慮し、各指標と俳句の良さの相関係数を求めた（表 1、表 2）。その結果、理解しやすさの評定平均値が 4 以上の句、すなわち理解しやすい句を対象とした場合、俳句データベースの句では良さと季語と対立語の関連度間に -0.2 ($p<.1$) という負の相関が、SeqGAN により生成された句では -0.47 ($p<.05$) という有意な負の相関が得られた。

6. まとめ

本研究では季節情報を付与した SeqGAN を用いて俳句を生成し、季語・季節を中心とした単語関連度と俳句の良さの関係性について検討した。その結果、SeqGAN を用いることで 1 句辺りに出現する季語数は増加してしまうが、季節情報を付与することで異なる季節に分類される季語を含む俳句の生成を抑制出来た。

また、評価実験結果から、俳句の良さと意味の理解しやすさ、季語と対立語の関連度、季語と他語の関連度と句全体の関連度の差分、季節と句全体の関連度の間の関連性に関し、俳句データベースに含まれる句と季節情報を付与した SeqGAN により生成された句において、同様の傾向が見られた。特に、意味の理解しやすい句において季語と対立語の関連度が低いほど良い俳句であると評価される傾向が示唆された。すなわち、先行研究（吉田 2005）で指摘された、良い俳句において季語と対立語の関連性が低いという傾向が、意味の理解しやすい句において確認された。しかし、俳句の良さと理解しやすさに関する評定値の間に強い相関がみられ、かつ、理解しにくい句を含めた場合、季語と対立語の関連度と良さの間に相関はみられなかった。この結果から、理解しやすさの指標として perplexity を用い、perplexity の値が域値以上の俳句に対し、季語と対立語の関連度を指標として用いることで、SeqGAN により生成された俳句の評価機構を構築できる可能性が示唆された。

文献

- [1] 皆川 直凡, (2005) “俳句理解の心理学”, 北大路書房
- [2] 横山 想一郎, 山下 倫央, 川村 秀憲 (2019) “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.
- [3] 吉田 靖, (2005) “俳句の創造性とアイデア探索空間の関係”, 立命館人間科学研究, Vol. 9, pp. 59-72.
- [4] Yu,L.,Zhang,W.,Wang,J.,and Yu,Y.,(2017)“SeqGAN: Sequence generative adversarial nets with policy gradient”. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), pp. 2852-2858.
- [5] Wu,X.,Klyen,M.,Ito,K.,and Chen,Z.,(2017) ”Haiku generation using deep neural networks”. 言語処理学会第 23 回年次大会 発表論文集, pp. 1133-1136.