

作風を反映させた小説の自動生成

大谷 紀子 研究室

1772028 鎌田 海央

1. 背景と目的

現在、コンピュータによる小説の自動生成システムの研究[1]が進められている。佐藤らはあらかじめ用意された文を切り貼りして本文として出力する方法を提案しているが、特定の作風を反映させることはできないため、「この人が書いたこんな話を読みたい」という需要には応えられない。特定の作風を反映させた小説の自動生成によって、他者が書いた小説の疑似的な生成が可能になると考えられる。

本研究では、従来とは異なった創作活動の楽しみ方が可能になることを目的とし、特定の作風を反映した小説の自動生成システムを構築する。

2. システム概要

本研究では、特定の動詞およびオノマトペの出現率により作風を表現する。作風の対象となる動詞は「言う」、「笑う」などの基本的な動作9個を表す単語とし、オノマトペは「くすくす」、「ふわふわ」などの、現在一般的に使用される擬音語・擬態語約4000語とする。また、本システムで自動生成する小説は、500~1000文字程度の短編小説とする。

本システムは、実行に先立って、小説本文の1文目、2文目…… n 文目の出力候補としてそれぞれ複数の短文を用意し、 i 文目と $(i+1)$ 文目の候補のすべての組み合わせに対して、それぞれ同じ小説に含まれるか否かを設定する必要がある。小説生成処理は、作風抽出フェーズ、原文生成フェーズ、および作風反映フェーズで構成される。

作風抽出フェーズでは、まず、特定の作者の小

説から、日本語形態素解析システム MeCab を使用して作風対象となる特定の動詞とオノマトペを抜き出して出現頻度を記録し、各動詞を意味により分類する。例えば「述べる」、「呟く」、「囁く」などは、「言う」という意味のカテゴリに分類される。次に、それぞれのカテゴリにおける各動詞と、オノマトペの出現頻度を基に、作風反映後の小説への出現確率を設定する。

原文生成フェーズでは、1文目はすべての出力候補からランダムに選び、2文目以降は前文と同じ小説に含まれる候補のみからランダムを選んで原文として出力する。例えば、1文目が「おはよう」だった場合には、2文目は「今朝は晴れたね」か「今朝は雨だったね」が選択されるが、「こんばんは」は選択されない。

作風反映フェーズでは、各単語の出現確率に基づき、原文の動詞の変換とオノマトペの挿入によって作風の反映とする。動詞は、原文全体での特定の動詞の出現頻度に各単語の出現確率を適用させてから変換し、同じ単語が連続して選択されないように都度並び替えることで文が単調になるのを回避する。また、作風元となる小説のすべてのオノマトペの出現頻度と、各単語の出現確率によって反映するオノマトペを決め、対応する語句の前に挿入する。

原文と作風反映後の文章の例文を以下に示す。

A はそう言った。 (原文)

A はすらすらとそう述べた。 (作風反映後)

3. 評価実験

原文と作者 A・作者 B・作者 C の作風をそれぞれ

れ反映させた小説 A・小説 B・小説 C を 3 つのストーリーで用意し、本学学生を含む 11 名を被験者に評価実験を実施した。なお、作者 A・B は現在執筆活動を行っている者であり、作者 C は 1900 年代に活躍した童話作家である。

作者 A・B 両名の作風を知る被験者 2 名には、原文と小説 A・B を読ませ、「どちらが作者 A の作風を反映した小説だと感じたか」などに関するアンケート調査を行った。

作者 A・B 両名の作風を知らない 9 名には、原文と小説 A・B・C を読ませ、「原文と比べて、小説 A・B・C のどの小説が一番好みか」を問うアンケート調査を行った。また、被験者にはシステム動作中の動画を見せ、小説の自動生成システムとしての満足度を 1~5 の 5 段階で評価させた。表 1、表 2 にそれぞれ実験結果の一部を示す。表 1 の「—」は、「わからない」と回答されたことを表す。

4. 考察

作者 A・B の作風を知る被験者に対して行った実験では、正答は 6 問中 2 問と、正答率 3 割程度にとどまった。正答した小説に関して、「作者 A は微笑むという表現をあまり使わない気がした」という意見も挙げられており、使用する動詞による作風を反映することはできたといえる。しかし、本システムでは 1000 文字以下の小説で、動詞・オノマトペの違いのみで作風を表現しているため、被験者が作風元の作者を特定するには情報が足りない部分が多く、正答率が上がらなかったと考えられる。特に、ストーリー I は 3 つのストーリーの中で、最も使用されたオノマトペの種類が少なかったため、被験者 2 名ともに「わからない」と回答したと推測される。

作者 A・B の作風を知らない被験者に対して行った実験では、ストーリー I・II・III のすべてで、被験者の 6 割以上が小説 A・B・C のいずれかが最も好みだと感じたという回答している。また、小説 A・B・C それぞれに人数が分散し、被験者ごとに

表 1 作者 A の作風であると感じた小説

	被験者 a	被験者 b	正解
ストーリー I	—	—	B
ストーリー II	A	B	B
ストーリー III	B	A	B

表 2 各小説を好みであると回答した人数

	小説 A	小説 B	小説 C	どれも変わらない
ストーリー I	3	2	1	3
ストーリー II	1	3	2	3
ストーリー III	2	4	2	1

好みと感じた小説が異なっていることから、動詞・オノマトペの出現率によって生じる作風の違いを表現できたといえる。「どれも変わらない」と回答した被験者からは「普段小説を読まないため、言い回しの違いに気が付きにくかった」という意見が挙げられた一方、「擬音が多く使われていると違いが見つかりやすい」、「オノマトペが入るとなんとなく柔らかい気がした」などの意見も挙げられ、普段小説を読まない人でも、オノマトペによる作風の違いはわかりやすかったと考えられる。

また、本システムの小説の自動生成システムとしての評価は平均 3.22 であり、中央値を超えているため有用性が示せたといえる。本システムの動作に対しての意見としては、「簡単な GUI があると良いと思う」、「web 上で動くようにしたら面白そう」などが挙げられた。

今後、動詞・オノマトペ以外の語句の出現率、地の文と台詞の割合、および句読点の位置などによる作風を反映させることで、普段小説を読まない人にも違いがわかりやすく、より元の作風に寄せた小説の生成が可能になると考えられる。また、web 上でも動作可能にするなど、誰でも楽しめるようなシステムへの改良が求められる。

参考文献

- [1] 佐藤理史, “コンピュータが小説を書く日—AI 作家に「賞」は取れるか”, 日本経済新聞出版, 2016