

超短編小説の自動生成機構



名古屋大学大学院 工学研究科

教授 佐藤 理史

コンピュータは、人間を楽しませる小説を作ることができるだろうか。

人工知能の守備範囲を理性から感生へ拡大する試みとして、グラントドチャレンジ「きまぐれ人工知能プロジェクト『作家ですよ』」が2012年秋にスタートした。このプロジェクトは、ショートショートの大匠である星新一に学び、コンピュータにショートショートを創作させることを目指している。我々は、発足当初からこのプロジェクトに参加し、主に文章生成を担当している。

コンピュータによるショートショートの自動生成（自動創作）には、多くの機能の自動化が必要である。その中で、最も大きな障害となると予想されるのが、文章生成、すなわち、「人間が読んで違和感を感じない、一段落以上のテキストを生成すること」である。日本語の文生成・文章生成の研究は、文解析・文章解析と比較して非常に遅れており、現状では、利用できる文生成ツールさえ存在しない。

このような背景より、本研究では、作成する小説の長さを、140字、400字と段階的に長くしていく戦略を採用した。研究の前半では、既存の小説の一部を利用して140字程度のテキストを生成することに取り組んだ。研究の後半では、日本語文生成ツール Haori を設計・実装するとともに、ストーリー文法を用いて400字程度のテキストを生成することに取り組んだ。

テキストの切り張りに基づく生成

新たなテキストを作成する比較的容易な方法は、模倣に基づく方法、すなわち、元となるテキストを用意し、その一部を改変して新たなテキストを作成する方法である。140字程度のテキストを生成する方法として、模倣に基づく2つの方法を実装した。

第一の方法は、元となるテキストの従属節を、別の従属節で置換するという方法である [1]。日本語の「節」は、述語を中心とした文法的単位であり、他の節で置き換えても、比較的意味が通りやすい。節の置換によって生成されるテキスト例を図1に示す。

第二の方法は、小説中の会話部分を抜き出し、別の小

若く美しく、つんとしていても楽しく、人間と同じに働く日。趣味での読書のとき、見た絶望という文字。いま、べたべたおせじを言うし、酒など飲んでいっばいの日々が、ますます人気ができるはずよ。だが、ぼろを出し、お客はこが笑いし、簡単なうけ答えができたロボットと気がつく時がある。今晚で終わりと思うのねとつぶやく。

図1：節の置換による生成例
元テキストは星新一『歲月』、置換先の節は『ボッコちゃん』より抽出。下線部が置換された部分。

「ねえ、ちょっと。起きてよ」
「どうしたんだ」
「いっかぜんで、あたしと別れてくださらない……」
「なんの話かと思ったら、そんなことか。ばかやかい。おまえはおれにとって、遊んで暮らすための大事な金づるだ。別れてやるわけにはいっかぜいぜ」

図2：会話の組み換えによる生成
星新一『泉』より抽出した4発話の後半部を、『小さくて大きな事故』より抽出した2発話で置換

説の会話と組み換えて、新たな会話を作成する方法である [2]。会話の組み換えによって生成されるテキスト例を図2に示す。

これらの2つの方法は、既存の小説のデータベースがあれば、いくらでも作れるが、意味が通るテキスト、あるいは、読んで面白いテキストが生成される確率はきわめて低い。その理由は、どのようなストーリーを作るかという計画性が全く存在しないからである。そのため、これ以上の長さのテキストの生成には、不向きである。

日本語の文生成器

文法的に正しく意味も通る文を、文の切り張りによって作成することには限界があるため、新たに日本語文生成ツール Haori を設計し、実装した [4,5]。

Haori は、入力として与えられた深い文法構造（抽象化された文節依存構造）を出発点として、辞書や文法な

どの知識を利用して語の活用形や機能語（助詞）等を段階的に定め、表層文字列化可能な文節依存構造を生成する。現時点までに中核部分がほぼ完成しており、今後、さらなる改良を行う予定である。

ストーリー文法に基づく生成

長いテキストを生成するためには、全体を組織化するための機構が必要となる。400字程度の超短編小説においても、ストーリーは、いくつかの部分から構成されるのが普通であり、さらに、それぞれの部分は、複数の文から構成される。このような構造を規定する一つの方法がストーリー文法である。

ストーリーの構成にはある種の典型がある。たとえば、導入部、展開部、結末部のような構成は、多くの小説にみられる。さらに、導入部は、「むかしむかし（1）、あるところに（2）、おじいさんとおばあさんがいました（3）。おじいさんは、山に芝刈りに（4）、…」のように、（1）時の描写、（2）場所の描写、（3）登場人物の導入、（4）登場人物の背景説明、といった要素から構成されることが多い。このような典型性のある程度抽象化した形でストーリー文法として規則化する。一方、個別のストーリーを生成するために必要な、時空間、登場人物、イベント系列などの情報は、仮想世界モデルとして、ストーリー文法から切り離す。このような構成を取れば、仮想世界モデルを入れ替えることにより、まったく別のストーリーを生成することが可能となる。図3にストーリー文法の規則例を、図4にストーリー文法と文生成ツール Haori の組み合わせにより生成されるテキスト例を示す。現時点では、生成できるテキストは限られているが、ストーリー文法の拡充と仮想世界モデルの詳細描写により、より多くの多様なテキストを生成できるように拡張していく予定である。

発表論文一覧

- [1] 緒方健人, 佐藤理史, 駒谷和範, 模倣と置換に基づく超短編小説の自動生成. 2014年度人工知能学会全国大会 (第28回), 1C3-OS-14b-2, 2014.
- [2] 高木大生, 佐藤理史, 駒谷和範, 会話を中心とした超短編小説の自動生成. 2014年度人工知能学会全国大会 (第28回), 1C3-OS-14b-3, 2014.
- [3] 高木大生, 佐藤理史, 松崎拓也, プロットと背景知識を用いた短編小説の自動生成. 情報処理学会第77回全国大会講演論文集, 2Q-03, 2015.
- [4] 佐藤理史. 「文生成器を作る」とはどういうことか. 現処理学会第21回年次大会発表論文集, pp.1080-1083, 2015.
- [5] 緒方健人, 佐藤理史, 松崎拓也, 文節木の段階的実

```
def_rule('導入部', # 左辺
[ :イベント, :アクション ], # 必須の引数
[ ], # この引数がtrueの場合はこの規則を適用しない
[ :時間, :場所, :主人公の背景, :アクションの付帯状況 ],
# 任意の引数
[ # 右辺の記述
node('イベント描写', { イベント: arg(:イベント) }),
node('時空描写',
{ 時間: arg(:時間), 空間: arg(:場所) }),
node('主人公導入',
{ 主人公: arg([:アクション, :agent]),
背景: arg(:主人公の背景) }),
node('アクション描写',
{ アクション: arg(:アクション),
付帯状況: arg(:アクションの付帯状況) }) ])
```

図3：ストーリー文法の規則例
導入部を生成する規則。導入部を、イベント描写、時空描写、主人公導入、アクション描写の4つの部分から構成することを規定している

```
スマホが震えた。 # イベント描写
深夜一時ころ。ここは研究室の中。 # 時空描写
鈴木邦男は、先月ここに配属されたばかりであるが、平均
帰宅時間まで深夜帯時を超えている。 # 主人公導入
邦男は大きなあくびをしながら、ポケットの中からスマホ
を取り出した。 # アクション描写
「もしもし」
「鈴木邦男さんですか？」
「はい、あなたは？」
「わたしは悪魔」
「イタズラならよしてくれ。僕もまレポートで忙しいんだ」
「なんでも一つ願いを叶えてみせましょう」
「バカバカしい、さあ、切りますよ」
「お待ちください、一度試してみても損はないでしょ
う？」
「それなら、このひどい眠気をなんとかしてくれ。レポート
が進みやしない」
「お女、御用です」
悪魔がスマホ越しに何やら呪文を呟いたと思うと、邦男の
眠気はさっぱりと消え飛んだ。レポートもぼつちり書けた。
しかしそれ以来、邦男は一睡もすることができなくなった。
```

図4：ストーリー文法による生成
ストーリーの原案は、緒方健人による。冒頭部は図3の規則により生成

- 体化による日本語文生成器の作成. 2015年度人工知能学会全国大会 (第29回), 3M3-1, 2015 (発表予定).
- [6] 佐藤理史. 小説生成器とはどんなシステムか. 2015年度人工知能学会全国大会 (第29回), 3M3-3, 2015 (発表予定).