

星新一のショートショートにおけるオチと状況描写を含むプロットの自動生成

豊澤 修平・村井 源（公立ほこだて未来大学大学院 システム情報科学研究科）

概要：星新一のショートショートにおいて「強盗・泥棒・詐欺・囚人」のパターンを利用し、シーンの状況に合う情景描写・心理描写を付与したオチを含むプロットの自動生成を行った。具体的には、筆者らが既に関与したオチを含むプロットの自動生成と同様のアルゴリズムを用い、個々の情景描写・心理描写が対象となるシーンに適合するかを種々の属性に基づいて判定することで、情景描写・心理描写の選定と統合を行った。結果としてオチを含み、シーンの状況説明を付与したプロットの自動生成ができることが明らかとなった。

キーワード：プロット、自動生成、構造分析

Automatic Generation of Plots Including Punchline and Situation Description in Shinichi Hoshi's Flash Fiction

Shuuhei Toyosawa / Hajime Murai (Graduate school of Future University Hakodate, System Information Science)

Abstract: In Shinichi Hoshi's flash fiction, the pattern of "robbery, thief, fraud, prisoner" was used to automatically generate plots including a punchline with scene description and psychological description of a character that suits the scene situation. Specifically, using the same algorithm as the automatic generation of plots including a punchline that we have already developed, it is judged based on various attributes whether each scene description and psychological description are suitable for the target scene. Therefore, the scene description and psychological description are selected and integrated. As a result, it became clear that the plot including the punch line and the situation explanation of the scene can be automatically generated.

Keywords: Plot, Automatic Generation, Structure Analysis

1. はじめに

近年、コンピュータを用いた自然言語生成の技術が発展してきており、深層学習などに代表されるような機械学習を用いた研究が数多くなされている。また、これら自然言語生成のアルゴリズムを物語自動生成に利用する発展的な研究も行われてきている。例えば、深層強化学習モデルを利用して既存のコーパスから特定の動詞を報酬としてプロットを自動生成する研究[1]や、登場人物の主語と述語の関係をそそえた文章が生成されるように述語構造から代用語に対して対応する主語の代入を学習するモデルの研究[2]などがある。しかし、これら深層学習などを用いた研究の多くが因果関係を処理することができないため、生成された物語には一貫性がなく、内容や前後関係に齟齬が生じてしまう。

このような問題を可決するためには物語の持つ構造を分析する必要がある[3]。そして、その分析結果を利用することで物語としての起伏を付けた物語自動生成が可能になると考えられる。

既存の物語を分析することで物語固有の構造を発見し、その物語構造を利用した物語の自動生成を行う研究は発展途上である。しかし、物語構造を用いることで、一貫性のある人間が理解可能

な物語を自動生成できる可能性がある。

物語の構造の一つとして意外性のあるオチを含んだ逆転型の構造が挙げられる。この意外性のあるオチに着目し、物語論の手法を用いて分析することで、星新一のショートショートに関していくつかのジャンルにおいてオチを含んだ物語の構造が既に明らかになっている[4, 5]。これにより、オチに至るまでの物語の流れを複数の要素の連続であるパターンとして抽出可能であることが明らかになった。しかし、発見されたパターンは物語の骨組みの部分に相当し、そのみではオチを含む物語のプロットを生成することができない。物語構造を用いたプロット生成には、例えばプロットを構成するシーンの各要素に代入することができる文章データ[6]を収集して、物語構造のパターンの各シーンに当てはめる手法が考えられる。実際に、シーンへの文章データの代入によってオチを含むプロットの生成が可能であることが確認されている[7]。

先行研究の手法で生成されるプロットは既存作品を分析したデータを用い、文章の持つ属性がプロットの条件に合うものを選択して生成されている[7]。しかし、情景描写や登場人物などの心理描写は先行研究では分析及び自動生成の対

象とされてきていない[6, 7]. そのため, 生成されたプロットはオチに至るまでの行動の羅列となり, どのような人物がどのような場所でその行動を行っているかというような状況の詳細が不明な出力になってしまっている点に課題が残っている. そこで本研究では, 先行研究でのオチを含んだプロットの自動生成に合わせて情景描写や心理描写を付与することでより詳細な情報を含む, 物語のプロットの自動生成を試みた.

2. 対象となる文章データの選定

星新一のショートショートにおいて代表的なジャンルとして SF が挙げられる. 星の SF 作品は「宇宙」, 「薬」, 「ロボット」, 「自動装置・発明・コンピュータ」の 4 種類に大別できる. また, SF ジャンルとは別に個性的なテーマとして「悪魔・魔人」も存在する. 加えて, 計量的な物語構造分析[9]において自動分類されたものを手動で修正・拡張を行った中で頻出のテーマであった「強盗・泥棒・詐欺・囚人」なども星新一のショートショートにおける特徴的なテーマとして挙げられる. しかし, これら作品を全て手動で分析をすることはできないため, 本研究では対象を限定した.

具体的には星作品の汎用的な特徴を把握するために, 頻出であったテーマである「強盗・泥棒・詐欺・囚人」から 133 作品を選択した. また, 代表的な SF ジャンルから星新一自身が製薬会社の社長をしていた経歴を持つのでテーマが「薬」の 47 作品, 星新一の個性的なテーマとして言及されることの多い「悪魔・魔人」から 20 作品の合計 200 作品からランダムサンプリングして 50 作品に限定した.

また, 本研究で使用する心理描写・情景描写の文章データは構造化された形で収集した. ただし, オチの構造分析を行った際に, オチにいたる展開上で必須の要素として抽出された箇所に含まれる情景描写・心理描写は除外した(図 1). 図 1 において太字の箇所が心理描写にあたる文である. しかし, オチの構造分析により前後の文も含めた該当部分が持つ機能としては「脅迫」にあたる要素が対応する. 従って心理描写が含まれているがオチの要素に対応しているため分析対象外となっている. また複数の機能がある文章や心理描写・情景描写と説明描写が複合的に用いられている文章はデータの対象外とした(図 2). 図 2 に関しては会話の応答と心理描写が複合的に利用されているため分析対象外となっている. またこれら 2 種類の場合を除き, それ以外で対象となるデータを機能ごとに整理した. 例えば, 図 3 の太字箇所にあるような場合は上記の条件に当ては

まらないためそれぞれ情景描写と心理描写として抽出可能となっている. 結果として情景描写と心理描写に相当する文を人手で 76 箇所抽出した.

猫と鼠

私は電話をかけた。
呼出音がとだえ、相手の声に変った。
それにむけて、私はていねいに、また、
にくにくしく話しかけた。

「おや、まだご在宅でしたね。わたしですよ。そう、お待ちしておりますよ。きょうは二十五日。まさかお忘れではないと思いましたがね。ちょっとご注意までに……」

相手はうんざりするような感じを、声
に露骨にあらわした。

「わかっています。もちろん、忘れてはいませんよ。しかし、ちょっと急用ができましたので、あしたにしていただけませんか。あしたの晩には、必ずおうかがいいたしますから」

星新一：「星新一ショートショート 1001」, 新潮社, 1998, 「猫と鼠」.

図 1 分析対象外の例 1

猫と鼠

ねぎらいの感情をこめてこう呼びかけたが、相手はかんだ苦虫をはき出すように答えた。

星新一：「星新一ショートショート 1001」, 新潮社, 1998, 「猫と鼠」.

図 2 分析対象外の例 2

猫と鼠

また二十五日の夜がやってきた。
かわいそうだが、督促の電話をかけるとしよう。
この冷酷な世の中に、同情は無用だ。

星新一：「星新一ショートショート 1001」, 新潮社, 1998, 「猫と鼠」.

図 3 分析対象外の例 3

3. 文章データの構造とアルゴリズム

収集した文章データを自動生成に利用するために構造化を行った。具体的には「抽象化された文章」、「文脈依存性」、「機能」の3つの属性を持つデータ[6]として構造化した。この「抽象化された文章」を抽出する際には登場人物や物などの具体名を抽象化し置換可能な箇所として“[]”で囲まれた形で表現した。置換可能であることをデータ中に明記することで、登場人物を星の作中で出てきた既存の名前や職業名などに置換することが可能となっている。次に、整理された情景描写・心情描写を文章データごとに「文脈依存性」を属性として付与し、属性ごとに0, 1, 空欄の3つの状態を持つビット列を持つ形で再度データ構造を整理した(図4)。

オチを含むプロットおよびプロット中の情景描写・心理描写の自動生成には豊澤ら[7]の既存のアルゴリズムとオチを含むパターンを利用した。文章データは属性ごとにビット列が付与されており、“1”の場合は属性を含む文章データを選択でき、空欄の場合はその属性を含んでも含まなくてもどちらでも利用でる。“0”が含まれる場合はその属性を含まない文章データを利用することができるようになっている。これらの属性の組み合わせによって、各場面の状況をデータとして表現し、各場面の状況に適合する文章データを論理演算で選択することを可能にしている。例えば図5の場合だと、何者かが「侵入」し「脅迫」を行っているような状況で適用可能な文章を選択する例である。これにより「侵入」と「脅迫」が“1”である場合は、選択された列と、ビット列が“1”または空欄の行にある文章が選択される。そのため「文章1」と「文章2」が選択されるようになっている。

本研究では、文章データを挿入するためのオチを含むプロットは「脅迫逆転」パターンを利用した。これは「強盗・泥棒・詐欺・囚人」から抽出されたパターンとなっている(図6)。登場人物の一人が脅迫をしてお金を強奪していたのが、脅迫し返されお金を奪い返されるという立場が逆転するのがオチとなっているパターンと脅迫されて仲間になるが最終的に裏切り脅迫していた犯罪者が警察に掴まるパターンの2つの分岐が存在する。本研究では「脅迫逆転」パターンの脅迫し返される分岐を生成するひな形として選定した。これらパターン以外にもオチの構造分析[4, 5]により複数のパターンの抽出も既に行われており、例えば7などがある。

4. オチを含むプロット生成の手法

プロットの各シーンの状況と矛盾のない文章

データを属性に基づいて抽出するアルゴリズム[7]を利用することで、シーンごとに特定の情景描写・心理描写が付与されるプログラムを開発した。

オチを含む物語の構造のパターンとそれらのシーンに代入可能な文章を構造化したデータに関しては先行研究の物を用いた[7]。

情景描写と心理描写のデータは3節の手法に基づき構造化し作成した。情景描写と心理描写のデータの一部抜粋をそれぞれ、表1・表2に示す。情景描写と心理描写のテキストは「抽象化された文章」、「機能」、「文脈依存性」の3種類の属性を有しているが、「機能」ごとに表を分けている。表中では「抽象化された文章」は「文章」列に記載している。また「文脈依存性」はその他の列が示す各属性列として表記している。

表1の情景描写の属性は「時間」に関する属性が5個と「場所」に関する属性が5個、そして動作の合間などに挿入可能な「動作」が1個を属性として採用している。また心理描写は各心理状態が属性として付与され、17個を属性として採用している。

次に既存のアルゴリズムに則りシーンごとに文脈依存性に適合するデータを選択し、情景描写と心理描写を含むオチを含んだプロットの自動生成を行った。ただし、シーンの文章中で心理描写と情景描写を付与する箇所の選定は事前に人手で行った。

ID	侵入	登場	脅迫	凶器	文章
1	1	1			文章1
2		1			文章2
3		0	0		文章3
4	0		0		文章4

図4 文章データの例

ID	侵入	登場	脅迫	凶器	文章
1	1	1			文章1
2		1			文章2
3		0	0		文章3
4	0		0		文章4

侵入
1

脅迫
1

「侵入」と「脅迫」の「1」が選択された

図5 文章データが選択される例

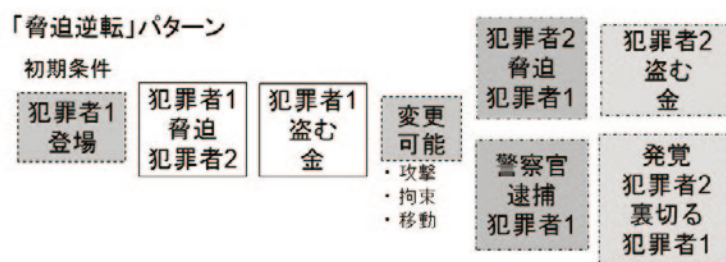


図6 「脅迫逆転」パターン

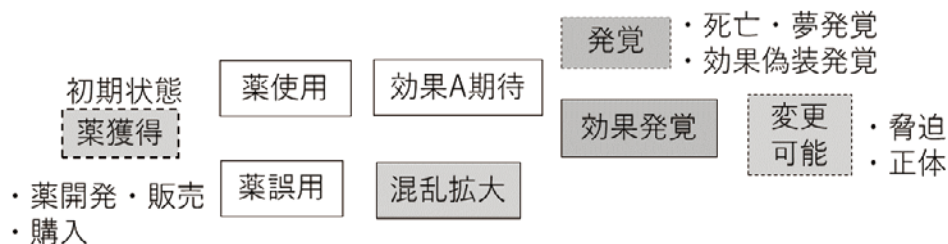


図7 「薬効果明示」パターン

表1 情景描写の文章データ例 (一部抜粋)

ID	title	夜	夕方	朝	昼	夏	マンション	ビル	部屋	屋外	店	動作	文章
1	模型と実物											1	さあ、急がなくては。
2	雪の夜					0		0	1	0		0	この[場所]はわりあいに広く、しかも大通りからはなれているので、なかには冬の静かさがただよっていた。
3	最高の悪事	0	1	0	0							0	翌日の夕方になった。
4	ひとりじめ	1	0	0	0							0	なにげなく時計をのぞくと、ずいぶんおそい時間だった。
5	運命のまばたき	1	0	0	0				1	0		0	ある晩、[人]は部屋の窓から、ぼんやりと、そとをながめていた。
6	雪の夜	1	0	0	0							0	夜ふけ。
7	雪の夜					0		0	1	0		0	この[場所]はわりあいに広く、しかも大通りからはなれているので、なかには冬の静かさがただよっていた。
8	雪の夜					0	0	0	1	0	0	0	暖炉の炎は時どきばちばち軽い音をたて、[人々]の会話の切れ目を埋めた。
9	ひとりじめ	1	0	0	0		0	0	0	1	0	0	夜ふけの道は人通りが絶え、少し先の街灯の光が、あたりを静かに青白く照らしている。
10	魔法の大金	1	0	0	0		0	0	0		1	0	夜の盛り場。八時半ごろ。

表2 心理描写の文章データ例 (一部抜粋)

ID	title	敵対 外	驚外	驚内	不安	奮	怒	恐怖	悲	無	疲労	緊張	苦	退屈	失望	諦	喜	安堵	文章
1	てが かり	0	1	0	0			0	0	0	0	0	0	0	1	0	0	0	いったい、どうなっ ているんだ。
2	模型 と実 物	1	0	0	0			0	0	0	0	0	0	0	0	0	0	0	あんな連中に対して、 良心なんかを抱くこと はない。
3	港の 事件	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	私はほっとした。
4	消え た大 金	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	いまの[人]は青ざめて いた。とんでもないこ とに、なってしまった のだ。
5	消え た大 金	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	ベンチにかけたとな ん、なぜか、ふっと気 が抜けた。
6	消え た大 金	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	感情の空白、そんなも のに包まれたようにな り、意識がずっと去っ ていった。
7	盗難 品	0	1	1		0	0	1	0	0	0	0	0	0	0	1	0	0	あまりの意外さ、あま りの突然さに、[人]は 呆然となり、抵抗する 気力を失っていた。
8	盗難 品	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	それを見て、[人]は目 を丸くした。
9	盗難 品	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	[人]は狐につままれた ようだった。
10	もと で	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	[人]はさらに足を早 め、ハンドバッグをに ぎる手に力をこめた。

5. 結果

50 作品を分析した結果、単一の機能に相当する情景描写・心理描写が合計で 76 箇所得られた。本論文で得られたデータを「脅迫逆転」パターンに適用して実際にオチを含むプロットの自動生成を行った例を図 8 に示す。また、比較のため既存の情景描写・心理描写がないオチを含むプロットを図 9 に示す。

図 8 中で太字部分は本研究の結果追加された情景描写・心理描写部分を示している。図 9 のオチに寄与する最低限のテキストのみの自動生成の場合に比べ図 8 にあるプロットの方がより実際の小説に近い文章が生成されていると言えよう。これにより前後の文脈や物語の設定に適合した情景描写・心理描写を含むオチを含むプロットの自動生成を生成可能となったと考えられる。

6. 考察

星新一のショートショート作品の特徴として、会話文や地の文等でシーンを抽象的に表現し、登場人物の状態や心情、場所の情報などを複合的にまとめて表現する箇所が多い。そのため本研究での文章抽出の対象外となる場合が多く、一作品から得られる情景描写・心理描写の文章データが少なくなっていると考えられる。

景描写の文章データは時間帯と場所に関するものが多数抽出された。また、これら文章データも同様に時間帯と場所それぞれに関する属性が付与される傾向にあることがわかった。

心理描写に関しては心情状態がそのまま属性として付与されるため、登場人物の「恐怖」を感じている描写の際に「喜」の様な異なる心理描写が描かれにくい。そのため制約が非常に強く、他の心理描写として挿入することが難しい。データの構造としても属性ビットの大部分が“1”と“0”で表現されて空欄となる箇所がなく、他の描写に

も適応可能となる場合が非常に少ないということがわかった。また、星新一の作品は人間側の無知や努力を皮肉る世界観[10]が知られている。そのため、人間の心理描写は「不安」、「恐怖」、「悲」といったネガティブな属性が多くあらわれた。そのためポジティブな心理表現を収集するためには他の作品を追加する必要があると考えられる。

7. 結論と今後の課題

本研究では、物語構造のオチを含むパターンを用いてシーンごとにテキストを付与することで、情景描写や心理描写も付加されたオチを含むプロットの自動生成が実現できることを明らかにした。

しかし、情景描写と心理描写を挿入する箇所の選定などをまだ人手で行っている。加えて語り手の視点や語り口に関してもまだ未統一の箇所が散見される。これらの問題点に関して、各データに属性の追加や文体変換の手法を取り入れることで改善することが今後の課題と言えよう。

利用可能な描写データを増やすことができればより多様なプロットを生成することが可能となる。また、データを星の作品に限定することなく別媒体などから設定や描写をデータ化することが可能となれば、描写だけではなく、作風や設定が新しいプロットの生成が可能になると考えられる。

ある晩、デール氏は部屋の窓から、ぼんやりと、そとをながめていた。
 ドアをあけると、四十歳ぐらいの、見おぼえない顔の男が立っている。
 拳銃がにぶく光って、彼の胸に迫ってきた
楽しい気分も、すっかり消えてしまった。
 ドリルで金庫をこじあげ、なかの金を盗み出す。
デール氏はそのようすを見て、心のなかで歯ざしりをした。
 デール氏はナイフを男に突き付ける。
 「それはいけません。このコピーはとってあり、私が死んだら、すぐ警察に送るように、友人にあずけてありますから」
心のふるえを声に出さないよう、注意して言う。
 デール氏は金を奪って言う。

図8 オチと描写を含むプロットの自動生成例

ドアをあけると、四十歳ぐらいの、見おぼえない顔の男が立っている。
 拳銃がにぶく光って、彼の胸に迫ってきた
 ドリルで金庫をこじあげ、なかの金を盗み出す。
 デール氏はナイフを男に突き付ける。
 「それはいけません。このコピーはとってあり、私が死んだら、すぐ警察に送るように、友人にあずけてありますから」
 デール氏は金を奪って言う。

図9 オチを含むプロットの自動生成例

参考文献

- [1] Pradyumna Tambwekar, Murtaza Dhuliawala, Lara J. Martin, Animesh Mehta, Brent Harrison, Mark O. Riedl: Controllable Neural Story Plot Generation via Reinforcement Learning, arXiv: 1809.10736.
- [2] Angela Fan, Mike Lewis, Yann Dauphin: Strategies for Structuring Story Generation, Association for Computational Linguistics, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 2650–2660, 2019.
- [3] ウラジーミル・ブロップ (北岡誠司, 福田美智代訳): 「昔話の形態学」, 水声社, 1987.
- [4] 豊澤修平, 村井源: 星新一のショートショートのおチの構造分析, 人文科学とコンピュータシンポジウム論文集, Vol. 2018, pp. 75-82, 2018.
- [5] 豊澤修平, 村井源: 星新一のショートショートにおけるSFジャンルのオチ構造分析, 人工知能学会全国大会論文集 第33回全国大会 2019.
- [6] 豊澤修平, 村井源: 物語自動生成のための文脈依存性を考慮した文章表現抽象化, 人文科学とコンピュータシンポジウム論文集, Vol. 2020, pp. 135-142, 2020.
- [7] 豊澤修平, 村井源: 星新一のショートショートにおけるオチを含むプロットの自動生成, 人工知能学会全国大会論文集 第35回全国大会 2021.
- [8] 星新一: 「星新一ショートショート1001」, 新潮社, 1998.
- [9] 村井源, 松本 斉子, 佐藤 知恵, 往住 彰文: 物語構造の計量分析に向けて-星新一のショートショートの特徴-, 情報知識学会誌, Vol. 21, No. 1, pp. 6-17, 2011.
- [10] 最相葉月: 「星新一 一〇〇一話を作った人」, 新潮社, 2007.