

テーマに基づく短編小説自動生成システム

松山諒平 佐藤理史 松崎拓也
名古屋大学大学院 工学研究科

1 はじめに

コンピュータプログラムで小説を生成する試みは、古くから行われてきたが [1]、まだ十分に実現されているとは言い難い。我々は、これまでに2つの小説生成システムを作成し、これらを用いて制作した作品を星新一賞に応募するとともに、システムをWEB上で公開している [2, 3]。

小説生成システムは、多くの異なる小説を出力することが求められる [4]。上記の2つのシステムは、表層的に異なる多くの小説を出力可能であるが、物語の骨格は1種類しか存在せず、内容的なバリエーションに乏しい。

本稿では、「内容的に異なる」小説を生成する方法として、テーマに基づく小説生成法を提案する。ここでの「テーマ」とは、「愛情」「裏切り」「努力」など、物語全体を貫くキーワードをさす。一般に、それぞれのテーマには物語の典型的な骨格が存在する。たとえば、テーマが「愛情」の場合は、典型的な骨格は「複数の登場人物が、何らかの試練を乗り越えて愛情を感じる」となる。一方、テーマが「裏切り」の場合は、典型的な骨格は「仲のよい登場人物のうち、どちらか一方が裏切る」となる。つまり、テーマを定め、そのテーマに典型的な骨格を利用して小説を組み立てれば、内容的に大きく異なる小説を生成できると考えられる。

以下では、まず、テーマに基づく小説生成法の概略を示す。次に、実際の小説の生成例に即して、作成したシステムを説明する。

2 テーマに基づく小説生成法

テーマに基づく小説生成とは、小説をトップダウンに作る方法の一つである。より具体的には、以下の5ステップで小説を生成する(図1)。

Step1 テーマの決定 テーマを決め、使用する骨格を定める。たとえば、テーマを「愛情」と決め、「複数の登場人物が何らかの試練を乗り越えて愛情を感じる」という物語の骨格を定める。

Step2 テーマの詳細化 物語の骨格は、肉付けを必要とする。たとえば、「複数の登場人物が何か試練を乗

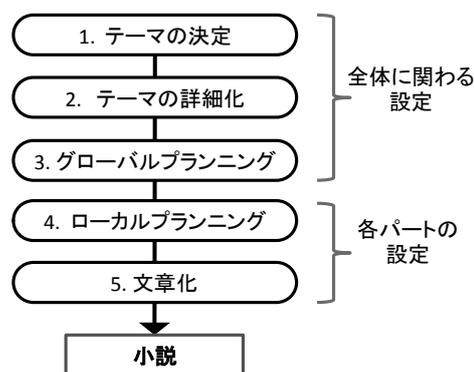


図 1: テーマに基づく小説生成のモデル

り越えて愛情を感じる」という骨格を用いる場合、物語開始時の登場人物同士の関係や、そこでどんな試練が生じるかなどを定める必要がある。ここでは、それらの情報を定める。

Step3 グローバルプランニング テーマ以外の、小説の全体に関わる設定を行う。具体的には、小説の舞台となる物語世界(時空、登場人物など)を設定し、小説全体をどのようなパートから構成するかを定める。

Step4 ローカルプランニング 小説のそれぞれのパートに対して、そのパートのみに関わる物語世界の情報を決定し、そのパートをどのようなアクションの並びとして構成するかを決定する。

Step5 文章化 それぞれのアクションを文字列として生成し、それらをつなげて小説を完成させる。

3 作成したシステム

上記の方法に基づいて小説を生成するシステムを作成した。このシステムは、現代のある場所を舞台とした、人間と動物が登場する500~700字程度の短編小説を生成する。システムの入出力例を図2に示す。この図の左側は、テーマ「愛情」の例、右側は、テーマ「裏切り」の例である。

本システムは、次の3種類の情報を入力として受け付ける。

1. テーマ

入力：
 テーマ = 「愛情」
 テーマ詳細情報 = {〈愛情の変化〉 = 「再確認」}
 物語世界情報 = {〈動物〉 = 「ゾウ」}

出力：

おれの趣味は動物にエサをやること。でも、ペットは飼っていない。だからおれは動物園などの動物が見えるところへ良く足を運ぶ。
 特に最近は楽しみが増えた。楽しみとは、よく行く動物園に新しく来たゾウのこと。名前は「ハナコ」と言うらしい。おれは今日もハナコに会いに行く。ハナコは、出会ったときからとてもかわいかったのだが、そのころは今ほど人懐っこくなかった。しかし、何度も顔を合わせているうちにだんだんとおれに近寄ってきてくれるようになり、今では大の仲良しだ。
 しかし、今日は出会ったときから様子が違った。おれの姿に気がつくなり、すっと檻の奥の方まで下がっていった。そして、おれが檻の前まで行くと、唸り声をあげながら助走をつけて飛びかかってきた。おれは怖くなって目を閉じた。いったいどうしてしまったのだろうか。何か怒らせるようなことをしたのだろうか。
 そして、足元でジャリッともものが落ちた音がして目を開けた。すると足元にあったのは、なくしたはずのおれの小銭入れだった。そういえばなくしたのは前回ここに来たときだった。顔を上げると、ハナコと目が合った。そこでようやく、小銭入れはハナコが拾って投げしてくれたものだとわかった。ハナコはまだこちらを見ていた。やっぱりハナコが大好きだ。

入力：
 テーマ = 「裏切り」
 テーマ詳細情報 = {〈裏切りの詳細〉 = 「騙し」
 〈裏切る側〉 = 「人間」}
 物語世界情報 = {〈人間〉 = 「少女」
 〈場所〉 = 「家」}

出力：

あたしのペットはポチ。とってもかわいいイヌだ。あたしの楽しみはポチへのエサやり。とってもかわいし、食いしん坊だから、ついエサをやり過ぎてしまう。あたしがポチと出会ったとき、まだポチはまだ小さかった。しかし、何度も顔を合わせているうちにだんだんとあたしに近寄ってくれるようになり、今では大の仲良し。きっとポチもあたしをとて信頼してくれているのだろう。そんなポチがあたしは大好きだ。
 今日のご飯はまだかとはばかりにポチがこちらを見ていた。しかし、家に常備しているエサが無くなってしまった。さらに、今からエサを買いに行くこともできなかった。というのも、あたしは頭が痛くて熱もありそうで、全然動けそうになかったのだ。ポチに、だからエサはあげられないと伝えた。それを聞くと、ポチは悲しそうな顔した。あたしは少し胸が痛んだ。
 あたしは、買い物にもいかず、のんびりと家でテレビを見ることにした。テレビではちょうど漫オグラブリが行われていた。次々におもしろい漫才が披露され、何度も大きな声で笑ってしまった。一方ポチは、その様子を怪訝な目で見ていた。ちょうど気分良くなっていたあたしは、体調なんて本当はたいしたことないのだけど、この漫オグラブリが見たくてねと、エサやりをサボって買い物にも行かなかった理由を正直に打ち明けた。ポチにこんなこと言ってもわからないだろうと思っていたのだが、どうやらあたしとポチは本当に心が分かってしまうようだ。ポチはみるみるうちに怒りだし、爪を立てて飛びかかってきた。

図 2: 小説生成システムの入出力例

- 2. テーマ詳細情報
- 3. 物語世界の中核項目の情報

現時点では、テーマとして、「愛情」、「裏切り」、「努力」の3種類が存在する。テーマ詳細情報は、それぞれのテーマに対して2~4種類存在する。物語世界の中核項目は、〈人間〉、〈動物〉、〈場所〉の3種類である。これらのすべてに対して、あらかじめ選択肢が定義されており、入力ではその中の一つを指定する。なお、指定しなかった場合は、生成過程において、選択可能な選択肢の中からランダムに一つが選ばれる。

以下では、小説を生成する過程を、図2の左側の小説を生成した場合を例に説明する。

3.1 テーマの決定

テーマが指定されなかった場合は、前述の3種類のテーマの中から一つを選択する。この例では、入力でテーマは「愛情」と指定されていた。

3.2 テーマの詳細化

テーマ「愛情」に対する詳細情報には、〈愛情の変化〉と〈試練の内容〉の2つの項目が存在する。〈愛情の変化〉は、「愛の芽生え」か「愛の再確認」のいずれかの値をとり、〈試練の内容〉は、「動物が脱走する」、「動物が襲ってくる」、「第三者が襲ってくる」、「物を紛失する」のいずれかの値をとる。

これらの値には、小説を生成できない組み合わせが存在する。そのため、そのような組み合わせを排除するための制約条件が、システムには組み込まれている。この例の場合、〈愛情の変化〉は、入力で「再確認」と指定されているため、もう一方の項目〈試練の内容〉の値は、制約条件により「動物が襲ってくる」か「動物が脱走する」に限定される。この例では、「動物が襲ってくる」が選択された。

表 1: 物語世界のグローバル項目

	項目	選択可能な値
中核	人間	少女、少年、青年*、老人
	動物	ゾウ*、ライオン、ネズミ、イヌ、ネコ
	場所	家、町、動物園*
非中核	動物との仲	仲よし*、顔見知り、初対面、険悪
	ペット	一般的、可能、不可能*
	動物の名前	ハナコ*、ハナオ、ボチ、シロ、...
	動物の体長	とても大きい*、大きい、普通、...
	動物の食べ物	草食*、雑食、肉食
	動物との関係	飼い主、管理者、その他*
	動物の管理	柵、檻*
	芸	得意、不得意*
	行動	散歩、エサやり*
	性格	穏やか、活発*
	一人称	ワシ、ぼく、おれ*、あたし、わたし
	性別	男性*、女性

3.3 物語世界のグローバルプランニング

本システムの物語世界のうち、小説全体に関わる項目(グローバル項目)は、表 1 に示す 15 種類である。これらの項目の値の決定を、次の 3 段階で行う。

1. テーマとテーマ詳細情報に基づき、それぞれの非中核項目に対して、選択可能な値を求める。
2. 中核項目の値を決定する。
3. 非中核項目の値を決定する。

物語世界は、テーマとテーマ詳細情報に適したものでなければならない。たとえば、非中核項目〈動物との仲〉は、「仲よし」、「顔見知り」、「初対面」、「険悪」の 4 つの選択肢を持つ。しかし、テーマ詳細情報の〈愛情の変化〉が「愛の再確認」であれば、〈動物との仲〉は「仲よし」でなければならない。反対に、「愛の芽生え」であれば、「顔見知り」か「初対面」のみが可能である。このように、テーマとテーマ詳細情報は、物語世界の非中核項目の選択可能な値を限定する。

テーマ詳細情報が 3.2 節のように定まった場合、以下の 2 つの非中核項目が制約され、選択可能な値は次のようになる。

項目	選択可能な値
動物との仲	仲よし
動物の体長	大きい、とても大きい

このような準備の後、中核項目の値を決定する。まず、中核項目のすべての値の組み合わせに対して、そのような値の割り当てが可能であるか¹を調査する。そして、この調査によって得られた可能な組み合わせ

¹すべての非中核項目が値を持ちうることに条件である。

の中から、ランダムに一つを選択する。図 2 の例の場合、以下の値が選択された。

項目	選択された値
人間	青年 (入力で指定)
動物	ゾウ
場所	動物園

最後に、非中核項目の値を決定する。選択された中核項目の値は、テーマやテーマ詳細情報と同様に、非中核項目の値を限定する。この例の場合、次の項目の値を限定する。

項目	選択可能な値
動物との仲	仲よし、顔見知り、初対面
ペット	不可能
動物の名前	ハナコ、ハナオ
動物の体長	とても大きい
動物の食べ物	草食
動物との関係	管理者、その他
動物の管理	檻
芸	不得意
性格	活発
一人称	おれ
性別	男性

こうして、それぞれの非中核項目に対して、可能な選択肢が最終的に定まり、システムはその中から一つを選ぶ。表 1 中の「*」は、図 2 の生成で選択された値を示す。

3.4 文章構造のグローバルプランニング

生成する小説は、第 1 パートから第 4 パートまでの 4 パートで構成される。システムは、それぞれのパートに対して定義された選択肢の中から、テーマとテーマ詳細情報に基づいて一つを選ぶ。たとえば、第 4 パートを決める場合、その候補は 20 種類存在する。図 2 の例では、テーマが「愛情」であることから、第 4 パートの候補は 4 つに絞られる。さらに、テーマ詳細情報〈試練の内容〉が「動物が襲ってくる」であることから、第 4 パートは「襲ったわけではなかった」に定まる。他のパートも、同様にして定める(図 3)。

3.5 物語世界のローカルプランニング

それぞれのパートは、パートを文章化するために必要な項目のうち、物語世界のグローバル項目には存在しない項目(ローカル項目)に対して、その選択肢を保持している。

ここでは、それらの項目の値を決定する。図 2 の例の、第 4 パート「襲ったわけではなかった」では、次の項目の値が決定された。

項目	値
怯え方	ちぢこまる
落としたもの	小銭入れ

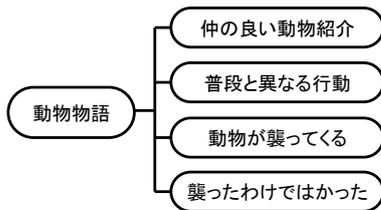


図 3: 文章全体のパート構成

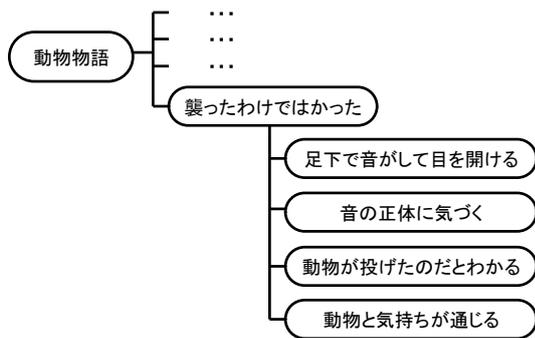


図 4: アクション列への展開

3.6 文章構造のローカルプランニング

それぞれのパートをアクションの列に展開する。システムは、パートをアクション列へ展開する規則を複数保持しており、物語世界の情報に基づいて、どの展開規則を用いるかを決定する。

図 2 の例の第 4 パート「襲ったわけではなかった」は、図 4 に示す 4 つのアクションに展開された。同様に、すべてのパートをアクションの列に展開する。

3.7 文章化

それぞれのアクションを、表 2 に示すようなテンプレートを用いて文字列化する。テンプレートの変数は、物語世界の項目に対応する。

こうして生成された文字列を結合し、各パートを 1 段落の文章とする。全てのパートを文章化したら、それらをつなげて、小説を完成させる。以上のような過程を経て、図 2 の出力例が生成された。

4 検討

以前に我々が作成したシステム [2, 3] が、「内容的に異なる」小説を出力できない理由は、文章のパート構成が 1 種類に固定されているからである。内容的に多様な小説を生成するためには、それらを動的に決定するプロセスが不可欠である。本システムでは、テーマとテーマ詳細情報に基づいて、文章のパート構成を決

表 2: アクションとテンプレートの例

アクション	テンプレート
足元で音がして目を開ける (小銭入れ)	そして、足元でジャリっとものが落ちた音がして目を開けた。
音の正体に気づく	すると足元にあったのは、なくしたはずの〈一人称〉の〈落としたもの〉だった。そういえばなくしたのは前回ここに来たときだった。
動物が投げたのだとわかる	顔を上げると、〈動物の名前〉と目があった。そこでようやく、〈落としたもの〉は〈動物の名前〉が拾って投げしてくれたものだとなった。
動物と気持ちが通じる	〈動物の名前〉はまだこちらを見ていた。やっぱり〈動物の名前〉が大好きだ。

定する。そのため、これらの情報を変更することにより、「内容的に異なる」小説を生成することができる。

多種多様な小説を生成するためには、さらに多くのテーマを実装する必要がある。しかしながら、新しいテーマの追加には、おおよそ次のことが必要であり、かなりの作業量を必要とする。

1. テーマとテーマ詳細情報に基づいて物語世界のグローバルプランニングを行うための制約条件の追加
2. テーマとテーマ詳細情報に対応する、文章全体のパート構成作成の実装
3. 新たに導入されるパートのローカルプランニングと文章化の実装

これらのうち最も作業量が多いのは、文章化のための、アクションに対応するテンプレートの実装である。この部分をできるだけ少ない作業量で実装できるようにすることが、今後の大きな課題である。

謝辞 本研究は、JSPS 科学研究費基盤研究 (B) 「文章の読解と産出のための言語処理技術」(課題番号 15H02748) の助成を受けている。

参考文献

- [1] Pablo Gervás. Computational approaches to storytelling and creativity. *AI Magazine*, 30(3):49–62, 2009.
- [2] コンピュータが小説を書く日.
<http://kotoba.nuee.nagoya-u.ac.jp/sc/gw/>
- [3] 人狼知能小説生成システム.
<http://kotoba.nuee.nagoya-u.ac.jp/sc/gw2016/>
- [4] 佐藤理史. 小説生成器とはどのようなシステムか. 2015 年度人工知能学会全国大会論文集, 2015.