

基盤モデルを用いた物語のセリフの自動生成手法の提案

村井 源・岩岬 潤哉・奥山 凌伍・松原 仁（はこだて未来大学 システム情報科学部）

概要：従来の物語構造に基づく物語自動生成システムと基盤モデルによる自然言語生成を組み合わせたハイブリッド型の物語自動生成モデルの実現可能性をケーススタディ的に検証するために、特定物語ジャンルのセリフの基盤モデルでの生成を行った。セリフを状況説明セリフと状況反応セリフに分類し、物語の状況を Few shot で提示することで、適切なセリフ生成が可能であることを示した。また生成されたセリフを組み合わせることで物語の 1 シーンを構成するセリフの連続的な表現の生成の可能性を示した。今後の課題として、一般的な物語でのセリフ生成を実現するために、複数ジャンル、多様な物語状況の設定、異なったセリフの例示などをより大規模に検証することが考えられる。

キーワード：物語自動生成、基盤モデル、セリフ生成

Proposal of automatic generation method for story lines using a foundation model

Hajime Murai / Junya Iwasaki / Ryogo Okumura / Hitoshi Matsubara
(School of Systems Information Science, Future University Hakodate)

Abstract: In order to verify the feasibility of a hybrid type automatic story generation model that combines the conventional automatic story generation system based on story structure and the natural language generation by a foundation model in a case study, we generated lines of a specific story genre utilizing a foundation model. By classifying lines into situation explanation lines and situation reaction lines and presenting the situation of the story with few shots, we showed that it is possible to generate appropriate lines. In addition, we showed the possibility of generating a continuous expression of the lines that compose one scene of the story by combining the generated lines. As a future task, in order to realize lines generation in general stories, it is possible to verify on a larger scale such as multiple genres, various story situations, and examples of different lines.

Keywords: Automatic Story Generation, Foundation Model, Line Generation

1. まえがき

コンピュータを用いた文学研究の大きな方向性の一つとして、物語の自動生成の研究が盛んになってきている。物語の自動生成をある程度の水準で達成するには、言語理解、物語理解に加え、自然文生成のアルゴリズムや、登場人物や読者の感情状態の推定など様々な要素技術が必要となる。そのため、物語の自動生成は研究推進のための一つのグランドチャレンジとしても有意義と考えられる[1]。

物語自動生成には様々なアプローチが検討されているが、その一つとして物語の構造[2]に基づきプロットを生成する手法がある[3]。既存の物語構造に基づくことで矛盾のない物語展開が生成可能だが、シーンなどの言語表現部分の自動生成は困難で手動のデータセット構築が必要となる。

他方で、近年様々な種類の大量なデータで訓練された汎用的に利用可能な大規模 AI モデル、基盤モデル(foundation models)[4-6]の研究が急速に進展しており、人間と同レベルの言語表現を出力することが可能になりつつある。基盤モデルでは与える入力情報の工夫によって状況に応じた自

然な文章を生成可能であり、この技法はプロンプトプログラミング[7]などと呼ばれている。プロンプトプログラミングによって生成文章を制御することで高精度の対話システムなどを作成することも可能である[8]。ただしこの種の言語モデル一般の特徴として長文を生成させる場合には論理的な一貫性担保のために外的な知識を用いるなどの工夫が必要となる[9]。

物語を自動生成するには、プロットと表現の双方を適切に扱える必要があるが、従来のプロット生成と基盤モデルによる言語生成を組み合わせることで、その目標が達成できる可能性がある。

2. 目的と手法

本研究は、既存作品から抽出した物語構造を用いたプロットなどの物語の枠組みに基づく従来の自動生成と、大規模なニューラルネットワークによる基盤モデルを用いた自然な日本語の生成を組み合わせることで、論理的な破綻がなく、かつ自然な言語表現を用いた物語を自動生成する手法の実現可能性をケーススタディ的に検証することを目的とする。

プロットの生成においては、既存作品に基づく物語のシーン系列の自動生成、およびシーン系列に対応する出来事の生成が必要である。本研究では利用可能な先行研究の多いミステリージャンルを対象とし、シーン系列及びトリックとそれを解明する探偵の言動の自動生成システム[10, 11]を用いてプロットの自動生成を行った場合に個々のシーンの表現を基盤モデルによって自動生成するケースを想定する(図 1)。

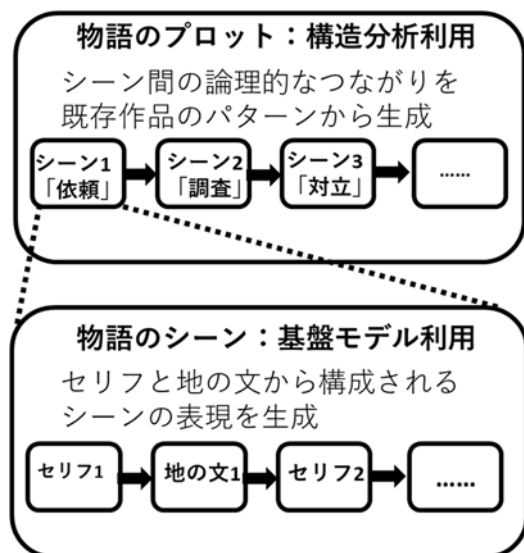


図1 プロット生成とセリフ生成のハイブリッド型物語自動生成の概念図

生成されたプロットに沿った物語の表現を AI によって自動生成する場合、小説などのテキスト中心の媒体の場合には登場人物のセリフ、およびセリフ以外の地の文の生成が必要となる。マンガや映画、脚本などの別の媒体の場合には、登場人物のセリフに合わせて、マンガは各コマの表現、映画はカットの表現、脚本であればト書き、柱書きなどがそれぞれ必要となる。そこで本研究ではまず様々な媒体で共通して必要となる登場人物のセリフを指定された状況に合わせて基盤モデルでの生成を目標とする。本稿では LINE と NAVER が共同開発している、日本語対応の基盤モデル HyperCLOVA[12]を用いる。

状況に応じたセリフを基盤モデルによって自動生成するためには、物語の状況に関する情報を何らかの形でプロンプトとして与える必要がある。しかしプロンプトの効率的な構成法は確立されておらず、試行錯誤によって探索的に目的に合致するプロンプトを記述する必要がある[7]。

基盤モデルに入力として与えるプロンプトは大きく、0 shot, One shot, Few shot に分けることができる。0 shot は基盤モデルに対して何の例も示さずにこの文章の翻訳をせよ、要約をせよなどの指示を出す形式である。One shot はある指示に対する正解例(例えばこの日本語を英語に翻訳せよ

という指示と正解の英訳のセット)を一つ示したうえで、別の類似の指示を出す形式である。同様に Few shot は複数の指示と正解例の組み合わせを提示したうえで、別の類似の指示を出す形式である。例外もあるが多くの場合、複数の指示と正解のセットを基盤モデルに提示した方がより自然な出力が得られると報告されている。そこで本研究でも、物語中の特定の状況の指示とその時の登場人物のセリフのセットを複数提示し、別の状況での登場人物のセリフを出力させる Few shot 形式での言語生成を採用する。

3. 物語のセリフの分類

登場人物のセリフを大まかに分類すると、物語の展開を読者が理解するうえで不可欠となる情報を提示する機能を持った状況説明的な内容を含むセリフと、省略されても物語の展開の理解には影響を及ぼさないセリフに分けることが可能である。

例として、国内のミステリーでシリーズ累計発行部数一億部を突破している『金田一少年の事件簿』[13]での死体発見のシーンにおけるセリフを図 2 に示す。

開かない・・・
 じーなってるんだ？
 わざわざ人呼び出しといて・・・
 ドアには内側から二つ掛け金がかかってる・・・
 彼女はとうとうつもりで・・・
 大変だ！
 人が倒れてる！
 くそっ開け～！
 影原さん！大丈夫ですか影原さん・・・
 ・・・し・・・死んで・・・る・・・
 く・・・首の動脈を切ったのか
 ・・・まだ少しぬくもりが残ってる
 死んでからまだ30分くらいってどこか？
 風呂場もドアが開いているが人の気配はない
 他に人が隠れられるようなお場所もないし・・・
 ドアにはしっかり掛け金がかかっていた
 このタイプのは外から鍵をかけるのはアリだ・・・
 窓も内側から鍵がかかってる
 影原は自殺したのか？
 でなけりゃー完全な密室殺人ー！！！！

図2 死体発見時のセリフ例

図2の例では、状況説明を含むセリフを太字斜体、それ以外を通常のフォントで示している。状況説明を含むセリフでは、ある登場人物の死体が発見されたこと、推測される死因、発見時の状況、密室であること、などの情報が読者に対して提示されている。一方で状況説明を含まないセリフでは発見者の心情吐露やリアクションを通じて臨場感を出したりキャラクターの性格を示したりするなどの効果を提示していると考えられる。

このように状況説明を含まないセリフは、粗筋を把握する上では必ずしも必要ではないが、登場

人物の個性や人間関係を描写したり、その場の雰囲気や詳細な設定を開示したりする機能を持っており、作品を鑑賞する上では有用と考えられる。

本稿では状況説明を含むセリフを状況説明セリフ、それ以外を状況反応セリフと呼称することにする。状況説明セリフの自動生成においては、読者に対して提示すべき情報が矛盾なく適切に含まれていることが重要であり、状況反応セリフの自動生成においては発話者が置かれている状況に対して読者が納得して受容可能な自然と思われる登場人物らしい反応が表現されていることが重要であると考えられる。そのため、本稿では状況説明セリフおよび状況反応セリフの生成用プロンプトの検討を個別に行うこととする。

4. 状況説明セリフの生成

ミステリーの状況設定としては、犯人、被害者、死因と凶器、動機、犯行現場、犯行時刻、犯行現場の状況（密室か否かなど）などの複数の属性が考えられる。本研究では物語構造分析の結果に基づくプロット生成の出力を基盤モデルのプロンプトの入力として利用することを想定し、ミステリー自動生成の先行研究で設定されている属性（動機、凶器、被害者、容疑者）をプロンプトに含めることとした[10]。登場人物の属性に関しても物語の登場人物の属性自動生成システムの出力における頻出の属性（性別、年齢、職業）を採用した[14]。これらに合わせて、一般的なシーンの状況設定である犯行現場と犯行時刻を加え、さらにセリフ生成を行うため、発話者、発話の意図、発話の種類、の三種類の属性を追加した。発話の種類は、既存の物語作品の発話を語用論的な観点から分類した定義表を用いた[15]。これらの属性を指定して発話を生成させるプロンプトは例えば図 3 のような形で表現することが可能である。

```
#ジャンル: ミステリー
#動機: 借金を返せない
#凶器: 鉄砲、銃殺
#犯行時刻: 昨日午前 11 時
#犯行現場: 会社のオフィス、施錠され鍵を紛失したので密室
#被害者: 平川啓一 (男、中年、サラリーマン)
#容疑者: 竹野知実 (女、中年、パートタイマー)
#発話者: 湯川真実 (男、少年、探偵)
#発話の意図: 死体発見
#発話種類: 陳述
#発話者のセリフ:
```

図 3 ミステリーの属性をプロンプト化した例

本章では検証のため、ミステリーにおける殺人事件での死体発見と調査のシーンを想定し、そのようなシーンで頻出する下記の 4 種類のセリフの基盤モデルによる生成が可能か確認する。

- ・ 死体発見に関する陳述

- ・ 凶器と死因の検討
- ・ 密室か否かの検討
- ・ 動機に関する質問
- ・ アリバイに関する質問

基盤モデルに Few shot 形式のプロンプトとして与えるため、図 3 と同様の形式で正解となるセリフを含めた、異なる状況でのプロンプトを追加した場合を図 4 に示す。この例では各状況は「。」で区切られている（基盤モデルで区切り文字は適宜設定可能）。正解例のセリフは太字箇所であり、斜体時部分は基盤モデルに生成を指示する物語上の状況に対応している。図 4 のプロンプトでは異なった状況に対応するセリフの生成可能性をケーススタディ的に確認するため、動機、凶器、犯行現場、犯行時刻等はすべて異なるように人手で設定した。

```
#ジャンル: ミステリー
#動機: 医療事故
#犯行時刻: 昨夜午後 9 時
#凶器: ナイフ、刺殺
#犯行現場: 病院の病室、入口が施錠された密室
#被害者: 梅本友則 (男、青年、医者)
#容疑者: 本橋達弥 (男、老年、患者)
#発話者: 杉江拓己 (男、中年、刑事)
#発話の意図: 死体発見
#発話種類: 陳述
#発話者のセリフ: 「し・・・死体・・・」。

#ジャンル: ミステリー
#動機: いじめ
#犯行時刻: 先週火曜午前 8 時
#凶器: 青酸カリ、毒殺
#犯行現場: 学校の教室、窓の鍵が閉まっていないので密室ではない
#被害者: 菅野実香 (女、老人、教師)
#容疑者: 仲野優哉 (男、少年、生徒)
#発話者: 小田島菜生 (女、中年、探偵)
#発話の意図: 死体発見
#発話種類: 陳述
#発話者のセリフ: 「・・・・し・・・・死んで・・・・る・・・・」。

#ジャンル: ミステリー
#動機: 借金を返せない
#凶器: 鉄砲、銃殺
#犯行時刻: 昨日午前 11 時
#犯行現場: 会社のオフィス、施錠され鍵を紛失したので密室
#被害者: 平川啓一 (男、中年、サラリーマン)
#容疑者: 竹野知実 (女、中年、パートタイマー)
#発話者: 湯川真実 (男、少年、探偵)
#発話の意図: 死体発見
#発話種類: 陳述
#発話者のセリフ:
```

図 4 生成に用いる Few shot 形式のプロンプト例（死体発見時の陳述）

図5は、図4で示したプロンプトによって異なった殺人事件の状況(刺殺と毒殺)での死体発見について述べる登場人物のセリフのセット2件を例示として与えた場合に、図3で指定した状況での登場人物の対応するセリフを5件生成させた結果である。

生成された5件のセリフ中4件は死体発見時のセリフとして適切と言えるが、1件(太字斜体字箇所)は不適切なセリフが生成されてしまっている。また例示したセリフは単に死体を発見したことを陳述する内容になっているが、生成させたセリフでは失敗した例を除き銃殺に対応する表現(「血」「倒れる」「撃たれる」)が含まれる形となっている。

「あ・・・血だ・・・」
 「う・・・撃たれて・・・いる・・・」
「き・・・気になるな・・・」
 「う・・・撃たれて・・・倒れてる・・・」
 「あ・・・あそこに・・・血まみれの・・・人が・・・」

図5 生成された5件のセリフ(死体の発見)

同様に、凶器と死因について言及する登場人物のセリフを生成するためにプロンプト中の「発話の意図」「発話種類」「発話者のセリフ」を正解のセリフを変更した。図4のプロンプトからの変更箇所を図6に示す。

.....
 #発話の意図:凶器と死因は何か?
 #発話種類:思考
 #発話者のセリフ:「刺し傷がある、そのナイフで殺されたのか?」。

 #発話の意図:凶器と死因は何か?
 #発話種類:思考
 #発話者のセリフ:「外傷はない、ひょっとすると毒殺か?」。

 #発話の意図:凶器と死因は何か?
 #発話種類:思考
 #発話者のセリフ:

図6 プロンプト変更箇所(凶器と死因)

図6のプロンプトを用いて凶器と死因に関して検討する台詞を5件自動生成させた例を図7に示す。生成された5件のセリフはいずれも、プロンプトで指定した「#凶器:鉄砲、銃殺」の内容を反映する形になっているが、例示として与えたセリフは刺殺と毒殺の場合に対応した文言であり(「刺し傷がある、そのナイフで殺されたのか?」と「外傷はない、ひょっとすると毒殺か?」)。

指定した状況に応じたセリフが基盤モデルによって問題なく自動生成されていることが確認できる。

「拳銃だ、弾も入っているぞ」
 「弾痕がある、ここに拳銃があったのか?」
 「弾痕がある、誰かに撃たれたのか?」
 「弾丸が残っている、これで撃たれたのか?」
 「弾痕がある、拳銃で撃たれたのか?」

図7 生成されたセリフの例(死因と凶器)

同様に、殺人現場が密室か否かを検討する登場人物のセリフを5件自動生成させるためのプロンプトの変更点を図8、生成した結果を図9に示す。図9中の太字斜体字のセリフは殺人事件ではなく盗難の場合のセリフになっているが、他の4件では指定した状況(「#犯行現場:会社のオフィス、施錠され鍵を紛失したので密室」)に沿ったセリフが出力されていることが確認できる。

.....
 #発話の意図:部屋は密室か?
 #発話種類:思考
 #発話者のセリフ:「病室のドアにはかぎがかかっていましたし、ほかには入口は何もない...これは完全な密室だ」。

 #発話の意図:部屋は密室か?
 #発話種類:思考
 #発話者のセリフ:「窓の鍵がかかっていないから出ることはできる、つまりこの部屋は密室とは言えない」。

 #発話の意図:部屋は密室か?
 #発話種類:思考
 #発話者のセリフ:

図8 プロンプト変更箇所(密室か否か)

「扉や窓などの出入り口はすべて内側からロックされているし、誰かが入ってきたらすぐわかるようになっているんだよ」
 「扉や窓などの出入り口はすべて閉じられていたため、犯人は外へ出られなかったはずだ」
「金庫の中にあったものはすべて持ち出されていたんだよね。それならば犯人はどうやって金庫を開けたんだろう?」
 「たしかにあのときぼくは机の上に置いたんだよね、でもそのあとすぐに失くしてしまったんだ!」
 「扉や窓などの出入り口はすべて施錠されているし、外へ出られる場所はどこにもないようだね」

図9 生成されたセリフの例(密室か否か)

同様に、殺人の動機に関する検討をする登場人物のセリフを 5 件自動生成させるためのプロンプトの変更点を図 10、生成した結果を図 11 に示す。図 11 中の 5 件のセリフはいずれも指定した状況設定（「#動機:借金を返せない」）を反映した形にはなっているが、5 件中 3 件が完全に同じセリフとなっており、妥当ではあるが他の場合に比べて表現の多様性が低くなっている。

```

.....
#発話の意図:容疑者の動機は何か?
#発話種類:質問
#発話者のセリフ:「この事件で何か考えられる動機はありますか?」。
.....
#発話の意図:疑者の動機は何か?
#発話種類:質問
#発話者のセリフ:「被害者に恨みを持っている人がいるとか、動機につながるようなトラブルはありませんでしたか?」。
.....
#発話の意図:疑者の動機は何か?
#発話種類:質問
#発話者のセリフ:

```

図 10 プロンプト変更箇所（殺人の動機）

```

「あなた以外にもお金を借りていた人はいましたか?」
「あなた以外にもお金を借りていた人はいましたか?」
「あなた以外にもお金を借りていた人はいましたか?」
「あなたから見て、彼らの中にお金を借りていて返さないまま行方不明になっている人はいませんか?」
「あなたもお金を借りていたそうですね。それなのになぜ殺されたのでしょうか?」

```

図 11 生成されたセリフの例（殺人の動機）

最後に、他の登場人物に対してアリバイに関する質問をするセリフを 5 件自動生成させるためのプロンプトの変更点を図 12、生成した結果を図 13 に示す。図 13 中の 5 件のセリフはいずれも指定した状況設定（「#被害者:平川啓一（男、中年、サラリーマン）#容疑者:竹野知実（女、中年、パートタイマー）」）を反映して適切に容疑者名と被害者名を使い分けられているが、5 件中 3 件が同じセリフ、残り 2 件も同じセリフであり、結果として 5 件生成して 2 種類しか台詞が出力されていない結果となっている。この台詞に

関しても妥当ではあるが表現の多様性が低くなっていると言える。

動機とアリバイに関するセリフの多様性が低くなった原因としては、より詳細な状況設定になるほど基盤モデル中で関連性が高いとしてピックアップされる文型が少なくなる可能性などが考えられる。

上記より、ミステリーにおける殺人事件での死体発見のシーンを想定したケーススタディとして、頻出すると考えられる 5 種類のセリフの自動生成を行った結果としては、8 割以上で適切な状況を反映したセリフが出力されることが確認できた。このため、指定した状況説明セリフを基盤モデルに出力させることは技術的には可能であると結論付けられよう。

```

「平川さんと最後に会ったのはいつですか?竹野さん」。
「平川さんと最後に会ったのはいつですか?竹野さん」。
「平川さんと最後に会ったのはいつですか?」。
「平川さんと最後に会ったのはいつですか?竹野さん」。
「平川さんと最後に会ったのはいつですか?」。

```

図 12 プロンプト変更箇所（アリバイ）

```

.....
#発話の意図:容疑者のアリバイがあるか
#発話種類:質問
#発話者のセリフ:「本橋さんは犯行時刻にはどちらで何をされていたか?」。
.....
#発話の意図:容疑者にアリバイがあるか
#発話種類:質問
#発話者のセリフ:「先週火曜日のアリバイはありますか? 仲野さん」。
.....
#発話の意図:容疑者にアリバイがあるか
#発話種類:質問
#発話者のセリフ:

```

図 13 生成されたセリフの例（アリバイ）

ただし、一部のセリフにおいては著しく多様性が低い結果も得られており、基盤モデルによるセリフの生成結果の品質や精度が、例示する発話やその種類、設定した状況やジャンル等に応じてどのように変化するかを確認するためにはより詳細かつデータ数を増やした検証が必要である。

5. 状況反応セリフの生成

状況説明的ではないが特定の状況に登場人物が反応する様子を描写するためセリフの生成が

可能であるかの検証を行った。本節では特定の状況への反応として発話者の感情状態に着目し、プロンプトで指定された状況および感情状態に対する台詞が生成可能かを確認するため、指定した同じ状況で発話者の感情状態のみを変化させてプロンプトを作成し基盤モデルに台詞生成を行わせた(図 14)。

図 14 のプロンプトは、図 2 と同様の形式に感情状態を追加し、殺人事件に遭遇した場合の反応のセリフのうちで「恐怖」と「怒り」に相当する反応のセリフを例示として与えたものである。

#ジャンル:ミステリー
#動機:医療事故
#犯行時刻:昨夜午後 9 時
#凶器:ナイフ、刺殺
#犯行現場:病院の病室、入口が施錠された密室
#被害者:梅本友則 (男、青年、医者)
#容疑者:本橋達弥 (男、老年、患者)
#発話者:杉江拓己(男、中年、刑事)
#発話の意図:恐怖
#発話種類:感情の表明
#発話者のセリフ:「病院で刺殺? 恐ろしい事件だ...」。

#ジャンル:ミステリー
#動機:いじめ
#犯行時刻:先週火曜午前 8 時
#凶器:青酸カリ、毒殺
#犯行現場:学校の教室、窓の鍵が閉まっていないので密室ではない
#被害者:菅野実香(女、老人、教師)
#容疑者:仲野優哉 (男、少年、生徒)
#発話者:小田島菜生(女、中年、探偵)
#発話の意図:怒り
#発話種類:感情の表明
#発話者のセリフ:「教師を青酸カリで毒殺するなんて絶対に許さない」。

#ジャンル:ミステリー
#動機:借金を返せない
#凶器:鉄砲、銃殺
#犯行時刻:昨日午前 11 時
#犯行現場:会社のオフィス、施錠され鍵を紛失したので密室
#被害者:平川啓一 (男、中年、サラリーマン)
#容疑者:竹野知実 (女、中年、パートタイマー)
#発話者:竹野知実(男、少年、探偵)
#発話の意図:悲しみ
#発話種類:感情の表明
#発話者のセリフ:

図 14 生成に用いる Few shot 形式のプロンプト例 (状況への感情的な反応)

本節では、「恐怖」と「怒り」の反応のセリフの例示に対して、それ以外の下記の感情状態 (エ

クマンの基本 6 感情のうちの例示以外) に相当すると解釈可能なセリフが出力されるか確認した。

- ・ 「悲しみ」
- ・ 「驚く」
- ・ 「喜び」
- ・ 「嫌う」

それぞれ、図 14 のプロンプトにおいて最後の発話の意図のみをそれぞれ対応する感情状態に変更して各 5 件のセリフの生成を実施した。「悲しみ」の結果を図 15 に、「驚く」を図 16, 「喜び」を図 17, 「嫌う」を図 18 にそれぞれ示す。

「もう少し早く来ていれば助けられたかもしれないな...」
「お父さんとお母さんと妹たちを殺されて悲しいよ」
「父さんと母さんから受け継いだこの会社を守りたかっただけなのに...」
「犯人は許せない!」
「僕たち二人だけでも犯人を捕まえてやる!」

図 15 生成されたセリフ例 (「悲しみ」の反応)

「私たち二人だけしかいないはずなのに誰かいるみたいない感じがしてならないんです」
「この人たち何なのよ!」
「どうしてこんなことになったのかわからないわ」
「まさかこんなことになるなんて...」
「まさかこんなことになるなんて...」

図 16 生成されたセリフ例 (「驚く」の反応)

「ついにやったぞ!これで借金地獄から抜け出せるんだ!」
「これでやっと解放されるよ!」
「ついにやった!これで俺も一人前の殺し屋だな」
「ついにやったぞ!これで借金も返せるし、あいつらにも仕返しができる」
「ついにやったぞ!これで借金から解放されるんだな」

図 17 生成されたセリフ例 (「喜び」の反応)

「私はあなたと一緒に仕事したくありません」
「あの人と一緒に仕事したくないんですよね」
「あの人と一緒に働くのは嫌なのよね」
「私たち家族からお金を借りておいて返さないなんて最低よ!」
「あんたなんか大嫌いよ!」

図 18 生成されたセリフ例 (「嫌う」の反応)

図 15 から図 18 中で太字斜体字で示したセリフは必ずしも指定した状況に適しているとは考えられないものである。いずれの感情の場合も過半数は適切と判断可能なセリフが生成されているが、5 件中に不適切と考えられるセリフも必ず含まれる結果となっている。

「悲しみ」(図 15)においては殺人事件が起こっているという状況は反映されているものの、太字斜体字のセリフが指し示す発話者の感情はどちらかというと「悲しみ」ではなく「怒り」ではないかと考えられる。「驚く」(図 16)においては、1 件ホラージャンルのようなセリフが生成されてしまっている。「喜び」(図 17)では、殺人の動機である借金(「#動機:借金を返せない」)が反映されたセリフが生成されているが、1 件は発話者ではなく犯人のセリフが生成されてしまっている。また、「嫌う」(図 18)の場合には、殺人事件に対して「嫌う」感情を示すのではなく、特定の相手に対する嫌悪感情の発露のセリフも出力されてしまっている。

これらのエラーと考えられる出力に関しては、プロンプトによる状況の指定を工夫することによって改善できる可能性がある。あるいは実用的に用いる場合には、適切な感情状態の表現になっているかを判定する分類器を組み合わせるなどの対応が有用とも考えられる。

実際の物語生成で用いる場合には精度を向上させる必要はあるが、過半数の生成結果で指定した状況に対応した適切な感情状態を示すセリフが出力されており、原理的に状況に応じた登場人物の反応に相当する表現を制御して出力すること自体は可能であると考えられる。さらに今後登場人物の属性や人間関係などを反映して適切な感情状態を指示できれば、同じ状況に対する個性的な反応も生成可能であると期待される。

6. シーンの表現自動生成に向けて

人工知能による物語の自動生成において、基盤モデルに各シーンを自然な言語表現として出力させるためには、上記のような生成結果のセリフを適切な順番で配置する必要がある。

上述の『金田一少年の事件簿』[13]中の短編作品から死体発見に相当するシーンをランダムに 5 シーン抽出し、セリフの連続における頻出のパターンを再現するように基盤モデルの出力結果を連続的に配置した結果を図 19 に示す。

セリフのみで構成されてはいるが殺人事件における死体発見シーンが理解可能なレベルで再現できていると考えられる。本研究はジャンル及びシーンを限定したケーススタディではあるが、基盤モデルを用いることでミステリーのシーンとして理解可能な表現が生成できる可能性は十分にありうると判断できよう。

なお、本研究では対象としていないが、実際に物語の全体を自動生成する場合には、台詞の時系列のパターンに合わせて、小説であれば情景描写や心理描写などによって構成される地の文、脚本などの場合には柱書きやト書きなどの個々の表現のパターンおよび、それらの複合的な配置のパターンも合わせて出力する必要がある。

「あ……あそこに……血まみれの……人が……」

「拳銃だ、弾も入っているぞ」

「扉や窓などの出入り口はすべて内側からロックされているし、誰かが入ってきたらすぐわかるようになっていているんだよ」

「もう少し早く来ていれば助けられたかもしれない…」

「なんでこんなことになったんだろう…」

図 19 生成されたセリフによるシーン構成例

7. 結論と今後の課題

本稿では従来の物語構造に基づく物語自動生成システムと基盤モデルによる自然言語生成を組み合わせたハイブリッド型の物語自動生成モデルを想定し、その実現可能性を検証するために、特定物語ジャンルの 1 シーンにおける頻出と考えられるセリフを基盤モデルにプロンプトを適用することで生成した。またセリフの生成においては、状況説明セリフと状況反応セリフに分類し、物語の状況および生成させたいセリフに関する情報を Few shot の形式で基盤モデルに提示することで、状況に応じたセリフの生成が可能であることを示した。また生成されたセリフを組み合わせることで物語の 1 シーンを構成するセリフの系列を生成することも可能であることを示した。

ただし、本研究は特定ジャンル、特定シーンを対象としたケーススタディであり、一般的な物語のセリフ生成が基盤モデル及びプロンプトプログラミングによって可能であるかは、さらに複数のジャンル、多様な物語状況の設定、異なったセリフの例示、などを多数検証することで確認する必要がある。異なったシーンにおけるセリフの時系列のパターンに関しても大規模なデータセットを構築し、汎用的な特徴を抽出する必要がある。また、生成されたセリフの精度や適切性に関しても、プロンプトの変更によってどのように制御が可能であるか、また読者はどのように受け取るかを被験者実験などを用いて評価する必要があると考えられる。

また、最終的な物語の自動生成を実施するためには、既存のプロット自動生成アルゴリズムの出力結果に対応したプロンプトを自動的に構築する手法の確立や、本研究では対象としていないセリフ以外の表現部分（地の文や柱書き・ト書きなど）に関しても基盤モデルによる自動生成が可能であるか検証が必要である。

謝辞

本研究は LINE 株式会社 AI カンパニーの支援を受けた。また、NEDO 人と共に進化する次世代人工知能に関する技術開発事業 インタラクティブなストーリー型コンテンツ創作支援基盤の開発、および科学研究費基盤研究 C「物語文と日常会話文における発話文体と話者属性の分析」の支援を受けた。

参考文献

[1] 松原仁, 佐藤理史, 赤石美奈, 角薫, 迎山和司, 中島秀之, 瀬名秀明, 村井源, 大塚裕子, コンピュータに星新一のようなショートショートを創作させる試み, 第 27 回人工知能学会予稿集, 2D1-1, 2013.

[2] ウラジーミル・プロップ (北岡誠司, 福田美智代訳). 昔話の形態学. 水声社, 1987.

[3] 村井源, 既存作品中の物語の基本パターンに基づく物語構造の自動生成, 情報処理学会論文誌, Vol. 63, No. 2, pp. 335-346, 2022.

[4] Bommasani, Rishi and Hudson, Drew A. and Adeli, Ehsan and Altman, Russ and Arora, Simran and von Arx, Sydney and Bernstein, Michael S. and Bohg, Jeannette and Bosselut, Antoine and Brunskill, Emma and Brynjolfsson, Erik and Buch, Shyamal and Card, Dallas and Castellon, Rodrigo and Chatterji, Niladri and Chen, Annie and Creel, Kathleen and Davis, Jared Quincy and Demszky, Dora and Donahue, Chris et. al., "On the Opportunities and Risks of Foundation Models", arXiv, <https://arxiv.org/abs/2108.07258>, 2021.

[5] Brown, Tom and Mann, Benjamin and Ryder, Nick and Subbiah, Melanie and Kaplan, Jared D and Dhariwal, Prafulla and Neelakantan, Arvind and Shyam, Pranav and Sastry, Girish and Askell, Amanda and Agarwal, Sandhini and Herbert-Voss, Ariel and Krueger, Gretchen and Henighan, Tom and Child, Rewon and Ramesh, Aditya and Ziegler, Daniel and Wu, Jeffrey and Winter, Clemens and Hesse, Chris and Chen, Mark and Sigler, Eric and Litwin, Mateusz and Gray, Scott and Chess, Benjamin and Clark, Jack and Berner, Christopher and McCandlish, Sam and Radford, Alec and Sutskever, Ilya and Amodei, Dario, "Language Models are Few-Shot Learners", NEURIPS2020, vol. 33, pp. 1877-1901, 2020.

[6] Radford, Alec and Kim, Jong Wook and Hallacy, Chris and Ramesh, Aditya and Goh, Gabriel and Agarwal, Sandhini and Sastry, Girish and Askell, Amanda and Mishkin, Pamela and Clark, Jack and Krueger, Gretchen and Sutskever, Ilya, "Learning Transferable Visual Models From Natural Language Supervision", Proceedings of the 38th International Conference on Machine Learning, vol. 139, pp. 8748-8763, 2021.

[7] Laria Reynolds, Kyle McDonell, "Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm", Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 314, pp. 1-7, 2021.

[8] 川本稔己, 山崎天, 坂田亘, 佐藤敏紀, "HyperCLOVA を利用したプロンプトプログラミングによるシチュエーションに適した応答生成" 人工知能学会研究会資料, 言語・音声理解と対話処理研究会, vol. 93, pp. 131-136, 2021.

[9] Hong Chen, Raphael Shu, Hiroya Takamura, Hideki Nakayama, GraphPlan: story generation by planning with event graph. Proceedings of the 14th International Conference on Natural Language Generation, pp. 377-386, 2021.

[10] 村井源, 複合的構造を持つ物語の自動生成に向けて, 情報知識学会誌, Vol. 30, No. 2, pp. 214-219, 2020.

[11] Hajime Murai, "Factors of the Detective Story and the Extraction of Plot Patterns Based on Japanese Detective Comics", Journal of the Japanese Association for Digital Humanities, Vol. 5, No. 1, pp. 4-21, 2020.

[12] AINOW, "世界初、日本語特化の大規模言語モデル「Hyper CLOVA」を開発", <https://ainow.ai/2022/01/05/261665/>, 2022/8/17 参照.

[13] さとう ふみや, 天樹 征丸, 金田一少年の事件簿『なぜ暖炉は燃えていたか?』, 講談社, 2016.

[14] Hajime Murai, Shuuhei Toyosawa, Takayuki Shiratori, Takumi Yoshida, Shougo Nakamura, Yuuri Saito, Kazuki Ishikawa, Sakura Nemoto, Junya Iwasaki, Shoki Ohta, Arisa Ohba, Takaki Fukumoto, Extraction and Automatic Generation of Characters' Attributes in Contemporary Japanese Entertainment Works, Conference Abstracts of Digital Humanities 2022, EP-01-111, pp. 661-663, 2022.

[15] 村井源, 松本齊子, 物語テキストにおける会話文の意図の話者属性による特徴, 情報処理学会論文誌, Vol. 60, No. 2, pp.135-143, 2020.