

# キャラクターシナリオ自動生成に向けた新規情報判定の検討

## Exploring Novel Information Detection for Character-Based Scenario Generation

高橋 聖弥\*  
Seiya Takahashi

岩田 伸治†  
Shinji Iwata

伊原 滉也\*,†  
Koya Ihara

加藤 昇平\*  
Shohei Kato

### 概要

アニメやゲームなどのキャラクターコンテンツでは、キャラクターの魅力を引き出すために、既存のシナリオや設定資料には記載されない新たな情報を含む追加シナリオが制作されることがある。本研究ではこうした追加シナリオの自動生成の創作支援を目的とする。ストーリー生成の枠組みでは、生成・批評・改善を繰り返す手法が有力であり、新たな情報の表出を目的とする場合には、生成の過程で新規情報の有無を適切に評価することが重要となる。しかし、大規模なシナリオを直接繰り返し処理することはトークンコストの課題がある。そこで本研究では、テキスト中に出現する設定情報を効率的に抽出し、新規情報検出に活用する手法を検討した。

### 1 はじめに

大規模言語モデル (LLM) の応用は多岐にわたり、キャラクター同士が対話するシナリオを自動で生成する方向にも活用されている [1, 2, 3, 4]。既存のキャラクターを用いたシナリオ生成の研究でも、キャラクターの既存設定に忠実に従いつつ新しいシナリオを LLM で生成する方法が着目されてきた [5, 6]。しかし、アニメやゲームなどのキャラクターコンテンツでは、キャラクターの魅力を引き出すために既存のシナリオや設定資料にはないキャラクターに関する新規情報 (以下、新規情報) を含む新しいシナリオを作成することがある。例えば、既存のシナリオにおいてかき氷に関する描写がなかったキャラクターが、新たに追加されたシナリオで「いちご味のかき氷を好む」と描写された場合、この「いちご味のかき氷を好む」が新規情報に該当する。

そこで、本研究では新規情報を表出するシナリオを LLM で自動生成することを最終的な目標とした。LLM を用いたシナリオ生成の枠組みでは、生成・批評・改善を繰り返し、シナリオを改善する手法が有力である [1, 4]。生成や改善においては、キャラクターの情報や新規情報や改善内容等に従い LLM がシナリオを出力する。一方、批評時には、生成・改善されたシナリオに新規情報が含まれているかを膨大な既存の全シナ

リオとキャラクターの設定情報を基に適切に繰り返し評価することが重要となる。しかし、大規模なシナリオを直接繰り返し処理することは、トークンコストの観点から難しい。そこで本研究では、既存のシナリオに出現するキャラクターの情報を抽出して、LLM の入力に用いることで少ないトークンで効率よく新規情報を判定する手法を検討した。

構築した情報抽出に基づく新規情報判定手法の性能を検証するために、評価実験を実施した。具体的には、スマートフォンゲーム『IDOLY PRIDE』\*1中のシナリオを用い、シナリオに含まれるキャラクターの新規情報を判定できるかを情報抽出の有無で条件分けて検証した。

結果として、シナリオ全文を入力に用いた時よりも情報抽出をした場合の方が入力トークン数を削減できただけでなく、新規情報判定の精度も向上した。このことから、シナリオから情報抽出を実施し新規情報を判定する本研究で構築した方法の有効性がトークン効率・判定精度の両面から示唆された。

### 2 関連研究

#### 2.1 LLM を用いたシナリオ生成

シナリオ生成に関連する研究として、Bae らの研究 [1] と Wen ら [4] の研究を挙げる。Bae ら [1] は、批評と改善を繰り返すことで生成するシナリオの一貫性と創造性を向上させた。具体的には、特定のペルソナを持つ三人の批評エージェントがシナリオを批評し、リーダーエージェントがその中から批評を選択する。選ばれた批評に基づいてシナリオを改善し、改善後のシナリオに対して再び評価する手法である。Wen ら [4] は、生成されたエピソードの不明確な部分を補完するプロセスを繰り返すことで、より複雑かつ創造的なシナリオ生成を試みた。具体的には、生成シナリオに不足している背景情報を指摘し、その理由を問い直すことで、内容の充実を図っている。このようにシナリオ生成の枠組みでは生成・改善を繰り返す手法が有力である。

既存のキャラクターを用い、作品の作風を維持したまま、既存作品の続編や新作を生成する取り組みも存在する。既存の作品の新規シナリオ生成の研究として、村井らの研究 [5] と Han らの研究 [6] を挙げる。村井ら [5] は、手塚治虫の『ブ

\* 名古屋工業大学

† サイバーエージェント

\*1 <https://idolypride.jp/>

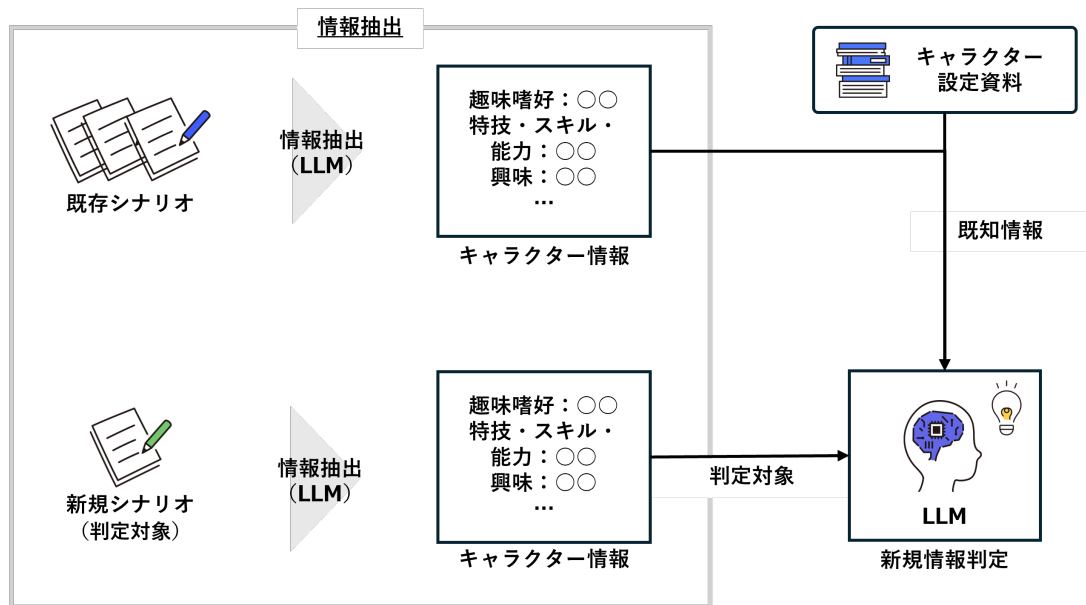


図 1: 本研究で構築した抽出情報に基づく新規情報判定手法の概要図。既存のシナリオと判定対象となる新規シナリオから抽出されたキャラクター情報と、元々のキャラクター情報を入力し、新規情報の有無を判定する。

『ブラック・ジャック』を対象に、既存エピソードの物語構造を分析し、その特徴をもとに新たな物語構造を自動生成する手法を提案した。生成された物語構造は専門家による評価で『ブラック・ジャック』らしさを備えていると判断された。Hanら [6] は、文体・キャラクター・プロット・表現力を段階的に学習するカリキュラム学習方式と、各物語要素に特化したモジュールを組み合わせた WriterAgent モデルを開発した。評価実験では、『ハリー・ポッター』や『紅樓夢』などの多言語作品を題材に、プロットの一貫性や文体の再現性、キャラクター行動の適切さなど、多面的な指標で高い性能を示した。

いずれの研究も作風の模倣やキャラクター再現に焦点を当てているが、本研究ではキャラクターの新規情報を表出する生成に焦点を当てる。

## 2.2 キャラクター情報に関する判定や抽出

対話履歴やキャラクター情報を入力として、特定の指標に関する判定や情報抽出の試みが報告されてきた。例えば、キャラクターをロールプレイするチャットボットの自動評価 [7]、キャラクター間の感情的な関係性の自動抽出 [8]、キャラクターが重大な決断をする点の判定 [9]、小説からのキャラクター情報のプロファイリング [10] がある。また、キャラクターの更新といった時系列情報を加味して、意図しない性格の検出 [11] も実施されてきた。プロフィール情報の抽出タスクはこれまでも研究されてきたが、シナリオの新規情報に関しては着目されていない。そこで、本研究では、ゲーム等のキャラクターコンテンツにおいて、既存のシナリオに加えて新規のシナリオが追加された状況を想定しシナリオの新規情報の有無を判定する。

## 3 抽出情報に基づく新規情報判定

本研究では新たに作成・生成されたシナリオがキャラクターに関する新規情報を含むかどうかを、少ないトークンで効率よく判定する手法を構築した。具体的にはキャラクターの設定情報に加え、既存シナリオと新規シナリオから抽出したキャラクター情報を LLM に入力し、新規情報の有無を判定する (図 1)。大規模なシナリオを直接 LLM に入力するとトークンコストの観点から課題が残るため、シナリオを LLM に入力して抽出されたキャラクターに関する小規模な情報を用いた。次節以降ではキャラクター情報の抽出方法と新規情報判定方法について詳細に述べる。

### 3.1 シナリオからのキャラクター情報抽出

シナリオを LLM に入力し、キャラクターに関する情報を抽出する。抽出対象は以下に示す基本的なプロフィール項目である。初期検討段階として、文面からの判定が容易と考えられる以下のプロフィール項目に限定した。

- 趣味嗜好：食べ物や飲み物、趣味、好きなこと
- 特技・スキル・能力：キャラクターが持つ特技やスキル、能力
- 興味：関心を持っていること
- コンプレックス・弱点：コンプレックスや苦手なこと、弱点

性格等の文面から高度な推論が必要と考えられる要素は本研究の対象外とした。プロンプトは、Yuan らのプロフィール抽出プロンプト [10] を参考に設計した。各項目は箇条書き形式で詳細に要約するよう指示した。

### 3.2 新規情報判定

既存シナリオから抽出されたキャラクター情報および基本プロフィール（以下、既知情報）と新規シナリオから得られた情報を比較し、既存情報に含まれていない内容が新たに記述されている場合、それを「新規情報」として判定する。具体的には、既知情報と判定対象となる新規シナリオから抽出したキャラクター情報を LLM に比較させることで新規情報の有無を判定する。プロンプト設計では、まず既知情報を提示し、その後新規シナリオから抽出したキャラクター情報を与えることで参照すべき情報を明示し、新規情報の有無を判定させる構成とした。

## 4 実験

3 章で述べた抽出情報に基づく新規情報の有無を判定する手法の有効性を評価した。以下に実験設定について述べる。

■**キャラクター** スマートフォンゲーム『IDOLY PRIDE』のキャラクターである『成宮すず』を用いた。

■**使用データ** ゲーム中で成宮すずが登場した 10 個のシナリオを用いた。これらは全て新規情報が存在することが複数の人手で確認されたシナリオである。

■**実験条件** 情報抽出の有効性を評価するため、既存シナリオ及び判定対象となる新規シナリオの抽出有無に関する 4 条件を設定した。

■**評価方法** Leave-one-out-cross-validation を参考に以下の手順でモデルを評価した。

- 10 個のシナリオから 1 個を除外シナリオ、9 個を既存シナリオとする。
- 除外シナリオから新規情報を含むシナリオを新しく作成する。
- 既存シナリオから 1 個選択し、そのシナリオから新規情報を含まないシナリオを新しく作成する。
2. で作成したシナリオを新規情報を含むシナリオ、3. で作成したシナリオを新規情報が含まないシナリオ、残りの 9 個のシナリオを既存シナリオとしてモデルが判定する。
1. から 3. をすべてのシナリオが一回ずつ新規シナリオと新規情報を含まない判定対象シナリオになるように繰り返す。

上記の 2. における新規情報を含むシナリオ、3. における新規情報を含まないシナリオは、シナリオから抽出したキャラクター情報を基に LLM に生成させたシナリオを用いた。LLM の出力は 1 意に定まらないことから上記の手順を 10 試行繰り返し、評価値の平均値を計算した。

■**評価指標** Accuracy, Precision, Recall, F1 を用いた。モデルの新規情報有無の判定結果を予測値、事前にラベル付した新規情報の有無を正解とした。

■**LLM** 抽出・判定時の全てにおいて OpenAI 社の gpt-4.1 を用いた。

表 1: 新規情報の有無に関する評価結果 (10 試行平均)

既存抽出	新規抽出	Accuracy	Precision	Recall	F1
		0.51	0.51	0.55	0.52
✓		0.56	0.64	0.28	0.39
	✓	0.53	0.52	<b>0.69</b>	0.59
✓	✓	<b>0.65</b>	<b>0.66</b>	0.63	<b>0.64</b>

既存抽出: 既存シナリオに対する抽出

新規抽出: 判定対象の新規シナリオに対する抽出

## 5 結果と考察

表 1 に新規情報の有無の判定に関する評価結果を示す。シナリオの全文をそのまま使用する場合や、既存及び判定対象の新規シナリオのどちらかを抽出する場合に比べ、既存と新規シナリオの両方から情報を抽出する場合が Recall 以外の評価指標において優れた性能を示した。この結果から、新規情報を判定する場合には、プロンプトの入力としてシナリオ全文を用いるよりもシナリオから抽出したキャラクター情報を LLM に与えることが有効であると示唆された。

既存シナリオに対する情報抽出のみを実施する場合や、既存シナリオ・新規シナリオの両方に対して抽出を実施する場合と比較して、新規シナリオに対する情報抽出のみを実施する場合は Precision が低いことが確認された。LLM が出力した判定理由から、既知情報中に含まれている内容を新規情報と誤って LLM が判断しているケースが確認された。また、ほかの 3 条件と比較し、既存シナリオに対する情報抽出のみを実施する場合は Recall が低いことが確認された。LLM が出力した判定理由から、既知情報に明記されていない内容が明記されていると LLM が判断していたことが確認された。これらの結果から、既知情報の形式と判定対象の形式が異なる場合、出力傾向に偏りが生じる可能性が示唆された。偏りの理由として、既知情報の形式と判定対象の形式が異なる場合、LLM が既知情報を適切に参照できないことが考えられる。

## 6 まとめと今後の展望

本稿では、ゲーム等のキャラクターコンテンツにおいて想定される、設定資料や既存シナリオには含まれていない新規情報を追加するシナリオの自動生成にむけて、新規情報の有無の判定に取り組んだ。新規情報の有無を LLM で判定する場合に、膨大な既存シナリオ全文を入力として用いる必要がありトークンコストの課題が考えられた。そのため、既存のシナリオに出現するキャラクターの情報を抽出して、LLM の入力に用いることで少ないトークンで効率よく新規情報を判定する手法を構築した。さらに、そのモデルの性能を評価する実験を実施した。

実験の結果、シナリオ全文を用いる場合と比較し、キャラクター情報を抽出するほうが入力トークン数を削減できただけ

でなく、新規情報判定の精度も向上した。そのため、本研究で構築した手法はシナリオ全文を用いる方法と比較し、トークンコストを抑えつつ、新規情報の判定精度も向上させる可能性が示唆された。

今後は、シナリオから抽出したキャラクター情報の正確性について人手による評価を実施する。また、今回はキャラクター情報の一部項目に限定して実験を実施したが、今後は性格や人間関係などの情報も対象に含め、キャラクタープロフィールを網羅した新規情報判定の実現を目指す。

## 7 謝辞

本研究において使用した『IDOLY PRIDE』のデータは、株式会社 QualiArts から提供されたものである。この場を借りて謝意を表す。

本研究は、一部、文部科学省科学研究費補助金（課題番号 JP24H00741）、ならびに、国立研究開発法人情報通信研究機構委託研究の助成により行われた。

## 参考文献

- [1] Bae, M. and Kim, H.: Collective Critics for Creative Story Generation, *arXiv preprint arXiv:2410.02428* (2024).
- [2] Wu, W., Wu, H., Jiang, L., Liu, X., Hong, J., Zhao, H. and Zhang, M.: From role-play to drama-interaction: An LLM solution, *arXiv preprint arXiv:2405.14231* (2024).
- [3] Huot, F., Amplayo, R. K., Palomaki, J., Jakobovits, A. S., Clark, E. and Lapata, M.: Agents' Room: Narrative Generation through Multi-step Collaboration, *arXiv preprint arXiv:2410.02603* (2024).
- [4] Wen, Z., Tian, Z., Wu, W., Yang, Y., Shi, Y., Huang, Z. and Li, D.: Grove: a retrieval-augmented complex story generation framework with a forest of evidence, *arXiv preprint arXiv:2310.05388* (2023).
- [5] 村井源, 青山美月, 大田翔貴, 大場有紗, 福元隆希, 奥山凌伍, 金刺智哉, 富田真生, 入船真誠, 坂本珠凜他: 物語の展開パターンの結合の特徴に基づく構造の自動生成-『ブラック・ジャック』 新作に向けて, じんもんこん 2023 論文集, Vol. 2023, pp. 11-18 (2023).
- [6] Han, X., Liu, Y., Li, M., Liu, W., Hu, S., Yan, R., Xu, Z. and Chen, X.: Pastiche Novel Generation Creating: Fan Fiction You Love in Your Favorite Author's Style, *arXiv preprint arXiv:2502.15616* (2025).
- [7] 木本晴久, 人見雄太, 佐藤大地, 跡部優吾, 小山正彦, 片田智大, 橋本圭, ジューストー沙羅, 守屋貴行: Character-LLM の評価指標の検討, 人工知能学会全国大会論文集, Vol. JSAI2024, pp. 4Xin2109-4Xin2109 (2024).
- [8] Kim, E. and Klinger, R.: Frowning Frodo, Wincing Leia, and a Seriously Great Friendship: Learning to Classify Emotional Relationships of Fictional Characters, in Burstein, J., Doran, C. and Solorio, T. eds., *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 647-653, Minneapolis, Minnesota (2019), Association for Computational Linguistics.
- [9] Tikhonov, A.: Branching Narratives: Character Decision Points Detection, in Madge, C., Chamberlain, J., Fort, K., Kruschwitz, U. and Lukin, S. eds., *Proceedings of the 10th Workshop on Games and Natural Language Processing @ LREC-COLING 2024*, pp. 70-75, Torino, Italia (2024), ELRA and ICCL.
- [10] Yuan, X., Yuan, S., Cui, Y., Lin, T., Wang, X., Xu, R., Chen, J. and Yang, D.: Evaluating Character Understanding of Large Language Models via Character Profiling from Fictional Works, in Al-Onaizan, Y., Bansal, M. and Chen, Y.-N. eds., *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 8015-8036, Miami, Florida, USA (2024), Association for Computational Linguistics.
- [11] 初野広夢, 鷹野孝典: 対話型 AI キャラクター更新学習における意図しない性格変化の検知手法, 人工知能学会研究会資料 知識ベースシステム研究会 130 回 (2023/11), pp. 55-61 一般社団法人 人工知能学会 (2023).