

LLM を用いた人物相関図の自動生成に関する基礎検討

吉田 真紘^{†, a} 藤川 雄翔^{†, b} 松下 光範^{†, c}

[†] 関西大学大学院総合情報学研究科 ^{††} 関西大学総合情報学部

a) k146344@kansai-u.ac.jp b) k920357@kansai-u.ac.jp c) m_mat@kansai-u.ac.jp

概要 本研究の目的は物語テキストから相関図を自動生成することである。登場人物間の関係性はセリフ文や地の文の行動描写などから間接的に示されることが多い。こうした暗黙的な関係性の抽出には文脈を踏まえた解釈が求められるため、キーワードや共起関係を手がかりとする従来のルールベース手法では抽出が困難であった。そこで本稿では文脈を考慮した解釈が可能な大規模言語モデルを用いて関係性を抽出し、それをもとに相関図を生成する手法を提案する。検証の結果、実験対象とした 4 作品において一定の精度で生成可能であることが確認された。

キーワード 登場人物間の関係性、相関図、大規模言語モデル

1 はじめに

近年、生成 AI (以下、AI と記す) 技術の急速な発展により、物語の創作を支援する多様なツールが登場している。例えば、プロットの構想、キャラクターのセリフ生成、物語の世界観の設定、文章の複数生成など、創作の各工程に AI を活用する試みが進んでいる [3, 7, 8]。実際に、創作アイデアの壁打ち相手となる AI ツール「ComicCopilot¹」や、自己創作物の詳細な閲覧データや読者の属性データを確認できる AI 分析ツール「マンガノアナリティクス²」などが登場している。こうした物語生成支援にとどまらず、AI を仮想的な編集者として機能させることで、創作物を客観的に評価・批評させたり、改善案を提示させたりする活用も期待される。例えば、読者をあえてミスリードする記述が意図通りに機能しているか、あるいは作者が意図せず誤解を招く箇所がないかを、第三者の視点から確認可能になる。今後、AI を用いた創作支援がより高度化すれば、プロからアマチュアまで幅広い創作者が、仮想的な編集者を通じて自らの作品を客観的に検証・改善できるようになると考えられる。こうした高度な創作支援技術の実現には、物語の構造やキャラクターの設定といった情報を AI が参照・処理可能なデータとして表現する必要がある。すなわち、物語の構造を記述した「設計図」や「中間生成物」としてのデータ形式が求められる。

物語の創作において、キャラクター間の関係性は物語の魅力や物語展開を左右する重要な構成要素である。物語はキャラクターが行動・会話することによって生じる変化の過程を描いたものであり [5]、新たなキャラクターの登場による関係の構築や、既存の関係性の変化といったキャラクター間の相互作用によって進行していく [2, 11]。キャラクター間の関係性は主人公を中心にネットワークを形成し [4, 9]、物語の展開や意味付けに深みを与える要素である。

キャラクター間の関係性を創作者が確認する手法の 1 つに相関図がある。相関図は、ノードにキャラクターの情報、エッジに関係性の情報を格納したグラフ表現であり、物語創作に必要なキャラクターの設定を統合的に管理できる。しかし、従来の相関図は AI による創作支援に十分活用されていない。これは、関係性が物語内で明示されることが少なく、セリフ文や地の文の行動描写といった断片的な情報から文脈を踏まえた解釈が求められるためである。こうした解釈は、従来のルールベース手法による計算機処理では困難であり、結果として相関図は人間が手作業で作成したデータ構造を持たない静的な画像形式で運用されることが一般的であった。そのため、AI を用いた創作支援に直接反映させることが困難となっている。

そこで本研究では、物語テキストから物語の構造やキャラクターの設定を確認可能な相関図を自動生成する手法の実現を目指す。生成した相関図を編集者の視点として活用することで、創作者が自身の創作物を客観的に確認できるツールとなることが期待される。本稿ではその端緒として、キャラクター間の関係性を対象に、物語テキストから大規模言語モデル (以下、LLM と記す) を用いて抽出し、相関図を生成する手法を提案する。

そこで本研究では、物語テキストから物語の構造やキャラクターの設定を確認可能な相関図を自動生成する手法の実現を目指す。生成した相関図を編集者の視点として活用することで、創作者が自身の創作物を客観的に確認できるツールとなることが期待される。本稿ではその端緒として、キャラクター間の関係性を対象に、物語テキストから大規模言語モデル (以下、LLM と記す) を用いて抽出し、相関図を生成する手法を提案する。

2 デザイン指針

本研究が目指す相関図の自動生成手法は、AI による創作支援へ活用可能にするため、以下の 2 つの条件を設計指針として定義する。

1 つ目に、物語テキストのみからキャラクター間の関係性を抽出可能にすることである。1 章で述べたように、物

Copyright is held by the author(s).

The article has been published without reviewing.

¹<https://lp.comic-copilot.ai/> (2026/06/15 確認)

²<https://manga-no.com/analytics> (2026/06/15 確認)

語内でキャラクタ間の関係性が「親子関係」や「恋人関係」などと直接記述されることは少なく、多くの場合、登場人物の言動を通じて間接的に示される。このような関係性を読み取るには、表面的な記述の解析に留まらず、文脈から読み取れる暗黙的な関係性を抽出可能な手法が必要である。

物語におけるキャラクタの関係性は、セリフの語調、人称代名詞の選択、敬語の有無、地の文による行動描写・心理描写などを通じて間接的に示される。こうした断片的な情報から読者が関係性を読み取れるのは、特定の言動パターンが特定の関係性に対応する“記号”として、制作者と読者の間の共通理解になっているためである。大塚は、作品に散りばめられた断片的な情報を手がかりに読者はキャラクタの設定を能動的に読み取ると論じた [13]。東は、キャラクタの記号の属性の組み合わせを読者は直接読み取る「データベース消費」へと移行したと指摘した [15]。これらの議論が示すように、キャラクタの言動を関係性の記号として解釈する能力は、コンテンツ受容において広く定着している。例えば、呼び捨てや愛称の使用は親密さの指標として機能し、敬語は社会的距離や上下関係の指標として機能する [16]。こうした記号性は、制作側においても意図的に活用されている。金水は、キャラクタの人物像 (e.g., 年齢, 性別, 社会的立場) に対応した典型的な言語表現を「役割語」として概念化し、制作者は役割語を用いることで読者がキャラクタの属性や立場を即座に把握できる記号として設計すると指摘した [17]。この設計原理は言語表現だけでなくキャラクタの言動にも働いており、師弟関係では教えを乞う・指導する言動、敵対関係では相手を妨害したり傷つけたりする攻撃的な言動がそれぞれ見られる。したがって、こうした記号性は制作者が意図的に設計し、読者がそれを読み取れる形で共有されている。

しかし、これらの記号性は、個々の言動が単独で意味を形成するのではなく、物語で繰り返される文脈の積み重ねによって解釈可能になる。例えば、ある場面での「冷たい態度」が敵対関係を示すのか、親密さの裏返しかは、他の場面における言動と照らし合わせて判断される。つまり、関係性の解釈には文脈を踏まえることが必要となるため、キーワードの出現頻度や共起関係を手がかりとするルールベース手法や統計的手法では、こうした関係性を抽出することが難しい [10]。したがって、物語テキストの文脈からキャラクタの言動に含まれる記号を解釈可能な手段が求められる。

そこで本稿では、LLM を用いてキャラクタ間の関係性の抽出を試みる。LLM は大規模なテキストコーパスを用いた事前学習により自然言語の理解・生成能力を獲得したモデルであり、文脈全体を考慮した解釈が可能で

ある [12]。実際に先行研究では、人物の経歴を間接的な表現や文脈から示した暗黙的なテキストに対して、LLM は経歴に関する情報を一定の精度で抽出可能であることが示されており [1]、物語テキスト中に含まれるキャラクタの言動から関係性を推論することが期待できる。以上を踏まえ、本稿では物語テキストのみを入力とし、LLM を用いてキャラクタ間の関係性を抽出し、そのデータをもとに相関図を自動生成する手法を構築する。

2 つ目に、計算機が参照可能なデータ形式で出力可能にすることである。抽出したキャラクタ間の関係性を創作支援に活用するには、ノードにキャラクタの情報、エッジに関係性の種類や方向性を付与したグラフ構造として出力し、後続の処理で参照可能にする必要がある。そこで本稿では、データの構造化に JSON 形式を採用した。抽出結果を JSON 形式で保持することで、キャラクタや関係性の情報をプロンプトに記述することなく、構造化されたデータとして参照可能になる。

3 提案手法

3.1 相関図自動生成手法

前章のデザイン指針に基づき、作成した提案システムの概観を図 1 に示す。提案システムは、画面左上部黄枠にテキスト入力領域、左下部赤枠に相関図表示領域、右部緑枠に関係性表示領域の 3 領域から構成される。テキスト入力領域では、物語テキストの直接入力やテキストファイルのアップロードにより物語テキストを入力する。相関図表示領域では、キャラクタをノード、キャラクタ間の関係性をエッジとして表示する。このとき、いずれかのノードを選択することで対象キャラクタに接続するエッジのみをハイライト表示する機能を設け、登場するキャラクタ数が多い場合でも特定キャラクタの関係性を容易に把握できるようにした。関係性表示領域では、LLM が抽出した全ての関係性を一覧表示する。このとき、相関図の表示領域におけるいずれかのノードを選択すると、そのキャラクタに関連する関係性のみがフィルタリングされて表示される。

提案手法では以下の手順で相関図の自動生成を行う。

1. ユーザが物語テキストを入力する
2. LLM が物語テキストから登場人物間の関係性を抽出する
3. 抽出した登場人物と関係性を構造化データとして整形する
4. 手順 4 で得られたデータをもとに相関図を出力する

手順 1 のテキスト入力において、ユーザが入力した物語テキストは、解析実行ボタンを押すことによりバックエ

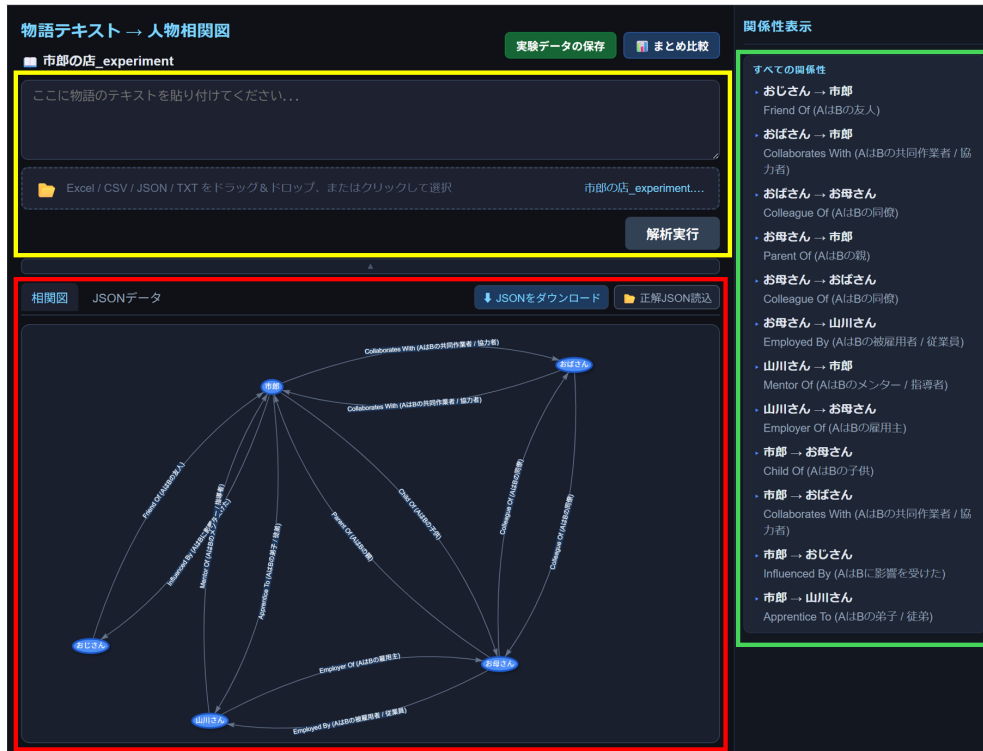


図 1: 提案システムの概観。提案システムは、テキスト入力領域（画像左上部黄枠）、相関図表示領域（画像左下赤枠）、関係性表示領域（画像右部緑枠）の 3 領域からなる。

ンドの LLM へ渡される。手順 2 では、LLM に対してフロントエンドから渡された物語テキストとシステムプロンプトから登場キャラクターの一覧と関係性を抽出する。手順 3 では、LLM によって抽出されたキャラクターおよび関係性を構造化データとして整形する。手順 4 では、手順 3 で得られた JSON 形式のデータをもとに相関図を生成し、最終的にユーザへ提示する。

本手法において、キャラクター間の関係性を抽出する際に用いたシステムプロンプトを図 2 に示す。本稿では、キャラクター間の相互の繋がりを把握可能にするため、関係性を有向グラフとして表現する。関係性の判断根拠については、Matsumoto らの「人間関係は会話内容に現れる」という指摘 [6] に基づき、セリフや対話といった直接的な相互作用を主たる根拠として採用する。ただし、キャラクター間に明示的な会話が存在しない場合であっても、地の文の行動描写や状況描写を通じて両者の間に関係性が示唆される場合がある。そこで、会話内容を主たる判断根拠としつつも、文脈から関係性が認められる場合も抽出対象に含めることとした。加えて、物語に登場するキャラクターは必ず他キャラクターと何らかの関係を持つと考えられる。そのため、LLM には各キャラクターについて少なくとも 1 つ以上の関係性を抽出するよう指示した。LLM が抽出する関係性の定義については、推論された関係性を計算機が読み取れる形式で表現するため、

本稿ではキャラクター間の関係性を端的に分類・記述可能な語彙群である「RELATIONSHIP³」の関係性ラベルを採用した。LLM はこれらの関係性ラベル以外の自由記述を禁止し、必ず定義済みの関係性ラベルの中から最も適当なものを選択して出力するよう制約を設けた。ただし、「Participant（参加者）」、「Participant In（～への参加者）」および「Relationship（関係性）」は人物間の具体的な繋がりを示すものではないため除外した。結果として、32 種類の関係性ラベルを推論対象として採用した。各関係性ラベルを表 1 に示す。

本稿では抽出されたデータを JSON 形式に整形する。以下に、本稿で定義した JSON の出力スキーマを示す。

```
{
  "characters": [
    { "id": "キャラクターを識別する id",
      "name": "キャラクターの名称" }
  ],
  "relationships": [
    { "from": "関係性の起点となるキャラクターの id",
      "to": "関係性の終点となるキャラクターの id",
      "label": "関係性ラベル" }
  ]
}
```

出力 JSON は、キャラクターの情報である `characters` 配列と、関係性の情報である `relationships` 配列から構

³<https://vocab.org/relationship/> (2026/06/15 確認)

あなたは物語テキストを解析する専門家です。以下の【絶対厳守の制約】に従い、与えられた物語テキストから登場人物の情報と相関関係を抽出してください。

【絶対厳守の制約】“relationships”（関係性）について：物語テキストに登場する全ての登場人物を抽出すること。登場人物間の関係性は、会話内容（セリフ・対話・やり取り）を根拠として判断すること。ただし、会話内容が無くても関係性が存在する場合は、それを含めること。すべてのエッジは有向（from → to）で表現すること。双方向の関係は必ず 2 エントリに分割すること。双方向に分割する場合、両エントリの label は同じである必要はない。それぞれの from 側の人物の視点に基づき、最も適した label を個別に選択すること。一方向の関係は 1 エントリのみ。label は必ず以下の【関係性定義リスト】の「英語（日本語）」形式のラベルのみを使用すること（リスト外の自由記述は絶対禁止）。characters に含まれるすべての登場人物は、必ず 1 つ以上の relationship エントリ（from または to のいずれか）に登場すること。

図 2: キャラクタ間の関係性を抽出するために用いたプロンプト例

成される。characters 配列の各要素は、キャラクタを一意に識別する ID とそのキャラクタに対応する名称を示す name のペアから構成され、ID は物語の登場順に c1, c2, c3, ... と付与される。relationships 配列の各要素は、from・to・label の 3 属性を持つ有向エッジとして表現され、それぞれ関係の起点となる人物の ID、終点となる人物の ID、および関係性ラベルを格納する。

3.2 実装

本稿で提案した相関図自動生成手法は、ブラウザで操作・閲覧が可能な Web システムとして実装した。本システムのフロントエンドは Nuxt (Ver.3.11.2)⁴および Vue.js (Ver.3.5.34)⁵を用いて実装した。相関図の描画は Vis.js に含まれるライブラリ vis-network (Ver.10.0.3) を用いて有向グラフを表示可能にした。バックエンドは Nuxt のサーバーエンジンである Nitro (Ver.2.13.4) を用いて TypeScript で実装した。LLM には Google Gemini API (モデル: gemini-3.5-flash) を使用し、フロントエンドから受け取った物語テキストをプロンプトとして構築して API リクエストを送信する。Gemini API から取得したキャラクタおよび関係性の抽出結果は、Nitro サーバーを介してフロントエンドへ送信され、相関図として描画される。このとき、出力の安定性を確保するため、LLM のランダム性を制御するパラメータである Temperature を 0.0, 乱数シード値を 42 に設定した。

表 1: 本研究で対象とする関係性ラベル一覧（「RELATIONSHIP」より引用）。本表はキャラクタ A から見たキャラクタ B への関係を示す。

関係性ラベル	意味 (日本語訳)
Acquaintance Of	A は B の知り合い
Ambivalent Of	A は B に複雑な感情を抱く相手
Ancestor Of	A は B の祖先
Antagonist Of	A は B の対立相手
Apprentice To	A は B の弟子 / 徒弟
Child Of	A は B の子供
Close Friend Of	A は B の親友
Collaborates With	A は B の共同作業 / 協力者
Colleague Of	A は B の同僚
Descendant Of	A は B の子孫
Employed By	A は B の被雇用者 / 従業員
Employer Of	A は B の雇用主
Enemy Of	A は B の敵
Engaged To	A は B の婚約者
Friend Of	A は B の友人
Grandchild Of	A は B の孫
Grandparent Of	A は B の祖父母
Has Met	A は B と会ったことがある
Influenced By	A は B に影響を受けた
Knows By Reputation	A は B を評判で知っている
Knows In Passing	A は B を少し知っている
Knows Of	A は B の存在を知っている
Life Partner Of	A は B のライフパートナー
Lives With	A は B の同居人
Lost Contact With	A は B と音信不通 / 連絡が途絶えた
Mentor Of	A は B のメンター / 指導者
Neighbor Of	A は B の隣人 / 近所の人
Parent Of	A は B の親
Sibling Of	A は B の兄弟姉妹
Spouse Of	A は B の配偶者
Works With	A は B と一緒に働く人
Would Like To Know	A は B と知り合いたい

4 実験

本手法を用いることで (1) 物語に登場するキャラクタ間の関係性を適切に抽出できるか、(2) 抽出された関係性の種類が適切に選択されているか、の 2 つの観点から実験を行い、提案手法の妥当性を検証する。なお、本実験では客観的な正解データの作成が困難であったため、筆者が手作業で作成した関係性データを正解として、提案手法の出力と比較することにより検証する。

4.1 対象データ

本実験で使用する物語テキストとして、一般投稿作品に類する 2 作品、商業作品 2 作品の計 4 作品を対象とする。本研究は創作した物語が読者にどう捉えられるかを提示する、編集者の代替となる AI ツールの提案である。そのため、対象ユーザは物語創作を楽しみたい一般

⁴<https://nuxt.com/> (2026/06/15 確認)

⁵<https://ja.vuejs.org/> (2026/06/15 確認)

ユーザからプロのクリエイターまで幅広いユーザが想定される。そのため、入力される物語テキストの完成度についても、Web 小説投稿サイト「小説家になろう⁶」に投稿されるような一般投稿レベルの作品から、プロが執筆した作品まで幅広いレベルを考慮する必要がある。以上を踏まえ、本実験では編集者による校閲が行われない一般投稿作品として筆者と AI の共作作品 2 作品と、実際に出版された商業作品として青空文庫の掲載作品 2 作品を対象とした。

筆者と AI の共作作品については、データセットの客観性を担保するため、AI が物語のベースとなる文章を生成する方針を採用した。筆者による介入は、物語の一貫性を保つための最小限の構造調整や微細な表現の修正にとどめた。物語の生成には Gemini（モデル：gemini-3.5-flash）を使用した。その際、生成したテキストがモデルの学習に利用され提案手法の評価に影響を与えることを防ぐため、学習利用をオフに設定した状態で使用した。以下、物語生成に使用したプロンプトを示す。

あなたは、Web 小説を書く作家です。Web 小説投稿サイト「小説家になろう」に投稿されているような物語を 1 作品作成してください。登場キャラクターは 5 名以上とし、物語の文字数は 3200 から 4800 文字の範囲で作成してください。また、作品のタイトル案も合わせて考えてください。

作成する物語の文字数は、「小説家になろう」における 1 話当たりの文字数の目安に合わせて設定した。一般的に、Web 小説の 1 話分は読者が隙間時間を利用して 10 分程度で読了できる分量に設定されることが多いとされる。そのため、一般的な読書速度である分速 400 文字を基準に、個人の読書速度の差を考慮して 3200 文字から 4800 文字の範囲に設定した。さらに、登場するキャラクター数が 5 名以上の作品に限定した。これは、登場キャラクターが少ない場合抽出される関係性の数も限られ、提案手法の抽出能力を十分に評価できないためである。以上の手順を経て、作品 A と作品 B の 2 作品を作成した。

青空文庫の掲載作品も同様に、文字数 3200 文字から 4800 文字の範囲かつ登場キャラクター数が 5 名以上の作品を対象とした。この条件を満たす作品として、「市郎の店」（豊島与志雄 著）[18] と「山さち川さち」（沖野岩三郎 著）[14] を選定した。本実験では、これら計 4 作品を用いて、提案手法の妥当性を検証する。

4.2 評価

本稿では、提案手法による相関図生成の妥当性を検証するため、評価実験を行った。評価は、関係性があるキャラクター同士を適切に結び付けられているかを評価するエッジレベルと、関係性ラベルが適切に選択されているかを評価する意味レベルの 2 つの観点から行う。評価に使用する正解データは、客観的なデータの生成が困難であるため、筆者が事前に 4 作品全てを通読した上で人手により作成した。正解データの作成方法については、今後の課題検討とする。評価指標について、エッジレベルでは適合率と再現率を採用し、意味レベルではエッジレベルで抽出された関係性に対して正誤判定を行うことから、適合率のみを採用した。各評価値は、各作品において抽出された全関係性の正誤をもとに算出した。

表 2: 作品ごとの評価結果。列見出しの括弧内は、E：エッジレベル、M：意味レベルを示す。値の括弧内は、（正解数/関係性の総数）を示す。

作品名	(E) 適合率	(E) 再現率	(M) 適合率
作品 A	1.000 (20/20)	0.833 (20/24)	0.700 (14/20)
作品 B	1.000 (18/18)	0.900 (18/20)	0.778 (14/18)
市郎の店	1.000 (12/12)	0.750 (12/16)	0.750 (9/12)
山さち川さち	1.000 (20/20)	0.800 (20/25)	0.800 (16/20)

るかを評価する意味レベルの 2 つの観点から行う。評価に使用する正解データは、客観的なデータの生成が困難であるため、筆者が事前に 4 作品全てを通読した上で人手により作成した。正解データの作成方法については、今後の課題検討とする。評価指標について、エッジレベルでは適合率と再現率を採用し、意味レベルではエッジレベルで抽出された関係性に対して正誤判定を行うことから、適合率のみを採用した。各評価値は、各作品において抽出された全関係性の正誤をもとに算出した。

表 2 に、評価結果を示す。エッジレベルの適合率は、全ての作品で 1.000 となり、提案手法によって結び付けられたキャラクター間の関係性は全て正解データと一致していた。一方で、再現率は 0.750～0.900 の範囲に留まり、正解データに含まれる一部の関係性は結び付けられなかった。さらに、意味レベルの適合率も 0.700～0.800 の範囲に留まり、LLM が選択した関係性ラベルに一部誤りが見られた。このことから、関係性ラベルの選択方法には改善の余地があることが示された。

以上の結果から、作品間で精度に差が観察されるものの、物語テキストから一定の精度で相関図を自動生成できる可能性が示唆された。

5 考察

5.1 本手法で適切に結びつけられた関係性

エッジレベルの適合率が 4 作品全てで 1.000 を示したことから、本手法が結び付けた全ての関係性は正解データと一致した。その理由として、3.1 節で設定したプロンプトの設計により、物語テキスト中で会話のやり取りが明示的に描写されているキャラクター間の関係性を主に結び付けていると考えられる。さらに、物語テキストでは「翌日、市郎は山川さんのところに、報告かたがた相談に行きました。」（市郎の店）や、「アレンが泥を払って立ち上がると、バルガスはふっと口角を上げ、手元の木剣を放り投げた。」（作品 B）のように、セリフ文の前後に地の文でキャラクター名が明示的に記述されているケースがある。そのため、LLM がキャラクターを正確に識別しやすく、誤ったキャラクター間で関係性が結び付くことがなかったと考えられる。

⁶<https://syosetu.com/>（2026/06/15 確認）

5.2 本手法で結び付けられなかった関係性

エッジレベルの再現率は 0.750~0.900 の範囲であることから、本手法では正解データに含まれる一部の有向関係は結び付けられなかった。結び付けられなかった関係性には、キャラクタ間の直接的な会話のやり取りがなく地の文によって関係性が示されているものや、特定のキャラクタが会話中に第三者のキャラクタについて言及することで関係性が示されているものである傾向があった。これらは間接的な描写による関係性であり、LLM がそのような関係性を読み取ることが困難であった可能性がある。例えば、「市郎の店」では、おばさんから山川さんへの「Knows By Reputation (A は B を評判で知っている)」という関係性ラベルが正解データに含まれている。これは「山川さんは、大人たちの間では変り者だ」という評判でしたが、子供たちにとっては、いっしょによく遊んでくれるおもしろい人でした。」という記述に基づくものであり、市郎のお母さんの代わりに店を預かるおばさんが、山川さんのことを評判を通じて知っていると解釈できる。しかし、この記述では誰が山川さんのことを評判で知っているか明示的されておらず、山川さんと他のキャラクタの関係性を直接結び付けることが困難である。そのため、本手法ではこうした関係性を結び付けることができなかったと考えられる。作品 A では、

(中略)「俺がなんとかしてやる。とりあえず、このお守りを持っておけ。一時的に霊の干渉を弾く結界符だ」
トウヤが懐から取り出したお札を渡すと、カナの背中に張り付いていた黒い霧がジュッと音を立てて消滅し、彼女の顔に少しだけ血の気が戻った。「ありがとうございます……！本当に、ありがとうございます！」
深々と頭を下げるカナを見送りながら、トウヤはコートのポケットに手を突っ込み、険しい表情を浮かべた。
「トウヤよ。あの呪いの気配、どうやら最近帝都を騒がせている「連続魂狩り」の仕業らしいぞ」
店の奥から、ゆったりとした足取りで現れたのは、白装束に身を包み、長い白髭を蓄えた老人、ゲンゾウだった。

とあり、トウヤがカナを見送った後、店の奥からゲンゾウが現れ、カナにかけられた呪いについてトウヤと会話している。この場面では、ゲンゾウとカナの間に直接的なやり取りはないものの、トウヤとの会話からゲンゾウがカナの存在を認識していることが示唆される。そのため、正解データではゲンゾウからカナへの「Knows Of (A は B の存在を知っている)」という関係性ラベルが正解データに含まれている。しかし、ゲンゾウの発言はトウヤに向けられており、カナは話題の対象として言及されているだけであるため、本手法ではこうした間接的な描写から関係性を結び付けることが困難であったと考えられる。

選択した関係性の種類の推論について考察する。意味レベルの適合率は 0.700~0.800 の範囲であることから、

本手法で結び付けた関係性の一部でラベルの選択に誤りが見られた。誤って選択された関係性の傾向として、大きく 3 つのパターンが観察された。

1 つ目に、定義の差異が小さい関係性ラベル間での誤選択が挙げられる。例えば、「Acquaintance Of (A は B の知り合い)」は、「A が B をわずか、あるいは表面的な知識以上に知っているが友人には至らない関係」を示すのに対し、「Knows In Passing (A は B を少し知っている)」は「A が B についてわずかな、または表面的な知識しか持っていない関係」を示す。両ラベルの定義の差異は微妙であり、関係性の解釈が読者の主観に依存する面もあることから、LLM が適切なラベルを選択することが困難であった可能性がある。

2 つ目に、LLM が地の文やセリフ文をもとに関係性を過大に解釈し、実際とは異なる関係性ラベルを選択するケースが見られた。例えば、「市郎の店」では、お母さんと山川さんの関係性について、本文中に「市郎のお母さんが、まだ独身でいる山川さんの着物の仕立直しなどを頼まれていましたので、市郎もよく山川さんを知っていました。」とあり、両者は仕立て直しの依頼を通じた「Acquaintance Of (A は B の知り合い)」であると推測できる。しかし、本手法では「Employed By (A は B の被雇用者)」および「Employer Of (A は B の雇用主)」という雇用関係を選択し、地の文の記述を過大に解釈したと考えられる。また、作品 A では、ソウマとコハク、トウヤのそれぞれの関係性について、

(中略)「ソウマか。お前もこの事件を追ってたのか。エリート様がこんな埃っぽい夜更けの神社に何の用だ？」
「決まっている。協会を舐め腐った呪詛師を、私の手で完全に浄化するためだ。貴様のような半端者が出しゃばる幕はない。大人しくそこで震えている」

(中略)

「ふん。相変わらず堅苦しい小僧じゃ。お前の術は綺麗じゃが、遊びがなさすぎる。あれでは「本物の悪意」には出し抜かれるぞ」
コハクがアクビをしながらソウマを挑発する。ソウマは青筋を立てながらも、コハクの底知れぬ妖気に僅かに後ずさった。「化け物め……。貴様らと馴れ合うつもりはない。足手まといになる前に消えろ」
ソウマはそれだけ言い残し、単身で廃神社の奥へと進んでいった。「さて、俺たちも行くか。エリート様がヘマをして死なれちゃ、夢見が悪いからな」
トウヤは微かに口角を上げ、鳥居の奥へと踏み出した。

とあり、ソウマが両者を挑発し対立する様子が描写されているものの、互いに危害を加えようとする意図は見られない。そのため正解データでは、ソウマからコハク、トウヤに対して、「A が B に反対し対立する関係」を示す「Antagonist Of (A は B の対立相手)」と定義している。しかし、本手法では「A が B に対して憎しみを抱き危害を加える意図を持つ関係」を示す「Enemy Of

(A は B の敵)」を選択しており、キャラクターの言動を LLM が必要以上に解釈した結果であると考えられる。

3 つ目に、一方が他方の意味を内包する定義を持つ関係性ラベルが混在し、誤りではないものの最も適切な関係性が選択されなかったケースが確認された。例えば、「山さち川さち」では、本文中に「今年六つになる信次といふ与兵衛の孫は、その子猿の頭を撫でながら泣きました。」とあり、与兵衛と信次の関係性は「Grandparent Of (A は B の祖父母)」および「Grandchild Of (A は B の孫)」であると推測できる。しかし、本手法では「Ancestor Of (A は B の祖先)」および「Descendant Of (A は B の子孫)」を選択した。祖父母と孫の関係は、祖先と子孫の関係を満たすため誤りではないが、3.1 節で設計したプロンプトでは最も適切な関係性ラベルを選択するよう指示しているため、より具体的な関係性を示す「Grandparent Of (A は B の祖父母)」および「Grandchild Of (A は B の孫)」が正解となる。このように、意味的に包含関係にある関係性ラベルが存在する場合、LLM がより抽象的な関係性を選択してしまう可能性があり、今後の検討が必要である。

6 おわりに

本研究では、物語テキストを入力として、登場キャラクターの相関図を自動生成する手法を提案した。提案手法では、キャラクターの言動に内包する記号性に着目し、文脈を考慮した解釈が可能な LLM を用いて関係性の抽出を行った。抽出された関係性は JSON 形式で整形し、AI が必要な情報を随時参照可能な設計図として保存可能な構造とした。得られた JSON 形式のデータをもとに、相関図を描画する web アプリケーションとして実装した。実験では、キャラクター間の関係性を適切に抽出可能か、抽出された関係性の種類が適切に選択可能か、の 2 つの観点から提案手法の妥当性を検証した。その結果、今回対象とした作品において、作品間で精度に差が観察されるものの、物語テキストから一定の精度で相関図を自動生成できる可能性が示唆された。

今後の展望として、選択対象とする関係性ラベルの見直しと、プロンプトの改善が挙げられる。本システムでは、意味的な差異が小さい関係性ラベルや、一方が他方の意味を内包する関係性ラベルが混在しており、LLM が適切な関係性ラベルを選択することが困難なケースが見られた。したがって、既存のラベル定義のうちどの関係性を採用し、どの関係性を統合・削除するかについて、選定基準を含めた検討が必要である。

謝辞

本研究の実施にあたり、JSPS 科研費 24K15255 の支援を受けた。記して謝意を表す。

参考文献

- [1] Alessandra, S., Andrea, S., Valentina, P., van Erp, M., Francesco, S. and Fabio, V.: Explicit vs. Implicit Biographies: Evaluating and Adapting LLM Information Extraction on Wikidata-Derived Texts, *arXiv preprint arXiv:2509.14943* (2025).
- [2] Bruner, J.: *Acts of meaning: Four lectures on mind and culture*, Vol. 3, Harvard university press (1990).
- [3] Chen, J., Zhu, X., Yang, C., Shi, C., Xi, Y., Zhang, Y., Wang, J., Pu, J., Feng, T., Yang, Y. et al.: Hollmwood: Unleashing the creativity of large language models in screenwriting via role playing, *Findings of the Association for Computational Linguistics: EMNLP 2024*, pp. 8075–8121 (2024).
- [4] Fludernik, M.: *An introduction to narratology*, Routledge (2009).
- [5] Jahan, L., Mittal, R. and Finlayson, M.: Inducing Stereotypical Character Roles from Plot Structure, *Proc. 2021 Conference on Empirical Methods in Natural Language Processing* (Moens, M.-F., Huang, X., Specia, L. and Yih, S. W.-t. (eds.)), Online and Punta Cana, Dominican Republic, Association for Computational Linguistics, pp. 492–497 (2021).
- [6] Matsumoto, K., Minato, J., Ren, F. and Kuroiwa, S.: Estimating human emotions using wording and sentence patterns, *2005 IEEE International Conference on Information Acquisition*, pp. 421–426 (2005). DOI: 10.1109/ICIA.2005.1635125.
- [7] Mirowski, P., Mathewson, K. W., Pittman, J. and Evans, R.: Co-writing screenplays and theatre scripts with language models: Evaluation by industry professionals, *Proc. 2023 CHI conference on human factors in computing systems*, pp. 1–34 (2023).
- [8] Reza, M., Laundry, N. M., Musabirov, I., Dushniku, P., Yu, Z. Y. M., Mittal, K., Grossman, T., Liut, M., Kuzminykh, A. and Williams, J. J.: Abscribe: Rapid exploration & organization of multiple writing variations in human-AI co-writing tasks using large language models, *Proc. 2024 CHI Conference on Human Factors in Computing Systems*, pp. 1–18 (2024).
- [9] Sugishita, K. and Masuda, N.: Social network analysis of manga: similarities to real-world social networks and trends over decades, *Applied Network Science*, Vol. 8, No. 79 (2023). DOI: 10.1007/s41109-023-00604-0.
- [10] Vincent, L. and Xavier, B.: Extraction and analysis of fictional character networks: A survey, *ACM Computing Surveys*, Vol. 52, No. 5, pp. 1–40 (2019).
- [11] Vogler, C.: *The writer's journey : mythic structure for writers*, Film Art, Inc. (2022).
- [12] Xin, Z. W., Kun, Z. et al.: A survey of large language models, *arXiv preprint arXiv:2303.18223* (2023).
- [13] 大塚英志: 定本物語消費論, 12174, 角川書店 (2001).
- [14] 沖野岩三郎: 山さち川さち, 金の星社 (1923).
- [15] 東浩紀: 動物化するポストモダン オタクから見た日本社会, 講談社 (2001).
- [16] 小井出慎, 古宮嘉那子, 佐々木稔, 新納浩幸: 物語における登場人物の親しさ推定, 言語処理学会 第 22 回年次大会 発表論文集 (2016).
- [17] 金水敏: ヴァーチャル日本語役割語の謎, 岩波書店 (2003).
- [18] 豊島与志雄: 市郎の店, 桜井書店 (1948).