

語り直しを目的とした大規模言語モデルを用いた Story Intention Graph の作成とその評価

吉川祐輔¹ 井上壮志¹ 東中竜一郎²

¹ インターメディアプランニング株式会社 ² 名古屋大学大学院情報学研究科
{y_yoshikawa,t_inoue}@ipi.co.jp higashinaka@i.nagoya-u.ac.jp

概要

人間同士の対話において、お互いの価値観や経験は物語によって表出されることが多い。対話システムにおいても、ユーザが話す物語を適切に理解することで、ユーザの価値観や経験を深く理解し、親密さを持ったやり取りを実現できるようになるとともに、より高度なやり取りも実現できるようになると考えられる。本論文では、大規模言語モデルを用い、物語文を Story Intention Graph (SIG) に変換する手法について述べる。SIG は、物語中の登場人物、および、その行動・意図・動機に焦点を当て、それらの関係性をグラフとして表したものである。今回、大規模言語モデルとして GPT-4 を用い、所定の物語文からの SIG の自動作成に取り組んだ。また、作成された SIG の妥当性を評価するとともに、SIG を用いた物語の語り直しを行い、その結果を確認した。

1 はじめに

人間同士の対話において、お互いの価値観や経験は物語によって表出されることが多い [1]。対話システムにおいても、ユーザが話す物語を適切に理解することで、ユーザの価値観や経験を深く理解し、親密さを持ったやり取りを実現できるようになるとともに、より高度なやり取りも実現できるようになると考えられる。人間同士の対話においても、narrative は多く見られる [2]。

これまで物語の理解には様々な手法が提案されてきた。例えば、イベントを検出し、それらの連鎖として物語を理解する方法が提案されている [3]。Lehnert は、プロットユニットと呼ばれる表現を提案した [4]。プロットユニットでは、話者の感情やイベントの因果関係が表される。Goyal らはプロットユニットの自動抽出を行うシステム AESOP を開発した [5]。本研究が着目する Story Intention Graph

(SIG) は Elson によって提案された物語の表現である [6]。SIG では、物語において重要とされる、登場人物、行動、意図・動機に焦点を当て、これらをグラフ構造として表す。本研究では、物語の要素を網羅的に扱うことができると考えられる SIG に着目し、物語文から SIG を自動作成することを試みる。この生成には大規模言語モデルを利用する。

物語の構造が得られれば、様々な応用が可能である。一つは物語の要約であるが、その他の用途として、語り直しが挙げられる [7]。語り直しとは、同じ物語に対し、内容は失われないように物語の方向性や印象を変化させた物語を出力させることを指す。たとえば、自身の経験などを踏まえたうえで、再構成して語ることも含まれる。語り直しは、相手の物語の深い理解を表出することに有用であり、聞き役となるエージェントやカウンセリングを行うエージェント [8] など、相手の発話を理解する対話システムにとって重要である。そこで、本研究では、SIG の自動生成に加えて、SIG から語り直しの生成についても取り組む。

2 Story Intention Graph

本研究が用いる物語の構造である Story Intention Graph (SIG) について述べる。SIG は、ノード、フレーム、エッジからなる有向グラフである。ノードは命題を保持し、ノードのタイプとして以下が存在する。

TE ノード 物語文の文章 (Text) の断片に対応するノードである。

P ノード TE ノードに対応する、命題を表すノード。Elson では、ノードの内容は Predicate-argument structure によって表現されるが、本研究では、大規模言語モデルの処理が容易と考えられることから、TE ノードの文章そのまま表現する。

I ノード P ノードを解釈して得られる要素. 物語中において文章として現れない, 意図や状況を表現する.

B ノード, G ノード, A ノード 登場人物の信念 (Belief), 目的 (Goal), 感情 (Affect) を表現する. これらのノードのうち B, G ノードには, 複数の I ノードがネストして含まれる. そのため, フレームともみなすことができ, 以降, この二種類のノードを B フレーム, G フレームと呼ぶ.

SIG は三層構造になっており, TE ノードが属するタイムライン層, P ノードが属する命題層, I ノード, B, G フレーム, および, A ノードが属する解釈層である.

これらのノードは, エッジで接続される. エッジの種類は多数あり, たとえば, ある P ノードが接続先の P ノードに物語の時系列的に続くことを示す「f (followed by) エッジ」, P ノードと I ノードの間の等価性が解釈されることを示す「ia (interpreted as) エッジ」, ある要素が別の要素の実現に必要な前提条件であることを示す「pf (precondition for) エッジ」, ポジティブな影響を与えることを示す「p (provides for) エッジ」, ネガティブな影響を与える「d (damages) エッジ」などがある. さらなる詳細については文献 [6] を参照されたい.

SIG は, 物語をノード (フレーム) およびエッジでグラフ化することで, 物語の登場人物の意図・感情・行動・動機および時系列についての構造を把握することを可能とする.

3 SIG の自動生成

本研究では, 大規模言語モデルを用いて, 物語文から SIG を自動生成する. 具体的には, 物語文が与えられたときの, SIG の自動生成のステップは以下の通りである.

1. TE ノード, P ノード, f エッジの生成: 元の物語文を句点など区切り文字を基準に TE ノードに分割したのち, 各文の断片について, 文章それぞれを命題化して P ノードを生成する. その後, TE ノードと P ノードを結ぶ f エッジを生成する.
2. エージェントの特定: 物語文そのものを解析し, その物語のエージェント (行動, 信念, 目的, 感情などの主体, つまり物語の登場人物な

私の趣味の一つとして、絵手紙で日記を書くというものがあります。

例えばここにある写真に写っている建物は、札幌にある時計台です。以前に家族と北海道へ行った時に、感動したので購入しました。

冬の夜になると電飾がいろいろな色に光ってとても綺麗なですよ。

毎日の日記に絵手紙を使うと安くない金額がかかりますけど、日記を書く時に一番苦労するのは書き続けることだと思います。いろいろな場所の絵手紙を使うことで、楽しみながら書けるのは利点ですね。

絵手紙で日記を書いていることを周りの友人に伝えると、旅行先で購入した絵手紙を友人から頂けるんです。

自分の知らない土地の風景や場所に、日記として書くというのは新鮮ですし、ペンも進むからお勧めですよ。

後は季節によって使い分けるのも良いですね。例えば冬なら北海道や東北、夏なら沖縄や九州。秋には紅葉が綺麗なところや、果物が美味しいところ。春には桜が綺麗に咲くところ、といったように使い分けるのも面白いですね。

図 1: 入力となる物語文の例

ど) となりえる名詞を抽出する. これは後の B フレーム, G フレーム, A ノードおよびエッジの解析に利用する.

3. I ノード, ia エッジの生成: 生成した P ノードを解釈し, 1 対 1 対応する I ノードと, 両者を結ぶ ia エッジを生成する.
4. B フレーム, G フレーム, 付随エッジの生成: 解釈により生成した要素が登場人物の信念・思考 (B フレーム), または目的・意図 (G フレーム) に該当する場合, それらの要素をグループ化してフレームを生成, また, 必要に応じて付随するエッジを生成する. フレーム生成の手順は一定回数繰り返し, これによりネストしたフレームの生成を行う.
5. エッジの追加生成: ここまで生成した各要素を分析し, 必要があればそのうち 2 者の関係性を示すエッジを必要に応じて追加生成する.
6. A ノード, p および d エッジの生成: ここまで生成した要素を分析し, 登場人物の感情を示す

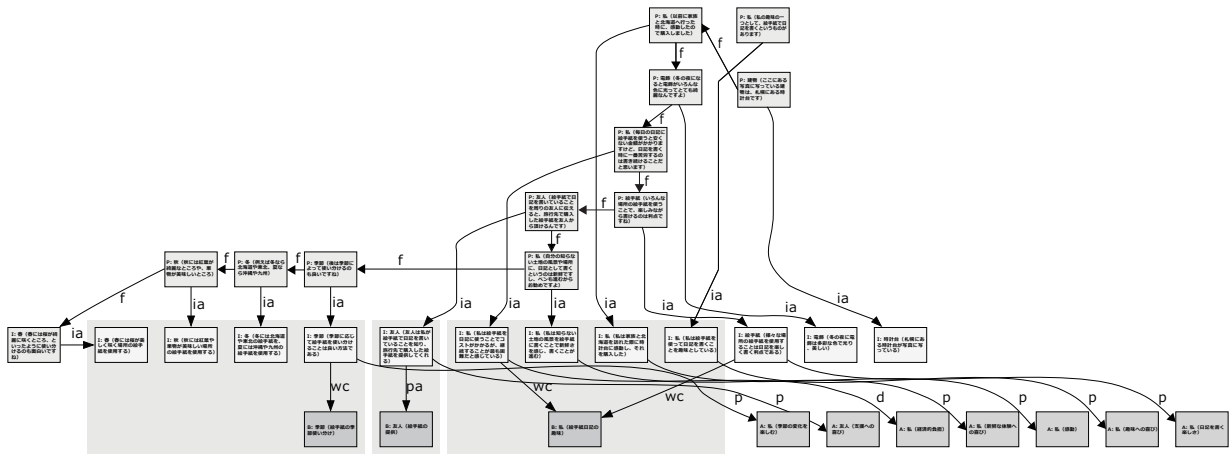


図 2: 図 1 の物語文を入力として自動生成された Story Intention Graph

A ノードが必要であれば生成し、その原因となったノードまたはフレームと A ノードを接続する p エッジまたは d エッジを生成する。

各手順の実施については、対応するプロンプトを用いて大規模言語モデルの出力を得ることで行う。各プロンプトでは、ノード、エッジ、フレームの定義を与え、入力の記事に対して、所定の出力を行うためのインストラクションを用いる。大規模言語モデルとしては、複雑な処理を行うことから、instruction-tune された、GPT-4 などを用いることを想定している。

3.1 SIG の自動生成結果

GPT-4 を用いて、所定の物語について SIG を自動生成した。なお、GPT-4 は英語のプロンプトを用いる方が妥当な結果が得られたため、プロンプトには英語を用いた。

具体的には、クラウドソーシングを用いて、高齢者の語りを模した物語文を 67 個収集した。高齢者の語りを対象とした理由は、後述する共想法への適用を念頭に置いたためである。そして、これらについて前節の手順に従い、SIG を自動生成させた。そして、SIG を著者らで目視で評価し、問題がないかどうかを検証した。ここで、検証には、SIG の定義や規則に反していないかどうかという観点について実施した。その結果、67 個中、45 個 (67%) について、妥当な SIG が生成されていたことを確認した。

図 1 は入力とした物語文の例である。この入力に対して得られた SIG が図 2 である。要素が多いため、ノードの内容が読み取りづらいかもしれないが、適切なノードが生成されているとともに、妥当なエッジでそれらが紐づけられている。例えば、知

らない土地の風景について日記を書くことについて新鮮さを感じるという解釈が得られていたり、日記を書くことが楽しさに寄与しているといった内容が含まれていたりする。

4 SIG に基づく語り直し

前節で得られた SIG のデータを用い、P ノードで保持される原文の内容に、解釈層で得られる、主体の意図・感情・行動・動機の要素の情報を加え、口語体で出力するプロンプトを与えて「語り直し」の実行を試行した。

具体的には、元の物語文と語り直し文を比較することにより、SIG が物語の内容を保持し構造を把握しているか、その構造を踏まえ、プロンプトに沿って適切に語り直せているかを評価した。妥当とされた SIG からサンプルされた 5 個の SIG について語り直しの評価を著者らが目視で行ったところ、三つの語り直しに対して妥当な情報追加を確認した。二つについては、元の物語文と同様の内容のみが生成されており、適切な追加情報が含まれていなかった。

図 3 は、図 2 に対して得られた語り直しの結果である。元の物語文と見比べると、語り直しの結果には、意図・感情・行動・動機などが含まれていることが分かる。例えば、季節の変わり目が楽しみであることは本文中には含まれていないが、その解釈を含んだ上での語り直しとなっている。ここでは、I ノードなどの解釈層の要素を含めることで語り直しを行ったが、SIG の構造は様々な形の語り直しを実現できると考えている。例えば、タイムラインを自由に移動しながら説明したり、それぞれの登場人物の主観から物語を再構成したり、質問応答について、感情を伴いながら回答することが可能になると

私の趣味の一つに、絵手紙を使って日記を書くというものがあるんです。
 実は、この写真に写っているのは札幌の時計台。家族と北海道へ行った時に感動して、思わず購入しちゃったんですよ。
 冬の夜になると、この時計台の電飾がいろんな色に光って、本当に綺麗なんです。
 絵手紙で日記を書くと、毎日のことだから結構お金がかかるんですけど、それ以上に日記を書き続けることの方が大変かなって思っています。
 でも、いろんな場所の絵手紙を使うことで、書くのが楽しくなるんですよ。
 友人にこの趣味を話したら、旅行先で買った絵手紙をくれることもあって、それがまた嬉しいんです。
 知らない土地の風景を日記にするのは新鮮で、ペンが進むんですよ。
 季節によって使い分けるのもポイントですね。
 冬は北海道や東北の絵手紙、夏は沖縄や九州。
 秋には紅葉が綺麗な場所や、果物が美味しいところの絵手紙を選んで、春には桜が綺麗に咲く場所の絵手紙を使うんです。
 季節の変わり目が楽しみの一つになっています。
 絵手紙で日記を書くことは、私にとってはとても喜びを感じる趣味なんです。
 感動したことや新鮮な体験を記録することで、日々の生活に彩りを加えてくれるんですよ。
 経済的な負担はありますが、友人からの支援や季節の変化を楽しむことで、その負担もずいぶん和らぎます。みなさんも、もし興味があれば、ぜひ試してみてくださいね。

図 3: 語り直しの結果の例

考えられる。SIG を活用した語り直し方法の模索は今後の課題としたい。

5 おわりに

本論文では、大規模言語モデルを用い、物語文を Story Intention Graph (SIG) に変換する手法について述べた。GPT-4 を用い、67 個の物語文について、SIG の自動作成を行った。これらについて、SIG の定義に合わせて妥当性を評価するとともに、SIG を用いた物語の語り直しを行った。その結果、SIG 生成手法の精度は 67%であった。また、語り直しについては適切な出力が得られつつも、適切な情報追加がなされないなどの問題が確認できた。

今後の課題として、SIG 生成の精度改善が挙げられる。作成された SIG を精査すると、I ノードが適切な解釈を含んでいないケースが散見された。プロンプトの改善やショットの追加等を通して、より適切と考えられる SIG を作成できるようにしたい。また、語り直しについても様々な観点からの語り直し

が考えられる。対話システムにおける利用を鑑みた時に、どのような語り直しがより効果的かを検討したい。加えて、実際のアプリケーション構築を通して、本取り組みの効果も検証していきたい。

我々が一つの適用例として想定しているものが「共想法」[9]である。共想法は、極度に低下すると認知症の症状を示すとされる記憶体験・注意分割・計画力の三つの認知機能を自然に活用する会話となるよう支援する手法である。複数人の参加者が決められたテーマに沿った写真等の画像データを持ち寄り、一人が話し手として画像データに関連する物語を話し、聞き手である残りの参加者が物語に関連する質問を行う。その後、話し手と聞き手が役割を入れ替えて進行する。話し手の対話システムによる代替はすでに研究が進んでいる[10]。「話し手の物語を理解する」や「語り直し」を行う技術ができれば、共想法の輪の中に対話システムが参加でき、役に立つのではないかと考えている。

参考文献

- [1] Roger C Schank. **Tell me a story: A new look at real and artificial memory**. Charles Scribner's Sons, 1990.
- [2] Yuya Chiba and Ryuichiro Higashinaka. Analyzing variations of everyday japanese conversations based on semantic labels of functional expressions. **ACM Trans. Asian Low-Resour. Lang. Inf. Process.**, Vol. 22, No. 2, 2023.
- [3] Jessica Ouyang and Kathleen McKeown. Modeling reportable events as turning points in narrative. In **Proc. EMNLP**, pp. 2149–2158, 2015.
- [4] Wendy G Lehnert. Plot units and narrative summarization. **Cognitive science**, Vol. 5, No. 4, pp. 293–331, 1981.
- [5] Amit Goyal, Ellen Riloff, and Hal Daumé III. Automatically producing plot unit representations for narrative text. In **Proc. EMNLP**, pp. 77–86, 2010.
- [6] David K. Elson. **Modeling Narrative Discourse**. PhD thesis, Columbia University, 2012.
- [7] Stephanie M. Lukin, Lena I. Reed, and Marilyn A. Walker. Generating sentence planning variations for story telling. **arXiv preprint arXiv:1708.08580**, 2017.
- [8] David DeVault, Ron Artstein, Grace Benn, Teresa Dey, Ed Fast, Alesia Gainer, Kallirroi Georgila, Jon Gratch, Arno Hartholt, Margaux Lhommet, et al. Simsensei kiosk: A virtual human interviewer for healthcare decision support. In **Proc. AAMAS**, pp. 1061–1068, 2014.
- [9] 大武美保子. 認知症予防回復支援サービスの開発と忘却の科学—共想法により社会的交流の場を生成する会話支援サービス—. 人工知能学会論文誌, pp. 568–575, 2009.
- [10] Seiki Tokunaga, Kazuhiro Tamura, and Mihoko Otake-Matsuura. A dialogue-based system with photo and storytelling for older adults: Toward daily cognitive training. **Frontiers in Robotics and AI**, Vol. 8, p. 644964, 2021.