

ストーリー型コンテンツ制作支援 AI フルコト

五木宏 1) 多和田紘希 2) 石綿陽一 3) 松原仁 4) 鈴木恵二 1)

1) 公立ほこだて未来大学 2) 東京大学 3) 株式会社 Ales 4) 東京大学

概要 近年、デジタル・コンテンツの配信形態や経路の多様化により、ストーリーを持ったデジタル・コンテンツ、すなわち映画やテレビドラマ、シリーズものドラマなどへの需要は大きく増大している。しかしながら、増大する需要に対して、現状としては供給が追いつかない状況である。そこで、ストーリー型コンテンツの制作工程のうち、脚本制作作業に着目し、これを支援する AI システム「フルコト」を提案する。フルコトでは、映画制作で多く用いられる方法を実装している。また、新規性や意外性を持った脚本を制作できるように、大規模事前学習モデルを利用して、人が思いつきにくい文章を生成することを実現している。

キーワード：Screen Play, Transformer, Movie

Story-based Content Creation Support AI FURUKOTO

Hiroshi Itsuki 1) Hiroki Tawada 2) Yoichi Ishiwata 3) Hitoshi Matsubara 4) Keiji Suzuki 1)

1) Chikodate University of the Future 2) The University of Tokyo 3) Ales Corporation 4) The University of Tokyo

Recently, a demand for digital content that has a storyline, i.e. Films, tv dramas and drama series, has increased significantly due to a development of content distribution system and channels. However, a supply of content cannot keep up with the increasing demand. We propose an ai system, Furukoto, to support screenplay writing work of story-based content. Furukoto implements methods often used in film production. In order to produce scripts with novelty and unexpectedness, the system uses a large-scale pre-learning model to generate sentences that are difficult for people to come up with.

1. はじめに

近年、インターネットの高速化やサーバの計算能力の向上により、デジタル・コンテンツの配信形態や配信経路が多様化している。すなわち、動画配信サービス、サブスクリプション形式の見放題サービス、ケーブルテレビを通じての世界規模でのコンテンツ・シンジケーションなどの伸展である。

それにともない、コンテンツへの需要が増大しており、制作効率の向上が課題となっている。特に、ストーリーをもったコンテンツにおいては、需要に対する供給の逼迫が顕著である。これは、ストーリーをどのように展開するかがコンテンツの品質に大きく関与している一方で、コンテンツ制作者が限られた時間のなかで、限られた経験から、多様なストーリー展開を案出す

ることが難しいことに起因していると考えられる。

そこで、本研究では、自然言語処理 AI を利用した、脚本生成支援システム「フルコト」を開発した。以下、第2章では、関連する研究および事例について、第3章ではフルコトで用いられている映画制作技法、それらの実装方法、および、フルコトの操作方法について、第4章ではフルコトを使用して制作された作品事例について述べる。最後の第5章ではフルコトの機能や事例を基に、フルコトについての評価、および、その意義について述べる。

2. 関連研究・関連事例

2.1. AI を利用した脚本生成

脚本の自動生成やプロットの制御に関する研

究は、以前から世界的に行われており、多くの論文や事例がある。ここでは、その中から、映画やテレビドラマの脚本、および、プロットの生成についての近年の研究や事例に絞ることとした。すなわち、Alyce と命名された脚本 (Screenplay) 生成 AI [1] や、ストーリーを構成するイベント間の因果関係に着目した試み [2]、「物語の構造」に基づいたプロットの自動生成システム [3] である。

Alyce では、本稿のフルコトと同様に感情値を利用している点においては共通性がある。しかしながら、フルコトが構造的に構築されたフレームワークにより物語全体での整合性を創出しようとするシステムであるのに対して、Alyce では局所的な整合性の創出にとどまっている。これは、Alyce が文を生成する際に利用している PPLM (Plug & Play Language Model) と呼ばれる、与えられたタイトル (トピック) を拡張することにより文を生成する仕組みを、物語全体ではなく、部分的にのみ適用していることに由来すると考えられる。フルコトでは、予め全体の構造となる文を設定し、そのすべての局面において PPLM と同様の、基本文からの拡張を行っている点が、Alyce との大きな差異である。

ストーリーを構成するイベント間の因果関係に着目した試み [2] においては、イベント間の因果関係を常識的範囲での柔軟な対応に留めることにより、物語の展開上の齟齬の縮小を企図したものである。この試みでは、有向無帰帰グラフ (DAG: Directional Acyclic Graph) と呼ばれる、グラフ構造を持ったプロットグラフを採用している。このグラフ構造においては、プロット中のイベントがノードとなり、イベントの遷移がエッジとして表わされる。無帰帰のグラフ構造とすることにより、イベントの繰り返しという、ディープラーニングによる文生成で起こりがちな問題を回避しているとしている。しかしながら、同論文では、DAG と柔軟な接続の許容による齟齬の回避について実験を行っているが、比較的短いパラグラフ内での齟齬の解決にとどまっており、物語全体での齟齬の解決については未知である。

先行するこれらふたつの研究のフルコトとの相違点は次の 2 つである。ひとつは、PLG 技法による主人公の感情値を活用した制御が検討・実装されていないことである。ふたつめには、大箱、中箱、小箱による全体構造の設定、および、再帰的構造を持たない点である。

「物語の構造」に基づいたプロットの自動生成システム [3] では、物語分析のひとつの手法である「13 フェイズ構造」を基本とし、登場人物の合流に着目してプロットの生成と矛盾点の解消を目指している。フルコトが大箱、中箱、小箱という再帰的構造を持つものに対して、13 フェイズ構造という単一の構造をもとにしている点、および、フルコトでは感情値による制御を中心としているのに対して、登場人物の合流に着目して物語の分割を試みている点が差異であると

いえる。

2.2. AI を利用した映画制作事例

AI による映画制作の事例としては、映画監督である Oscar Sharp と AI 研究者の Ross Goodwin らが共同して開発した AI 脚本家 Benjamin がある。Benjamin が出力した自動的に出力した脚本に基づいて、9 分間の短編映画「Sunspring」(2016) が制作された。また、同じく Benjamin が、脚本のみならず、編集や画像合成までを実行した短編映画「Zone Out」(2018) がある。

これら Benjamin を用いて制作された映画は、ストーリー展開やセリフに不自然な部分が多いという課題がある。例えば、Sunspring の脚本の一部では次のように物語が展開する。「私達は H が本棚から本を取り出して、話しながらパラパラとめくって。そして、それを元に戻す。H 大量の失業者がいる未来では 若者は血を売ること余儀なくされる。血を流す。それがまず、私にできることをする。H2 少年たちを見て、黙る。私は、この先 百歳になる H また見た。」(脚本中、「H」は主人公であり、「H2」は別の登場人物である。)

これらの映画については、さまざまな評価がある。すなわち、AI が制作した初の作品であるという、映画史上の意義を高く評価するものから、不条理な展開が続く映画、支離滅裂であり評価不能であるというものまでである。ただし、毎年ロンドンで開催される映画祭「SCI-FI-LONDON Film Festival」の「48-Hour Film Challenge」(48 時間以内に映画をつくる) 部門で採択されていることから、映画史的には評価されているといえる。なお、採択はされたが受賞はしていない。

2.3 コンテンツの創造性評価方法の提案

コンテンツを評価するにあたってのひとつの評価軸として、創造性がある。創造性についての定義の方法は様々あるが、当研究を進めるにあたり、ひとつの定義として、「新規性」、「意外性」、「諒解可能性」を満たすという定義を提案する。新規性とは、従来に類をみないものであること、意外性とはオーディエンスの発想を超えた部分を持っていること、そして、諒解可能性とはオーディエンスが物語の世界観、登場者の心理や行動、場面や状況のつながりなどについて一定の理解を示すことができる内容であることであると定義する。芸術的映画といわれるジャンルにおいては、諒解可能性は低めであっても評価が高い場合があり、娯楽映画などの広く一般に受け入れられる必要のある映画にあっては、諒解可能性は高いことが求められる。また、新規性および意外性と、諒解可能性はトレードオフの関係にある。新規性や意外性が高まると、諒解可能性は低下する傾向にある。

Benjamin が生成し、また、制作した映画を、本稿で提案する創造性の定義に照らしてみると、AI が生成した映画は、新規性や意外性について

は満たしているが、諒解可能性についての部分において、改善の余地があると言える。

3. 提案手法

フルコトはストーリーを持ったコンテンツの制作支援を行う自然言語処理システムである。特に映画やテレビなどで放映されるドラマの脚本を制作することを主たる目的として開発された。フルコトは、最終的には脚本の完全自動生成を目標としているが、現時点では、脚本生成を支援するシステムとし、人とAIの協働により物語を完成させるためのシステムとしている。

3.1. 映画制作技法の導入

フルコトの特徴のひとつとして、4つの映画製作技法を実装していることが挙げられる。すなわち、「ログライン」、「箱」の概念、「プロットライングラフ」および、「柱書き」である。次に、それらの技法の内容を記す。

(1) ログライン

「ログライン」とは、コンテンツ全体を60文字から80文字程度の単文で物語を表現した短文である。この語は、映画やドラマの制作において一般的に使用される用語である。ログラインの例としては、「母や姉に日頃虐げられている村娘が、魔法の力で舞踏会に行き王子と出会い、一旦は離れ離れになるが、王子の努力により二人は再会して結婚し幸福になる。(以上、72文字)」という文がある。このログラインを基に、関連する要素を付加し、ストーリーを膨らませてゆくことにより、一貫性を持ち、さらには物語を制作することができる。映画の企画書では、このログラインが冒頭に提示される。

(2) 箱の概念

「箱」とは、脚本を構成する要素単位である。「箱」や「箱書き」という語は、映画やドラマの制作において一般的に使用される用語である。シナリオ制作の指導書[4]では、「ストーリーがだいたいこう発展してゆくという地図に相当する覚え書」と説明されている。フルコトを設計するにあたり、脚本家の意見から、この箱の概念を実装することとした。

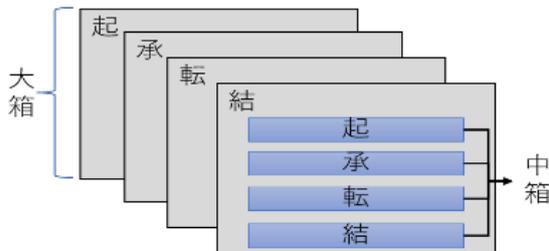


図1 箱の入れ子構造

箱は入れ子構造となっている。(図1) 物語全体を三分割、もしくは四分割した箱をそれぞれ「大箱」と呼ぶ。日本で多用されている物語構造としての「起承転結」での「起」「承」「転」「結」

は、それぞれ大箱にあたる。各大箱は、その内部に中箱と呼ばれる要素を内包している。中箱は、大箱の内容を三分割、もしくは四分割したものである。例えば、大箱を四分割した構造とする場合であれば、「起」「承」「転」「結」の四つの中箱が、各大箱に内包され、合計16個の中箱で物語全体が構成されることになる。なお、フルコトでは、分割数として4を採用している。(図2)

さらに、16個の中箱に対して、それぞれ4個ずつの起承転結の小箱を記述することにより、合計64個の小箱からなる脚本が完成する。なお、30分程度までの短編映画の脚本は中箱16箱程度で構成され、90分から120分程度の映画の脚本は64箱から90箱程度で構成されている。この箱と呼ばれる構成要素の入れ子構造は、再帰的処理であることから、システムとの親和性及び拡張性が高い。

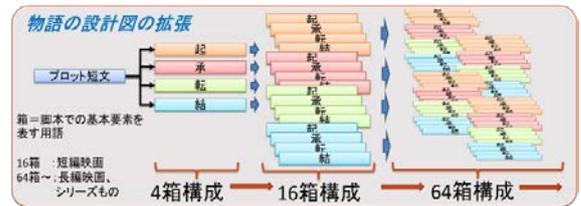


図2 箱の拡張方法

(3) プロットライングラフ

「プロットライングラフ (PLG)」とは、視聴者に共感や感動を引き起こすことができる物語の制作技法として、特に米国の映画制作において多用されてきた技法であり、近年では国内でも利用されるケースが増加している。(図3) なお、PLGは、感情曲線 (Emotion Arc) と呼ばれる場合もある。

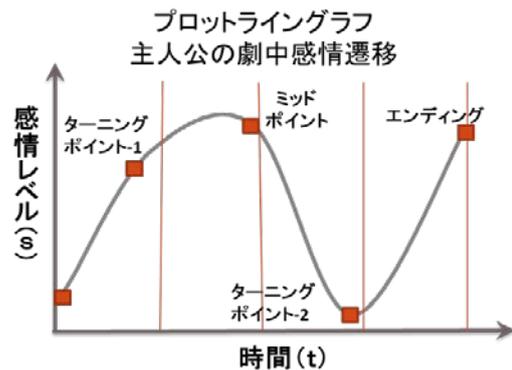


図3 プロットライングラフ

PLGでは、主人公の心理的状況や主人公を取り巻く環境の状況の良し悪しを縦軸に、物語の時間遷移を横軸とした2次元のグラフで表す。(図X) 標準はN字型の遷移である。N字型の場合、主人公の心的あるいは環境的状态は当初、すなわち起の時点では低いレベルもしくは中程度のレベルにある。3.1で示したログラインの例では、母と姉に虐げられている村娘の部分である。その後なんらかのエピソードにより、主人公の状

態は上昇する。ログラインの例においては、魔法により王子と出会う部分である。しかし、その後、主人公の状態は物語中で最悪の状況に陥ることとなる。ログラインの例においては、一旦離れ離れになる部分である。そして、最終的には主人公の状況はもっとも良い状態、もしくは、よい状態になることを示唆するエピソードで終結する。すなわち、ログラインの例での再会と結婚である。

以上の N 字型がもっとも多く用いられる傾向にあるが、逆 N 字型や V 字型、さらには、右肩上がりの直線型、右肩下がりの直線型まで、多様な遷移型が存在する。一般に使われている表現に例えば、N 字型や V 字型はハッピーエンド、逆 N 字型や右肩下がりの、逆 V 字型などはバッドエンドのストーリーと言える。

(4) 柱書

「柱書き」とは、脚本において、ある場面の場所と時間に関する記述部分のことである。例えば、「学校（朝）」、「オフィスビル通用口（夜）」などと表記される。柱書は、映画やドラマの作成におけるシーンを表している。この柱書の内容を、候補文の属性情報として付加することにより、ストーリーの一貫性保持、および、流れの制御を容易にすることができる。

3.2. 候補文生成過程

フルコトは三段階で脚本の候補文、および、その属性情報を生成する。すなわち、学習段階、候補文生成段階、および、評価指標の付加段階である。以下、各段階でのシステムの動作を説明する。

(1) 学習段階

学習段階では、教師となる文章群をもとに、学習アルゴリズムを利用して学習する。この学習タスクの実行により、文章生成モデルができる。フルコトには、多階マルコフモデル、LSTM (Long-Short Term Memory) および、深層学習 (Deep Learning) を利用した大規模事前学習モデルである Transformer モデルの 3 種類を実装したバージョンが存在する。本稿執筆時のフルコトにおいては、Transformer モデルを用いている。

この学習段階の実施により、フルコトは、任意のある単語列に対して、次に接続される単語を確率的に推定することができるようになる。なお、現在利用している Transformer モデルについては、すでに大規模言語セットによる学習済みのモデルであることから、フルコト内部での学習は行っておらず、外部で提供されている Transformer モデルを利用している。

(2) 候補文生成

候補文生成段階では、ログライン、すなわち、人が作成しフルコトに入力したあらすじの文をもとに、フルコトが大箱と中箱の候補文群を生

成する。以下に、Transformer モデルを使用したバージョンのフルコトにおける動作の詳細を記す。(図 4)

まず、フルコトは、入力されたログラインを、機械的に起承転結の 4 つの部分に分割し、それらを起承転結の各大箱の基礎となる文として設定する。このように、ある文章を任意の数で分割したそれぞれの文を部分文と呼ぶことにする。ログラインの分割は、字数および読点をもとに行われる。なお、フルコトにおいて入力されるログラインの文は、60 字から 80 字を想定している。

次に、4 つの大箱それぞれに対する候補文群として各 256 文、4 つの大箱の合計で 1,024 文を、Transformer モデルによって生成する。Transformer モデルには多くの種類がある。フルコトでは、そのうち、Rinna 社が提供する GPT-2 極小規模日本語モデル、GPT-2 小規模日本語モデル、GPT-2 中規模日本語モデル、13 億パラメータ GPT 日本語モデルおよび、BUET CSE NLP Group が提供する mT5_multilingual_XLSum のいずれかを選択して、文生成を行うことができる仕様となっている。

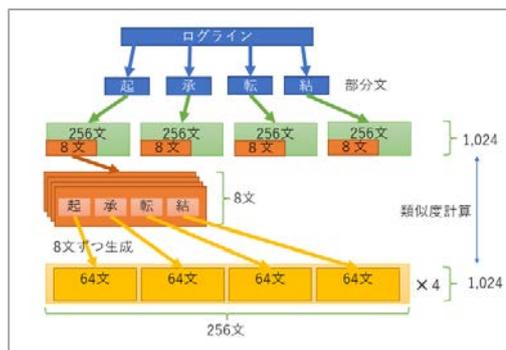


図 4 候補文の生成過程

各大箱向けに生成された箱あたり 256 の候補文のなかから、ログラインから生成された部分文との類似尤度を基準に、各箱の候補から 8 文字ずつを選択し、合計 32 文を中箱生成用の文とする。ここでいう類似尤度とは、生成された候補文と、ログラインが、共通のトークンを含む割合をいう。トークンとは、Transformer モデルで使用される単語の最小単位であり、それぞれの Transformer モデルごとに、トークン化を実行するためのトークナイザとよばれる機能を有している。トークナイザは教師文から、単位要素を抽出するものである。なお、トークナイザは、従来の自然言語処理で利用されている形態素解析手法と同様に、文を語素に分割するものではあるものの、それぞれの Transformer モデルに最適化されている。このため、トークナイザが出力する分割結果とは異なる場合がある。なお、大箱、および、中箱用に生成する文のサイズは、文字数ではなく、トークン数を用いて指定している。大箱の候補文については 50 トークン以上、60 トークン以下、また、中箱の候補文については 25 トークン以上、30 トークン以下を指定している。

©2022 Information Processing Society of Japan

クン以上, 30 トークン以下としている。

上記の処理により, ひとつの大箱あたり 8 つの候補文, 合計 128 文が大箱候補文として選択される。次にフルコトは, それら 128 文すべてについて, 再び, 字数および読点をもとに 4 分割し, それらの部分文に対して 8 文ずつの候補文を生成する。これにより, 大箱ひとつあたり 256 文, 4 つの大箱合計で 1,024 文が生成されることとなる。これらトークンに関する条件は, 60 字から 80 字で書かれたログラインを四分分割する方法を, 大箱文や中箱文に適用できるように, おおよそその文字数がログラインと同等となるようにするために定められた制約条件である。

(3) 評価指標の付加

この段階では, 各文への評価指標の付加を行う。人が考えて文章を生成する場合とは異なり, 自然言語処理 AI による文章生成では, システムの設定次第で, 短時間に大量の文章を生成することが可能である。しかしながら, 大量の文章が生成されたとしても, そのなかから, なんの手がかりもなく, 物語をかたちづくるのに適した文章を人が抽出することは困難である。そこで, フルコトでは, 候補文に対して評価指標を付加することにより, ストーリーの各局面に対して適切な候補文を抽出し, 提示することを可能としている。

フルコトでは, 2 つの評価指標を使用する。すなわち, 類似度と感情値である。

類似度 (Similarity) とは, 文章同士がどの程度似ているかの指標である。ここで類似度を設定し, 利用するのは, 次項で述べる人とフルコトの対話的な脚本生成作業において, 大箱の文に対して, 大きく異なる文の提示順位を下げるためである。選択計算は, TF-IDF 法とコサイン類似度によって計算される。

TF (Term Frequency) は, ある形態素が全候補文の中に現れる頻度で, 多く使われる形態素ほど高い値となる。IDF (Inverse Document Frequency) は, ある形態素を含む候補文が現れる頻度の逆数で, 含まれる候補文が少ない形態素ほど高い値になる。TF と IDF を掛けた値 (TF-IDF) を, その形態素の重要度とする。各文を, その文に出現する各形態素の出現回数を用いたベクトルとし, その形態素の TF-IDF により出現回数を加重したうえで, 類似度を計算する。類似度計算は, 大箱の候補文として生成された 1,024 文と中箱の候補文として生成された 1,024 文のすべての組み合わせに対して, コサイン類似度計算, すなわちベクトル同士の内積計算により行う。

感情値 (Sentiment Score) は, その文章がポジティブかネガティブかを数値で表す指標である。この指標は, 生成する脚本の各局面において, プロットライングラフへの合致度の高い文の選択を支援することを目的とした指標である。例えば, 物語序盤で主人公の感情や環境が高揚している場面に対しては感情値の高い文章を選

択し, 逆に, 物語中盤で感情や環境が悪化した状況を呈する場面では感情値の低い文を選択するという具合に利用する。感情値の付与は, Google 社の感情解析 API (CNL API) を利用している。

3.3. フルコトの操作

前項で述べたフルコトの機序を利用して, 人がどのようにこれを利用するかについて以下述べる。

フルコトの操作は大きく次の 3 ステップでの実施となる。すなわち, 言語モデルの選択, 候補文生成の指示操作, および, 対話的な候補文選択と修正作業である。

(1) 言語モデルの選択

フルコトでは, 本稿 3.1-(2)に記したように, Transformer モデルのうち, 5 つの事前学習済みモデルから, ひとつを選択して実行することができる。また, Transformer バージョンに移行する以前の, 多階マルコフモデル, もしくは, LSTM モデルを選択することが可能となっている。

ここで, 多階マルコフモデルおよび LSTM モデルと, Transformer モデルでの, 脚本生成における差異について概略しておくこととする。

多階マルコフモデル, および, LSTM モデルを使用したフルコトにおいては, 比較的少ない教師データを用いて学習を行う。これは, 限られた計算資源の下で, 大規模な学習データによる学習を実施しようとするとき, これらのアルゴリズムでは並列処理ができないため, 現実的な時間内で計算が終了しないことや, 計算資源の制約から計算が異常終了してしまう場合があることが理由である。外部の大規模事前学習モデルは, 膨大な計算資源と計算時間を費やして生成されたものであり, それらを利活用することにより, 学習のために自らの計算資源を費やすことなく, より高精度, かつ, より多様な出力を得ることが出来る。また, Transformer モデルは学習データの多さから, 出力される文の流暢さが比較的高いことも挙げられる。さらに, Transformer モデルのうち, フルコトにおいて利用可能な mT5_multilingual_XLSum は, 多言語, 多地域の文章により事前学習が行われていることにより, 出力される文の多様性を享受することができる。

しかし, 多階マルコフモデル, および, LSTM モデルにも良い点がある。すなわち, 学習データの範囲が絞られていることにより, 脚本生成のために提示される候補文に収斂傾向があることである。新規性, 意外性よりも, 諒解可能性を重視した脚本制作に向けた技法であると言える。例えば, 次章で述べる短編映画「少年, なにかが発芽する」の原作であるところの, フルコトを使用した脚本「とまと」の制作過程での効果が挙げられる。とまとの出力を目的として, 11 の文章が 8 人の制作者によって制作され, それらを学習データとして, 多階マルコフモデルを学習させた。異なる制作者による脚本では, それぞれの制作者の個性に応じて, さまざまなストーリー展開

が記述されたが、共通の語として、「とまと」、「少年」、「ビニールハウス」、「おじさん」を使うこととした。これにより、学習済みのモデルからの出力にあっては、大きな齟齬や突飛すぎて諒解不能な文が出現する頻度が下がった。

なお、本稿 2.2 に示した Benjamin では、学習データとしては、「ゴースト・バスターズ」、「フィフス・エレメント」、「2001 年宇宙の旅」といった SF 映画が使用され、アルゴリズムとしては、LSTM が使用されているとのことである。

(2) 候補文生成の指示操作

フルコトは、クラウド上に実装され、ブラウザベースのインタフェースが提供されている。ユーザは、インタフェース上で、考案された 60 字から 80 字程度のプロットのテキストデータを、ログラインとしてフルコトに入力する。次にユーザは、フルコトに対して、入力されたログラインをもとにして候補文を生成し、さらに、生成された候補文の類似度計算処理と感情値情報の付加処理の実行を指示する。これにより、フルコトは上記の処理を開始し、処理が完了すると、完了した旨の表示がなされる。前項のモデルの選択、ログラインの入力、および、実行の開始処理は、ブラウザベースのインタフェースから対話的に実行することができる。(図 X)

なお、ログラインを入力し、候補文生成、および、評価指標の付与にかかる時間は、ログラインに依存し、数時間から数十時間までばらつきがある。参考に、具体例を次に示す。投入したログラインが「透明人間になる能力を得た少年は、世界紛争の引き金となった皇帝を暗殺するため旅立ち計画を実行する、世界には平和が戻る、大きな出来事の影には名もなきヒーローがいる。」とし、利用言語モデルは rinna13 億パラメータ GPT モデル、実行環境は Google Kubernetes Engine (limits=cpu:"1000m"/memory:"8Gi"/ephemeral-storage:"512Mi") において、処理に要した時間は 6 時間 7 分であった。

(3) 対話的な脚本制作作業

前項の処理が完了すると、システム内には、前項で入力したログラインに合致する、大箱および中箱向けの候補文が、評価指標付きで格納される。そこで、ユーザは、それら候補文のなかから、4 つの大箱、および、合計 16 の中箱に対して、ユーザが最適と考える候補文を選択することにより、脚本を完成させる。この作業も、ブラウザベースのインタフェースから対話的に実施することができる。以下、脚本制作作業の実施の詳細について述べる。

インタフェースでは、ログライン、大箱、および、中箱のボックスが表示される。同時に、各箱の下部には、目標感情値、および、感情値重視度を変更するためのスライダが表示される。

ここで、目標感情値は、その箱に対してフルコトが提示する候補文を抽出する際の感情値の値を調整するものである。また、感情値重視度は、

感情値と類似度のふたつの評価指標のうち、類似度に対して、感情値をどの程度考慮するかのバランスを調整するものである。(図 5) フルコトは、デフォルトの目標感情値、および、感情値重視度の値をもとに、大箱・中箱それぞれに対して 10 文ずつを感情値および類似度をもとに、候補文のなかから自動的に選択し、提示する。



図 5 インタフェース画面例

ユーザは、箱ごとに提示された候補文のなかから、最適と考える文を選択することができる。また、提示された候補文のなかに、最適と考えられない文章がない場合は、ユーザが目標感情値や感情値重視度のスライダを調整することにより、新たな候補文 10 文のセットの提示を受けることができる。さらに、ユーザは、ユーザが選択した候補文の一部、もしくは、全部を修正し、保存することも可能である。これらの機能により、ユーザは対話的に、多様でありながらも、ストーリーに沿った文章候補の中から、各箱に最適な文を選択し、場合によっては修正することが可能となっている。

上記の操作を経て、最終稿となる 4 個の大箱文と 16 個の中箱文が選択されると、ユーザはそれを一連のシナリオとして保存し、自由にそれらを取り出し、さらに改変を重ねることもできる。なお、同一の脚本についても、フルコト上のインタフェースからバージョン情報を付加することにより、派生的脚本の生成と保管が可能となっている。

4. 事例

4.1. 短編映画のストーリー生成

多階マルコフモデルを利用したバージョンのフルコトを用いて、短編映画の脚本を生成した。以下に、生成の過程とその結果、および、得られた知見について述べる。

まず、共通のテーマに基づき作成された 15 の文章をフルコトに学習させた。次に、学習済みモデルを実装したフルコトに対して、さまざまなログラインを与え、大箱および中箱の候補文の生成と、適合する候補文の推薦を行わせた。候補文からの文の選択は脚本家が行い、最終稿として中箱 16 個からなる脚本「とまと」を完成させた。最終的に入力されたログラインは、「母親と一緒にトマト畑に来た少年が、働いている母親から離れひとりで萎れたトマトの木に水やりし、やがて仕事を終えた母親と一緒にトマト畑を去る」というものである。

脚本「とまと」をもとに、30分の短編映画「少年、なにかが発芽する」が制作された。(図6)映画の制作にあたっては、プロデューサ、監督、脚本家らが協議し、最終稿を完成させた。なお、映画の題名「少年、なにかが発芽する」は、フルコトが出力した候補文のひとつである「少年、発芽する」から、制作陣が発想を得て命名したものである。



図6 「少年、なにかが発芽する」映像例

完成した映画は、大阪アジア映画祭2022にて特別上映、また、ショート・ショート・フィルム・フェスティバル(SSFF & ASIA2022)において招待上映された。

5.2 短編アニメーションの脚本生成

Transformerモデルを利用したバージョンのフルコトを用いて、テレビ番組向けの短編アニメーションの脚本生成を行った。テレビ番組のプロデューサが、フルコトを対話的に利用して脚本を作成し、3分間の短編アニメーションが制作され、放映された。(図7)



図7 アニメーション映像例

5. 考察

(1) フルコトの開発

ストーリー型コンテンツの需要の増大に対して、供給が逼迫している状況がある。この問題を解決するために、ストーリー型コンテンツ生成支援AIシステム「フルコト」を開発した。

ユーザは、フルコトを利用することにより、ログラインと大箱・中箱による再帰的構造をベースに、プロットライングラフに沿った感情の遷移を持つ脚本を、対話的に生成することができる。また、大規模事前学習を文章生成に利用し、さらに評価指標を活用することにより、創造性のひとつの指標としての「新規性」、「意外性」と

「諒解可能性」のトレードオフにおいて、バランスのとれた文の提示と選択が可能になったと考えられる。

このシステムを使って生成された脚本をもとに、映画やアニメーションが制作され、映画祭での上映や、テレビでの放映が実施された。

(2) 成果物の評価

映画「少年、なにかが発芽する」は、公益財団法人芳泉文化財団、および、文化庁ARTS for the future! (コロナ禍を乗り越えるための文化芸術活動の充実支援事業)の支援の下に、渡辺裕子監督、東京藝術大学、および、トリクスタ社により制作されたものである。制作に関わりたいずれの団体、個人ともに、これまでに映画制作、および、文化活動支援の実績を多数有している。それら、映画制作についての経験を持った団体・制作者が、フルコトを利用して制作された脚本「とまと」について、これを原作として映画化するに足る脚本であると判断し、実際に制作し、さらに、国際的映画祭において上映されたという点で、フルコトの支援により制作された脚本がストーリー型コンテンツの制作に貢献するものであるとの評価が可能であると考えられる。また、本稿2.2において示した創造性の定義に照らしてみても、「少年、なにかが発芽する」にあつては、

「新規性」、「意外性」、「諒解可能性」いずれもが、実用に耐えるレベルで満たされていると考える。

テレビ番組向けアニメーションの制作において、プロデューサが当初考案したログラインは「ゾンビとなっても人間の意識が消えない男がゾンビ狩りから逃げた先で目の見えない少女と出会い恋をしてしまう、男はゾンビである事を言えず、人間の心を悔い世界の果てへと旅立つ。」というものであった。入力されたログラインに対して、フルコトからは「(世界は)ゾンビの化石に悩まされている」、「(少女の)週に一度の目の見えない日」などの表現を含む候補文が提示された。いずれも新規性、意外性がありながらも、諒解可能性も備えている候補文であるとえられる。プロデューサは、それらフルコトが生成した文に触発され、元のログラインからは逸脱しているものの、より新味のある物語へと発展させることに成功した。新規性、意外性については、製作作業中のプロデューサの発言であるところの、「自分はこれまで常識をやぶる発想で制作してきたつもりだが、AIの発想は自分の発想を超えている」(大意)というものが、フルコトの脚本制作における支援の能力を表現していると言える。

(3) フルコトの効果と影響

フルコトの特徴は箱書きによる再帰的構造性と感情値に基づく物語制作のためのPLG技法を複合的に利用する点にある。以下、この特徴が、どのような効果を生み出したか、どのような影響があったか、について考察する。なお、考察は、映画、および、アニメーションの脚本生成を実務

として行う脚本家、プロデューサらと実施することにより得られた知見を基にしている。

脚本制作におけるフルコトの効果や影響は3つあると考えられる。すなわち、新規性や意外性のある脚本生成に資すること、整合性のあるストーリー生成に資すること、および、制作の効率向上に資すること、の三点である。以下、それぞれについて詳細を記す。

フルコトで実装されている自然言語処理技法では、生成される候補文におおきな「ゆらぎ」がある。ここで言うゆらぎとは、日常の会話や、ストーリーの考案過程では発生しにくい、語の連接や共起を持った候補文が提示されること、および、内容的に離れた文候補も提示されること、すなわち、システムの観点からは言語空間内の距離が大きい文やベクトルの角度の開きが大きい文や単語が提示されることである。脚本生成においては、このゆらぎが、新規性や意外性の源泉となっていることが、ユーザである映画やテレビ番組の制作者らの意見から明らかになった。ユーザは、それらの単語や文章の組み合わせから、新たな世界観やキャラクタ、および、ストーリーを創出し、脚本を制作することができるようになる。

いっぽうで、ユーザに対して、全体としても、また、部分としての箱の内部にあっても、整合性のある脚本の制作に適した素材の提供が可能となっていると考えられる。これは、提示される文についてのある程度の整合性を実現するための指標であるところの類似度の導入、および、広汎に使用されている作劇法に則った感情曲線を実現するための指標であるところの感情値の実装が寄与しているといえる。

三点目として、上記の効果により、フルコトの利用は、脚本制作に要する時間の短縮等、コンテンツ制作の効率向上に貢献すると考えられる。現在、新規性や意外性を生み出すために、ユーザである制作者は、資料や情報の収集と、それらの整理、そして、従来の形態を超えるための思考に大きな時間と労力を費やしている。

ここで、観点を変えてフルコトを見てみることにする。上記のユーザの行動と思考を軸として考えると、Boden [6] は創造性についての、三種の分類、すなわち、組み合わせ型の創造性 (Combinational Creativity)、探索的創造性 (Exploratory Creativity)、変革的創造性 (Transformational Creativity) を適用することができる。この分類に照らしてみても、フルコトでは、それらすべての分類項目について、対応しているといえる。つまり、言語モデルが提示する語や事象の多様な組み合わせ、および、大規模事前学習モデルが提供する広汎な探索範囲、および、ゆらぎがもたらす変革性を、フルコトは提供していると考えられる。

なおあわせて、フルコトに実装されている映画制作技法、すなわち、ログライン、箱書き、PLG、

柱書の機能が、ユーザであるプロフェッショナルなコンテンツ制作者にとって、受け入れ易いものであることを付記しておく。これは、用語法が映画制作技法と合致していることのみならず、コンテンツ制作の現場において、コンテンツ制作者たちが平素から制作作業での基本としている技法や思考方法を、システム上に実装できていることがその理由であると考えられる。

以上より、ストーリーを持ったコンテンツ制作支援ツールとしてのフルコトにあっては、構造的にはユーザへの親和性があり、また、新規性のあるストーリー創出を十分に支援できる機能を保有していると考えられる。

6. 今後の展望

ひとつには、より長いコンテンツの脚本の制作を支援するために、取り扱える箱数の64箱への拡張を予定している。また、制作支援対象を、ト書き部分のみから、柱書やセリフ部分に拡張する予定である。

また、意外性のある単語や文章から創造する力量には個人差がある。すなわち、フルコトが出力する候補文は、ログラインからかけ大きく離れた文章を出力することがある。その際に、それらの突飛な文に対応して、より新規性や意外性のあるストーリー、世界観へと飛躍させることができるユーザと、突飛な候補文を自らの内部でストーリーの一部に昇華させることが難しいユーザが存在することが観察された。それら、さまざまなユーザレベルに対応する方策についても今後検討を行う予定である。

参考文献

- [1] "Alyce: An Artificial Intelligence Fine-Tuned Screenplay Writer"; Eldhose, K.A., Jose, C., Siddharth, S., Geejo, S.S., Sreedevi, Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT, volume 59), 2021
- [2] Automated storytelling via causal, commonsense plot ordering."; Ammanabrolu, Prithviraj, et al., *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 7., 2021.
- [3] "シナリオ自動生成のための映画脚本ストーリー展開構造分析とプロット生成"; 川野 陽慈, 宇都宮 悠輝, 高屋 英知, 山野辺 一記, 栗原 聡, 人工知能学会全国大会論文集 JSAI2020 (0), 3D5OS22b04-3D5OS22b04, 2020
- [4] "シナリオの基礎技術"; 新井一, ダヴィッド社(1985)
- [5] 「AIはどこまで創造力を持つか」(北海道道) 著作制作; 日本放送協会, 初回放送日: 2022年5月13日
- [6] Margaret A. Boden, "Creativity and artificial intelligence", *Artificial Intelligence* 103, 1998