

# 物語のコンセプトに基づく情報アクセス手法の基礎検討： 短文要約文とあらすじ文との比較

Information Access Method Based on Story Concept:  
Comparing Short Summary and Outline

藤川 雄翔 \*1  
Taketo Fujikawa

畑 玲音 \*1  
Reon Hata

松下 光範 \*1  
Mitsunori Matsushita

\*1関西大学  
Kansai University

This study aims to realize the information access method based on story concepts. A story concept is a short description of the bones of a story structure. For example, “This story is about a man and a woman who belong to each other’s unfriendly groups fall in love.” In our previous study, we collected short summaries of respondents’ favorite stories. Analyzing those summaries revealed that they can be used as story concepts. In this paper, we used the collected short summaries as story concepts, comparing them with outlines of stories collected from the Internet, to clarify features of story concepts: We made sentence vectors from story concepts and outlined each, then examined differences between the story concepts and the outline of the same work by calculating the similarity of their vectors. The result of the study suggested that a story concept can be made from the outline.

## 1. はじめに

物語作品におけるストーリーには基本的なコンセプトが存在する。コンセプトとはストーリーの土台となるアイデアであり、物語を創作する際の起点としてだけでなく、他者に物語作品を推薦するための簡単な説明や、漠然としたストーリーの表現から物語作品を探索する際のクエリとして活用することができる。

本研究では、例えば“敵対する組織に属する男女二人が恋に落ちる話”のように、ストーリーを端的に表現した文章をコンセプト文と定義する。コンセプトに基づいた物語作品への情報アクセスが実現すれば、従来の物語作品の検索に用いられるジャンルや著者などといった書誌情報よりもストーリーの内容を踏まえた検索が可能になることが期待されるが、それらを可能にするためには、個々の物語作品に対するコンセプト文が必要になる。

コンセプト文に類似するものとしては、Wikipedia などの Web 百科事典や Amazon や楽天といった EC サイトに記載されるあらすじ文が挙げられる。あらすじ文は物語作品の要約と捉えることができ、作品全体の要約や 1 巻ごとの要約、作品の“世界観”に言及した要約、作品のメタコンテンツが付与された要約 [山西 22] といったように多様な粒度で存在しているが、その中には、売上や掲載紙の情報といったようなストーリーとは無関係のノイズが含まれていることもある。

一方で、コンセプト文は端的にその物語作品のストーリーの内容のみを伝えることを目的にしているため、重要な情報だけが記述される文章であるといえる。ただし、コンセプト文は、見どころを伝えるハイライトやストーリーの流れを示すダイジェストと異なり、コンセプト文そのものを読む行為のエンタテインメント性は高くないことから、全ての作品に対して必ずしも用意されているものではない。コンセプト文を機械的に自動生成するアプローチも考えられるが、生成モデルの構築には学習対象となる物語作品とその物語作品に対応するコンセプト文のデータセットが必要となる。

こうした背景の下、筆者らは先行研究として、読者によって生成された物語作品のコンセプトを表現した短文を収集し、それらを用いてコンセプト文のデータセットの作成を試みてきた [藤川 23]。この研究では、収集した短文に含まれる単語の分散表現ベクトルの平均を文章ベクトルとして算出し、同一物語作品の文章ベクトルの類似度から、物語作品が要約された短文が物語作品のコンセプトとして利用可能であるかについて検討を行った。その結果、同一物語作品同士の短文の文章ベクトルに類似度がある程度確認されたことから、読者が物語作品を要約した短文がコンセプト文として利用可能であることが示唆された。しかし、主に若い年代である学生を対象として短文を収集したことにより、収集できた物語作品やそのジャンルが偏っていた。また、コンセプト文に類似するものとしてあらすじ文が挙げられることから、あらすじ文を基に機械的にコンセプト文を自動生成することが考えられるが、読者によって生成されたコンセプト文があらすじ文に含まれる情報を有しているかの検証は行っていない。

そこで本稿では、人手により漫画作品を対象としたコンセプト文のデータセットを新たに構築する。物語作品が要約された短文が物語作品のコンセプトとして利用可能であるかについて、文脈を踏まえた文章ベクトルの類似度を用いて再度検討する。さらに、コンセプト文とあらすじ文、各々の文章ベクトルを作成し、物語作品間の類似度からあらすじ文とコンセプト文の差異を観察することで、あらすじ文からコンセプト文を生成できるかについてを検討する。

## 2. 先行研究

本研究のコンセプトに類似した考え方である、物語作品の“機能”や“モチーフ”に着目した研究がある。プロップ [プロップ 87] は 100 篇の魔法昔話を分析し、それらの物語に共通して観察可能な登場人物の行為である 31 の機能を明らかにした。小野ら [小野 19] は個々の昔話の特徴づける登場人物とその行為をモチーフと呼び、モチーフを用いた物語の自動生成を試みている。これらの機能やモチーフは物語のストーリー構造についての個々の記述であり、ストーリーの内容全体を記述したコンセプト文とは異なる。

連絡先: 藤川雄翔, 関西大学大学院総合情報学研究科, 〒569-1095 高槻市霊仙寺町 2-1-1, Tel:072-690-2437, Fax:072-690-2491, k920357@kansai-u.ac.jp

表 1: コンセプト文同士のコサイン類似度上位 5 作品 (一部). 各 ID に対応するコンセプト文は表 3 を参照.

ID	SD1	SD3	SD5	SD6	SD7
1	SD8	SO1	SD8	HH5	HH5
	0.8184	0.6436	0.8087	0.6970	0.6641
2	SD4	KO1	SD4	HK2	KT1
	0.8044	0.5829	0.7906	0.6257	0.6529
3	SD9	YK2	SD9	SD7	SD6
	0.7715	0.5598	0.6973	0.6137	0.6137
4	SD2	GT1	SD2	SD9	HI1
	0.7404	0.5509	0.6958	0.6002	0.6070
5	SD5	KJ1	SD1	YH1	BH1
	0.6679	0.5464	0.6679	0.5930	0.5947

### 3. あらすじ文の収集

コンセプト文との比較に用いるあらすじ文は、漫画全巻ドットコム<sup>\*1</sup>内の「作品概要」とし、歴代発行部数ランキングの上位 150 作品を対象として、人手で収集した。作品概要の一部に、商品説明や同名作品の他シリーズの作品説明が記述されているものは漫画作品のストーリー内容に直接的な関係がないため、該当箇所を除いた。また、本稿では漫画作品を対象としていることと、物語作品のストーリーを要約したコンセプト文との比較を行うことから、作品概要の大部分に商品説明やメタ情報が記述されているものやライトノベルの作品を除いた。この処理の結果、条件に当てはまった 135 作品の漫画作品のあらすじ文を対象とした。

### 4. コンセプト文の収集

本研究ではコンセプト文を収集するために、Yahoo!クラウドソーシング<sup>\*2</sup> (以下、クラウドソーシング) を用いて、自身の好きな漫画作品を他者に推薦するという状況設定のもと、195 名の回答者から 3 作品ずつ、計 585 文の漫画作品を要約した短文を収集した。短文の記述に際し、(1) 100 字以内で要約すること、(2) 文の末尾を「～物語」で締めくくることが、(3) その漫画作品の固有名詞 (e.g., キャラクター名, 組織名) を使わないこと、の 3 つの条件を課した。アンケート収集の際、漫画作品を要約した短文に加えてその漫画のタイトルとジャンルについても回答を課した。

収集された 585 文の短文のうち、上記の条件を満たし、かつ収集したあらすじ 150 作品と同一作品の短文である 126 文を本稿の分析対象とした。このとき収集した短文を人手で確認し、明らかな誤字は修正し、読み仮名や、短文の先頭に記述されたその漫画作品のタイトル、末尾の「です。」「。」は削除を行った。

### 5. 収集した短文の分析

前章の方法で収集した短文は、その物語作品に対する捉え方が回答者によって異なる場合がある。そこで本稿でも、収集した短文各々の文章ベクトルを算出し、先行研究 [藤川 23] と同様に同一物語作品の文章ベクトルの類似度から、物語作品が要約された短文が物語作品としてのコンセプトとして利用可能であるか検討する。

\*1 <https://www.mangazentan.com/> (2024/1/10 存在確認)

\*2 <https://crowdsourcing.yahoo.co.jp/> (2024/1/30 存在確認)

表 2: 同一漫画作品におけるあらすじ文とコンセプト文の類似度 (降順一部). 各 ID に対応するコンセプト文は表 3 を参照.

タイトル	ID	類似度
ジョジョの奇妙な冒険	JJ2	0.9036
はじめの一步	HI1	0.8214
...	...	...
ONE PIECE	OP6	0.1430
ビー・バップ・ハイスクール	BH1	0.0133

本稿では、sentenceBERT [Reimers 19] を用いて収集した短文のベクトル化を行い、それぞれの短文同士のコサイン類似度を算出した。表 1, 3 は“SLAM DUNK”を代表とした結果の一部とそれぞれの ID に対応する短文である。以下、個々のコンセプト文に対しての結果、考察を述べる際にも、表 3 の ID を用いて説明を行う。

SD1, SD5 はコサイン類似度の上位 5 作品すべてが同一の漫画作品の短文となった。これは短文内にその漫画の主題である“バスケット”や“バスケットボール”が含まれていることと、主人公を中心としたキャラクターがバスケットボールを軸にどのようなか記述されていることによると考えられる。

一方で、SD3, SD6, SD7 は上述の 2 文とは異なり、コサイン類似度が最も高い値でも、0.70 を下回る結果となった。SD3 はコサイン類似度の上位 5 作品には同一の漫画作品の短文は含まれなかった。SD3 では、その漫画作品の内容を示すタグ情報となるような単語を中心に記述しており、“根性”や“学園”、“友情”といった単語と類似する文脈で用いられる単語が含まれる短文が上位に入る傾向にあった。SD6 では、漫画作品内で主人公が成長することを軸に記述されていることから、同様に主人公が成長することを軸に記述された短文の類似度が高くなったと考えられる。SD7 では、スポーツを通じてキャラクターが成長することを軸に記述されていることから、同様にスポーツが絡んだキャラクターの成長を軸に記述された短文の類似度が高くなったと考えられる。

### 6. コンセプト文とあらすじ文の類似度測定

あらすじ文も前章のコンセプト文と同様に、sentenceBERT を用いてベクトル化を行い、コンセプト文とあらすじ文のコサイン類似度を算出した。その結果の一部を表 2 に示す。

同一漫画作品において最も類似度が高い上位 2 作品のあらすじ文とコンセプト文は“ジョジョの奇妙な冒険 (JJ2)”と“はじめの一步 (HI1)”であった。どちらのコンセプト文も、あらすじ文と類似したキャラクターの特徴や、そのキャラクターがストーリーにおいてどのような行動をとるのか、またどのような変化が生まれるのかが記述されていたことから、あらすじ文との類似度が高くなったと考えられる。また、JJ2 のコンセプト文と類似度の高いあらすじ文の漫画作品を確認したところ、“DEATH NOTE”、“北斗の拳”、“るろうに剣心 —明治剣客浪漫譚—”、“バカボン”が上位 5 作品に含まれた。これらの作品は全てライバルたちと死闘を繰り広げていく漫画作品であることがあらすじ文から読み取れることから、キャラクターのストーリーにおける主軸となる行動について記述されたコンセプト文は、物語のコンセプトに基づく情報アクセスに有用であると考えられる。

一方で、あらすじ文とコンセプト文の最も類似度が低かった“ビー・バップ・ハイスクール (BH1)”と“ONE PIECE (OP6)”は表 3 より他のコンセプト文と比較すると、キャラ

表 3: クラウドソーシングによって収集されたコンセプト文 (一部)

ID	タイトル	コンセプト文
BH1	ビー・バップ・ハイスクール	仲間同士で学校を守ったり仲間を助ける物語
GT1	銀魂	主人公が日常を面白く過ごしたり、ある時は真剣に悪人を倒したり、ギャグがある物語
HH5	HUNTER × HUNTER	少年が仲間とともに成長していく冒険物語
HI1	はじめの一步	いじめられっ子がある日ボクサーに助けられ、ボクシングに魅了され成長していくスポーツ物語
HK2	ハイキュー!!	低身長主人公がバレーの世界で成長していく物語
JJ2	ジョジョの奇妙な冒険	吸血鬼や未知の能力を持った者が死闘を繰り広げる物語
KJ1	寄生獣	恐ろしい生物が出てくるもユーモラスな表現とキャラクターで魅了される社会派な物語
KO1	今日から俺は!	ツッパリ少年たちの笑いあり涙ありの青春物語
KT1	キャプテン翼	少年時代から仲間やライバルと共にサッカーを通じて成長していく物語
OP6	ONE PIECE	仲間を大切に作る物語
SD1	SLAM DUNK	バスケット初心者の赤髪男がチームメイトと協力して成長していく物語
SD2	SLAM DUNK	バスケットボール少年たちのバイブルである主人公の成長物語
SD3	SLAM DUNK	単純なスポコンでなく学園物でもあり、かといってなあなあ友情物でもない、熱い男の物語
SD4	SLAM DUNK	不良の高校生がバスケットに出会い、色々な仲間と協力し合いながら全国大会出場を目指す物語
SD5	SLAM DUNK	高校のバスケットボール部で、色々なキャラクターがバスケットボールに打ち込む物語
SD6	SLAM DUNK	初心者の落ちこぼれが仲間とともに成長して大活躍する物語
SD7	SLAM DUNK	個性豊かな若者達が一つのスポーツを通じて先走ったり、その場に留まったりしながらも少しずつ成長する過程に親近感がわく物語
SD8	SLAM DUNK	不良少年がバスケットボールで仲間達と活躍する物語
SD9	SLAM DUNK	バスケットボールが未経験だった主人公が、努力して成長していく物語
SO1	魁!!男塾	男しか登場しない、根性や友情、死闘を繰り広げる物語
YH1	幽☆遊☆白書	子供を助けるために事故死してしまった主人公が生き返る為に事件を解決しながら仲間とともに成長する物語
YK2	有閑倶楽部	お金持ちなどの男女の学生の冒険や事件に巻き込まれたりしたりしたエピソードを面白おかしく描写した物語

タの特徴に関する記述がないことがわかる。さらに、ストーリーにおいてキャラクターの主軸となる行動についての記述も、多くの作品に当てはまるような抽象的な記述であるため、あらずじ文との類似度が低くなったと考えられる。このことから、物語のコンセプトに基づく情報アクセスに有用であるコンセプト文の記述にはキャラクターの特徴に関する記述も含めることが望ましいといえる。

以上のことから、キャラクターに関する情報やキャラクターのストーリー上での主軸となる行動についての記述があるコンセプト文が、物語のコンセプトに基づく情報アクセスに有用であると示唆された。あらずじ文にはこれらの記述が含まれていることから、これらの情報を抽出しつなぎ合わせることで、あらずじ文からコンセプト文が生成可能であると考えられる。

## 7. まとめ

本稿では、読者によって生成された物語作品のコンセプト文から文脈を踏まえた文章ベクトルを算出し、同一物語作品におけるコンセプト文同士の文章ベクトルの類似度を用いて、コンセプト文が物語作品のコンセプトとして利用可能か検討を行った。さらに、漫画の作品概要に記載されたあらずじ文についても同様の手法で文章ベクトルを算出し、物語作品間の類似度からコンセプト文とあらずじ文の差異を観察することで、あらずじ文からコンセプト文を生成可能かについて検討を行った。その結果、キャラクターの特徴に関する記述やそのキャラクターのストーリーにおける主軸となる行動の記述がコンセプト文に含まれることで、あらずじ文からコンセプト文が生成可能であることが示唆された。

今後は、あらずじ文からのコンセプト要約の自動生成を行い、コンセプト文間の類似度を用いた情報アクセス手法の実現

を試みる。

## 謝辞

本研究は、科研費 22K12338 の支援のもと行われた。記して謝意を表す。

## 参考文献

- [Reimers 19] Reimers, N. and Gurevych, I.: Sentence-bert: Sentence embeddings using siamese bert-networks, in *arXiv preprint arXiv:1908.10084* (2019)
- [プロップ 87] プロップ ウラジーミル (著), 北岡 誠司, 福田 美智子 (訳): 昔話の形態学, 水声社 (1987)
- [山西 22] 山西 良典, 西原 陽子, 松下 光範: 語彙の標本化と量子化によるあらずじの特性表現に関する基礎検討, 第 7 回コミック工学研究会, pp. 31–38 (2022)
- [小野 19] 小野 淳平, 小方 孝, 伊藤 拓哉: 昔話のモチーフを物語生成へ利用するための基礎研究, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, 1F2-NFC-1-05 (2019)
- [藤川 23] 藤川 雄翔, 松下 光範, 山西 良典: 物語作品の短文要約によるストーリーの特徴分析, 第 10 回コミック工学研究会, pp. 34–36 (2023)