

物語の登場人物を把握しやすくするシステムの提案

謝涵^{†1} 西田健志^{†1}

概要: 長い小説などを読む際、登場人物が前にどのような場面で出てきたどのような人物なのか思い出せず、読みづらいつと感じることがある。そこで本研究では、物語の登場人物を把握しやすくするシステムを提案する。提案システムには、登場人物を記憶しやすくすることと思い出しやすくすることの両面から、初登場シーンにジャンプする機能と人名に絵文字を追加するなどの機能を実装している。

キーワード: 読書支援, 登場人物, 形態素解析, ユーザインターフェース

A Read-Support System for Grasping Characters of the Stories

HAN XIE^{†1} TAKESHI NISHIDA^{†2}

Abstract: Along with the development of the plot, it is difficult for us to remember the character's name or the scene that the character appeared before, which presents obstacles for understanding the plot. In this paper, we propose a system to support readers to grasp the characters of the novel easier. In this system, we implement functions to achieve this both in memorizing and recalling the character using several methods, e.g. jumping to the first scene that the character appeared before, adding emoji and colors in the character's name, etc.

Keywords: Reading Support, Characters, Morphological Analysis, User Interface

1. はじめに

長い小説を読む時、物語の進展に伴い、登場人物が増える。それと同時に、人間関係も複雑になり、読者は登場人物が前にどのような場面で出てきた、どのような人物なのか思い出せず、読みづらく感じることもある。そこで、読者が長い文章を読むときに登場人物を思い出しやすくする支援機能があれば、物語をより簡単に理解することができ、読み返し時間を減らせることが期待できる。

ニュース記事や歴史に関する文章といったノンフィクションでも、背景にあたる前後の出来事などを含めて把握したほうがよい、ストーリー性のある文章についても同様に人物の把握を支援することで文章の理解がより良くなる可能性がある。

このような文章中の人間関係把握を支援するような研究としては田中らの物語の想起支援[1]や西原らの人物関係抽出[2]などが行われている。しかし、ユーザ自身が感じた登場人物に対する印象を活用するなど、ユーザが作品に能動的に関わりながら理解を深めることを支援するユーザインターフェースの検討はあまり行われていない。

本研究は、物語の登場人物に着目して、テキストから登場人物を抽出し、記憶しやすくすることと思い出しやすくすることの両面から、拡張機能を追加し、支援システムを提案する。本稿では、第2章では物語における内容理解の関連研究についてまとめ、第3章ではシステムの概要を説明、第4章ではインターフェースの詳細を解説、第5章でインターフェース機能の予備実験を紹介、そして、第6章でまとめと今後の課題を述べる。

2. 関連研究

読者が物語の内容を容易に理解するのをサポートするために、テキストを解析する様々な研究が行われてきた。

物語の内容を想起する関連研究として、田中らはユーザが自動で抽出された人物情報をマウス操作で編集することで、登場人物の関係図を閲覧できるシステムを開発した[1]。神代らは物語中に含まれる各発話文について話し手と聞き手を同定し、その会話の中身に暗黙的に示された情報から、キャラクタ同士の関係を推定し、キャラクタの関係図を出力するシステムを開発した[3]。縣らは人物間の友好敵対関係に着目して、登場人物の登場判定と退場判定を行い、物

^{†1} 神戸大学
Kobe University

語の読み進み具合に応じて、可視化する手法を提案した[4].

登場人物の抽出手法の関連研究としては、米田らは未知の人物名に対しても抽出できるように、動作するよう各単語の主語としての出現頻度と述語の性質から登場人物を推定するという手法を提案している[5]. 馬場らは形態素解析を利用して、「身体を表す語」や「基本的な性格表現用語」などのリストを作成し、係り受け解析の導入により、身体特徴や性格などの人物の属性を抽出することができるシステムを開発した. 西原らは形態素解析と係り受け解析を前処理して、関係辞書と関係抽出パターンを利用し、人物関係を構築するという手法を提案している[6]. 松尾らと伊藤らはノードとエッジを付けることによって人間関係ネットワークを構築するという手法で人間関係を可視化する[7][8].

3. プロトタイプシステムの概要

本プロトタイプは、読みたい文章をアップロードすると人名を抽出し、人物が最初に出るシーンにジャンプするや人名に絵文字を付けるなどの機能を追加することで読みやすく表示することができるウェブアプリである.

本システムの概要を図1に示す. ユーザがテキストファイルを上ロードした後、システムは自動的に人名抽出し、人物リストを作る. 人物リストはボタングループの形式で、本文の隣に貼っている. ユーザがテキストを読む時、気になる登場人物があると、すぐ隣のリストから人名を対応して操作できる. 人名を選択すると拡張機能のオプションが出てくる.

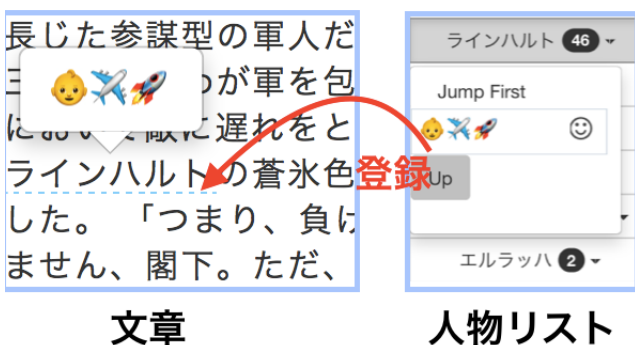


図1 プロトタイプ概要

ユーザが登場人物に対して前に出たことあるかどうか迷う時に、最初に出るシーンを提示すると人物を思い出すことができると考えている. そこで、最初のシーンにジャンプする機能を追加する.

次に、登場人物を覚えやすくするために、その人物の性別、性格などの特徴が分かれば、人物像が想像しやすいと考える. そこで、システムには人物に絵文字追加機能を追

加している. 絵文字は性別、職業、場所また感情などの特徴的なアイコンがあり、スマートフォンの普及によって人々の日常生活に親しむから、共感しやすい. 読者は登場人物に対して自分の理解を基づいて絵文字をつけることで、登場人物を覚えやすくできると考える.

4. インターフェース詳細

4.1 主画面

ユーザは、インターフェースの送信画面(図2)からテキストファイルを選択して、「アップロード」ボタンを押すとファイルをアップロードする. テキストの分析が終わったら、読む画面(図3)に進む. 主画面にテキストと人物リストのボタングループがある.

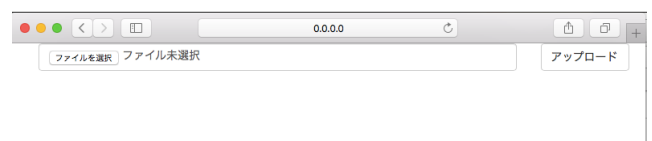


図2 送信画面

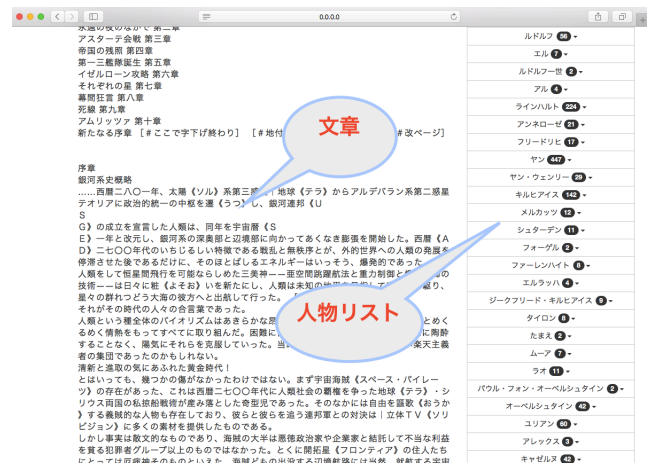


図3 主画面

4.2 人名の抽出

アップロードしたテキスト文書から、人物抽出を行う. 本研究は、形態素解析ツール MeCab[a]を利用する. 品詞が「名詞-固有名詞-人名-姓」もしくは「名詞-固有名詞-人名-名」と判定された単語を抽出して人物リストに入れる. また、姓と名を分けて抽出しているため、姓と判定した単語の次の単語が名と判定すると、フルネームとして抽出している.

4.2.1 人名辞書の拡張

MeCab による人名抽出には、システム辞書に基づいて判断するため、辞書に登録していない単語は人名として抽出することができない. 例えば、「森岡耕介」という人名が出

a) <http://taku910.github.io/mecab/>

る時、辞書には「耕介」という名前が登録していないため、「森岡 固有名詞-人名-姓」は「耕 固有名詞-人名-名」「介 名詞-サ変接続」と判断している。そこで、本研究では人名辞書を追加することで、抽出の精度を上げるよう工夫している。

4.3 人物リスト

抽出した人名が、ボタングループの形式で、テキストの隣に貼っている (図3)。リストには人物がテキストにおける出場順に並んでいる。一つの人物ボタンには人名とテキストに出る回数が並んでいる。



図3 人物リスト

人物ボタンを押すと機能リスト (図4) が出ている。



図4 機能リスト

4.3.1 最初のシーンにジャンプする機能

インターフェイスで機能リストの「Jump First」という選択肢を選ぶと、この登場人物が最初に出るシーンにジャンプする。また、テキスト内に出現するこの人物の名前をすべてハイライトにする。そして、人物リスト最後の「Back」ボタンを押すと、ジャンプする前の読んでいたところに戻る (図5)。

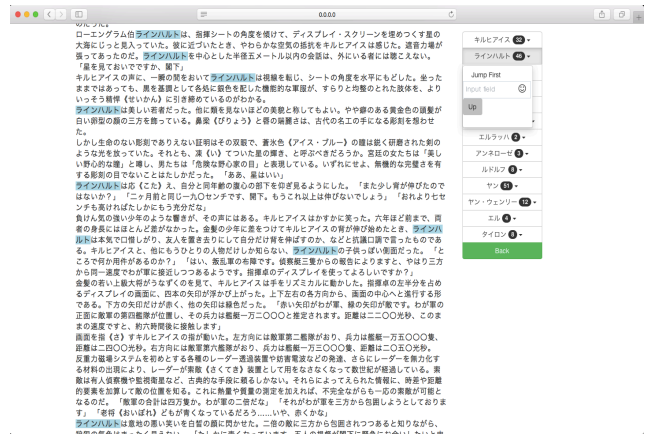


図5 最初のシーンにジャンプさせる

4.3.2 絵文字を追加する機能

インターフェイスに、絵文字マークをクリックすると絵文字テーブル (図6) が出て、ユーザは登場人物に対して自分が付けたい絵文字を選択し、「Up」ボタンをクリックすれば、テキスト内にこの人物の名前の部分が点線で表示される。また、その名前をクリックすると、選択された絵文字が現れる。(図7)



図6 絵文字の追加

的要素を加算して敵の位置を知る。これに熱量や質量の測定を加えれば、不完全なが、
「敵軍の合計は四万隻か。わが軍の二倍だな」「それがわが軍を三方から...」
「おいばれ」どもが青くなっているだろう.....いや、赤くかな」
ラインハルトは意地の悪い笑いを白哲の顔に閃かせた。二倍の敵に三方から包圍され、
狼狽の気色はまったく見えない。「たしかに青くなっています。五人の提督が閣下に
し込んで来られました」「ほう、おれの顔も見たく...」
「いや、会ってやるさ.....奴らの蒙を啓(ひら)く九...」
ラインハルトの前に現れたのはメルカツ大将、シュターデン中将、フォーゲル中将、
将、エルラッハ...だった。ラインハルトの言う「老将」たちである。しかし、
もしれない。最...カツでもいまだ六〇歳には達しておらず、最年少のファ...
しかなかった。ラインハルトたちのほうが若すぎるのである。「司令官閣下、意見具
ありがとうございます」
...が述べた。ラインハルトが生まれる遙か以前から軍籍...
も豊富な知識と...している。中背で骨太の体格と眼そうな両眼をはぶいては特
の実績と声価はラインハルトなどよりずっと大きいであろう。「卿らの言いたいこと
メルカツの示した儀礼に形ばりの答礼をして、ラインハルトは先手を打った。「
る、そのことに私の注意を喚起したいというのだろう」「さよう、閣下」

図7 絵文字提示

5. 予備実験

本インターフェースの絵文字機能に関して、容易に絵文字を付けることができるか、利用者がどのような絵文字を付けるか、その他どのような改善点があるかを調べるために予備実験を行った。

5.1 実験内容

本インターフェースを利用して、実験参加者は登場人物に絵文字を付けながら、事前に準備したテキストを読む。その後、インターフェースの絵文字機能に対する感想を尋ねた。

今回は、絵文字機能が物語を読む際に効果があるかどうか、また他の種類の文書（例えば新聞記事など）に応用できるかどうか、違う種類のテキストの使用感はどう違うのかなどを確認するために、フィクションとノンフィクション2種類のテキストを用意した。

フィクションの方は、田中芳樹「銀河英雄伝説」の第一巻の第一章を用いた。「銀河英雄伝説」は登場人物の名前が全部カタカナで、数が多いという特徴があり、本システムに有用性が高いと考えられる。ノンフィクションの文書は、大学生6人の18歳選挙権に関する対談記事である[b]。

実験参加者には、二つのテキストをそれぞれ15分間ずつ読み、読みながら登場人物に自分がふさわしいと思う絵文字をつけるように指示した。

5.2 実験対象

関西圏の国立大学の大学生3名が実験に参加した。そのうち、2人が日本人、1人が中国人留学生であった。日本人の一人は「銀河英雄伝説」を読んだことあった。

5.3 実験結果

実験者3人がノンフィクションに対してつけた絵文字を表1、フィクションに対してつけた絵文字を表2に示す。被験者ごとに読書のスピードに差があるので15分間で絵文字を付けた登場人物の数には差がある。中国人のCは少なく、フィクションの作品を一度読んだことがあったAは多くなっている。

ノンフィクション

	A	B	C
原田			
田中			
大隈			
菅			
町田			
塚田			
松浦			
川合			
黒石			

表1 ノンフィクションの絵文字

フィクション

	A	B	C
キルヒアイス			
ラインハルト			
メルカッツ			
シュターデン			
フォーゲル			
ファーレンハイト			
エルラッハ			
アンネローゼ			
ルドルフ			
ヤン			
ヤン・ウェンリー			
エル			
タイロン			

表2 フィクションの絵文字

結果から見ると、登場人物の理解や絵文字の選択標準は人それぞれである。例えばノンフィクションの文章については司会者である「原田」にカメラ絵文字をつけた人、男性を表す絵文字をつけた人がいた。一方で、フィクションの文章の「ラインハルト」という人物に3人も男性を表す絵文字を選択するなど共通の絵文字を付ける場合もあった。そして、人物に対して「男性」「年長の男性」「男の子」「女性」という絵文字を選択する率が高く、絵文字は性別

b <http://blogos.com/article/189716/>

を表しやすいと考えられる。

5.3.1 絵文字選択基準

どういった基準で絵文字を選択するかという質問に対しては、ノンフィクションは会話文だけなので、ぼんやりしたイメージから自分の理解の人物像を基準として絵文字を選択した。また記事の話による選挙で投票したことがあるかどうかという事実に基づいてマルとバツを選択し、メモとして絵文字機能を利用している人もいた。

フィクションの方では、テキストに出た言葉から絵文字を選択する(例えば飛行機)、またテキストに書かれている容貌の特徴などの情報から得た人物像によって絵文字を選択できる。

比較として、フィクションは容姿の紹介や姿があることが多く、人物像をイメージしやすいので、ノンフィクションより絵文字の選択がしやすいという声があった。

5.3.2 絵文字機能の評価

絵文字機能の使用についての感想を聞くと、3人ともフィクション文章に対して、登場人物に絵文字付くことは読書体験にプラス効果があるという評価がある。特に外国人として分かりにくいカタカナ人名に絵文字提示があれば、テキストがわかりやすいと述べている。そして、どのような絵文字を選択するかを考えることで、文章を理解することができるということが分かった。また、絵文字を追加した名前の下に点線を引かれ、その後出場してもこの人物が前に出たことがあるということが提示できる。

問題点として、絵文字テーブルから絵文字を探しにくいという指摘がある。今の絵文字テーブルが、大まかに五種類に分けて、一種類に100ぐらいの絵文字が入るので、適切な絵文字を探すことは容易ではない。普段どれくらい絵文字を利用するかによって、その難易度には大きな個人差がある可能性もある。これらは今後の課題として検討したい。

6. まとめと今後の課題

本研究では物語を読む時、登場人物を把握しやすくするシステムを提案した。初登場シーンにジャンプする機能と人名に絵文字を追加する機能を実装した。

予備実験の結果と関連研究に基づいて、今後の課題としては以下のものが挙げられる。

6.1 絵文字機能の改良

絵文字を探しにくいという問題点に、まず絵文字の量と分類方法を改善すべきである。今のインターフェースの絵文字テーブルはIOS用の絵文字機能のテーブル分類である。テキスト読む時に使いやすいように分類することが今後の課題である。

6.2 人物抽出の精度

本システムでは、MeCabの辞書に基づいて人物の抽出を

行なっている。そこで、辞書に載っていない人名も正しく抽出するために、テキストを係り受け分析などの手法で解析すれば、抽出精度を上げることが期待できるかもしれない。他にも、澤田らが提案したN-gramを用いて抽出モデルを構築する手法[9]やTf-idfなどの抽出方法がある。

6.3 機能の追加

インターフェースとして、機能を追加することも今後の課題である。登場人物を把握するために、人物の容姿の特徴や性格の特徴などを把握することは無論、登場人物の間の人間関係をわかりやすくすることも期待している。関連研究は人間関係図を構築するに関して研究が多いが、読む途中にインターフェースでどのように人物の関係をわかりやすく提示することが今後の課題である。また、機能を追加する同時に、登場人物に関しての特徴などを提示するとき、物語をネタバレするかどうかとも考えべきだ。

参考文献

- [1] 田中翔太郎, 岡部誠, & 尾内理紀夫. (2014). 物語の内容想起支援インタフェースの開発. 第55回プログラミング・シンポジウム予稿集, 2014, 7-15.
- [2] 西原弘真, & 白井清昭. (2015). 物語テキストを対象とした登場人物の関係抽出. 言語処理学会第21回年次大会発表論文集, 21, 628-631.
- [3] 神代大輔, 高村大也, & 奥村学. (2008). 物語テキストにおけるキャラクタ関係図自動構築. 言語処理学会第14回年次大会発表論文集, 14, 380-383.
- [4] 縣啓治, 伊藤雄一, 高嶋和毅, 北村喜文, & 岸野文郎. (2010). 物語テキストから進行状況に応じて登場人物の存在状態と関係を推定する手法. 第18回インタラクティブシステムとソフトウェアに関するワークショップ, Dec.
- [5] 米田崇明, 篠崎隆宏, 堀内靖雄, & 黒岩真吾. (2012). 述語情報を利用した小説の登場人物の抽出. 言語処理学会第18回年次大会発表論文集, 18, 855-858.
- [6] 馬場こづえ, & 藤井敦. (2007). 小説テキストを対象とした人物情報の抽出と体系化. 言語処理学会第13回年次大会発表論文集, 13, 574-577.
- [7] 松尾豊, et al. "Web上の情報からの人間関係ネットワークの抽出." 人工知能学会論文誌 20.1 (2005): 46-56.
- [8] 伊藤正彦, & 赤石美奈. (2011). 史料データにおける活動履歴に基づく人間関係変化の可視化手法. じんもんこん 2011 論文集, 2011(8), 21-28.
- [9] 澤田真吾, & 浜中雅俊. (2013). Web中の文章とリスト構造体を用いたバンドメンバー自動収集手法に関する研究. 研究報告エンタテインメントコンピューティング (EC), 2013(9), 1-6.