

小説内の動的人物相関図を用いた読書システム

太田 彩[○] (東京工業大学), 脇田 建 (東京工業大学)

A reading system using dynamic character correlation in novels

Aya OTA and Ken WAKITA

ABSTRACT

Books are reliable and excellent sources of information. However, reading requires long time constraints and the need to supplement the information from the text. Therefore, even if you have the desire and need to read, you may give up. For this reason, researches have been conducted to automatically extract information from novels and to assist the reader by representing the content of the story in a single page. However, there is a risk of spoiling. In this study, we propose a method to support reading by presenting a character correlation. In order to reduce the risk of spoilers, we propose a character correlation chart that changes as the reading progresses. We used two methods to add information: automatic extraction using natural language processing techniques and manual addition. In addition, by sharing the input information with others, we were able to reduce the time and effort required to input information. In order to verify the effectiveness of the proposed method, we constructed a reading support system using Aozora Bunko data. Finally, we discussed the degree of achievement and issues in comparison with the objectives.

Keywords: visualization/dynamic graph/natural language processing

1. はじめに

1.1. 背景

読書はメジャーな趣味の一つであり、信頼性の高い優秀な情報源でもある。しかし近年インターネットの普及により読書を習慣的に行う人と全く読書をしない人に二極化してきた。実際に全国大学生生活協同組合連合会が大学生を対象に行った第 55 回大学生生活実態調査 概要報告によると、一日の読書時間が 0 時間である人が 48.1 %、その一方で毎日平均して 1 時間以上読書を行っている人は 26.8%であった。世の中では読書の効用を訴える書籍が数多く出版されており、読書する意欲はあるが実行に至らないという人々も存在する。その原因を推測すると、文字のみから内容を理解するという性質が他の情報入手媒体に比べて煩わしく感じやすいという点や、日々の忙しさから読書の時間が断片化し内容を思い出すための労力がかかるという場合が考えられる。

これまでも自然言語処理の技術を使って読書の補助を行うという研究はされてきた。本文から自動で登場人物や人物間の関係性を自動で抽出しユーザーに提示するものもあった。しかし小説全体に出てくる人物と関係性を抽出するため、まだ読んでいない部分の情報を意図せず獲得し小説による感動が薄れてしまう可能性があった。そのため小説内の情報を一枚の絵で表すのではなく、ユーザーによって取得する情報の範囲を決めてもらう必要がある。加えて先行研究では小説内の暗に示された関係などは抽出することができず、抽出した情報に不足が出てしまうという問題もあった。そのため情報抽出の方法も工夫する必要がある。

小説を対象とした可視化研究も行われてきた。しかしその研究は文学者による小説の分析を支援するものであり、

一般の読者の理解を促進するものではなかった。そこで関係性の内容を詳しく記述したり、全体の話の流れが一目で分かるようにするなど、読書に特化した設計を行う必要がある。これらの問題に着目し読書の補助に特化した可視化手法というのは今までに提案されてこなかった。そこで本研究では本の内容を提示することで読書の補助を行う手法を提案する。

1.2. アプローチ

本研究では**時間発展する人物相関図**を提案する。本研究では特定のページまでに出てきた登場人物・関係のみを人物相関図に表示することで、意図せず先の情報を取得することを防止する。加えて人物相関図を作成するために必要な登場人物と関係性の情報の追加は、計算機による自動抽出と人間による手作業の二通りで行う。計算機により自動抽出することでユーザーの情報追加の労力を大幅に削減する。それに補足して人間が足りない情報の追加・情報の修正を行うことで高い精度の登場人物と、繊細な関係の描写を実現する。しかし新しい小説を読むたびに情報を個人で修正することは作業量が多く現実的では無い。また同じ小説を読む人にとって登場人物や関係性の情報は共通であるということができる。そのため登場人物や関係の情報を他者と共有し、個人の作業量の削減、読書体験の共有を行う。本研究では以上の手法を提案し、1.1 で述べた問題を解決したいと考える。

1.3. 貢献

読書を支援するために小説の内容を表す研究は以前にも行われてきた。それらの研究では人物相関図を一枚の絵で表したり、登場人物の時系列ごとの位置を図によって示してきた。本研究の貢献はネタバレの可能性に着目し、**読書の進捗に合わせて内容が変化する人物相関図**を提案したことである。読書の進捗をページ数で管理し、現在のページ

までに出てきた登場人物と関係性を人物相関図に表示した。以前にも小説の登場人物や関係を計算機によって自動抽出する研究は行われてきたが、本研究では日本語の曖昧性や登場人物の定義の曖昧性に着目し、人間と計算機が共同で情報を抽出する方法を提案した。加えて多数の人間が小説の情報を共有することで抽出の手間を削減させた。最後に提案手法を元に読書システム構築することができた。

1.4. 残りの構成

本論文の以降の章では以下の議論を行う。2章では関連する研究についての紹介を行い、本研究の立ち位置を明らかにする。3章では提案手法の詳細の紹介と提案手法の効用を確認するためのシステムの要件定義を行う。4章では実際に開発したシステムの紹介を行う。5章では目標の達成度について議論を行い、今後の課題について言及する。

2. 関連研究

2.1. 小説からの情報抽出

小説の内容を理解する上で**登場人物の名前と人物間の関係**は大切な要素である。そのため以前から小説からそれらの情報を抽出するための研究がなされてきた。米田(米田崇明 et al. 2012)では人物候補の**局所出現性**と**述語情報**を利用し、小説の本文から登場人物を自動抽出した。西原(西原弘真 and 白井清昭 2015)は辞書から手作業で作った関係辞書と、10編の小説を元に人手で作ったパターンと半自動抽出して作ったパターンを元に、小説から関係性を抽出した。この研究では青空文庫の小説10編を対象に人手で抽出した関係性のうち、どれだけ多くの関係性を抽出できたか実験した。関係性の完全一致は適合率は53.6%、再現率は17.9%、F値は26.8%と、低い値が出ている。この値の低さの原因としては正しいパターン生成の難しさ、人物抽出の精度の低さなどがあげられている。

機械による自動抽出によって人手で抽出した情報と全く同じ情報を抽出することは難しい。加えて小説においてどこまで登場人物や関係性として認識するかには議論の余地がある。例えばアガサクリスティーの「そして誰もいなくなった」では最初に兵隊島に関する記事について述べられた部分が存在する。その部分では兵隊島を新たに買った人物は誰かという議論が行われており、「ガブリエル・タール」や「若き貴公子L卿」などの人物の名前を挙げられている。しかしこれらの人物は物語の本筋に関係のない人物であり、このような重要度の低い人物を全て抽出することは逆に読みにくさを助長する可能性がある。また登場人物の飼犬など人間でない動物が重要な働きをする場合も存在するが、動物を登場人物として扱うか否かを一概に決めることはできない。

以上の議論より読書の補助という目的において、全ての登場人物や関係性を計算機を用いて自動抽出することは得策ではないと考えた。そこで本研究では計算機による情報抽出に加えて、人間の手により情報の追加・修正を可能にすることにより、読書に有用な情報の提示を実現する。

2.2. 物語の内容の可視化

物語の内容を一眼で理解したい、関係を図を用いて整理したいという願望は度々生じる。そのためテレビ番組や映

画の内容を一枚の絵で表す研究は今までにも行われてきた。Filipovら(Filipov et al., n.d.)はテレビ番組「Game of Thrones」の中の84人の登場人物と216個の関係を放射状に配置することで一枚の絵で表した。Tanahashiら(Tanahashi, Hsueh, and Ma 2015)は時間と共に変化する映画内の登場人物の位置情報を複数の曲線を用いて表すアルゴリズムを提案した。以上の研究では物語の内容を一枚の絵で表しているが、まだ読んでいない部分の内容も意図せず知り読書の楽しみが損なわれてしまう恐れがある。本研究ではネタバレを防止しながら読者の理解を補助するために、**読み進めた時点までの関係を表示するインタラクティブな可視化**を提案する。

2.3. 文章に関する可視化

テキストに付随する情報をテキストに関連づけて表示する研究は行われてきた。Wuら(Wu, Zhao, and Amiruzzaman 2020)では音声の長さや音声認識の信頼度などの情報をテキストに関連づけて可視化をした。Johnら(John et al. 2019)は文学者向けに登場人物の関係性の分析を可視化で支援するシステムを提案した。本研究では(John et al. 2019)で用いられたNode Link diagramとAdjacency matrixを参考に情報の提示を行う。

3. 提案手法

読書をする際に読む意欲や必要性があっても、「登場人物が多すぎて覚えきれない」や「歴史的な名前や外国の名前が覚えにくい」「前回読んだ場所までの内容を忘れてしまう」などの理由から、意欲が下がり読書を断念してしまうことがある。本研究ではこれらの読書の意欲を下げる要因を**読書の手間**と名付ける。読書の手間とは具体的には以下のような例が挙げられる。

- 前回読書をした時から時間が経ち内容を忘れてしまったので、また読み直す必要がある
- ずっと前に読んだ本の内容が思い出せない
- 登場人物が多い
- 名前が複雑で覚えられない
- 登場人物間の関係が覚えきれない
- どこで出てきた関係なのか忘れてしまう

これらの大半は記憶力を補助することで解決できると考え、本研究では小説の内容を可視化することで読書を支援することを決めた。手法を考える上で注目したのは**人物相関図**である。人物相関図とは物語内の登場人物や関係性をネットワークとして一枚の絵で表したものである。人物相関図は映画やドラマの宣伝や紹介を行うために、よく公式サイトや広告などに用いられている。加えて話数が多いアニメや小説において、登場人物や人物間の関係を整理して確認するためにも人物相関図は有効である。内容について理解を深めるために個人で相関図を作っている人もいる。本研究では**話の整理や表示に有効である**という性質に着目し、人物相関図を用いることを決めた。

3.1. 要件収集

人物相関図を用いて読書を支援するための要件を考えるために、実際に手書きで人物相関図を描きながら読書を行った。実験に使用した小説は以下の通りである。

- アガサ・クリスティー
「そして誰もいなくなった」
- イワン ツルゲーネフ 「はつ恋」

ネタバレの防止が研究課題であるため、特にネタバレが忌避されるジャンルであるミステリー小説の「そして誰もいなくなった」を選定した。加えて感情の機微や関係性の詳細が重要である恋愛小説である「はつ恋」を選定した。

続いて読書の方法について説明を行う。どちらも紙の書籍で読書を行い、一章読み終わるごとに人物相関図をiPad上で手書きで描く。また「そして誰もいなくなった」ではどの章でその関係が生まれたか把握するために、登場人物や関係を章ごとに色分けして描く。その結果プロローグとエピローグも合わせて、「そして誰もいなくなった」では18章分、「はつ恋」では10章分の人物相関図が出来上がった。

その結果3つの問題を発見することができた。

1点目は**1枚の絵で表すと重大なネタバレを生じてしまう**という問題である。

ネタバレとは一般に物語の展開や結末といった重要な部分を暴露すること、ネタバレをばらしてしまうことを意味する。先に重要な部分を知ってしまうと、コンテンツを体験したときの感動や楽しさを減らしてしまう可能性がある。ネタバレを回避するために、ネタバレを含むレビューコメントを検出する研究なども行われている(前田恭佑 et al. 2019)。

人物相関図にはネタバレをひき起こす可能性がある。例えば「そして誰もいなくなった」ではエピローグ部分で殺人犯が誰であるか明らかにされる。しかしFig. 1のように小説内の全ての関係を人物相関図に表示すると、ヴォーグレイヴ判事が犯人であることが一目で分かってしまう。その他にも物語の性質上重要な情報を意図せず取得してしまう。これは明らかに読書を読む楽しみや感動を減少させてしまう。本研究では人物相関図におけるネタバレを改めて小説の先頭から一定の位置までの内容を示す人物相関図が、小説のそれ以後にしか記述されていない登場人物・関係性を包含することと定義する。今後の議論では人物相関図におけるネタバレを**ネタバレ**と表記する。本研究ではネタバレを防止することを目的とする。

2点目に**人間関係の描写は複雑なものが多い点**である。小説には親子や恋人など単語で言い表せる関係もあれば、簡潔に言い表せないものも存在する。例えば「はつ恋」の中には様々な関係が存在する。小説のヒロインであるジナイダには取り巻きのいるが、その中の一人マイダーノフとの関係は「詩的な素質の相手はしてもらいが信用はしていない」など単語で表せない。同様に主人公の母は父に対して「ジナイダとの不貞を問いたです」ことをしており、この関係を表すには長い記述を要する。このように文学において繊細な感情の変化や関係性は重要な要素であり、全ての関係性を単語で記述することは難しい。そのため詳細な関係を記述し、それを人物相関図にて確認できる必要がある。

3点目は**読書を進めながら人物相関図を手書きすることは難しい**ということである。登場人物や人間関係の数が多い場合、それらを線の交差や湾曲が少なくなるよう配置することには非常に手間がかかる。加えて読書を進めながら人物相関図を書くと、その先の情報が分からない中レイアウトを決定するため、最終的に読みにくいレイアウトが出来上がる可能性がある。実際に「はつ恋」を読んだ際にもFig. 2のように配置に問題が生じ、8章から9章にかけて大幅にレイアウトを変更した。手書きの人物相関図の作成には非常に労力がかかるため、読書を進めながら人物相関図を手書きすることは現実的ではない。

以上の考察により提案手法の果たすべき小目標を以下のように設定した。

- ・ (P1) 小説のどの部分を読んでも、人物相関図によってネタバレを起こさないこと
 - ・ (P2) 小説内の長文でのみ表される人間関係を取扱い、ユーザーがそれを確認できること
 - ・ (P3) 入力した情報をもとに人物相関図を自動でレイアウトしてくれること
- この小目標をもとに議論を進めていく。



Fig. 1 Hand-drawn character sketches for "And Then There Were None."

3.2. 提案手法の概要

(P1)の実現のために本研究では**時間発展する人物相関図**を提案する。比較のために人物相関図を「与えられた小説の登場人物のすべて、あるいは一部について、その存在をノードで、登場人物間の関係をリンクで表現したもの。」と定義する。一方で時間発展する人物相関図は「小説の先頭から始まる任意の一定の範囲の内容についての人物相関図を生成する能力を有するインタラクティブな可視化システム」と定義する。時間発展する人物相関図を用いて、読んできた中に登場した人物・関係のみを人物相関図に表示することで、ネタバレのない可視化を実現する。

(P3)の実現のために人物相関図の自動レイアウトを行った。レイアウトに必要な情報は現在までのページに出てきた登場人物と関係のリストである。それらの情報をもとに**人物相関図**と**人物相関表**という2種類の表示を行う。人物相関表はノードリンクを用いたグラフで、各ノードは登場人物を表し各リンクは人間関係を表している。人物相関表は隣接行列を用いたグラフで、登場人物は縦と横の軸に並んでいる。関係は有向であると考え、関係の元の人物と行かつ関係の先の人物の列であるセルを着色することで表現する。

3.3. ユースケース

大学生の桃子さんは授業で読書の重要性を説かれ、教養を深めるために読書をすることにした。また今まで読んだことのある小説は主に日本の娯楽小説だったため、外国文学を読むことにした。そこで読む本に迷ったため、友人であり読書家である桜さんにおすすめの本を聞くと、イワン・ツルゲーネフの「はつ恋」を勧められた。桃子さんは「はつ恋」を読み始めたが、登場人物の名前が覚えきれず中々読み進めることができない。その時以前紹介された読書システムを思い出し、システムを使って読書をすることにした。まず、システム上で「はつ恋」を読めるようにした。

1週間後花子さんは前回読み終わったところから読書を再開することにした。しかし期間が空いたために、登場人物の名前を忘れてしまった。そこでシステムにおいて人物相関図を表示して、内容を思い出すことにした。人物相関図を見ると今までに出てきた登場人物の一瞥を見ることが

でき、特に重要な人物が「ウラジーミル・ペトロヴィチ」と「ジナイダ」であることが確認できた。その後読み進めていくと「ルーシン」という人物が出てきた。「ルーシン」が誰か思い出せなかったのも、システム上で「ルーシン」が本文中のどこに出ているのか確認し、ルーシンは医者であり取り巻きの中で最初に会った人物だと思いつくことができた。

桜さんにそのシステムを使っていることを話すと、桜さんは小説の考察をすることが好きなので、システムを使って考察をまとめてみることにした。先に桃子さんが登場人物の情報を加えてくれていたので、桜さんは手軽に情報を追加することができた。また各々自分が関係だと思った情報を追加していき、二人とも最後まで本を読み情報を入力することができた。その後二人でシステムを見ながら感想を語り合うことにした。システムを見ることで二人の解釈の違いや、印象に残ったシーンを気軽に見つけることができた。このように桃子さんと桜さんは読書体験を楽しんだ。

3.4. システムの要件設定

ユースケースの内容からシステムの要件を以下のように定義した。

- (N1) システムの上で読書ができること
- (N2) 小説に関する重要な情報を手軽に追加できること
- (N3) 情報を後から編集できること
- (N4) 特定のページまでの内容の人物相関図を表示すること

これらの要件を詳しく順番に確認していく。

3.4.1. システムの上で読書ができること

ユースケースにて桃子さんは小説を読み進めながら、情報の追加・編集、人物相関図の確認を行なった。このようにシステムを使用するためには、システムと小説の本文を交互に見る必要がある。しかし小説の本文とシステムが分離していると視線移動が多くなり、システムを利用することが負担になってしまう可能性がある。そこで本研究ではシステム上での読書を可能にすることで、システム利用の利便性を向上させる。具体的な読書機能としては小説のインポート、本文の表示、ページの移動、しおりをつける、などの機能を実装する。今回は限定的に著作権の消滅した作品を中心に電子書籍として公開しているサイトである青空文庫上の書籍をこのシステムで読めるようにした。

3.4.2. 小説に関する重要な情報を手軽に追加できること

2章において、小説における登場人物や関係性の認識は非常に曖昧であり機械による自動抽出が難しいことを議論した。そのため本研究では機械で抽出した情報を人間の手で修正し、小説の理解に適した登場人物と関係のリストを作成する手法を提案した。その手法を実現するためにシステムには小説の本文から情報を自動抽出する機能と人間による情報追加を可能にする機能の2つの機能が必要である。この機能を実装する上で最も重視すべき事項はこの機能がユーザーによって使いやすい機能であるかという点である。本研究の目的は読書を円滑にすることであるため、情報の追加に手間取り読書を停滞させることは本意ではな

い。そのためなるべく少ない手順で情報を追加できるように機能の詳細を決定していく。

3.4.3. 情報を後から編集できること

この機能は誤った情報を入力したとき、機械による自動抽出に誤りが存在したとき、読書が進むうちに状況が変化した時におこる問題を解決するために必要である。

情報を手動で追加したり、アルゴリズムで情報を自動抽出する時にミスが起こらないと考えることはできない。しかし誤った情報をそのまま保存しておくことと映し出された人物相関図も誤った情報を含み、人物相関図を見ることで読解を阻害する可能性がある。そのため誤りを見つけた際には即座に修正できる必要がある。

加えて小説は読書が進むうちに状況が変化する場合がある。例えばユースケースにおいて序盤に出てきた「背の高い少女」は後に「ジナイダ」という名前であることが判明した。名前が明かされた後、人物相関図における人名の表示を「背の高い少女」ではなく「ジナイダ」に変えたいと思う場面も出てくるはずだ。そのためこの場合にもすでに入力した情報を修正できる必要がある。

以上の理由から情報を後から修正できることをシステムの要件として設定した。その要件を達成するためにも、情報を追加するときと同様にユーザーが情報を修正しやすいような画面の配置、機能の設計が重要である。

3.4.4. 特定のページまでの内容の人物相関図を表示すること

この要件は (P1) と (P3) を達成するために非常に重要であり、この4つの中で最も大切な要件である。この節では可視化において対象となるデータと想定される動作や動機について考える。今回可視化に使用データは小説の本文と登場人物のリストと人物関係のリストと人物間の関係のリストと登場人物のページごとの出現回数である。登場人物は名前という属性をもち、人物間の関係は誰から、誰への、どんな関係、好感度の値という属性をもち、両方共通で初めて登場したページ数を持つ。次にユースケースをもとに想定される動作を考える。ユースケースでは主に登場人物を追加するとき、内容を思い出すとき、特定の人物を思い出すとき、情報を修正するとき、二人で感想を話すときにシステム上で情報を確認していた。各シチュエーションにおいて必要な可視化要件を考える。登場人物を追加するときには機械によって抽出された単語の一覧とそれらの単語が文中のどこに出てきたのかの表示が必要である。内容を思い出すときや二人で感想を話すときには関係の全体像を確認できること、人物と人物間の関係が確認できること、1番重要な人物を見つげられること、時系列によって変化する関係図が眺められること、人物相関図からその登場人物・関係が出てきた箇所を確認できることなどがあげられる。特定の人物を思い出すときにはその人に関する情報を収集するために人物を出てくる時間や関係する人物によって絞り込むことができること、特定の人物と関連する人物だけを取り出せること、特定の人物が登場するページを見つげられることなどの絞り込み機能が必要である。情報を修正するときは (N3) に関わってくるが、現在の登場

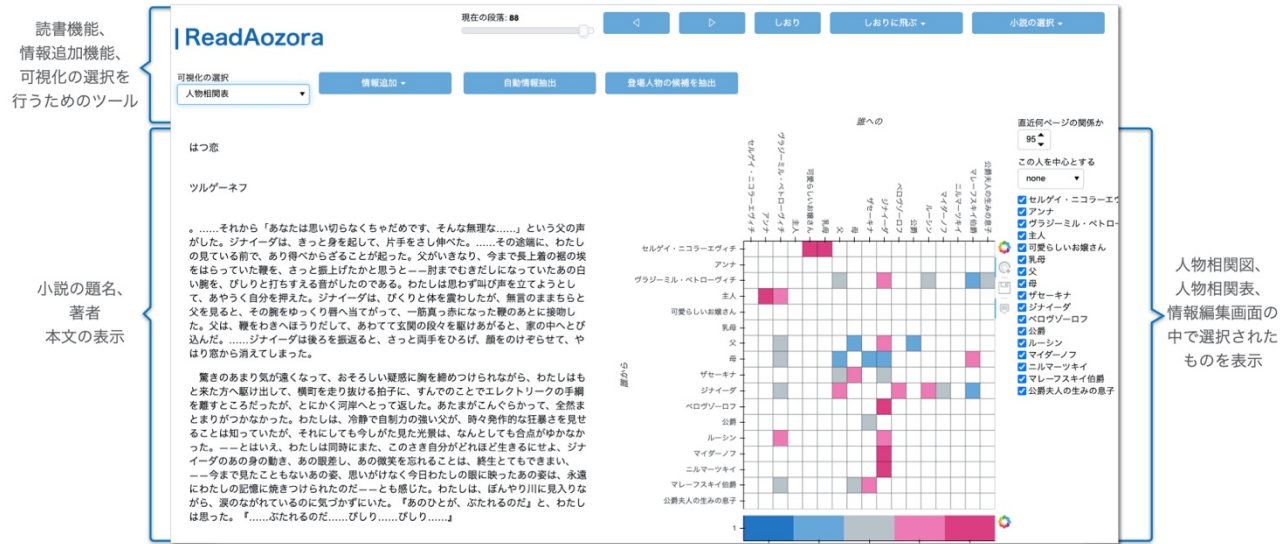


Fig. 3 A view of "First Love" on 'ReadAozora'.

人物と人物間の情報の一覧が確認できることが必要である。これらの細かい要件をもとに人物相関図など様々な可視化を実現する。

本章では要件分析をもとに、小目標を設定した。加えてその小目標をもとに提案手法の概観を紹介し、ユースケースからシステムの要件を設定した。4章においては実際のシステムについて詳しく説明を行う。

4. 可視化システム

4.1にて設定した要件をもとに Fig.3 のような読書システム ReadAozora を開発した。この節では要件に沿ってシステムの説明を行う。

4.1. (N1)システムの上で読書ができること

様々な電子書籍リーダーを参考にし、読書に必要な機能を最低限実装した。読書画面には、小説の選択機能、本文の表示、しおりなどの機能が実装されている。「小説を選択」ボタンを押すと新しく小説を青空文庫のサイトからダウンロードするか、既にダウンロードした小説から選ぶか選択することができる。小説をダウンロードする時は、青空文庫の作品の図書カードページのリンクを貼り付ける。貼り付けると、青空文庫のデータを表示しやすい形へ成形しシステム上に保存する。ダウンロード済みの作品を選ぶ際は、作品の一覧から自分の読みたい作品を選択する。選択された小説の本文はシステムの左下に表示され、最初に読むときは1ページ目が、2回目以降に読む場合は前回最後に読んだページが表示される。元々の青空文庫で公開しているデータには区切りがないが、そのままシステム上に表示すると読んだ部分の把握が難しくなるので、一定文字数で区切りページとして管理している。システムの上部には読書を進めるためのツールが位置している。左端のスライダーではつまみを動かして、見たいページを選択しそのページに飛ぶことができる。「◀」を押すと前のページへ、「▶」を押すと次のページへとめくることができる。次の「しおり」と書かれたボタンを押すと現在表示しているページをしおりづける。最後に「しおりに飛ぶ」ボ

タンを押すと、以前に追加したしおりのページを選択しそのページに飛ぶことができる。

4.2. (N2)情報の追加方法

この節では読んだ小説内の情報をシステムに追加する方法を紹介する。このシステムで入力できる情報は議論した通り、登場人物と関係性の二つである。システム上で入力された情報はシステムを閉じてでも保存され続ける。情報追加は手動と自動の二通りで行えるので順番に紹介する。

4.2.1. 自動抽出

自動抽出は「自動情報抽出」ボタンを押すことで、小説内の登場人物と人間関係の候補を自動で全て抽出することができる。抽出された情報は自動で保存され、その情報はシステム内で確認することができる。

4.2.2. 手動抽出

登場人物の追加方法は主に二つである。一つ目の方法は完全に手作業で追加する方法である。1番上の「情報追加」ボタンを押すと、人物を追加するか、関係を追加するか選ぶことができる。人物の追加を選ぶと、情報入力画面が出る。テキストボックスに追加したい登場人物を入力し、「以上の人物を追加」ボタンを押すと、入力した人物が登場人物のリストに追加される。二つ目の方法は計算機を用いて登場人物の候補を提案し、その中から人物を選んでもらう方法である。「登場人物の候補を抽出」ボタンを押すと、現在表示している文章内の登場人物らしい単語を抽出して表示する。候補のうちチェックを付けたものは文中で色が変化するので、その単語が本当に登場人物を指すのか確認できる。最後に登場人物として追加したいものだけチェックをつけて、「追加」ボタンを押すとまとめて登場人物のリストに追加される。

次に関係性の追加方法について紹介する。システムの上部の「情報追加」ボタンを押し、関係性追加を選択すると関係性の入力フォームが出現する。そこでその関係が誰から、誰への、どんな関係でどのくらい好意的なのかを入力

する。好感度は5段階評価で高い値ほど好感度が高いことを示す。以上の情報を入力し終わったら追加ボタンを押すことで入力した関係性はシステムに保存される。

4.3. (N3) 情報を後から編集できること

この節では既に追加した情報を修正する方法について紹介する。情報編集は情報編集画面を開くことで行うことができる。情報編集画面は Fig.4 のように登場人物編集表と登場人物の出現度のヒートマップと関係編集画面の3つから成り立つ。登場人物の情報を編集する場合は表において変更したい部分をダブルクリックし直接書き換えることができる。しかし書き換えただけではシステムに保存された情報は変わらないので右にある「変更を保存」ボタンを押す必要がある。加えて登場人物の左横にチェックをつけると登場人物が選択され、選択された人物はヒートマップ上に表示される。右の「消去」ボタンを押すとチェック付きの人物をまとめて消去することができる。これは関係編集画面も同様に操作することができる。ヒートマップにおいて縦軸は登場人物、横軸はページ数を表している。各セルはそのページにおいて登場人物の単語が何回出てきたかによって色分けされており、濃い色であるほど登場回数が多い。カーソルを各セル上におくと、そのセルが「誰のどのページでの出現回数」を表していて、正確な出現回数を確認することができる。加えて各セルをクリックするとそのセルの表すページへ飛ぶことができる。

4.4. (N4) 小説の進行度に合わせて関係図を表示すること

最後に関係図を表示する機能について紹介する。本研究では登場人物と関係を表すために人物相関図と人物相関表という2種類の表示を行う。

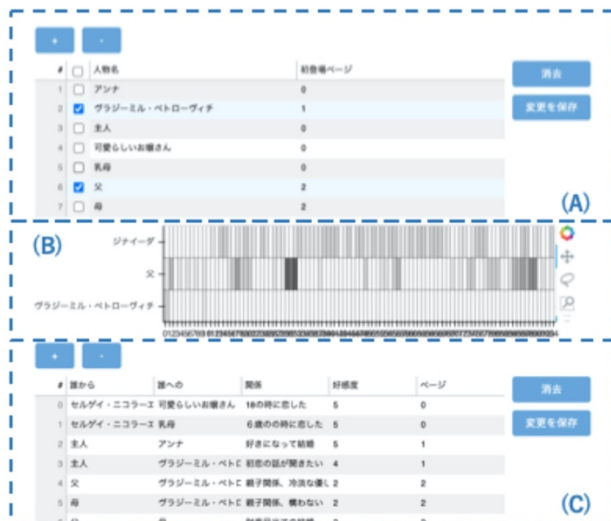


Fig. 4 Information editing screen. By editing the table, you can edit information about the characters and their relationships. (A) Table for editing characters (B) Heat map of the frequency of appearance of characters (C) Relationship editing screen

4.4.1. 人物相関図

人物相関図はノードリンクである。人物相関図は上の可視化を選択から人物相関図を選択することで Fig.5 のように表示できる。人物相関図には現在のページまでに出てきた登場人物と関係の情報が表示される。各ノードは登場人物を表し、各リンクは1番直近の関係性を表している。各ノードの大きさはその人物の持つ関係性の多さを表し、大きいほど小説において重要な人物である。ノード上でクリックをするとそのノードに繋がったリンクのみが強調され、関係する人物を一眼で確認できる。加えて各リンクはその関係の持つ好感度の値によって色付けされている。各リンクの上をホバーすると関係についての正確な情報を得ることができる。人物相関図は登場人物が少ない場合に一目で重要な人物や関係性を確認することができ、多くのユーザーに馴染み深い表示手法を採用しているため内容を理解しやすいという利点がある。

4.4.2. 人物相関表

次に人物相関表の紹介を行う。人物相関表はシステムの上部の可視化を選択から人物相関表を選ぶことで Fig.6 のように表示できる。人物相関表では現在のページまでに出てきた人物の名前が縦軸と横軸に並ぶ。各セルは縦軸の人物から横軸の人物への直近の関係を表しており、その関係の好感度によって色分けされている。またセル上をホバーすることでその関係の詳細を確認でき、クリックすることで過去の関係の一覧表を見ることができる。人物相関表の隣には表示する範囲を制限するためのボタンが位置する。これらのボタンでは、小説のどのページの関係を表示するのか、どの人物に注目した人物相関表を表示するのか、人物相関表にどの人物を表示するのかの3つを調整することができる。人物相関表の機能は以上の通りである。人物相関表の長所は登場人物の多い長い小説に対しても、機能を使えばある程度対応することができることである。このように人物相関図と人物相関表では長所が異なっているため、人物の数や知りたい情報によってビューを使い分けることを想定している。

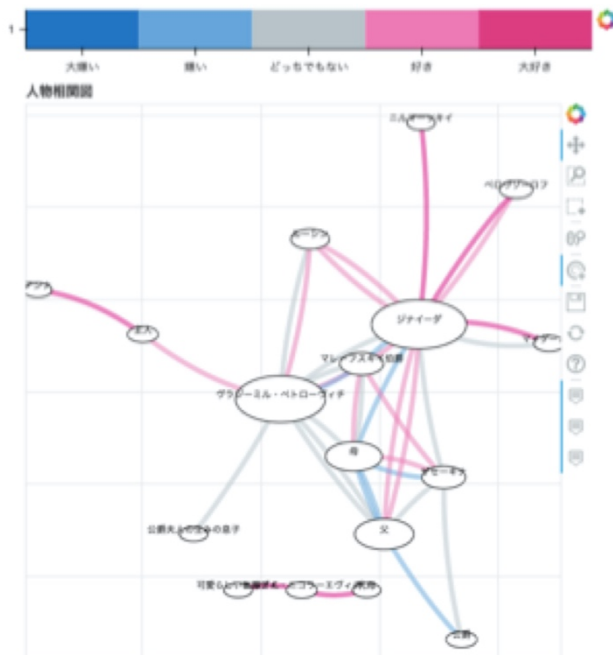


Fig. 5 A character correlation chart showing the relationships in "First Love" up to page 94.

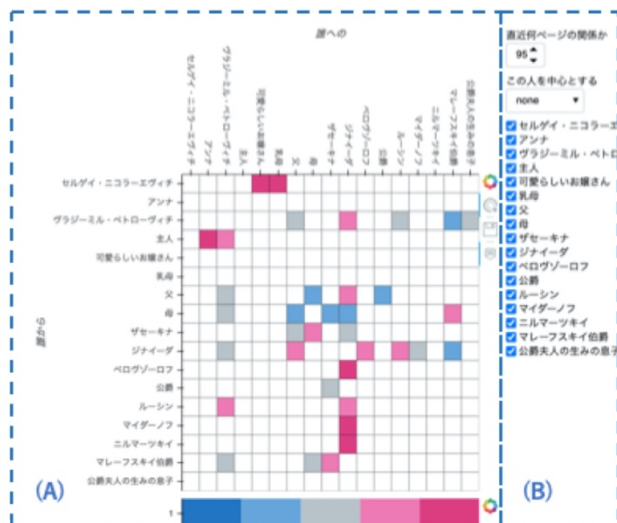


Fig. 6 A character correlation matrix showing the relationships in "First Love" up to page 94.

(A) Person correlation matrix (B) Tool for selecting the display range

5. 議論

この章では3章で定義した提案手法の果たすべき小目標を達成できているのか議論を行う。小目標は以下の通りであった。・(P1) 小説のどの部分を読んでいても、人物相関図によってネタバレを起こさないこと・(P2) 小説内の繊細な人間関係を記載し、ユーザーがそれを確認できること・(P3) 入力した情報をもとに人物相関図を自動でレイアウトしてくれること

5.1. (P1) 小説のどの部分を読んでいても、人物相関図によってネタバレを起こさないこと

本研究ではこの小目標を達成するために時間発展する人物相関図を提案した。提案手法では登場人物や人物間の関係性は小説のページが進むにつれて登場しその後持続して存在するものとして扱う。そして現在表示しているページまでに出てきた登場人物と関係性を元に人物相関図と人物相関表という2種類の可視化を行う。これによりユーザーは表示するページ数の範囲を指定して人物相関図を表示することができ、ネタバレを受けずに書籍の内容を整理することができる。

課題点は2点ある。1点目はシステムの誤操作によって意図せずまだ読んでいない部分の情報を取得してしまう可能性があることである。今回人物相関図表示の内容は「現在表示しているページまでに出てきた情報」を表示しているため、スクロールバーの操作を間違えてまだ読んでいない部分を選択するとその時点までの登場人物と関係を人物相関図に表示してしまう。今後の課題としては意図しない表示が行われないための仕組み作りがあげられる。2点目は人物相関図に表示された登場人物からネタバレを起こす可能性がある点である。現在一人の登場人物に対し名前が一つ対応しており、それは何度でも修正することができる。加えて小説内の関係情報は他者と共有されているため、先に読んだ人が自分の読んだ内容に合わせて登場人物の名前を修正している場合、登場人物の真の名前が人物相関図に表示される可能性がある。登場人物の名前がわざと序盤において隠されたり展開に大きく影響している場合があるため、登場人物の真の名前を知ることによってネタバレにつながる可能性がある。そのため今後の課題として人物の名前を複数保存できるようにし、その名前がどのページまで使われているのかを確認し、ネタバレが起こらないようにする。

5.2. (P2) 小説内の繊細な人間関係を扱えること

この小目標を達成するためにシステムの要件として、本研究では計算機の自動抽出を人間の手によって補完する方法を提案した。またその手法を実現するためにシステムの要件として(N2) 小説に関する重要な情報を手軽に追加できること、(N3) 情報を後から編集できることを設定した。(N2)の達成のために小説内の全ての登場人物と関係を自動抽出する機能と、人間の手によってそれらの情報を追加する機能を実装した。そのことにより機械による完全自動抽出に比べて小説にそった内容であり、人間による完全手動より少ない手間での情報の追加を実現させた。(N3)では情報編集画面を用意し、人間による情報修正をシステムによってサポートした。加えて登場人物の出現頻度のヒートマップを表示したことで、登場人物の出現する場面へ容易にアクセスできるようになり、情報の妥当性の検証のサポートを行った。以上よりシステムにおいて小説内の複雑な関係をシステム上で扱うことができるようになった。

現在の課題点は2点存在する。1点目は情報入力機能に必要な操作の数が多くことである。現在システム上で登場人物を入力するためには登場人物の名前を実際に打ち込むなど4ステップが必要である。情報追加ではその関係が誰か誰へのどんな関係か打ち込むなど7ステップが必要であ

る。どちらも自分で文章を打ち込む必要がありユーザー側の負担が大きいため、本文上を選択して追加したり、情報の追加を何らかの形でアシストする必要がある。2点目は情報の自動抽出の精度に改善余地があることである。現在小説の本文から登場人物の自動抽出を行っているが、登場人物ではない物体も登場人物として抽出してしまうことがある。実際に宮沢賢治の「銀河鉄道の夜」に適用した際には物語に出てきたりんどうの花が「りんどう」として登場人物の中に抽出された。今後は抽出アルゴリズムを改良し抽出精度を向上させたい。

5.3. (P3) 入力した情報をもとに人物相関図を自動でレイアウトしてくれ、それを適切なタイミングで確認できること

この小目標を達成するために提案手法は**人物相関図**と**人物相関表**という二種類の図で現在のページまでの人物関係の表示を行った。どちらもシステムに保存された登場人物のリストと関係性のリストをもとに、人物相関図はノードリンクで人物相関表は隣接行列で表示した。これらの可視化は小説の本文の右側に配置され、読書をしながら確認でき、小説の内容を整理することができる。人物相関図はノードの大きさで登場人物の重要性を表現し、全体の構造を直感的に理解することができる。一方人物相関図はフィルタリング機能が豊富で、ページ数の範囲や登場人物のリストから人物相関表に映し出す範囲を絞り込むことができる。加えて2人の人物を選択することで、その二人の過去の関係も参照することができる。人物相関表を用いることで登場人物の多い小説でも関係の詳細を確認することができる。

一方課題点は2点存在する。1点目は**人物相関図の時間変化が捉えにくい点**である。現在人物相関図のレイアウトはページを移動するたびに再計算している。そのため人物や関係が大きく変わる時にレイアウトが大幅に変更され、構造の変更点を確認することが難しくなる。今後は動的ネットワークの既存アルゴリズムを参考にして、レイアウトの変化が少ないアルゴリズムを設計したい。2点目は登場人物の多い小説では人物相関図が複雑で認識にくくなる点である。登場人物が多いと線の交差は多く、個々の人物を表す頂点は小さくなってしまふ。そうなると人物相関図の理解にかかる労力が増してしまう。今後は人物相関表に用いたフィルタリング機能を人物相関図にも適用したいと考える。

以上のように各小目標に対し解決策を考え、システムで必要な機能を実装することができた。今後は各小目標で述べた課題点を解決したいと考える。

6. まとめ

読書はメジャーな趣味であり信頼性の高い情報源だが、様々な理由から読書を断念してしまう人々が存在する。本研究では計算機による支援によって人々の読書が円滑に行えるように支援を行う。その上で人物相関図の**物語の内容を整理するのに有効である点**に着目し、人物相関図の提示によって読書を支援することを決めた。そして具体的な手法を考えるために実験を行い、ネタバレの危険性やレイアウトの難しさなど様々な問題を発見し、目的達成のための小目標を3つ設定した。その上で本研究では**時間発展する**

人物相関図を提示することで読書の支援を行う手法を提案した。提案手法の有効性を検証するために青空文庫で公開されている小説を関係図を確認し読み進められる読書システムを開発した。開発にあたりユースケースを考え、システムに必要な4つの要件を定義した。4つの要件とは具体的に(N1) **システムの上で読書ができること**、(N2) **小説に関する重要な情報を手軽に追加できること**、(N3) **情報を後から編集できること**、(N4) **小説の進行度に合わせて関係図を表示することを設定した**。(N1)の達成のために最低限の読書機能をつけ、(N2)の達成のために登場人物と関係性の追加を機械による自動抽出と手作業での追加の2通りで実現した。(N3)のためにユーザーが容易に情報を編集できる情報編集画面をつけ、(N4)のために人間関係を**人物相関図**と**人物相関表**という2種類の可視化で表現した。以上のように要件をもとにシステムを実装することができた。また議論において小目標ごとに振り返りを行い研究の達成度を確認することができた。

今後は6章で述べた課題の解決とシステムの評価を行う。本研究では小目標の達成度によってシステムの評価を行った。今後はユーザー実験など提案手法の有効性を評価するための実験を準備し、実行したいと考える。

参 考 文 献

Filipov, Velitchko, Davide Ceneda, TU CCAST, Michael Koller, Alessio Arleo, and Silvia Miksch. n.d. “The Circle of Thrones: Conveying the Story of Game of Thrones Using Radial Infographics.”

John, Markus, Martin Baumann, David Schuetz, Steffen Koch, and Thomas Ertl. 2019. “A Visual Approach for the Comparative Analysis of Character Networks in Narrative Texts.” In *2019 IEEE Pacific Visualization Symposium (PacificVis)*, 247–56. <https://doi.org/10.1109/PacificVis.2019.00037>.

Tanahashi, Yuzuru, Chien-Hsin Hsueh, and Kwan-Liu Ma. 2015. “An Efficient Framework for Generating Storyline Visualizations from Streaming Data.” *IEEE Transactions on Visualization and Computer Graphics* 21 (6): 730–42.

Wu, Tsung Heng, Ye Zhao, and Md Amiruzzaman. 2020. “Interactive Visualization of AI-Based Speech Recognition Texts.” *EuroVA*.

前田恭佑, 土方嘉徳, 中村聡史, and 酒田信親. 2019. “ストーリー文書を用いたレビュー文書のネタバレ判定.” システム制御情報学会論文誌 32 (3): 87–100.

米田崇明, 篠崎隆宏, 堀内靖雄, and 黒岩眞吾. 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* 18: 855–58.

西原弘真, and 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* 21: 628–31.