小説内の動的人物相関図を用いた読書の支援

太田　彩 (東京工業大学），脇田　建（東京工業大学）

An Example of the Manuscript for  
the Visualization Symposium Japan (Times New Roman 16pt)

―Subtitle (Times New Roman 14pt)―

Taro KASHIKA and Hanako JOHO

##### ABSTRACT

This template is a guide to prepare manuscript for Visualization Symposium Japan. Here, please write the abstract serving as an index and as a summary of the present paper. It should be as long as approximately 150 words. \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\*\*\*\*\*《150語程度の英文ABSTRACT及び5語以下のKeywordsを記入してください．文字の大きさは10pt，字体はCentury，行間は１行分 (シングルスペース) です．》

Keywords: Visualization, Digital image processing, \*\*<5語以下>\*\*, \*\*\*\*\*, \*\*\*\*\*

# 序　論

本原稿は，可視化情報シンポジウムの原稿テンプレートです．原稿執筆にはこのテンプレートをご使用ください．原稿サイズはA4版とし，ページ数は任意ですがファイルサイズは5MB以下としてください．ページ余白は，上25mm，左20mm，右20mm，下20mmとし，ヘッダーおよびフッダーは入れず，表題は3行目から始めてください．

# 文字及び改行

## 文字の大きさと改行

使用言語は日本語または英語とします．和文原稿の場合，題目，著者名，所属を和文で書き，続けて題目，著者名の英文名を記入してください．英文原稿の場合，題目，著者名の英文名を記入してください．なお，和題・英題は講演申込時と同じであることが必要です (異なる場合，プログラムや講演論文集の目次には講演申込時の題目が使用されます)．

著者名は11pt，1行に36字以内で右寄せとし，これを越える場合は改行してください．本文は9ptの2段組みとし，片側25字×50行，段間は10mmを基準としてください．章間の行間は2行分 (ダブルスペース) とし，その中央に章名を10ptで記し，節間は行を空けないようにしてください．

## 字　体

題名，章・節名はゴシック体，本文は明朝体，英文題名，著者名はTimes New Roman, 本文中の英文はCenturyとします.参考文献1) は8ptで表記ください．

* *強調*
* **とても強調**

# 図　表

本文中の図表の表記には，下記Fig.1と Table 1のように，図表には通し番号をふり，図表番号はゴシック体，図表題はCenturyで書きます．キャプションは表の場合は上部に，図の場合は下部に記し，説明文は英文で表記してください．

Table 1: Sample of table

| Table | Table |
| --- | --- |
| 1 | 2 |

表 [Sample of table] を参照のこと。



Figure 1: PacificVis PC meeting

# 原稿提出

本テンプレートで作成した講演原稿 (PDF形式) と，調査票 (PDF形式，必ず押印したものをお願いします) をシンポジウムHPの案内に従って提出して下さい．

# 数式

二点を結ぶふたつの制御点列とについて。。。

macros.tex に定義したようなTeXマクロにも対応しています。

二点を結ぶふたつの制御点列とについて。。。

以上で定義した三つの指標がいずれも一定値以上出会った場合に限り，束化圧があるものと評価する．このことは閾値 () を用いて以下のように定式化される：

# その他

## 文献の引用

この例題では、library.bib 文献データベースとし、文献引用のスタイルにシカゴスタイルを指定して文献リストと文献の引用をしている。.cslファイルを別のものに指定すれば、他の引用形式にも対応できる。文献の引用方法は LaTeX の \cite {...} コマンドとは異なる。以下を参考にすること。

(Kashika 1997; Kashika and Hanko 1999) は英語の文献で(可視化 and 情報 1998)は日本語の文献である.

# はじめに

# 関連研究(書きおわり)

## 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)は人物関係を表す語を集めた関係辞書と、人物間の関係を抽出する関係抽出のパターンの集合を作成し、人物間の関係を抽出した。しかし自然言語処理の技術だけを用いて正確な情報を抽出することは難しく、読者によって必要とする情報の範囲は異なる。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。

## 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。FCK(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。THM(Tanahashi, Hsueh, and Ma 2015)は時系列に沿って変わる映画内の登場人物の位置関係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、本研究では読書の支援が目的であるため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。

## 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。WZA(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音声認識のサポートを行った。IBS(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サポートするシステムを提案した。本研究ではこれらの論文で採用されているいくつかの可視化手法を参考にしシステム設計を行う。

# 提案手法（一旦書きおわり）

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「前回読んだ場所までの内容を忘れてしまった」などの理由から、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r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った

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。

## 実験

人物相関図を用いて読書を支援する手法を考える上で、実際に手書きで人物相関図を書きながら読書を行った。その際に３つの問題を発見することができた。1点目は**一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。人物相関図に物語内の登場人物や人間関係を表示すると、物語内の展開上重要な情報を事前に知ってしまう可能性がある。そのため読書をしたときの驚きや感動などの新鮮な感情が阻害され、読書の楽しみを減らす恐れがある。2点目に**人間関係の描写は複雑なものが多い**ということである。 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、**〇〇さんの好きな人だと気づく**など簡潔に言い表せないものも存在する。文学において繊細な感情の変化は重要な要素なので詳細な関係を記述し、それを人物相関図にて確認できなければならない。3点目は**手書きで美しいレイアウトの人物相関図を描く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婉曲が少なくなるようにレイアウト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、読みやすい人物相関図を書くことのハードルは跳ね上がる。そのため読書の度に人間の手でレイアウトを行うことは現実的でない。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

## 提案手法の概要

3.1で設定した小目標をもとに提案手法の概要を紹介する。 (P1)の実現のために本研究では**読書の進捗状況に応じて変化する人物相関図**を提案する。小説内の登場人物と人間関係を動的ネットワークとして扱うために、登場人物や関係の出てきた位置をページ数で管理する。そして今まで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。(P2)の実現のために登場人物と人間関係の情報追加を手動と自動の２通りで実現した。自動抽出では独自のアルゴリズムにより登場人物を表す可能性の高い単語を抽出し、登場人物の共出現した文を関係として抽出した。手動での追加では登場人物と人間関係に関する様々な情報をユーザーの手によって登録する。(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。加えて本研究では**読書体験を共有するために、提案手法では登場人物や関係の情報を共有する**。そもそも登場人物のリストなどの読書に必要な情報は誰にとっても共通である。その情報を共有することで、新たに読む人は入力する手間を削減することができる。また、登場人物の間の関係は人によって受け取り方が異なることもあり、そこが議論の中心になることもある。 したがって関係を共有することで自身の細かい解釈についても人に伝えることができる。以上が提案手法の概要である。次の章では手法をもとにシステムの設計を行う。

# システムの構成

## システムの要件

### ユースケース

以下のようなユースケースを設定した。

大学生の楓さんは友人の紅葉さんから勧められてイワン ツルゲーネフの「はつ恋」を読むことにした。彼女は読書 まず、青空文庫のサイトで「はつ恋」の図書カードを開く。そしてそこからテキストのリ ンクをコピーして、システムに貼り付けて、システム上で「はつ恋」を読めるようにする。 システムで小説の選択からダウンロード済みを選んで、「はつ恋」をシステム上に表示する。 楓さんは読み進めると 1 ページ目で3人のおじいさんが 出てくることに気づく。主人とセ ルゲイ・ニコラーエヴィチと、ヴラジーミル・ペトローヴィチの3人で、それらの人物を登 場人物に追加したい。そこで登場人物の候補を抽出ボタンを押すことで候補を出した。候補 には「イヴァーノヴナ」「客」「お嬢さん」など様々な候補が出されたので、ハイライトされ た部分を確認しながら人物を絞って追加した。 この調子で気になるところで人物候補の抽出を行っていく。5 ページ目に「背の高い少女」が出てきたのでその人物を追加する。この時、入れたい人物は一人だけだったので情報追 加で登場人物追加を選び入力した。また、それと同時に小説内で関係だと思うものを情報追 加から関係性追加を選んで追加していく。例えばウラジーミル・ペトローヴィチは6ページ 目で背の高い少女と知り合いになりたいと思っているのでその情報を足していく。その後 12 ページめでその少女の名前が「ジナイーダ」であることが明かされたので、可視化の選択か ら編集画面を選ぶ。選ぶと登場人物の表が出てくるので該当箇所を修正して変更を保存ボタ ンを押した。 その後も読み進めて 22 ページでジナイーダの取り巻きであるマレーフスキイ、ルーシン、 マイダーノフ、ニルマーツキイ、ベロヴゾーロフが登場したのでそれらの人物も追加した。 そこで楓さんはロシア人の名前がたくさん出てきて読むのに疲れたためいったん読むのをや めた。 楓さんは 1 週間開けてまた途中から読み直すことにした。しかし、1 週間期間を開けてし まったので、小説の中で誰が重要人物か思い出せなくなってしまった。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。22 ページにおける人物相関図を見 ると、「ウラジーミル・ペトローヴィチ」と「ジナイーダ」が大きな頂点として描かれていた ので主要な人物が誰か確認できた。(図 5.1) 読み進めていくうちに「ルーシン」という人物が出てきた。「ジナイーダ」の取り巻きで あることは覚えているが、実際に誰なのかはわからない。そこで編集画面で登場人物表の 「ルーシン」をクリックしてどのくらいの頻度で確認しているかを確認した。 友人の紅葉さんにそのシステムを使っていることを話すと、紅葉さんは小説の考察をすることが好きなので、自分もそのシステムで自分の考察をまとめてみることにした。1 週間後 に楓さんが読み終えて、紅葉さんも考察が終わったので、二人で感想を言い合うことにした。 二人で入力した内容を見ている時に、楓さんは 51 ページの「ジナイーダがウラジーミルに塀 から降りるようにいった」ところの捉え方が異なっていることに気づいた。楓さんはこの関 係の好感度が低いと思ったが、紅葉さんは「ジナイーダは本当に飛び降りて欲しいとは思っていない」からだと好感度を高めに設定していた。自分の読んだ内容を思い出しながら会話を楽しむことができた。

### 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) 小説の進行度に合わせて関係図を表示すること

## 登場人物の抽出

Filipov, Velitchko, Davide Ceneda, TU CVAST, 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>.

Kashika, T. 1997. “Visualization Technique.” *VSJ Press*, 1111–25.

Kashika, T., and J. Hanko. 1999. “An Example of the Manuscript for VSJ Journal.” *VSJ Journal* 10 (5): 10–15.

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*.

可視化太郎, and 情報花子. 1998. “可視化情報学会における論文の書き方.” *可視化情報* 100 (1000): 11–12.

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集* 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集* 21: 628–31.