



## 2011 年度 未踏 IT 人材発掘・育成事業 採択案件評価書

### 1. 担当PM

首藤 一幸 PM

(東京工業大学 大学院情報理工学研究科 数理・計算科学専攻 准教授)

### 2. 採択者氏名

チーフクリエイター: 田中 翔太郎

(電気通信大学 電気通信学部情報工学科)

### 3. 委託金支払額

1,792,000 円

### 4. テーマ名

文章理解を支援するインタフェースの開発

### 5. 関連Webサイト

なし

### 6. テーマ概要

本を読んで内容を理解することは、多大な労力を要する。例えば、難しい語が多く含まれる本を読むとき、事前知識がないと 1 度読んだだけでは内容を理解できない。読者は何度も読み返したり、分からない語に線を引いたり、重要な語をノートにうつしたりして少しずつ内容を理解していく。ここで、効率よく本の知識の抽出を行うことができれば、短時間で多くの本の内容を理解できるようになる。

本提案では、同義語の探索や階層構造の表示を行うなどして、自然言語処理の結

果をユーザに伝わりやすいようにビジュアライゼーション(可視化)する。さらに、ユーザインタフェースを使用してユーザとシステムが対話を進めることで、効率良い文章理解を可能とする。このようなインタフェースは今までに存在せず、学術的にも大きな新規性が期待できる。また、ビジュアライゼーションの方法やユーザインタフェースには様々な形が考えられるため、本提案の応用性は非常に高い。

本提案の実現により、既存の自動要約技術では不可能なユーザのニーズに基づいた要約作成が可能となる。また、それらの要約をデータベースに保存しておくことで、今までにない内容ベースの本の検索システムが実現できる。さらに、既存の電子書籍リーダーと組み合わせることで、ユーザの負担を軽減し、より素早く内容を理解できるようになると期待される。

## 7. 採択理由

文章(例:小説や理工系参考書)の理解を支援するソフトウェアの提案である。指定した文に関連する語をハイライトしたり、また、長文の要約を自動生成したりする。応用としては、国語の学習支援や、要約に基づいた書籍検索などを想定している。

結果の提示方法やユーザインタフェースにはいくらかでも工夫の余地がある。また、要約の仕方によって文学作品に対して定説とは異なる解釈を与えたり、さらには、文章の読みやすさを評価したりといった発展もありそうである。様々な可能性のどれを見せてくれるか、大変楽しみにしている。

## 8. 開発目標

目標は、大量の文章を理解容易な形にして利用者に示すことである。

当クリエイターは、長文からの重要語や重要文の抽出に取り組んできており、それを物語やエッセイの自動要約に応用することで一定の成果をあげていた。例えば、文庫本程度のエッセイ 1冊から数十程度の文章を抽出し、それを眺めるとおよそエッセイの内容が理解できる、というくらいの自動要約を達成していた。

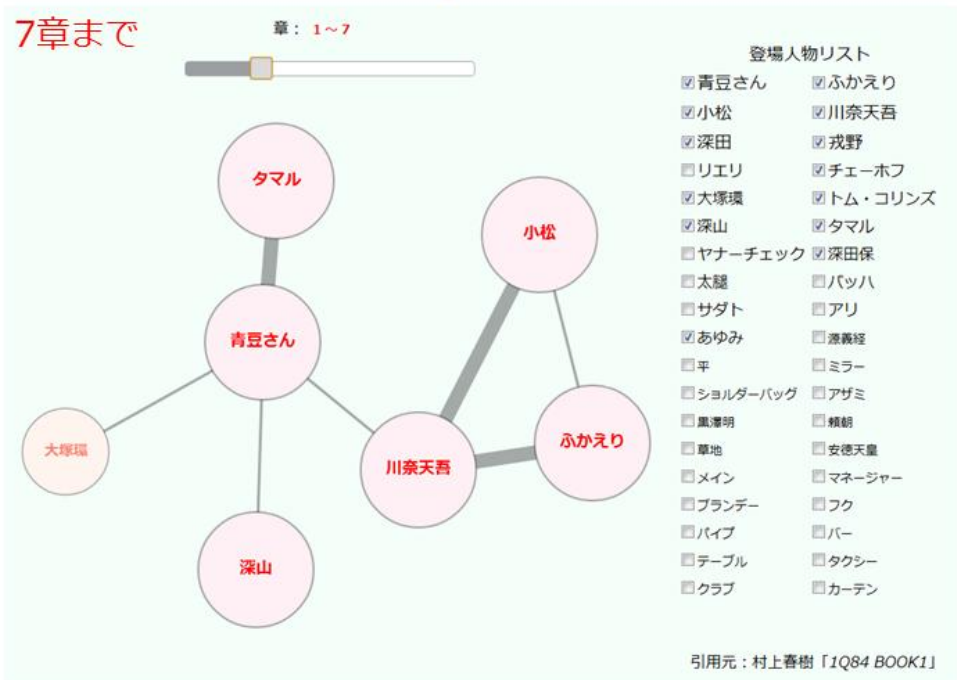
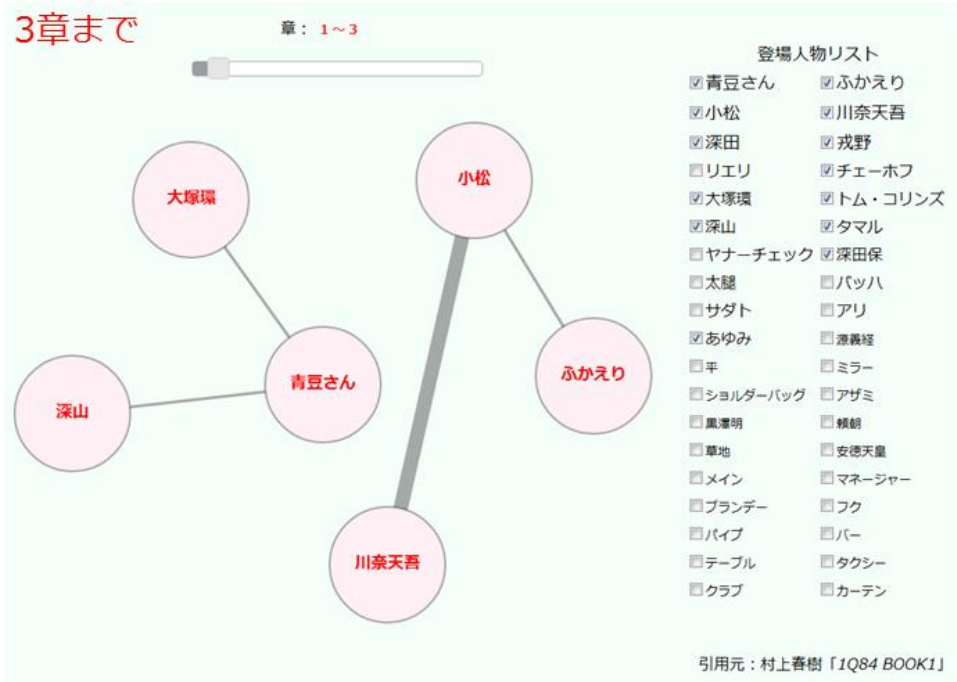
本プロジェクトでは、その要約技術を発展させて、利用者に新しい読書体験を提供することを狙う。

## 9. 進捗概要

プロジェクトでは、まず、それまで試みてきた自動要約技術をどういった分野の長文にどのように適用するかを検討した。物語、エッセイ、技術文書、試験問題などへ

の適用を試み、学校での国語の試験に回答する、といったグランドチャレンジ規模の案も出て、実験も行った。

最終的には、「物語」の「登場人物」にフォーカスすることとした。長文を入力として、登場人物の相関図を表示する。物語の各時点に対応した相関図を表示することで、物語の進行に応じた相関の変化を見て取ることもできる。



図：登場人物の相関図(1Q84 の例)

## 10. プロジェクト評価

開発を始める時点までに試みてきたことを未踏でどう発展させるか？を定めるところに時間を要した。もともと、ただ 1 つのゴールに向けて疾走するというよりは、自身の興味に従って、様々な可能性・seeds を、手を動かしてどんどん試していく、というタイプのクリエイターである。開発期間のあるところまでは、様々な可能性の追求を行ってもらった。

プロジェクトの成果物によって、人物相関図を見て物語を振り返る、もしくは、傍らに置いて読む、という読書体験が、あらゆる物語に対して可能となった。

## 11. 今後の課題

機能面では、人物の関係抽出の裏付けとなっている重要文抽出機能の統合や、表示レイアウトの改善などを予定している。では、読書体験という面でどう変わったか、変わるのかを考察して、今後につなげて欲しい。