

研究報告書

「物語構造に基づく情報編纂基盤技術」

研究タイプ: 通常型

研究期間: 平成 21 年 10 月～平成 25 年 3 月

研究者: 赤石 美奈

1. 研究のねらい

現代の知識化社会において、人々は膨大な情報の中から、それぞれが抱えている問題に関連する情報を収集・分析し、迅速かつ的確にまとめて、己自身や所属する組織等の必要とする知識を獲得して意志決定に供することが求められている。

検索エンジン等で代表される情報検索技術は、利用者が必要とする情報が、(文書のような)検索単位として、予め存在しているような場合には、有効な検索手段である。しかしながら、利用者の要求を満足する情報が存在するかどうか分からない場合、単一の文書では情報要求に応えられず複数の文書の組み合わせが必要な場合、あるいはそもそも利用者自身が情報要求を分節できていないような場合には必ずしも有効であるとはいえない。

このような情報要求を満たすには、情報を検索しては解釈し、また検索するといった試行錯誤が必要となる。このような探索的情報アクセスを支援するためには、試行錯誤の際に、より有用な方向へ進めるように、情報洪水の中に道標を示すことが必要であろう。

本研究では、「物語」に着目し、大量な情報の中に潜在する多重・多層文脈を顕在化し、多様な視点や文脈に応じて有益な情報を編纂することを支援するシステムの構築を目指す。これにより、知識創造につながる探索的情報アクセスを支援する情報アクセス環境の実現を図る。本システムは、単にユーザが直面している問題の解答を探索するのみではなく、場合によっては、問題や要求そのものが変化し、その解決法を探索するプロセスにより、新しい概念世界を創造することを目指している。このために、既存の情報を発掘するだけではなく、新しい物語を生成するための機構を実現しなければならないと考えている。つまり、検索技術が(過去に生成された)蓄積物を対象にしているのに対して、ナラティブ連想情報アクセスでは、これから生成される新しい物語を産出す技術の研究開発を目指している。

2. 研究成果

(1) 概要

本研究では、文書の持つ『物語構造』に着目し、大量な情報の中に潜在する多重・多層文脈を顕在化し、多様な視点や文脈に応じて有益な情報を編纂することを支援するシステムの構築を目指した。これにより、知識創造につながる探索的情報アクセスを支援する情報アクセス環境の構築を進めた。

基盤技術としては、物語構造に基づき、新しい文脈を生成しながら、物語を生成するナラティブ連想情報アクセスのフレームワークについて研究・開発を行った。まず、文書における物語構造モデルを定義し、ナラティブ連想情報アクセス・フレームワークの基本を提案し、「語の共起依存性に基づく語彙連鎖構造」を基に、基盤技術となる文書分割方法、断編連結方法の

研究を進め、応用システムであるナラティブ連想情報アクセス・システムの開発に関する研究に取り組んだ。

ナラティブ連想情報アクセスのフレームワークは、新しい文脈を生成しながら、物語を生成するための枠組みである。基本的な構造は、文脈に応じて、既存の知識を再構築するために、既存の文書を分割し、再構成する仕組みである。

さらに、このフレームワークに、概念空間の概念を導入し、物語の表面には明示されていない関係を見つけるための仕組みについて研究し、潜在文脈の顕在化に焦点をあてて、研究を進めた。一つの文脈に沿って表現されている文書群を別の観点から捉えなおすことで、新しい文脈に基づく物語が顕在化される。

また、概念空間に、物語の構成場面を連結するための場のポテンシャルと連結力に加わる外力の導入を検討し、概念空間の変化により物語構造を変形される仕組みについても研究を行った。

(2) 詳細

研究テーマ「ナラティブ連想情報アクセスのフレームワークの構築」

本研究では、物語構造に基づき、新しい文脈を生成しながら、物語を生成するナラティブ連想情報アクセスのフレームワークについて研究を行った。本研究では、「語の依存性に基づく語彙連鎖構造」を基に、基盤技術となる(i)文書分割方法、(ii)断編連結方法の研究を進め、応用システムであるナラティブ連想情報アクセス・システムの開発に関する研究に取り組んだ。

(i) 文書分割方法: 語彙連鎖に基づく物語分割

ひとつの文書には、複数の主題が含まれている。文脈に応じた再構成を考えた場合、既存の文書を主題毎の断編に分割する必要がある。本研究においては、「語の共起依存性に基づく語彙連鎖構造」に基づく系列的文書分割法と場面的文書分割法などの文書分割手法を開発した。

(ii) 断編連結法: 語彙連鎖に基づく文脈生成

本研究で提案するナラティブ連想情報アクセスは、対象文書を、物語構造に基づき分節し、ユーザが選択した文脈に沿って生成される新しい物語の候補を結果として出力し、既存文書集合を横断的に再構成して得られる新しい知識獲得を支援するものである。本研究においては、各種の文書集合に対して、トピック遷移パターンの解析を行い、応用分野において適切な物語生成を可能とする遷移ルールの確立と、支援システムの構築を目指した。

研究テーマ「動的に変化する概念空間の構築」

多様な物語を生成するための母体となる概念空間は、固定化されるものではなく、状況に応じて変化可能な仕組みを持たせる必要がある。状況に応じて、拡張・収縮可能な概念空間の設計を行った(論文 4 参照)。

(i)概念語ベクトルに基づく概念空間

概念空間の基本軸となる概念語ベクトルを動的に生成することにより、物語構造が動的に変化できる概念空間を実現する。図 1 は、同一の歴史史料を異なる概念語の配置された概念空間にて配置した例である。

(ii) 時間軸と概念語ベクトルを軸とする概念空間

時間変化が重要である編年型データにおいては、時間軸を基にした概念空間が必要である。これに特化した概念空間の特性に関して研究を進め、多様な物語を生みだす仕組みの基盤を形成した。図 2 は、時間軸と概念語(人名)に基づくデータ分布を示す。このような概念空間から、一連の関連情報を抽出し、歴史解釈を展開させる基本プロットの構築を図った(論文 2, 3 参照)。

研究テーマ「潜在文脈の顕在化」

(i)潜在的な物語を抽出するための潜在文脈の顕在化

概念空間に配置された物語構成場面間の関連を基に、潜在する文脈を顕在化する手法について研究した。各構成要素場面同士の連結は、語の吸引力遷移パターンとして定義される。図 3 は、人工衛星の設計議事録を用い、衛星の運用上の問題解決のための情報へアクセス可能なことを示した Topic Tracer システムである。図 4 は、様々な視点に基づき形成される概念空間から抽出される一連の関連情報の例である。ここでは、人間関係の変遷が、視点となる概念や人物によって異なることを示している。図 5 は、プレゼンテーションに利用されたスライドを語の連想パターンに基づき、再構成し、新たな文脈におけるプレゼンテーション資料作成のための支援ツールを示す。

(ii) 物語生成のためのミッシングリンクの補間
概念空間内の物語構成場面のみでは、有意な文脈を形成できない場合がある。その際に、概念空

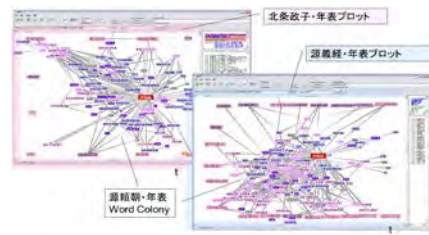


図 1. 多視点に基づく歴史史料の捉え方



図 2. 時間軸と概念語の出現例

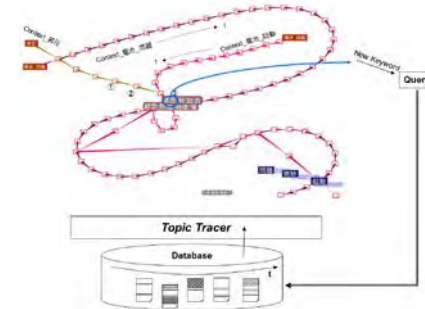


図 3. 設計議事録からの知識獲得

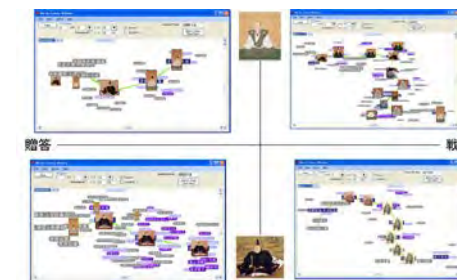


図 4. 多視点からの人間関係の変遷

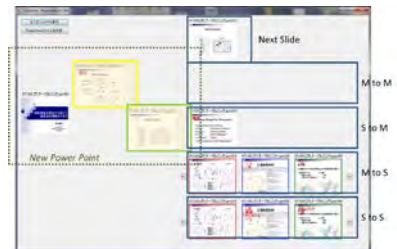


図 5. プレゼンテーションの動的再構成

間の構造自体を変化させ、物語構造を形成する方法と、不足している場面情報を概念空間内に取り込み、物語構造を形成する方法が考えられる。これらの場面連結の欠損(ミッシングリンク)を補間する技術について、トピックブリッジング手法として研究を進めた(論文 1, 5 参照)。

以上、様々な対象に対して、ナラティブ連想情報アクセスのフレームワークを適用し、応用可能なことを示した。

3. 今後の展開

本研究においては、「物語」を意図的に関連付けられた一連の出来事としてとらえ、「編集」を出来事間の関連付けることとしてきた。動的な物語編集技術とは、状況に応じた物語(知識)の生成技術を意味し。これにより、視点を少しずつ変化させながら関連情報を辿る技術の研究を進めてきた。

今後は、実際の文学作品としての物語生成支援へと研究を展開させていく予定である。このためには、一見、関連のなさそうな概念を結び付けたり、想像していなかった結果へと物語を展開させていく編纂技術が必要となる。また、さらに日常生活で得られる情報(ライフログ)から生成される物語を通じ、相互に影響しあう、現実世界と情報空間から物語を生み出す仕組みを考案していきたいと考えている。

4. 自己評価

情報探索手法としてのナラティブ連想情報アクセス手法の研究に関しては、いくつかの応用分野において研究を進めることができた。しかしながら、未だ、「物語」に関する編纂基盤技術としては、まとめきれておらず、開発された基盤技術を統合環境として提供するには至らなかった。また研究成果の発表が滞っており、今後、整理して、まとめあげることが必要である。

5. 研究総括の見解

文書中の単語の共起依存関係だけから文書に書かれた文脈(物語)を抽出し、DB 化や新規文書構成、文脈の検索などを可能にすることを目指した研究課題であった。

連想情報アクセス手法の情報探索手法としての研究について、いくつかの応用分野において成果をあげている。「物語」に関する編纂基盤技術として、さらに開発された基盤技術を統合環境として提供できるようにしていくのがよいであろう。今後、文学作品等に適用し、物語生成支援へと研究を展開させていってほしい。

6. 主な研究成果リスト

(1) 論文(原著論文)発表

1. Makoto Sato, Mina Akaishi, Koichi Hori: Topic Bridging by Identifying the Dynamics of the Spreading Topic Model, Intelligent Interactive Multimedia: Systems and Services, Proceedings of the 5th International Conference on Intelligent Interactive Multimedia Systems and Services (IIMSS 2012), Smart Innovation, Systems and Technologies, Vol. 16, pp. 619-628, 2012.

| | |
|----|---|
| 2. | <u>Masahiko ITOH</u> , <u>Mina AKAISHI</u> , Visualization for Changes in Relationships between Historical Figures in Chronicles, Proc. of Int. Con. on Information Visualisation 2012, pp.283-290, 2012 |
| 3. | <u>Mina Akaishi</u> , Makoto Sato, <u>Koichi Hori</u> , Tetsuya Ishikawa: A Visual Analysis Tool for Finding Embedded Relations in Chronicles, In Proceeding of 16th International Conference on Information Visualisation (IV), pp.302-307, 2012. |
| 4. | <u>Mina Akaishi</u> , Makoto Sato and <u>Koichi Hori</u> : A Visual Analysis Tool for Amplifying Story Generation Cycle, Proc. Of IADIS International Conference, Computer Graphics, Visualization, Computer Vision and Image Processing 2011, pp.349-352, 2011 |
| 5. | Makoto Sato, <u>Mina Akaishi</u> , <u>Koichi Hori</u> : Topic Bridging for Story Generation, Information Modelling and knowledge Bases XXII, pp.247-257, 2011 |

(2)特許出願

研究期間累積件数:0件

(3)その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)