

Wikipedia 記事再編集支援のための履歴可視化システム

塩谷 圭甫 †

金井 秀明 ‡

† 北陸先端科学技術大学院大学 知識科学研究科

‡ 北陸先端科学技術大学院大学 知識科学教育研究センター

1 はじめに

本研究では Wiki による百科辞典『Wikipedia』において、履歴の流れを可視化することで記事編集を支援する「編集履歴可視化システム」を提案する。

近年、ネットワーク環境の整備が進んだことで、WEB 上でのコラボレーションがますます盛んになっている。たとえば、いつでも・どこからでも・だれでもが記事の加筆訂正を行える Wiki は、インターネット上で知識を集積する場として頻繁に用いられるようになった。しかし複数のユーザーが分散・平行して記事の編集を行える Wiki においては、その性質ゆえに編集の流れを追いかく場合がある。

例えば、Wikipedia のベースとなっている Wiki エンジン『MediaWiki』は、ページの履歴を見るのにごく単純な仕組みしか提供しない。そこでは編集履歴の日付・ユーザー名・簡単なメモが並ぶだけで、編集の内容や箇所に関してはまったく情報を得られない。履歴同士を比較するモードで差分を表示することは可能だが、そこでは 1 ページずつの比較しか行えない。そのため、複数回の編集が重なったときに個々の変更を把握することは困難である。長期にわたる連続的な変化に表れる「編集の流れ」を見るのに、MediaWiki のインターフェースは向いていないと言える。

この弱点を補うため、本研究では記事の連続的な変化を把握するための「編集履歴の可視化システム」を提案する。具体的には、時系列に沿って編集差分を並列表示することで、編集の意図や趨勢を把握しやすくなる。

2 関連研究

Wikipedia の編集履歴に関する研究は以下のようなものがある。

Michael D. Ekstrand らは版の削除や差し戻しなどによる版の分岐について可視化を行い、分析を行った。^[1]

その成果は、版の分岐を履歴ページに表示する MediaWiki のプラグインとして提供されている。Peter Pirolli らは、編集者に注目して分析を行い、記事に対する貢献をグラフ化するサービスを作成した^[2]。Mikalai Sabel らは、編集履歴の量的な変化について可視化を行い、編集合戦が起きた場合などに表れる特徴について分析した^[3]。一方、本研究では、編集プロセスにおいて必要な機能について分析し、編集の流れを掴むための、履歴本文を中心とした整理・可視化を行う。

3 記事編集のプロセス

Wiki において記事を編集するとき、他の編者からまったく独立して加筆訂正を行う場合と、協調作業として流れを踏まえた上で記事を編集していく場合がある。誤字訂正のような細部の編集であれば該当箇所のみを注目すれば良い。しかし、ある程度まとまった量の加筆訂正をするには本文全体を読む必要がある。可能ならば過去の編集履歴や議論を読んでおくことが望ましい。ただし、自分が書いた記事を再編集する際には、注目する箇所の変更履歴を追えれば十分である。

現在、MediaWiki で実現されているユーザインタフェースはごくシンプルなもので、編集の流れを追うためには個々の履歴ページを、また議論を追うためには別に用意されたノートページを順に読んでいく必要がある。そのため、記事の規模が大きくなったときには記事の流れを追うことが困難になる。この記事履歴とノートページを編集者の視点に合わせて整理・可視化できれば、編集支援が可能になると考えた。

たとえば編集プロセスにおいては、以下のような視点があると考えられる。

- (1) ある節がどのような変遷を辿ってきたのか
- (2) 全体の論理構造がどのように変化してきたか
- (3) ある文章がどの版で初めて登場したのか
- (4) ある人物はどのような編集を行ってきたか
- (5) 本文記事に関連する議論はあるか

本研究ではこれらの視点に合わせて編集履歴のデータを抽出・整理する、いくつかのモードを用意した。そして、それらを見通しの良いインターフェースにまとめ、編集者の助けとなるよう配置した。

Revision History Flow: Page History Viewer for Wikipedia

†Keisuke SHIOYA †Hideaki KANAI

†Japan Advanced Institute of Science and Technology School of Knowledge Science

‡Japan Advanced Institute of Science and Technology Center for Knowledge Science

4 システム構成

本システムは Wikipedia の Wiki エンジンである MediaWiki に、Java アプレットの形で実装される。図 1 にシステム構成図を示した。具体的には変更履歴のページに埋め込み、記事のアウトラインや履歴を表示する。

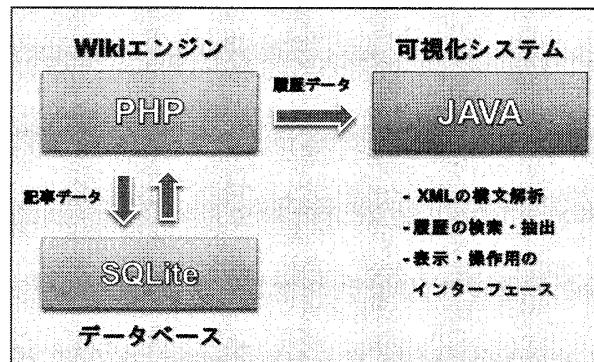


図 1: システム構成図

履歴データの関しては MediaWiki に用意された「ページデータの書き出し」機能で当該記事およびその履歴を XML で取得する。それを Java で構文解析・検索・整理をして、WEB ページにアプレットとして出力する。これら処理はページ単位で行い、複数ページに渡ってクロールすることはない。履歴は最近 50 件を取得し、必要に応じてさらに遡って読み込んでいく。

5 システムの実行

システムの表示部分では節・項というアウトラインに注目した。これは節・項が Wikipedia の最小の編集単位であるのと、編集する際には全体の構造が重要であると考えたからだ。図 2 に示したように、本文とは別に文章全体のアウトラインを常に表示しておき、見出しの色を変えることで差分の有無などを提示する。また、見出し部分をダブルクリックすることで MediaWiki の節・項の編集画面へ移る。

基準となる版は左から 2 番目に表示し、他版との差分に色を付けて表示する。長文などをスクロールする際には全ての版がベースラインを揃えて動き、比較が容易になるように工夫してある。

また、3 章の編集プロセスにおける編集者の視点に対応する 5 つのモードを用意した。

- (1) 節・項に注目して変更が合った過去の版を並べる
- (2) アウトラインに関わる変化のみを抽出して並べる
- (3) 特定の文章に注目して初出の版へと遡ってゆく
- (4) 特定の人物による対象ページの編集履歴を並べる
- (5) 版の編集日時に近い議論を表示する

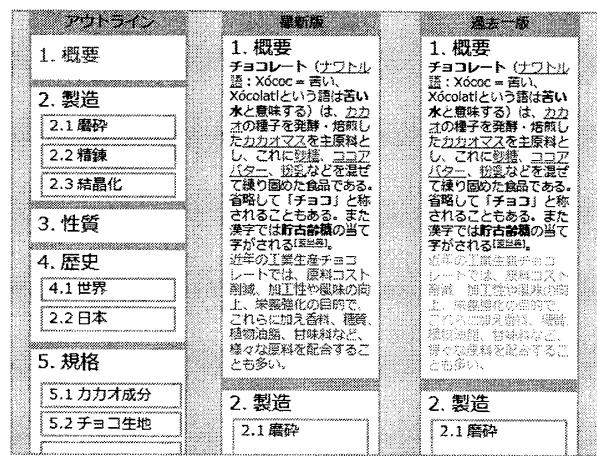


図 2: 実行画面の一部

また、全てのモードでマウス・キーボードによる操作が行え、ドラッグ・ドロップ・スクロール・拡大縮小などができる。

6 おわりに

本稿では、編集履歴の流れの可視化による Wiki の編集支援システムについて紹介した。今後、本稿で提案したシステムの実証実験を行い、システム使用時と非使用時の編集行為をそれぞれ比較する。また分析結果に基づいたシステムの改善を行ってゆく。

参考文献

- [1] Michael D. Ekstrand and John T. Riedl. rv you're dumb: identifying discarded work in wiki article history. In *WikiSym '09: Proceedings of the 5th International Symposium on Wikis and Open Collaboration*, pp. 1–10, New York, NY, USA, 2009. ACM.
- [2] Peter Pirolli, Evelin Wollny, and Bongwon Suh. So you know you're getting the best possible information: a tool that increases wikipedia credibility. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pp. 1505–1508, New York, NY, USA, 2009. ACM.
- [3] Mikalai Sabel. Structuring wiki revision history. In *WikiSym '07: Proceedings of the 2007 international symposium on Wikis*, pp. 125–130, New York, NY, USA, 2007. ACM.