

研究課題別研究評価

1. 研究課題名: 利用目的に応じた情報の組織化と自動編集

2. 研究者名: 佐藤 理史

3. 研究の狙い

インターネット上に存在する膨大な情報を有効に使いこなすためには、計算機による支援が不可欠である。本研究は、「自動編集＝編集の自動化」という新しいアイデアに基づき、インターネット上の情報を自動的に整理することにより、情報をわかりやすく、かつ、使いやすくする方法について研究する。

4. 研究結果及び自己評価:

研究結果

1) レジュームの自動編集という課題を設定し、これを実現するための方法論としてカテゴリにガイドされた自動編集という方法論を提案した。

2) 上記の方法論をサーチエンジンの高度化に応用し、探索結果をレジューム形式で出力するサーチエンジンを、3つのカテゴリ(住所情報、人物情報、用語説明)を対象に実現した。これらのシステムは、いずれも、ウェブから素材情報を探し出し、それらを整理(パッケージング)することにより、典型的な情報要求(「住所を調べたい」、「ある人物について知りたい」、「用語の意味を知りたい」)に対して、その答えと詳しい情報へのポイントを提示することができる。これらのシステムの実現には、情報収集、情報抽出、情報分類、情報統合等の技術が必要であり、これらの技術の開発と統合を行なった。

3) レジュームの自動編集を利用してウェブディレクトリを自動的に作成する方法を示し、水族館、動物園、美術館等の特定カテゴリのウェブディレクトリを自動作成するシステムを実現した。また、これとは別に、有用な情報源を探し出し、そこから得られる情報を再組織化することによってウェブディレクトリを作成する方法を明らかにし、日本の地域情報ディレクトリを自動生成するシステムを実現した。

4) これらのシステムの作成を通して「編集と何か」について考察し、編集の説明理論として「編集＝デパッケージング＋パッケージング」という考えを示した。本研究で作成した5つのシステムで行なわれている編集は、いずれもこの考え方に基づいて明快に説明できる。

自己評価

1) この3年間の研究は、紆余曲折があったものの、「情報の自動編集」というアイデアの具体化とその実現手法の検討、および、デモンストレーションシステムの実装などを行なうことができた。これらにより、自動編集という新しい研究領域の大枠を示せたのではないかと考えている。その意味において、本研究は、十分な成果をあげることができた。

2) 自動編集の具体的課題は、計画時の予定を変更して、研究途上で「レジュームの自動編集」という野心的な課題に見直したが、これが結果的に研究を大きく前進させることとなった。本研究によって、レジューム自動編集の第一段階(「切り貼りによる編集」)は達成できた。しかしながら、本課題の完全な達成のためには、より高度な編集(「加工による編集」)の自動化が不可欠である。これがこれからの研究目標となる。

3)本研究で作成した自動編集システムは、単なるデモンストレーションのためだけでなく、いずれも、現在のニーズを考慮し、実際的に使えるシステムを指向したものである。作成したシステムは、さらに改良してソフトウェアシステムとしての完成度を高めれば、いずれも実用システムとして利用できると考えられる。

4)本研究の目に見えない重要な成果は、編集に対する理解が大きく進んだ点にある。「編集＝デパッキング＋パッキング」と考え方により、自動編集システムの構成が明確になり、ソフトウェアシステムの部品化が促進されることになった。

5)計画時には、自動編集の要素技術として「テキストの言い換え」の自動化についての研究を予定していたが、この点については、研究はそれほど進まず、不十分な成果しか上げることができなかった。「加工による編集」の自動化を実現するためには、言い換えの自動化が不可欠であり、今後、この研究に再びチャレンジする予定である。

5. 領域総括の見解:

大規模情報の検索と編集というテーマに取り組み、大規模カテゴリーにガイドされた自動編集という方法論を提唱、編集という多面的な活動に一貫性を与えただけでなく、実際に自然言語データベースの検索や編集に利用できるシステムを設計、実装し、日本の地域情報ディレクトリを自動生成・編集することに成功したことは高く評価できる。

自動編集や自動要約のためのコンピュータシステムは、大量の情報が Web のホームページ等で流されるようになった今日、多くの人々にとって喫緊に必要なシステムとして研究の的になりつつあるが、一方で基礎から実用化まで相当の研究レベルが必要とされるテーマであった。佐藤理史はさきがけ研究21に採用される以前からきわめて創造的でアクティブな若手の自然言語処理研究者として知られていたが、「情報と知」領域第1期生への応募時にこの重要なテーマを自ら選び、採用されて3年の間、果敢に研究に取り組んで、編集という活動の理論的な基礎から実用化システムのプロトタイプ設計、実装、評価に至るまで、広範にわたって極めて優れた成果を挙げた。

第1期生の研究期間を通じて見ていたところでは、佐藤理史は、この3年間の成果を踏まえ、知的情報処理および自然言語関連分野においてさらに新たな研究プロジェクトをリードしていくことにより、多くの新しい成果を世にもたらすことは確実であり、そうした活動を通してわが国を代表する研究者になることを確信している。

6. 主な論文等:

Satoshi Sato and Madoka Sato. Rewriting Saves Extracted Summaries. In *Intelligent Text Summarization*, Technical Report, SS-98-06, American Association for Artificial Intelligence, pp76-83, 1998.

佐藤理史. 論文表題を言い換える. 情報処理学会論文誌, Vol.40, No.7, pp2937-2945, 1999.

Satoshi Sato and Madoka Sato. Toward Automatic Generation of Web Directories. *Proc. of International Symposium on Digital Libraries 1999 (ISDL '99)*, pp127--134, Tsukuba, September 28-29, 1999.

Satoshi Sato. Automated Editing of Hypertext Résumé from the World Wide Web. *The 2001 Symposium on Applications and the Internet (SAINT-2001)*, San Diego, USA, Jan. 9-11, 2001. (in press)

佐藤理史. ワールドワイドウェブを利用した住所探索. 情報処理学会論文誌, Vol.42, No.1, 2001. (掲載予定)