

ユーザの閲覧情報を利用したWWWの情報検索支援

6N-8

佐々木 祐介 松永 賢次

専修大学経営学部情報管理学科

1. はじめに

現在, WWW空間上をネットサーフをするにあたり, ユーザはハイパーリンクをたどることで情報を得ている。ユーザがある目的をもって, WWW空間上の情報を渡り歩いている時は, ハイパーリンクが指している情報だけではなく, その目的に合致する情報を見たいという要求がある。それらの情報は単純にハイパーリンクをたどるだけでは得られないことが多く, ユーザはサーチエンジンなどを利用して, 関連する情報を探している。ここで, 目的にあった関連ページをユーザが利用している最中に自動的に収集し, ユーザに提示するシステムを用意することにより, ユーザの検索の負担を減らすことができる。本稿で述べるシステムは, これまで読んできたページの単語に, 単語シソーラスを参照し分野シソーラスを分野に振り分けることにより優先順位をつけ, そこからユーザが次に閲覧するURLの候補群を予測し, 先読みしたリンク先のページの中から選択候補を提示するものである。

WWW Browsing Support using User's History

Yuusuke Sasaki and Kenji Matsumaga

Senshu University

2-1-1 Higashi-mita, Tama-ku, Kawasaki, Kanagawa 214-80,
Japan.

2. 本研究における手法

本システムは次のような特徴を持つ。

- (1) ユーザの閲覧履歴を用い, その文書内の単語から必要な情報を推測する。
- (2) ページの対象範囲として, URLの先読みを行うことで, 読んだ情報から数リンク先を含める。
- (3) ユーザに補助情報を入力してもらうことで, 対象範囲やキーワードの設定を制御する。

以下ではそれぞれについて述べていく。







2.1. 閲覧履歴からの推測

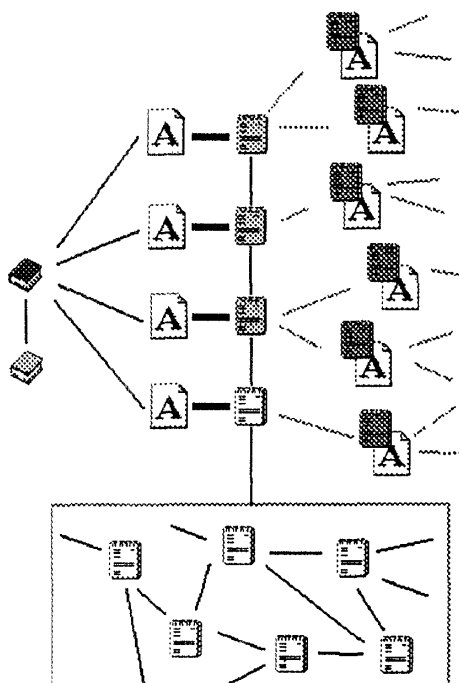
ユーザがそれまで辿ってきたHTML文書(閲覧情報)の内容から, どのような興味を持って文書を選んできているか推測し, 以後読むべきHTML文書の候補群を予測する。そのための手順は以下の通りである。

- (1) ユーザがそれまで閲覧してきたページから単語(名詞)を切り出す。そしてそれらに重み付けの計算をし[1], 単語の優先順位を決める。
- (2) 前もって用意した単語シソーラス(現段階ではコンピュータに関連するものに限定する。本研究では, 単語の表記を整える単語シソーラスと, 単語を特定な分野に指定する分野シソーラスの2つを用いる。)と参照する。
- (3) 単語シソーラスと参照した単語をコンピュータ関連の中でいくつかの分

野に分けるために更に分野シソーラスと参照する。

- (4) 分野にも重み付けの計算をする。この場合の計算方法は、単純に分野毎に分けられた単語を分野別に数えることによって、分野に優先順位をつける。この時、閲覧履歴もページ毎に分野に分ける。これにより、分野がまたがる単語が含まれていても、ページに分野が指定されていればどの分野であるかが特定できる。

-  現在閲覧しているURL
-  過去に閲覧したURL
-  過去に先読みしたURL
-  URLを単語に切り出したもの
-  単語シソーラス
-  分野シソーラス



先読みしたURL

2.2.URLの先読み

2.1で優先順位を定めた単語を、ユーザが利用している最中にあらかじめ先読みした

ページの中から探し出す。検索する場合、現在はサーチエンジンを使用して単語を検索する場合がほとんどであるが、大量の情報を取得してしまうため、かえって混乱をきたす事がある。したがって、先読みの対象範囲を数リンク先と決めた方が良い結果を生むと考えた。

2.3.インターフェース

GUIによりインタラクティブ性を持たせている。クライアントプログラムに対して検索支援をする最も有利な点は、ユーザによるキーワードの設定や、先読みの深さの調整などをユーザに与えてもらうことができる点である。本システムではそのような機能を持たせている。このGUIはJAVAで実装している。

3.終わりに

シソーラスを構築する際に、単語を分類するのに人の手を介して行っているのがおこる。従ってこれを解消するのにユーザが自分でシソーラスをカスタマイズできる必要がある。また、シソーラスの種類も増やせるように柔軟性のあるシステムとユーザインターフェースが必要と思われる。

参考文献

- [1] 平嶋宗, 柏原昭博, 豊田順一: 文脈情報を利用したハイパーテキストのブラウジング支援とその実験的評価, 人工知能学会研究会資料 SIG-HIDSN-9602-04(1996), pp. 19-24.