

# ブックマークエージェントによる WWW の URL 情報の共有

Information Sharing of WWW Using Bookmark-agent

森 幹彦  
Mikihiko Mori

山田 誠二  
Seiji Yamada

東京工業大学 大学院 総合理工学研究科  
Tokyo Institute of Technology

An idea of bookmark-agent was studied on the basis of the file of bookmarks, which is registered user's preference URL addresses and is one of derivatives of the cooperative information filtering. The bookmark-agent we studied is more effective in searching information than existing search engines such as yahoo, altavista and so on. When an user tries to find for certain URLs (i.e. the home pages which is concerned with artificial intelligence) by browsing hyperlink, the agent begins to search for similar web pages to what is dealt with by user at present. At the same time, the agent requests for other agents to search for their own bookmarks, because one agent is stood up as one user in a small scale group, and can correspond with each other. As a result, the user can obtain similar pages by his agent as hyperlinks on his web browser. The information which he got is filtered beforehand and reduced.

## 1 はじめに

インターネットの急速な広がりとともに、ユーザは膨大な情報の渦中におかれるようになった。そこで、ユーザは欲しい情報を得るために多大な労力を必要とする。このため、検索エンジンなどの開発も進められているが、必要以上の情報を与えられてしまうのが現状である [1]。

そこで我々はブックマークエージェントと呼ばれるシステムを提案する。すなわちブックマークエージェントと呼ばれるプログラムは、ユーザのブックマーク\*を参照し、さらに他のエージェントとの通信により、他のユーザのブックマークの URL をも参照して、現在ユーザのしている Web ページと類似したページを提示する。このように、ブックマークエージェントは、ユーザの明示的要求を必要とせず、ユーザの欲する URL 情報を提示できる。また、ユーザは、ブックマークエージェントに、明示的にキーワードを提示することも可能である。

ここで、留意して欲しいことは、ブックマークエージェントが類似ページを検索する範囲は、興味がある程度共通していると想定される小規模なグループのユーザのブックマークに限定されていることであり、ユーザが個々に持つ既にフィルタリングされた URL に関する情報であるブックマークを共有することである。このことは、あるキーワードによる検索において、共通の興味という文脈での絞り込みが行われていることを意味し、一般の検索エンジンよりも効果的な情報フィルタリングが期待できる。また、この分野の先行研究 [2][3] では、ユーザが次にたどるページの予測の学習に主眼が置かれているが、本研究では、グループ間の情報共有に主眼を置いている。

## 2 ブックマークエージェント

### 2.1 システム構成

ブックマークエージェントのシステム構成を図 1 に示す。

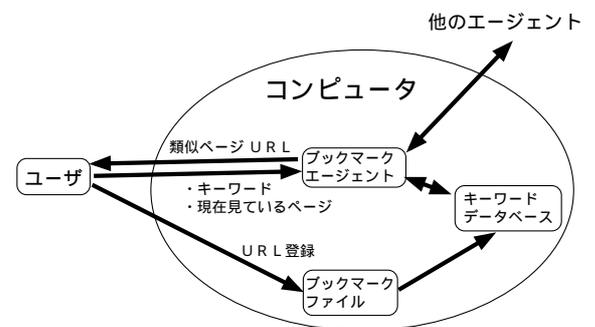


図 1 ブックマークエージェントのシステム構成

### 2.2 基本機能

ブックマークエージェントは、1 ユーザあたりに 1 つ起動される。つまり、ブックマークエージェントは、一人のユーザに対し情報検索支援を行う。ブックマークエージェントの機能を以下に挙げる。

- 新規 URL のキーワード獲得: ユーザのブックマークを監視し、新たな URL の追加があった場合、その URL の HTML ファイルからキーワードを抽出してキーワードデータベースに追加する。
- 類似ページの提示: ユーザが現在見ている Web ページからキーワードを抽出し、それを元にそのページの類似ページの URL をユーザに提示する。さらに、ユーザが明示的にキーワードをブックマークエージェントに与えて、類似ページを提示させることも可能である。ここで、類似ページとは、後述する類似度があるしきい値以上のページを意味する。類似ページの検索方法は、以下の 2 通りがある。

- ブックマークエージェントが担当しているユーザの URL データベースを検索する。

連絡先: 森幹彦 東京工業大学大学院総合理工学研究科

〒226 横浜市緑区長津田町 4259 番地 TEL,FAX: 045-924-5218

email: mori@ymd.dis.titech.ac.jp

\*一度見た興味ある Web ページを記録するファイルのことである。

- 他のブックマークエージェントに URL 情報検索を依頼する。
- 他のブックマークエージェントから URL 情報検索の依頼があった場合、担当ユーザの URL データベースで検索を行い、類似ページを返す。

## 2.3 キーワード抽出と類似度

テキストからのキーワード抽出は、出現頻度を用いるものを始めとして、さまざまな方法が提案されている [4] が、HTML という構造化された文書を扱うことから、本研究では、以下に述べるようなタグ構造に注目したキーワード抽出を用いている。

HTML ファイルにおいて、下記のタグ中に含まれる単語（名詞）に対し、括弧内の重みを加算していき、その重みの上位 5 つをその URL のキーワードとする。

・<META>(10) ・<TITLE>(10) ・<Hn>(6, 4) ・<EM>(1)

また、ページ間の類似度はそれぞれのキーワードの積集合の要素数とする。よって、類似度は 1 から 5 の値をとる。

## 2.4 実行例

以下に実行例を示す。ここではブックマークファイルには人工知能関連の国際会議のホームページが 30 個程度登録されている。

図 2 は、ユーザの現在見ているページである。図 3 は、ブックマークエージェントの Web ページであり、ユーザの現在見ているページの類似ページを提示しているものである。この提示例では、ユーザは”Second International conference on Multiagent Systems ICMAS '96”のホームページを閲覧しており、その類似ページとしてエージェントは”International Conference on Evolvable System(ICES '96)”と”ICMAS '96”自身を提示しており、直観的にも類似したページを提示している。

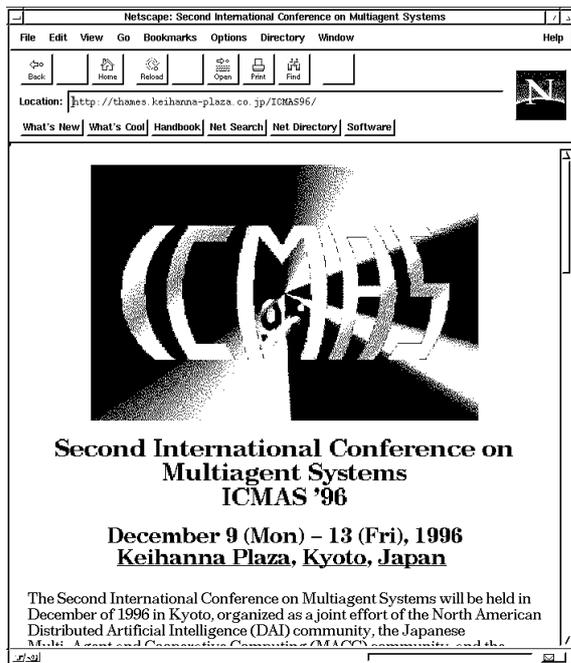


図 2 ユーザの現在見ているページ

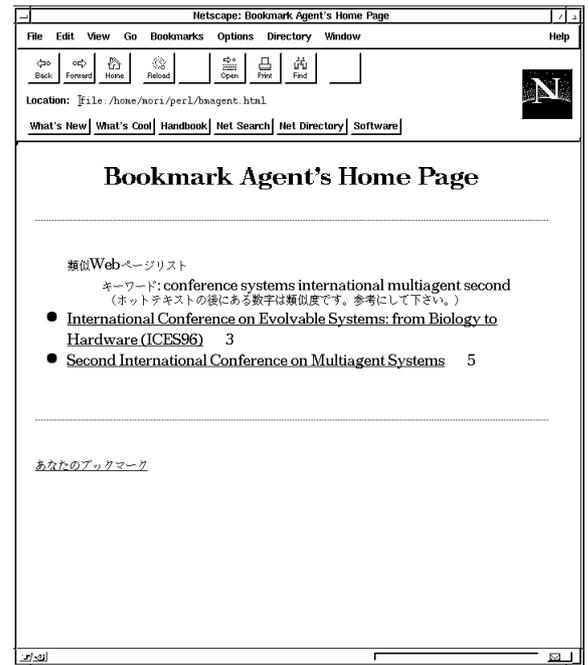


図 3 ブックマークエージェントによる提示

## 3 まとめ

本報告では、URL 情報であるブックマークを用い、情報検索を支援するブックマークエージェントを提案した。ブックマークは、ユーザによってあらかじめフィルタリングされた URL 情報であると考えられ、その情報を用いているブックマークエージェントは、単なる探索エンジンとは違う、適度にフィルタリングされた情報をユーザに提供できることが期待される。

現在、ブックマークエージェントは、開発中である。今後、実際に使用してみて、エージェントの提示する URL が、ユーザの要求する URL に見合ったものか否かを調査する必要がある。

## 参考文献

- [1] 森田, 速水: 情報フィルタリングシステム, 情報処理, Vol. 37 No.8, pp. 751-757 (1996)
- [2] Lieberman, H.: Letizia: An Agent That Assists Web Browsing, *IJCAI-95*, pp. 924-929 (1995)
- [3] Joachims, T., Mitchell, T., Freitag, D., and Armstrong, R.: WebWatcher: Machine Learning and Hypertext, *Fachgruppentreffen Maschinelles Lernen* (1995)
- [4] Salton, G. and McGill, M. J.: Introduction to modern information retrieval, *McGraw-Hill* (1983)