

記憶を中心とする人生の記録

- ユーザの知識空間の作成による Web ブラウジング履歴の想起支援 -

村上 晴美[†] 平田 高志[‡]

[†] 大阪市立大学大学院創造都市研究科 [‡] 防衛庁陸上自衛隊

本研究の長期的な目的は個人の人生の記録である。研究の一部として、人間の記憶を外化記憶と呼ぶコンピュータ上の疑似的な記憶として蓄積するシステム Memory-Organizer を開発してきた。本稿では、Memory-Organizer を利用して Web ブラウジング履歴からユーザの知識空間を作成して、過去の履歴の想起を支援するシステムについて述べる。

Record of Personal Life

- Recall of Web Browsing History by Generating the User's Knowledge Space -

MURAKAMI Harumi[†] HIRATA Takashi[‡]

[†] Graduate School for Creative Cities, Osaka City University [‡] Japan Ground Self Defense Force

The long-term goal of this research is to record personal life. As a part of this research, we have developed a system called "Memory-Organizer," which helps users to construct "externalized-memory." Externalized-memory is a concept that aims to externalize and store the contents of human working memory. In this paper, we describe a system that generates the user's knowledge space from his/her Web browsing history, and helps users to recall their history by using Memory-Organizer.

1. はじめに

人文科学とは広く人類の創造した文化を対象として研究する学問と定義される [1]。人文科学は、遺跡や書物などの人間が残したものを手がかりに研究されてきた。コンピュータが発明されたことにより、コンピュータを用いて人間の思考などの精神生活を蓄積し、後世に残すことが可能となった。

本研究の長期的な目的は個人の人生の記録である。その研究の一部として、人間の記憶を支援するシステム Memory-Organizer を開発してきた。Memory-Organizer では人間の記憶を「外化記憶」と呼ぶコンピュータ上の疑似的な記憶として蓄積する枠組を検討してきた。

Memory-Organizer は、アイデアなどから外化記憶の作成、編集、想起を支援する「思考空間ブラウザ」、Web 閲覧時における外化記憶の作成、編集、想起を支援する「オーバーレイ Web ブラウザ」、ユーザの時系列の興味空間を作成して外化記憶を配置し、想起を支援する「興味空間ブラウザ」、位置情報に基づき外化記憶を地図上に配置し、想起を支援する「地理空間ブラウザ」、日付情報に基づきカレンダー上に外化記憶を配置し、想起を支援する「カレ

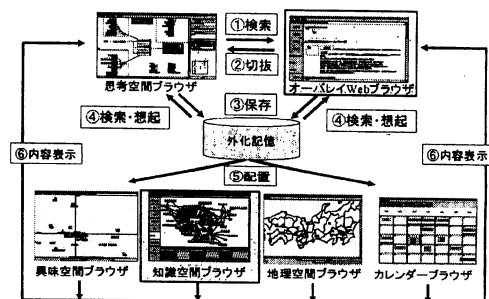


図 1: Memory-Organizer の概要

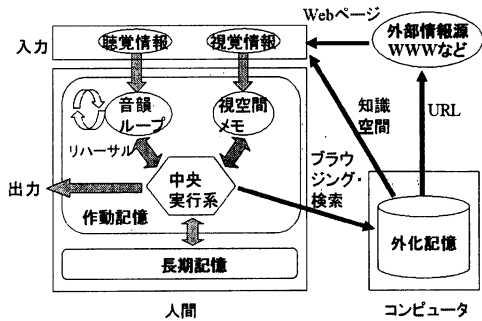


図 2: 外化記憶モデル

ンダーブラウザ」，ユーザの知識空間を作成し，外化記憶の想起を支援する「知識空間ブラウザ」から構成される．図 1 に Memory-Organizer の概要を示す．各ブラウザに関する詳細は [2], [3], [4] などを参照されたい．

以下では，まず，外化記憶モデルの概要について述べる．次に，Memory-Organizer の機能の中から，知識空間の作成による Web ブラウジング履歴の想起支援について述べる．

2. 外化記憶モデルの概要

外化記憶とはコンピュータ上に構築される疑似的な記憶である．Baddeley[5] らの作動記憶モデルに基づき，人間の認知過程において処理する情報を作動記憶を通じて蓄積することを目指している¹．

図 2 に外化記憶モデルの概要と，本システムの関連を示す．Web ブラウジング(リンク選択)時にアンカーテキストを中心とするユーザが処理する情報が外化記憶へ蓄積される．ユーザの指示により，外化記憶から知識空間が作成され，ユーザの入力情報となる．ユーザが履歴を選択すると，外化記憶を経由して URL を送り Web ページがブラウザに表示され，再びユーザの入力情報となる．このように循環的な情報処理を行う．

3. 知識空間の作成による Web ブラウジング履歴の想起支援

本節以降では，Memory-Organizer を利用して Web ブラウジング履歴からユーザの知識空間を作成して，過去の履歴の想起を支援するシステムについて述べる．

図 3 にシステムの概要を示す．ユーザがオーバーレイ Web ブラウザでリンクを選択すると，日時，リン

¹ 実際の作動記憶の内容を蓄積することは現在の技術では不可能と考えられるため，作動記憶の内容の近似的な蓄積を目指す．

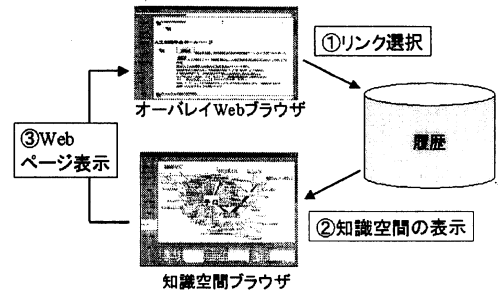


図 3: システムの概要

ク先 URL，アンカーテキストから抽出されたキーワードの組からなる履歴が蓄積される．ユーザのキーワード選択に基づき，知識空間ブラウザ上に，履歴から作成される知識空間と，キーワードを含む履歴一覧を表示する．ユーザが履歴一覧から URL を選択することにより，オーバーレイ Web ブラウザ上に該当の Web ページを表示する．

3.1 オーバレイ Web ブラウザ

オーバーレイ Web ブラウザは，通常の Web ページ表示に加えて，リンク選択時の履歴作成，Web ページへの上書き等の機能を持つ [2][3]．ここでは本稿に関連する独自の履歴作成機能について説明する．

オーバーレイ Web ブラウザの履歴の特徴は，ユーザのリンク選択時の文脈の想起を支援する目的で，アンカーテキストから抽出したキーワードとリンク先の URL を関連付けることである．既存の Web ブラウザの履歴では，日時，URL，該当 URL の Web ページのタイトルが組として蓄積されるが，オーバーレイ Web ブラウザの履歴では，日時，リンク先 URL，アンカーテキストから抽出されたキーワード群が組として蓄積される．なお，本研究では，アンカーテキストからキーワードが抽出されない場合は履歴として蓄積しない．

図 4 に既存の Web ブラウザとの履歴の違いの例を示す．Yahoo!ニュースにおいて，「円急騰 1 2 1 円台，日銀が介入(読売新聞)」というリンクを選択すると，既存の履歴では URL として「headlines.yahoo.co.jp/hl?a=...」，タイトルとして「Yahoo!ニュース - 経済総合 - 読売新聞社」が蓄積される．本研究の履歴では同じ URL に対して「円急騰 1 2 1 円台」「日銀」「介入」「読売新聞」というキーワードの組が蓄積される．

3.2 知識空間ブラウザ

知識空間ブラウザは，ノード(キーワード)とリンクからなるユーザの知識空間を作成することに

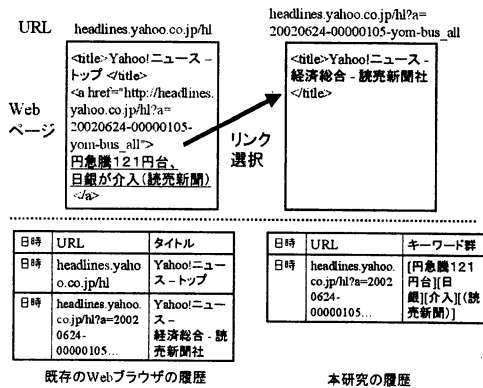


図 4: 既存の Web ブラウザの履歴との違い

より、ユーザの知識の理解と外化記憶の想起を支援する。

知識空間ブラウザは、ユーザが各種操作を行う操作部、知識空間を表示する知識空間表示部、ユーザが選択したキーワードを含む履歴を表示する履歴表示部から構成される。

著者らは 2001 年 4 月からシステムを利用している。図 5 は第一著者の履歴から生成された知識空間ブラウザの画面例であり、2003 年 12 月 17 日に「テロ」「アルカイダ」「自衛隊派遣」を順番に選択して近傍のキーワードを表示した例である。「テロ」という文字が大きく表示され、対米同時多発テロ事件やパレスチナのテロに関するキーワードが見られ、断片的な知識が関連付けられていることがわかる。近傍のキーワードを次々と表示していくため、拾い読みするような検索ができる。履歴表示部において、「自衛隊派遣」を含む URL 一覧がキーワードと共に表示されるため、「自衛隊派遣」「基本計画決定」「陸自 500~700人」などのキーワードを参考にしながら、イラクへの自衛隊派遣に関する 2003 年 12 月 9 日に閲覧したニュース記事を見つけることができる。

知識空間ブラウザのインタフェースは、人間の意味記憶のモデルである活性化拡散モデル [6] を参考にしている。キーワードを選択すると、そのキーワードに関連するキーワードが表示される。現在のところはノード間の距離は活性化拡散モデルにおける意味的な距離を反映していない。記憶の活性化の状態を表現するためにノードの大きさを変更している。

4. システムの流れとアルゴリズム

システムの流れは、(1) Web ブラウジング、(2) 知識空間の探訪、(3) 履歴選択、の 3 段階からなる。この中で、(1) Web ブラウジングはオーバーレイ Web

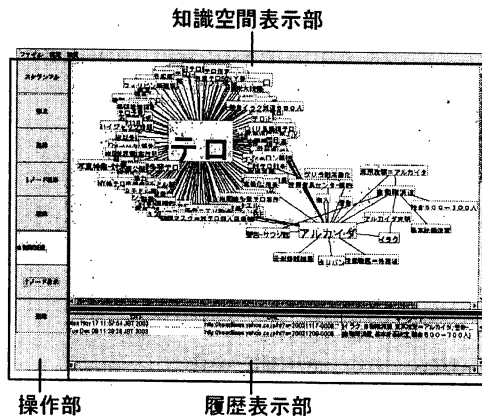


図 5: 知識空間ブラウザの画面例

ブラウザで、(2)、(3) は知識空間ブラウザで行う。図 6 にシステムの流れを示す。

4.1 Web ブラウジング

ユーザがオーバーレイ Web ブラウザ上でリンクを選択すると、日時と、リンク先 URL と、アンカーテキストから抽出されたキーワード群が、組として履歴に保存される。ブラウジング時の履歴作成方法の詳細は [4] を参照。

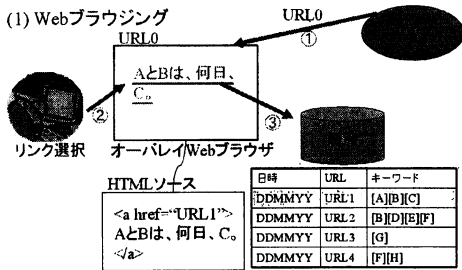
図 6 の例では、ユーザが URL0 の Web ページを表示しているオーバーレイ Web ブラウザ上で、「A と B は、何日、C。」というリンクを選択すると、HTML ソース「A と B は、何日、C.」を解釈し、アクセスした日時、URL1、キーワード群 [A][B][C] の組を履歴として、履歴に保存する。

4.2 知識空間の探訪

ユーザのキーワード選択に応じて、キーワードを履歴から検索し、知識空間表示部に、キーワードと共に共起するキーワードをリンクで結びつける知識空間を表示し、履歴一覧表示部に、キーワードを含む履歴を表示する。

より詳細には以下のとおりである。

- step 1: ユーザがキーワードを入力する。
- step 2: ユーザがキーワードを選択して展開すると
 - 履歴からキーワードを検索し、検索されたキーワードを含む履歴を「限定蓄積履歴 (後述)」に追加する。



- 限定蓄積履歴を用いて知識空間表示部を描画する。
- 履歴一覧表示部に、該当のキーワードを含む履歴を表示する。

本研究における知識空間の特徴は、(a) あらかじめ静的なネットワーク表現の知識空間を作成することではなく、ユーザの操作に応じて動的にネットワーク表現を作成すること、(b) 履歴中の共起するキーワードをリンクで結ぶという単純な手法であること、(c) 人間の認知過程を模倣してノードの大きさをかえることである。

本研究では、図2に示されるような人間の内的機構を仮定し、作動記憶における刺激の頻度に応じて記憶が活性化されると考え、活性化の状態に応じて、キーワードのフォントを大きくする表示を行っている。

作動記憶における刺激には記銘時と想起時の2回が考えられる。履歴中のキーワードの頻度に応じてフォントサイズを変える場合は、記銘時の刺激の頻度に相当すると考える。本研究では「記銘時の刺激+想起時の刺激」を計算する方法を検討するため、限定蓄積履歴という概念を考案した。限定蓄積履歴は、履歴検索の都度、キーワードを含む履歴を追加するものである。

なお、履歴の検索時は、まず、完全一致検索を行い、検索結果が得られなかった場合にのみ前方一致検索を行う。

以下に、図6の例を示す。

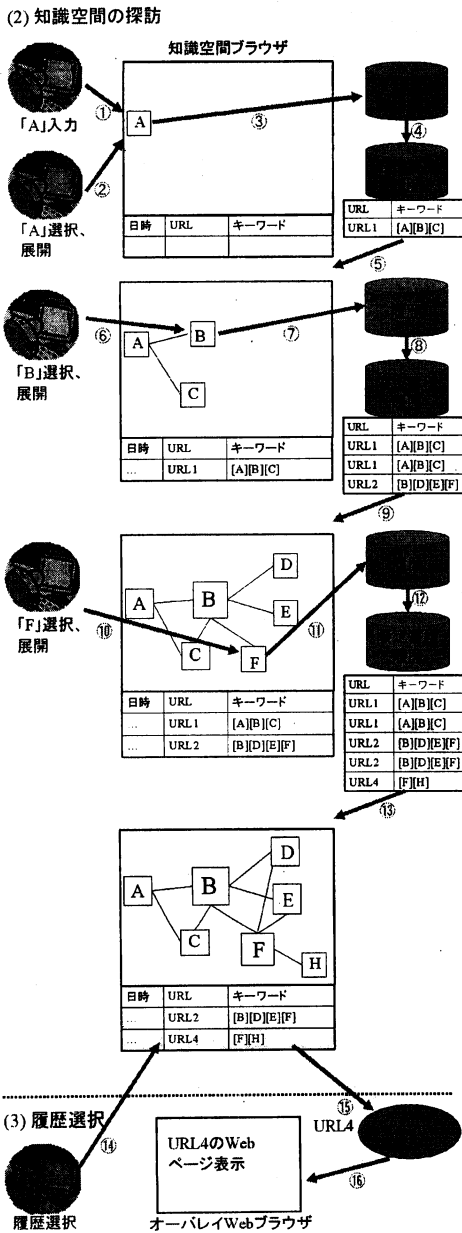


図6: システムの流れ

- ユーザが「A」を入力すると「A」が表示される。
- ユーザが「A」を選択して、近傍のキーワードの表示を指示(以下、展開と呼ぶ)すると、履歴から「A」を検索し、「A」を含むURL1が限定蓄積履歴に追加されて(URL1)となる。「A」と共起するキーワードの「B」「C」を表示し、「A」からリンクがはられる。このとき、限定蓄積履歴における「A」「B」「C」の頻度はそれぞれ「1」であるため、それに応じたサイズでノード表示される。
- ユーザが「B」を選択して展開すると、同様に、限定蓄積履歴に「B」を含むURL1, 2が追加され、限定蓄積履歴は(URL1, 1, 2)となる。「B」から新たに「D」「E」「F」を表示し、「B」から「C」「D」「E」「F」にリンクをはる。限定蓄積履歴における「A」「B」「C」「D」「E」「F」の出現頻度はそれぞれ「2」「3」「2」「1」「1」「1」であるため、それに応じたサイズでノード表示される。
- ユーザが「F」を選択して展開すると、同様に、限定蓄積履歴に「F」を含むURL2, 4が追加

され、限定蓄積履歴は (URL1, 1, 2, 2, 4) となる。「F」から新たに「H」を表示し、「F」から「D」「E」「H」にリンクをはる。限定蓄積履歴における「A」「B」「C」「D」「E」「F」「H」の出現頻度はそれぞれ「2」「4」「2」「2」「2」「3」「1」であるため、それに応じたサイズでノード表示される。

このようにして、ユーザの操作により「A → B → F → H」というパスの表示が可能となる。

本研究では、外化記憶からのユーザの意に沿わない削除 (長期記憶における忘却に相当) は行わない。ただし、ユーザが指定したキーワードの非表示機能や、見やすくするための自動的な非表示機能は必要であると考えている。

4.3 履歴選択

ユーザが履歴一覧表示部から URL を選択することにより、オーバーレイ Web ブラウザに選択した URL の Web ページを表示する。

図 6 の例では、URL4 の履歴を選択すると、インターネットに接続し、オーバーレイ Web ブラウザに URL4 の Web ページを表示する。

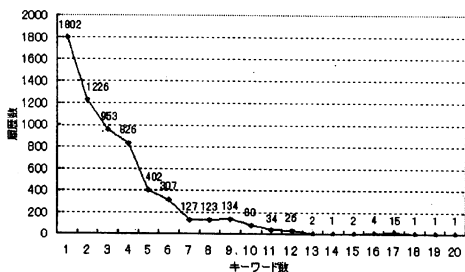
5. 実験

本手法において Web ブラウジング履歴から作成される知識空間が妥当であるかどうかは、(a) キーワードと共起するキーワードがあるか、(b) 別の履歴に一致するキーワードがあるか、(c) 抽出されるキーワード群の内容が妥当であるか、などに関わる。今回は、(a) と (b) に関して予備的な実験を行った。

5.1 Web ページを対象とした実験

第一著者は Memory-Organizer を日常的に使用している。使用条件の詳細は [4] 参照。ここでは、2001 年 4 月 9 日から 2003 年 12 月 17 日までの自由な利用による履歴を調べた。履歴総数は 6081、URL の重複を除いた履歴数は 4364、抽出されたキーワードのべ数は 20158、重複を除いたキーワード数は 9466 であった。1 履歴あたりの平均キーワード数は 3.3 (20158/6081) である。図 7 にキーワード数と履歴数のグラフを示す。

キーワードが 1 つの履歴は 1802/6081 (30%) である。すなわち、この履歴においては、キーワードが抽出できた履歴の 70% が知識空間作成のための候補となりうる。



(注) キーワード数が 20 を越える履歴が 14 存在するがグラフからは省いている。

図 7: 履歴とキーワード数

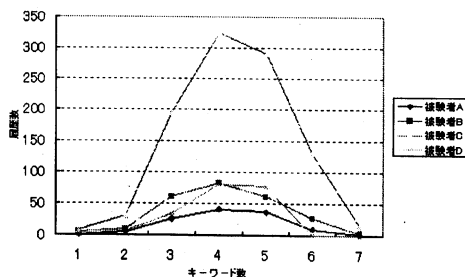


図 8: 履歴とキーワード数

5.2 ニュース記事を対象とした実験

被験者 4 名の毎日インタラクティブの記事 3 か月分の履歴を分析した。月毎に記事のリストを提示し、見たい記事の見出しを、1 日につき 1 つ以上選択させる。実験の詳細は [4] 参照。履歴総数の平均は 396、抽出されたキーワードのべ数の平均は 1735、1 履歴あたりの平均キーワード数の平均は 4.4 (1735/396) である。履歴には 8 つ以上のキーワードは存在しなかった (図 8 参照)。

キーワードが 2 つ以上の履歴の平均は 99% であった。すなわち、ほとんど全ての履歴が知識空間作成のための候補となりうる。

これらの結果は、アンカーテキストからキーワードを抽出する Web ブラウジング履歴が知識空間作成のための情報源となりうることを示している。特にニュース記事の場合は、ほとんどすべての履歴 (見出し) から知識空間を作成できる可能性を示している。

6. 関連研究

個人の人生の記録、Web ブラウジング履歴の可視化の二つの観点から関連研究を述べる。

6.1 個人の人生の記録

個人の人生の記録に関しては、いくつかの研究やプロジェクトがある。

MicrosoftのMyLifeBits[7]では、日常生活におけるあらゆる情報の記録を目的としている。米国防総省は、人間の生活に関するあらゆる情報を記録しデータベース化するLifeLogという実験の構想を発表している[8]。これらは、本研究の長期的な目的とほぼ同じであると考えられる。

Time-Machine Computing[9]やLifestreams[10]では時間軸に基づくデスクトップ上の文書などの個人の情報の組織化を行っている。これらは個人の知的生活の想起を支援するという点で本研究と類似する。

本研究のMemory-Organizerにおける研究の特徴は、人生の記録の中でも、記憶を中心とする人間の認知過程の視覚化に焦点をあてている点である。現時点では日本語のテキスト処理を中心に行っている。

6.2 Web ブラウジング履歴の可視化

Web ブラウジング履歴の可視化に関する研究の多くは、既存のWebブラウザの履歴を対象としている。代表的な研究に、Webページ(またはサイト)をノードとして、ノード間の関連を木やグラフ構造を用いてネットワーク表示することによりナビゲーションを支援する、MosaicG[11]やPadPrints[12]等がある。

既存の研究と本研究との主な相違点は、既存のWebブラウザの履歴とは異なり、ユーザのリンク選択行動に基づく新しい履歴を作成すること、アンカーテキストから抽出したキーワードをノードとする空間を作成し、履歴選択の手がかりとすることである。

7. おわりに

本研究の長期的な目的は、個人の人生の記録である。研究の一部として、人間の記憶を、外化記憶と呼ぶコンピュータ上の疑似的な記憶として蓄積するシステムMemory-Organizerを開発してきた。本稿では、Memory-Organizerを利用してWebブラウジング履歴からユーザの知識空間を作成して、過去の履歴の想起を支援するシステムについて述べた。

参考文献

- [1] 新辞林, Yahoo!辞書検索の結果による。
<http://www.yahoo.co.jp/>
- [2] 村上 晴美, 平田 高志: WWWからの情報獲得・整理支

援 - 思考・興味空間ブラウザ -, 情報処理学会研究報告, Vol.2001, No.20, pp.167-174 (2001).

- [3] 村上 晴美, 平田 高志: Memory-Organizer: 個人の外化記憶構築システム, 2001年度人工知能学会全国大会(第15回)論文集, 3F1-03 (2001).
- [4] 村上 晴美, 平田 高志: Webにおけるリンク選択行動からユーザの時系列の興味空間を作成するシステム, 日本認知科学会テクニカルレポート JCSS-TR-47, pp.1-12 (2003).
- [5] Baddeley, A. D., and Hitch, G.: Working memory, *The Psychology of Learning and Motivation*, Academic Press (1974).
- [6] Collins, A. M., Loftus, E. F.: A spreading-activation theory of semantic processing, *Psychological Review*, 82, 407-428 (1975).
- [7] Gemmell, J., Bell, J., Lueder, R., Drucker, S., Wong, C.: MyLifeBits: Fulfilling the Memex Vision, *ACM Multimedia '02*, 235-238 (2002).
- [8] 米国防総省の『ライフログ』プロジェクト, 一部修正してスタート, *WiredNews*, <http://www.hotwired.co.jp/news/news/culture/story/20030715201.html>
- [9] Rekimoto, J.: Time-Machine Computing: A Time-centric Approach for the Information Environment, *UIST'99* (1999).
- [10] Fertig, S., Freeman, E., Gelernter, D.: Lifestreams: An Alternative to the Desktop Metaphor, *CHI'96* (1996).
- [11] Ayers, E. Z. and Stasko, J. T.: Using Graphic History in Browsing the World Wide Web, *Proceedings of WWW4* (1996).
- [12] Hightower, R. R., Ring, L. T., Helfman, J. L., Bederson, B. B., and Hollan, J. D.: Graphical Multiscale Web Histories: A Study of Padprints, in *Proceedings of ACM Hypertext'98*, pp.58-65 (1998).
- [13] Armstrong, R., Freitag, D., Joachims, T. and Mitchell, T.: WebWatcher: A Learning Apprentice for the World Wide Web, *Proceedings of AAAI Symposium on Information Gathering from Distributed, Heterogeneous Environments*, pp.6-12 (1995).
- [14] Lieberman, H.: Letizia: An Agent That Assists Web Browsing, *Proceedings of IJCAI95*, pp.924-929 (1995).

(注) URLは2003年12月18日にアクセス確認した。