

Web リソースを対象とした プロセス想起型リファインディング支援

Process-Recollection-Based Refinding Support for Web Resources

西本 一平*
Ipppei NISHIMOTO†

戸田 真志‡
Masashi TODA§

Abstract 情報化が進む近年、過去に一度見た情報のリファインディングはますます日常かつ困難な問題になりつつある。このリファインディングという活動において重要な役割を果たしているのが、それを行う当人の記憶や想起である。しかし現存するほとんどのリファインディング支援ではこの特性が十分に考慮されておらず、文脈情報の欠損など多くの問題を引き起こしている。通常の情報検索とリファインディングはユーザにとって全く異なる活動であり、これらは明確に分離して考えるべき問題である。そこで本研究ではプロセス想起型リファインディング支援システムを提案する。プロセス想起型アプローチでは、ユーザ自身が過去のブラウジングプロセスを反復することでリファインディングが行われる。システムの役割は、プロセスにおける Waypoint (目印) を視覚的に提示することで、プロセス反復を手引きすることである。これにより、ユーザが保持している視覚的な手がかりを利用可能にすると同時に、反復されるプロセスに埋め込まれた情報による想起の喚起が期待できる。

1 はじめに

近年の検索技術分野の発展により、ユーザは莫大な Web リソースの中から、求めている情報をすばやく発見することが可能となった。このように情報検索が日常化するなか、新たな問題として重要視されはじめているものがリファインディング、すなわち「過去に一度見た情報をもう一度検索する」[11] というケースである。

我々はブックマークやメール、プリントアウト等の手段を用いて、日常的にリファインディング活動を行っている [11]。またこうした既存のシステムを拡張することで、情報への再アクセスを効率化するタイプの研究も既に数多く行われている。しかしこれらのリファインディング支援システムにはいくつかの重大な欠点が存在する。その中で最も重要な点は、既存手法には「想起を含む活動としてのリファインディング」という視点が欠けているということである。

通常の情報検索が「未知の情報」を対象とする一方で、リファインディングは「既知の情報」を対象としたものである [5]。すなわち、リファインディングとは必然的にユーザの想起によって支えられている活動である。ここに通常の情報検索とリファインディングとの

決定的な差異がある。効率的にリファインディングを支援するシステムを作成するには、この「想起」を有効に機能させるような設計を行うべきである。しかし現在のところ、そのような設計指針を採用したリファインディング支援システムはほとんど存在しない。

そこで本研究ではプロセス想起型のリファインディング支援システムを提案する。プロセス想起型とは、ユーザ自身が初回検索時時のページ遷移のプロセスを再現することでリファインディングを行うというものである。これはブックマークをはじめとする従来の手法、すなわちターゲット情報への直接的なアクセスを提供するものとは根本的に異なるアプローチである。プロセス想起型のメリットは、この一見煩雑な反復作業を利用することで、その過程に埋め込まれた豊富な想起の手がかり情報を利用できるという点にある。それと同時に、従来のリファインディング手法が引き起こすいくつかの問題、PVR や文脈情報の欠損なども同時に解決することができる。

以降、本論文では 2 節でリファインディング分野における研究の現状と問題点を述べ、3 節でその解決に取り組んでいる類似研究を紹介する。4 節ではそれまでの議論をベースとして、リファインディング支援に有効なシステムについて社会的側面・認知的側面の双方から検討を行う。5 節では本研究が提案するシステムのアウトラインを示す。そして 6 節で以上の内容をまとめ、今後の展望を示す。

*公立はこだて未来大学大学院 システム情報科学研究科

†Graduate School of System Information Science, Future-University Hakodate

‡公立はこだて未来大学

§Future-University Hakodate

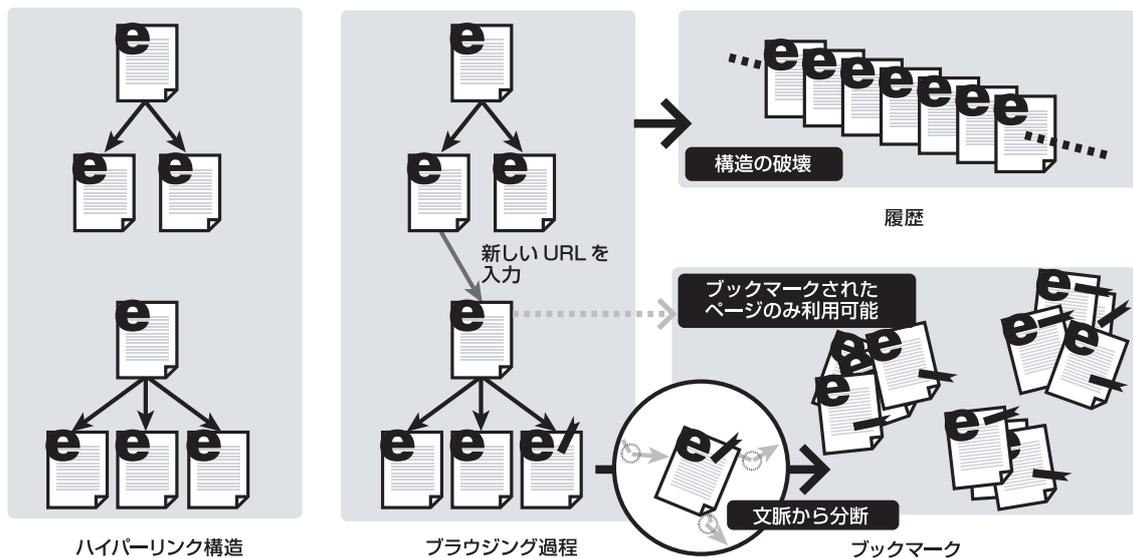


図 1: 各情報空間の特性の違い

2 背景

ブックマークは Web の初期段階からブラウザの標準機能として提供されてきた。Pitkow らの調査 [15] も示している通り、我々は日頃の Web ブラウジングにおける最も身近なリファインディング手段として、ブックマークを日常的に利用している。しかし Abrams ら [3] など多くの研究者によって指摘されている通り、ブックマークには管理困難性、アクセス効率の低さなど多くの問題点がある。このような問題を解消するため、PowerBookmarks [12] をはじめとするブックマーク拡張システムが数多く提案されてきた。

Jones ら [11] はこうしたリファインディング一般の困難性を KFTF (Keeping Found Things Found) としてとりあげた。彼らは人々が日々どのように Web 情報のリファインディングを行っているかをフィールド調査し、それらの手法の機能分析を行った。その結果、ブックマークや履歴に比べて、プリントアウトやメールといった手法は、可搬性などいくつかの点で優れていることを明らかにした。

しかし以下に挙げる通り、Jones らの取り上げたリファインディング手法にはほぼ共通して、根本的な問題がいくつか存在する。

- PVR (Post-Valued Recall)
- 文脈情報の欠損
- 体験の差異

1 つめの問題 PVR (Post-Valued Recall) とは「はじめでその情報を見たときに、その情報が将来必要になる

かを正確に判断できない」というものである [20]。ブックマークを代表とする取捨選択型のリファインディング、すなわちユーザの情報選別を前提とした方法では本質的にこの問題に対応できない。

2 つめの問題は文脈情報の欠損である。この問題はリファインディングシステムのデータ構造によって引き起こされる。そもそもあるページを閲覧するという体験は、ユーザの Web ブラウジングの文脈に意味的・時系列的に埋め込まれたものである。しかし図 1 が示す通り、ブックマークや履歴はこのプロセスから情報の多くを削除し、全く別の情報空間を形成してしまう。しかし実際には、この失われた文脈情報は想起のための手掛かりを豊富に含むものである。Capra ら [6] の実験は、リファインディングにおいて Waypoint が重要な役割を果たしていることを明らかにした。これは、利用可能な文脈情報の量がリファインディングの難易度を決定するという意味を意味する。

3 つめの問題は体験の差異である。この問題はリファインディングシステムのインターフェースによって引き起こされる。ユーザにとって取捨選択された情報空間からのリファインディングは、WWW 空間における以前の検索と根本的に異なる体験である [5]。ユーザの活動という視点で見れば、それは厳密には“リ”ファインディングではなく、新しい空間を対象とした全く新しい検索行為である。このような場合、初回検索時の文脈が再現されないという理由から、リファインディングはより困難なものとなる [19]。文脈情報の中には、例えばユーザがそれを意識しなかったとしても、想起の手がかりとして有効に機能するものは数多く存在する。しかしユーザにとってリファインディングと初回検索

が異なる体験である限り、明確に意識・言語化されないこれらの情報（図2）を手がかりとして利用することはできない。

以上で述べた問題点のいくつかは、次節で述べる通り、既に先行研究によって解決されているものである。しかし現在、これら全ての問題を同時に解決するようなリファインディング手法は存在しない。本研究はグラフデータ構造とプロセス反復を採用することで、これら全ての問題を同時に解決するための方法論を提案するものである。

3 関連研究

一般的なリファインディング手法の中で、唯一取捨選択を行わない、すなわちPVRに対応可能な手法が履歴機能である。しかし履歴機能はリファインディングに効果的なインタフェースを持たず、その検索機能は非常に貧弱なものである。またデータを時系列のみで管理するため、ページ間の関係が破壊されてしまうという文脈情報欠損問題も含んでいる。図1からもわかるように、ブラウジングのプロセスは、ブックマークや履歴はもちろん、WWWのハイパーリンク構造とも異なる独自の情報空間を形成する。より適切なリファインディングを行うためには、ブラウジングのプロセスにおいて生じる情報空間をより正確に保持するべきである。

こうした問題を解決するため、WebNet[7]やMosaicG[4]は履歴を拡張し、データをグラフ構造として管理している。一方、Lifestreams[8]やTimeScape[16]は履歴をベースとした、個人情報全体を対象とするリファインディング支援を提供している。これらは独自のインタフェースによってPVRや文脈情報の欠損を解決しつつ、通常の履歴機能よりも効率的なリファインディングを可能にしている。しかしこれらのシステムでは、初回検索とリファインディングにおける体験差異の問題は解決されていない。ユーザにとって検索とリファインディングが異なる体験である限り、初回検索時のプロセスに埋め込まれた豊富な手がかり情報の多くが切り捨てられてしまう。

これに対し、本研究の提案するプロセス想起型リファインディング支援システムでは、ユーザ自身が以前のプロセスを反復することによってリファインディングが行われる。これはリファインディングと初回検索の状況を一致させ、体験差異の問題を解決するものである。これはまた、本システムと従来のリファインディングシステムの役割が反転しているという特徴をもつ。従来、システムはユーザから受け取ったリファ

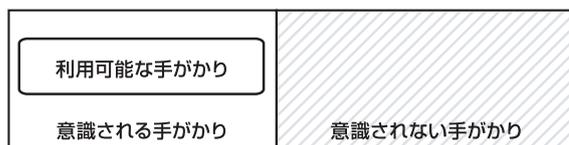


図2: 記憶された手がかりの分類

表1: リファインディングシステムの対応能力

| | PVR | 文脈情報の欠損 | 体験の差異 |
|----------------|-----|---------|-------|
| ブックマーク | | | |
| 履歴 | ✓ | | |
| 構造化履歴 | ✓ | ✓ | |
| 構造化履歴 + プロセス反復 | ✓ | ✓ | ✓ |

インディングターゲットについての手がかりを元に、ターゲット候補を直接提示するものであった。それに比較して、本システムは過去のブラウジングプロセスについての手がかりをユーザに出力し、ユーザ自身が過去のプロセスをスムーズに反復できるよう支援するものである。

なお本システムは、このプロセス反復によるリファインディングアプローチとともに、データ構造としてWebNetのようなグラフ構造を採用する。これにより、表1に示すように、これまで述べてきたPVR、文脈情報の欠損、体験差異の問題を全て同時に解決することが可能となる。

しかしこのようなプロセス反復によるリファインディングは、直接ターゲットにたどり着く従来手法と比較すると、どうしても過程が煩雑にならざるを得ない。本手法よりも従来手法を用いたほうが容易にリファインディングを完了できるようなケースも多く存在する。実際に、プロセス反復の煩雑さをいかに軽減すべきかという点は、本手法が抱える大きな問題点の一つである。しかし本手法は従来手法に取って代わることを目指すものではない。むしろ、従来の手法では断念せざるを得なかった、困難なケースをターゲットとしたものである。

プロセス反復型に似たアプローチの例としては、パーソナライズド検索が挙げられる。パーソナライズド検索では、キーワードを入力することで、過去の利用状況を考慮した検索結果を得ることができる。これは過去に検索したキーワードについてもう一度調べなおしたいというケースにおいて効果的に機能する。しかしこれには二つの問題点がある。まず一つは、「検索エン

ジンでキーワード入力を行い」「直接ターゲットとなる情報に遷移した」という、非常に限られたケースでしか利用できない点である。すなわち、検索エンジンを使わずにたどり着いたようなページのリファインディングには全く利用できない。そしてもう一つの問題は、ユーザが過去に利用したキーワードを明確に想起できなければいけないという点である。ユーザはリファインディング時において、以前利用したキーワードを必ずしも明確に想起できるわけではない。逆にユーザがどれだけ手がかりを持っていても、それが単純なキーワードに還元できないもの、例えば視覚的な情報などであれば、それらは全くリファインディングに役立てる事はできない。実際に Capra らの実験によれば、リファインディングにおいて検索エンジンが利用されたケースはわずか 44.60 %であった [5]。また Teevan らの実験においても、個人情報空間におけるリファインディングにキーワード検索が利用されたケースはわずか 39 %であった [18]。これまで以上にリファインディングを効率化するには、単なるキーワード検索だけではなく、より広い範囲での手がかり情報をリファインディングに利用できるような枠組みを提供する必要がある。

パーソナライズド検索に限らず、基本的に既存のリファインディングシステムは、ユーザが手がかり情報を入力することで初めて支援を行うという仕組みになっている。図 2 が示すように、これらのシステムにおいて利用可能な手がかりは、ユーザに意識され、なおかつ入力インタフェースが提供されているものに限られる。これに対してプロセス反復型アプローチでは、ユーザが明示的にシステムに手がかりを入力する必要はない。それどころか、ユーザが意識していなかったような手がかりについても、プロセス反復の過程においてユーザの脳内で活性化され、有効な手がかりとして機能することが期待できる。すなわち本アプローチにおいて、ユーザは図 2 の全領域の手がかりをリファインディングのために有効活用することができる。

なお、本研究のようにユーザに意識されていない手がかりに着目したシステムの例としては、Remembrance Agent[17] が挙げられる。このシステムでは現在利用しているドキュメントの情報を元に、システムが関連情報だと判断したものを自動的にユーザに提示する機能を備えている。しかしこれは特定の情報をリファインディングしたいというユーザのニーズに応えるものではない。また本研究の目指すプロセス反復型アプローチとも趣旨を異にするものである。

4 設計理論

我々は前節においてプロセス反復型リファインディングの持つメリットを明らかにした。本節では、プロセス反復型の構想を具体的なシステムとしてデザインするにあたり、その背景となる理論を認知的視点から議論する。

なお本研究の提案するシステムは、プロセス反復の手がかりとしてブックマーク情報を利用するという特徴をもつ。その背景について、社会的背景を元に議論する。

4.1 記憶理論とリファインディング

リファインディングは、それをを行う人間の想起と密接に関係したものである。本小節でははじめに符号化特定性原理について述べ、本研究が提案するプロセス反復型リファインディングの有意性を認知的側面から議論する。次に、人間の記憶・想起の特性について検討し、その結果に基づいてシステムのデータ構造設計に関する議論を行う。最後に、人間の想起を喚起するいくつかの効果を参照しつつ、システムがどのようにプロセス想起を支援すべきかについて議論する。

4.1.1 文脈再現の重要性

符号化特定性原理 [19] は想起の成否に関わる重要な理論の一つである。符号化特定性原理によれば、想起の手がかりとして何が有効であるかは、個々のターゲット情報の記銘時における手がかりの符号化状況によって決定される。このことから、記銘時の文脈情報は想起時において非常に重要なものであるということが導かれる。この原理は、2 節で述べた経験差異の問題の重要性を理論的に裏付けるものである。従来の取捨選択型リファインディングでは、リファインディング時（想起時）に初回検索時（記銘時）の文脈が再現されないため、対象情報の想起は困難であった。これに対し、プロセス反復型のリファインディングでは、想起時に記銘時の文脈がほぼ再現されるという特徴をもつ。よって符号化特定性原理の観点から考えても、プロセス反復型リファインディングは想起に有効なアプローチであることがわかる。

4.1.2 記憶構造とデータ構造

次に、リファインディングプロセスと人間の想起のプロセスを比較することで、想起の仕組みをリファインディング支援に生かす方法について検討する。情報

の記録から想起にいたるまでの主要要素を単純化すると以下のようにまとめられる:

1. ターゲット情報の記録と同時に、そのときの文脈情報が手がかりとして符号化される
2. 符号化された事象と類似した事象が起こると、符号化された情報はその事象とともに再符号化される
3. 符号化された文脈情報は想起の手がかりとして利用される
4. 想起とは単に情報を取り出すものではなく、検索手がかりと貯蔵情報を元に検索結果を新たに構築する行為である

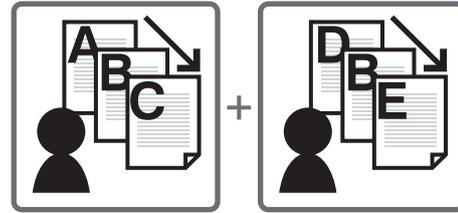
これらが示す通り、いったん脳内に貯蔵された情報は、符号化や想起によって常に再構成され得る動的なものである。それに対しコンピュータに記録されたデータは静的なものであり、通常は読み出しによってその構造が変化することはない。この構造の違いは、リファインディングにおけるユーザの期待とシステムの挙動との間にしばしば齟齬を引き起こす。

LifeStreams[8]をはじめとする時間指向インタフェースでは、データは最終アクセス日時によってソートされる。これは一見、想起による貯蔵情報の再構成機能を模したものであるように見える。しかしこれは単に該当情報のみが再活性化されただけに過ぎず、人間の記憶内部で起こる意味的な再構成・再符号化とは異なる働きである。

そこで、我々は人間の記憶と想起の特性を考慮したデータ構造として、Webの閲覧履歴のグラフ表現を採用する。まずWebにおける1ページを1ノードに割り当て、ユーザのブラウジングセッションで起こったページ遷移に沿ってノード間を接続していく。そして他のブラウジングセッションにおいて過去に閲覧したページが現れた場合、共通ノードを統合することで二つのセッションを結合する。この逐次的な結合機構により、セッション内の文脈情報を保持したまま、セッション間を意味的に構造化することが可能となる。これによりシステムは、過去のセッションを全て取り込んだ巨大な履歴ネットワークを保持することになる。このネットワークは、時系列的な流れを保持すると同時に、ユーザが同一ページを閲覧するたび、その共通部分によって構造を再構成するという特徴をもつ。これはまさにユーザの記憶内容を模したものである。

なお、ここで注意すべき点は、このように構造化されたデータは、単にWebのHyperlink構造を部分的に再現したものではなく、ユーザ固有のものであるとい

ブラウジングの過程:



想起された過程:

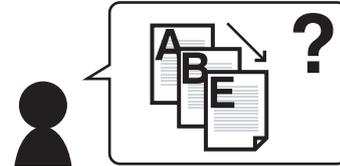


図 3: 記憶混同の例

うことである(この違いは図1においても既に示されている)。

脳内で起こる再符号化はときに、現在記録しようとしている情報と、それに類似した貯蔵情報を混合させてしまうという現象を発生させる。これにより人間はしばしば、生起時期の異なる二つの事象を明確に分離できなくなってしまう。例として、ユーザがまずページA,B,Cの順にアクセスし、後日、今度はページD,B,Eの順にアクセスしたとする。この場合、情報Bが重複していることでユーザの記憶の再符号化が行われ、後に想起において情報Eの手がかりとして情報Aを思い浮かべてしまう可能性がある(図3)。従来のように複数セッションを独立して取り扱う、もしくはセッションの情報を分割する手法では、Aを手がかりとしたEのリファインディングは不可能であった。それに対して本システムの採用するデータ構造は、既に述べたように記憶の再構成機能が再現されている。これにより、こうしたケースにおいてもリファインディングを完了することが可能となっている。

前節で述べたように、履歴をグラフ(ツリー)表現するという仕組み自体はWebNet[7]やMosaicG[4]などで既に実現されている。以上の議論は、認知的側面からこれらのシステムの有用性を明らかにするものである。これらのシステムはそのデータ構造を可視化表現し、ユーザに提示するというインタフェースを採用している。こうしたインタフェースは、ユーザに効率よく大量の情報を伝えるという点においては優れたものである。しかしそれでは過去の経験の想起を喚起することはできない。なぜならユーザにとって、これらのシステムが提供する可視化表現は、リファインディング時にはじめて目にするものだからである。過去にユーザが体験したプロセスに埋め込まれた手がかりを

有効活用するためには、何度も述べている通り、ユーザ自身が以前のプロセスそのものを繰り返す必要がある。本研究では候補を提示することでユーザのプロセス反復を支援するというアプローチを取っており、ここにこれらのシステムとの大きな違いがある。

4.1.3 プロセス想起支援の方法

プロセスにおける中間点・目印として機能する Waypoint は、リファインディングにおいて重要な役割を果たす。我々のシステムは人工的に Waypoint の候補をつくりだし、それをサムネイル表現でユーザに提示する。これはユーザのプロセス想起を支援するものである。Waypoint の候補をつくり出すためには、ユーザがどのページを強く記憶しているかを推測する必要がある。そこで本研究では、ブックマーク、遷移先ページ数、アクセス数、最終アクセス時刻の4要素に着目した。これらの要素に着目した理由、また各要素の優先順位とその根拠について、以降の段落で解説する。

まず、ページがブックマークされている場合、そのページは無条件に Waypoint であると判断する。ブックマークは人工的な手がかりの一つであり、ユーザがそのページを重要なものと判断しているのもっとも明確な証拠である。さらに自己選択効果 [14] の理論に従えば、「ブックマークする」という選択行為自体が記憶促進を引き起こしていることも期待できる。これらの理由から、ブックマークされたページは Waypoint の最も有力な候補であると見なすことができる。なお人工的な手がかりとしてブックマークを採用した経緯については次小節で議論する。

遷移先ページ数およびアクセス数は、互いに密接に関連するものである。遷移ページ数が多いということは、すなわちアクセス数も多いことも意味する。しかし同一アクセス数のページ群の中では、遷移先ページ数が多いものほどより記憶されやすい。なぜなら、遷移先が多いということは、それだけ多くの文脈から該当ページにアクセスされたと考えられるためである。これはそのまま、対象情報を思い出すのに有効な手がかりが多く存在するということの意味する。よって本システムでは、アクセス数が同数のページ群の中では、遷移先ページ数が多いものをより重要視する。

以上3つの要素がすべて同値であった場合、最終的には最終アクセス日時からの経過時間で順序付けを行う。これは単に、古い情報ほどユーザに忘却されている可能性が高いという理由からである。

上述した手順によって Waypoint 候補と判断されたページは、リファインディングの過程で必要に応じて

ユーザに提示される。なお、提示方法としては、ページのタイトルとともにサムネイルを表示する。サムネイルを利用する理由は、サイトのレイアウトや色味といった視覚的な要素からの想起喚起を支援するためである。また一般に画像情報は文字情報よりも優先して想起されるという画像優位性効果 [13] の理論に従ったものでもある。

4.2 ブックマークの変化

2節で述べたように、ブックマークはリファインディング手段としては大きな欠点を持つ一方で、ユーザに日常的に利用されているという特徴がある。それに加え、ここ数年、ブックマークのありかたは大きく変化しつつある。Abrams が分析したブックマークの持つ役割 [3] のうち、個人情報空間的な側面が突出して強くなりはじめていたのである。このパラダイムシフトの中心にあるものが Del.icio.us [1] をはじめとする SBS (ソーシャルブックマークサービス) である [10]。SBS の登場によるブックマークの位置付けの変化は、単なる機能拡張 すなわち複数マシン間での同期や、階層構造に取って代わる柔軟な分類 [9] だけにとどまらない。例えば日本で最もユーザ数の多い SBS であるはてなブックマークでは、コメントやタグによってコミュニケーションの手段としてブックマークが用いられている。また Digg [2] の爆発的普及などは、協働的フィルタリングによる集合知、ネット上における新たな情報伝播モデルなど、今後の Web の可能性を示している。この新しいパラダイムの中で、ブックマークするという行為はモデレーショナルな感覚に近くなり、それに伴いブックマーク数も、従来の個人的なブックマークに比べ飛躍的に増加している。これにより、ブックマークは、ネット上における活動の軌跡、あるいは選別されたライフログという意味合いを強めていると言える。ブックマークは単なる個人的なメモ機能、リファインディング支援の一手法という領域を越え、日常の Web 活動の中にこれまで以上に自然に埋め込まれた重要な要素になりつつある。

2節においてプロセス反復型リファインディングは、従来の取捨選択型リファインディングとは排他的なアプローチとして提案された。しかし上述したような現状を考慮すると、ブックマークを排除してしまうようなシステムは現実的ではない。また、Capra らの実験 [5] でも、ブックマークのような人工的な目印は、Waypoint としてプロセス想起に有効に機能するということがわかっている。

そこで本研究では、ブックマークをリファインディ

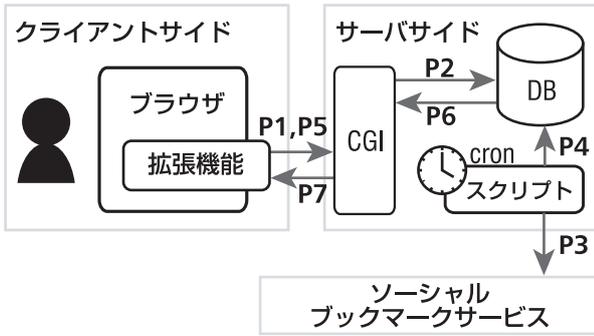


図 4: システムアーキテクチャ

ングの手がかり、あるいはリファインディングの対象という形でシステムに取り入れる。ブックマークの活用により、文脈に埋め込まれた手がかりと人工的な手がかりを相補的に生かしたリファインディングが可能になる。これはまた、本研究の提案するリファインディング支援が二つの側面を持つことを意味する。一つは既に述べた通り、ブックマークがこれまでサポートしてこなかった範囲のリファインディングの支援である。そしてもう一つの側面は、これまで独立した情報群であったブックマークを履歴によって文脈の中に位置付けることで、その発見を支援するというものである。今後、本小節が述べた新しいブックマークのありかたがより普及し、ますますその管理の困難性が問題になるにつれ、この二つ目の側面はますます重要なものとなるだろう。

5 システム概要

5.1 システムアーキテクチャ

本システムのアーキテクチャを図 4 に示す。本システムはサーバアプリケーションおよびクライアント用アプリケーションによって構成される。サーバアプリケーションは LAMP (Linux/Apache/MySQL/Perl) で構成され、クライアントアプリケーションはブラウザ拡張機能 (Firefox Extension) およびユーザスクリプトとして実装される。以下、図に従ってシステム動作のプロセスを説明する。

ユーザがブラウザ上でページ遷移が起こるたび、ブラウザ拡張機能によって図中 P1, P2 のプロセスが実行される。まずサーバ CGI に遷移元ページと遷移先ページについての情報が送信される。このときサーバに送信される情報は、両ページの URL とタイトル、遷移時刻である。CGI は受信したデータをもとに、DB に格納されたデータ構造を更新する。まず URL をキー



図 5: リファインディング候補提示の様子

として DB から対応するノード情報を読み込む。該当するノードが存在しない場合、それは新しいノードとしてデータ構造に追加される。そしてノードに必要な更新処理 (例: Waypoint 評価値の更新やアクセス回数 of インクリメント) を行い、DB にその結果を再度保存する。

サーバはまた一定時間おきに、ユーザによってあらかじめ設定された SBS の RSS を適宜読み込み (P3)、ユーザの新着ブックマーク情報をインポートする。システムはインポートされたデータをもとに Waypoint 評価値の再計算を行い、その更新結果を DB に保存する (P4)。このときブックマークに付加されたコメントやタグといったメタデータも同様に保存する。

ユーザがブラウザからリファインディング支援を要求すると、(図中 P5)、ユーザスクリプトはサーバ CGI に現在閲覧しているページの URL を送信する。CGI は DB から Waypoint 候補を抽出し (P6)、クライアントから与えられた URL 情報を元に選別を行った結果を JSONP 形式でクライアントに返却する (P7)。クライアントは返却された Waypoint 候補の URL を元に、外部のサムネイル作成サービスを呼び出し、その結果画像をユーザに提示する。

5.2 インタフェース

本システムのインタフェースイメージを図 5 に示す。図が示すように、ブラウザウインドウ内の端に Waypoint 候補提示エリアがオーバーラップして表示される。各 Waypoint はタイトル、およびサムネイル画像として表示される。ユーザは必要に応じて Waypoint のサムネイルをクリックすることで該当ページにジャンプすることができる。

図中では右端のエリアに通常の Waypoint 候補が表示されているが、逆に左端のエリアに逆方向候補を表示することも可能である。逆方向候補とは「逆方向の」プロセス反復支援に利用されるもので、例えば「以前、このページの前に見たページを見たい」といったケースなどで有効である。

なお Waypoint 候補はプロセス再現に行き詰まった場合の補助と同時に、煩雑になりがちなプロセス再現の途中過程をショートカットとしての役割も持つ。さらにプロセス内の中間地点のページが消えてしまった場合においても、Waypoint を利用して消えたページを飛び越えることで、プロセス反復の続行が可能となる。

5.3 Waypoint 推測アルゴリズム

Waypoint 候補として提示される URL は、現在ユーザが閲覧中のページからのノード距離が 3 以内のものに限られる。この条件のもと、以下のアルゴリズムによって Waypoint 候補判定が行われる。

まず、該当ページがブックマークされているかどうかを調べる。ブックマークされていれば、そのページは無条件で Waypoint 候補と判断される。この候補群 A は、閲覧中のページからのノード距離を第 1 キー、最終アクセス時刻からの経過時間を第 2 キーとしてソートされる。

次に、該当ページがブックマークされていない場合、そのページからの遷移先ページ数を調べる。このとき遷移先ページ数が 3 以上であれば、そのページは Waypoint 候補と判断される。この候補群 B は、遷移先ページ数を第 1 キー、アクセス数を第 2 キー、ノード距離を第 3 キー、最終アクセス時刻からの経過時間を第 4 キーとしてソートされる。

最後に、該当ページがブックマークされておらず、なおかつ遷移先ページ数が 3 未満の場合、該当ページのアクセス数を調べる。過去に 4 回以上そのページがアクセスしていれば、やはりそのページは Waypoint 候補と判断される。この候補群 C は、候補群 B と同様の手順でソートされる。最終的にユーザには候補群 A,B,C の順に表示される。

問題は、このアルゴリズムのみでは、履歴が蓄積されるにつれ、Waypoint 候補の数が莫大なものになってしまうという点である。この問題を解決するため、本システムではキーワードとアクセス日時による Waypoint 絞り込みのインタフェースを提供する。

Waypoint 一覧が表示されているときにユーザがキーワードを入力すると、一覧の中から、そのキーワード情報を保持しているページのみが抽出される。キーワー

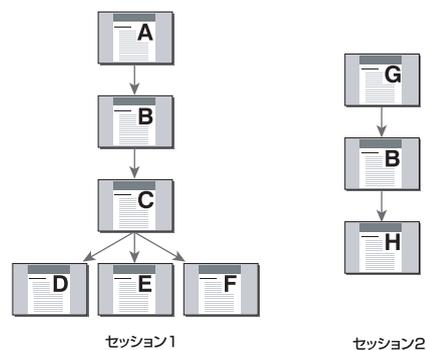


図 6: ページ遷移

ド検索される対象は、URL、タイトル、ブックマーク時につけたタグ、およびユーザがそのページで過去に入力したテキスト情報である。

アクセス日時による絞り込み機能も同様、ユーザが指定した時間範囲内にアクセス記録があるページのみが抽出される。

5.4 利用例

次に、本システムを利用する際の具体的なインタラクション例を示す。このとき、ユーザは以前のセッション 1,2 で図 6 のようにページ遷移したと仮定する。まず、セッション 1, および 2 が終了した時点で、サーバ内で保持されているデータ構造は図 7 のように表現される。そして後日、ユーザは情報 D をリファインディングする際、「ページ A を基点とした」という手がかりが利用可能であったと仮定する。まずユーザはページ A をブラウザで開き、リファインディングモードを有効にする。すると、図 8 のような画面が表示される。ここではページ B,C が Waypoint 候補として提示されている。次に候補ウィンドウからノード B を選択すると、画面は 9 のようになる。ここではノード C が Waypoint 候補として表示されている。ここで今度はノード C をクリックすると、画面は図 10 のようになる。Waypoint 候補であったノード B は現在のノード C から見て逆方向に相当するため、画面左端の領域に移動していることがわかる。そしてユーザは現在ブラウジングしているページ C の中からターゲット情報 D へのリンクを発見する。以上の過程によってリファインディングが完了する。

6 結論

本論文ではプロセス想起型リファインディング支援の概要、およびその具体的なシステム実装例を示した。

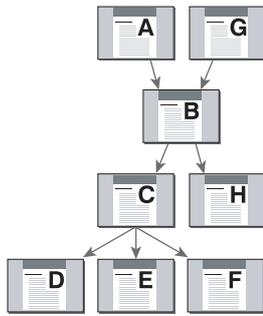


図 7: データ構造

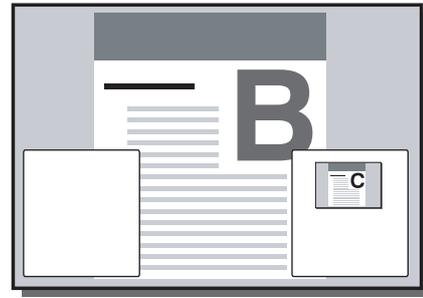


図 9: リファインディング過程-2

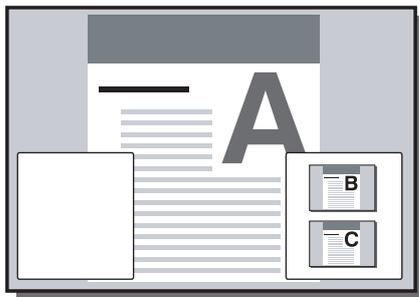


図 8: リファインディング過程-1

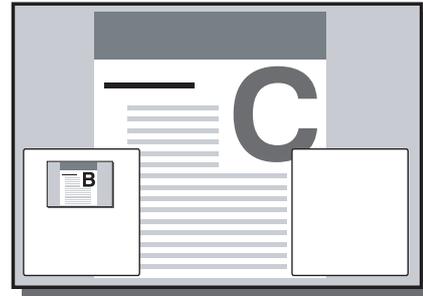


図 10: リファインディング過程-3

近年リファインディング支援に対する注目は高まりつつあるが、従来の支援の多くは取捨選択型のアプローチをメインとしたものであり、「想起」という認知的側面からのアプローチは行われてこなかった。本研究では主に認知的側面から従来の手法の問題点、および本研究の設計方針を明らかにした。

今後の展開としては、想起を支援する認知的効果をより効率的に利用したシステムの機能拡張、便宜的にパラメタを定めた評価関数の改良、インタフェースの洗練などを予定している。具体的な機能拡張例としては、虚偽の記憶を考慮したプロセス支援が考えられる。ブックマークの管理に用いられている意味的構造をプロセス想起に取り入れることで、ユーザが再符号化の過程で、本来は接点を持たない複数のプロセス（あるいは複数のページ）の記憶を混合してしまったような場合においても、適切にプロセス想起を支援できるようなシステムを目指す。

また本システムにて構築したデータ構造を利用することで、本論文で述べたアルゴリズム以外にも様々なリファインディング支援アプローチが可能である。たとえば現在、システムのリファインディング候補算出は、1URLを入力することで、それに対応する独立した候補群を出力するというものである。これを拡張することで、例えば複数URLを入力として与え、そのURL群の中間に閲覧したページを候補として提示した

り、あるいは逆に過去の遷移プロセスそのものを候補として提示することも可能となる。このように、よりさまざまなアプローチからリファインディング支援が行えるよう、システムを発展させていく予定である。

また本研究で得られた結果を、コンピュータでの活動全般におけるリファインディング、さらには実世界でのリファインディングに応用することも考えている。

参考文献

- [1] Del.icio.us. <http://del.icio.us>.
- [2] Digg. <http://digg.com>.
- [3] D. Abrams, R. Baecker, and M. Chignell. Information archiving with bookmarks: personal Web space construction and organization. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 41–48, 1998.
- [4] E.Z. Ayers and J.T. Stasko. *Using Graphic History in Browsing the World Wide Web*. Graphics, Visualization & Usability Center, Georgia Institute of Technology, 1995.
- [5] R. Capra, M. Pinney, and M.A. Pérez-Quñones. Refinding Is Not Finding Again. Technical re-

- port, Technical Report TR-05-10, Computer Science, Virginia Tech, 2005.
- [6] R.G. Capra III and M.A. Pérez-Quiñones. Mobile refinding of web information using a voice interface: an exploratory study. *Proceedings of the 2005 Latin American conference on Human-computer interaction*, pages 88–99, 2005.
- [7] A. Cockburn and S. Jones. Which way now? Analysing and easing inadequacies in WWW navigation. *International Journal of Human Computer Studies*, 45(1):105–129, 1996.
- [8] E. Freeman and D. Gelernter. Lifestreams: a storage model for personal data. *ACM SIGMOD Record*, 25(1):80–86, 1996.
- [9] S. Golder and B.A. Huberman. The Structure of Collaborative Tagging Systems. *Arxiv preprint cs.DL/0508082*, 2005.
- [10] T. Hammond, T. Hannay, B. Lund, and J. Scott. Social bookmarking tools (I). *D-Lib Magazine*, 11(4):1082–9873, 2005.
- [11] W. Jones, H. Bruce, and S. Dumais. Keeping found things found on the web. *Proceedings of the tenth international conference on Information and knowledge management*, pages 119–126, 2001.
- [12] W.S. Li, Y. Hara, R. Ito, Y. Kimura, K. Shimazu, Y. Saito, Q. Vu, E. Chang, D. Agrawal, K. Hirata, et al. PowerBookmarks: a system for personalizable Web information organization, sharing, and management. *Proceedings of the 1999 ACM SIGMOD international conference on Management of data*, pages 565–567, 1999.
- [13] DL Nelson, VS Reed, and JR Walling. Pictorial superiority effect. *J Exp Psychol [Hum Learn]*, 2(5):523–8, 1976.
- [14] LC Perlmutter, RA Monty, and GA Kimble. Effect of choice on paired-associate learning. *Journal of Experimental Psychology*, 91:47–53, 1971.
- [15] J.E. Pitkow and C.M. Kehoe. Emerging trends in the WWW user population. *Communications of the ACM*, 39(6):106–108, 1996.
- [16] J. Rekimoto. Time-machine computing: a time-centric approach for the information environment. *Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 45–54, 1999.
- [17] B.J. Rhodes. The wearable remembrance agent: A system for augmented memory. *Personal Technologies*, 1(4):218–224, 1997.
- [18] J. Teevan, C. Alvarado, M.S. Ackerman, and D.R. Karger. The perfect search engine is not enough: a study of orienteering behavior in directed search. *Proceedings of the 2004 conference on Human factors in computing systems*, pages 415–422, 2004.
- [19] E. Tulving and D.M. Thomson. Encoding Specificity and Retrieval Processes in Episodic Memory. *Psychological Review*, 80(5):352–73, 1973.
- [20] J. Wen. Post-Valued Recall Web Pages: User Disorientation Hits the Big Time. *IT & Society*, 1(3):184–194, 2003.