

WWW 検索行動における「戻る」行動と検索方針の変化との関係

臼澤 基紀*

新垣 紀子†

野島 久雄†

石崎 雅人*

*北陸先端科学技術大学院大学

†NTTコミュニケーション科学基礎研究所

WWW 検索行動の中で「前の画面に戻る」という行動の持つ意味を明らかにするため、戻った後どのような行動を選択するのかということ进行分析し、何のために戻るのかを分類した。まず、戻る行動を記述するモデルを作るため、時間の経過と検索行動の状態を記述するチャートを作成した。次に、検索方法がどのように変化するのかを明らかにするために、どこに戻るのかということと、戻った後選択した行動の関係を分析した。その結果、検索エンジンの選択や検索方法の選択に戻る際には方針変更が見られることが多く、検索結果表示に戻る際には方針の変更が見られないことが示唆された。しかし、キーワードを分析し方針変更の内容を分類したところ、方針変更はキーワードの言い換えのレベルで生じていることが示唆された。

Preliminary study on the changes of the search strategies in the case of “turning back” in WWW information retrieval

Motoki Usuzawa* Noriko Shingaki† Hisao Nojima† Masato Ishizaki*

* Japan Advanced Institute of Science and Technology

† NTT Communication Science Laboratories

In order to understand “turning back” behavior in searching WWW, which means to browse previous pages, we classified the reasons why users often turn back to previous pages by analyzing users’ actions followed by “turning back”. First, in order to make a model which can describe “turning back”, we made a chart which showed relation between time and searching behavior. Second, in order to show how users’ strategies change, we analyzed the relation between “turning back” and the action followed. We suggest a) when subjects selected search engines or operators for searching, they often changed their search strategies, b) but when they selected search results, they did not change their strategies. We analyzed key words selected for searching WWW to classify the pattern about how users changed their strategies. The results suggest that they often generated new key words by paraphrasing old key words.

連絡先: 臼澤 基紀

〒923-1211 石川県能美郡辰口町旭台 1-8 北陸先端科学技術大学院大学 知識科学研究科
e-mail: musuzawa@jaist.ac.jp

1. 背景

1.1. なぜ WWW 検索が難しいのか

国内のインターネット人口は 1998 年 2 月に 1000 万人を超え、今後も着実に増加することが見込まれている[1]。

しかし、ユーザの満足度は必ずしも高くない。WWW 上の情報に関する利用者の評価は厳しく、全体の 80.5% が「情報を探するのに苦労している」と評価している[2]。

WWW 検索の難しさは、web 空間の持つ次のような特徴によっていると考えられる。(1)情報量が膨大で、検索に手間がかかる。(2)情報発信が簡単で、内容についてのチェックなしで公開可能なため、情報の質にばらつきがあり、選択に手間がかかる。(3)どのような情報にアクセスできるかは、ハイパーリンクがどのようにはられているのかに依存しており、全体を見通すことは困難である。

このような WWW 空間を検索する有効な手段として検索エンジンの利用があげられる。WWW ユーザの 9 割以上が検索エンジンを使っている[1]。検索エンジンには、検索式による検索などさまざまな機能があり、それらは検索のノウハウとして紹介されることが多い。しかし、われわれが観察した限りでは、ユーザは検索エンジンの機能をあまり活用していないようである。

1.2. WWW 検索支援に関する研究

WWW 検索行動を支援するアプローチは、情報の取捨選択に伴う負荷をどのように低減させるかに関するものが中心である[3][4]。このアプローチは、情報をいかに提示すべきかといういわば開発者の視点に立ったものである。

しかし、WWW 検索支援を効果的に行うためには、情報の取捨選択に伴う負荷を低減させる方法を開発するだけでなく、検索を行うユーザの行動にも注目する必要がある。なぜなら、行動を分析することでユーザがいつ、どのように困っているのかを明らかにすることができるからである。本研究では、WWW 検索行動を分析するために、行動を問題解決として捉え、特に「前の段階に戻る」という行動に着目した。問題解決は問題空間での探索であり、WWW 検索は問題解決としての側面を持つ。問題解決において、前の段階に戻るという行動は次の 2 つの要素から成る[5]。

- ・現在の方法が失敗であったという判断
- ・新たな行動系列の選択

本研究では、前の段階に戻るという行動を次のように定義した。

- ・ブラウザのバックボタンや検索履歴を利用して前の画面を表示させる行動
- ・キーワードを入力し直す行動

これらの行動は、検索がうまく行かないとユーザが判断した結果生じる行動で、検索支援に有効な手がかりを発見できる行動であると思われる。そこで、われわれは WWW 検索において見られるさまざまな行動の中で、前の段階に戻る(以下「戻る」と略す)という行動に着目した。

2. 目的

本研究の目的は、WWW 検索支援に有効な示唆を与えることである。今回は、以下の 2 点について明らかにするためにユーザの WWW 検索行動を観察した。

- (1)「戻る」行動がいつ、どのように生じるのかを分析する。分析を行うために、SPROT[6]を参考にして行動の系列を記述した。SPROT は、電気製品の操作など適切な手順の存在する行動について、ユーザーの状態を記述することのできる方法である。
- (2)「戻る」行動を分析することで、検索に困っている状態を明らかにする。そのために「戻る」行動の分類を行った。分類の視点として、検索のどの段階で「戻る」行動が見られるのか、どの画面へ「戻る」のか、「戻る」行動の後どのような行動が見られるのか、という点に着目した。

3. 実験

被験者に WWW 検索の課題を行ってもらい、その行動を記録した。概要を以下に示す。

3.1. 方法

実験には Windows95 が動作するノートパソコン(SONY 社製 VAIO PCG729)および 19 インチモニターを利用した。ブラウザには Netscape Communicator4.5 を利用した。WWW 検索の履歴を記録するために、AT マシン(OS: Linux)においてプロキシサーバの一種である Delegate を動作させ、外部ホストへの接続は Delegate を経由して行った。検索のスタートページには実験者が作成したページを使用した。このスタートページには、Yahoo Japan や goo など国内の主要なサーチエンジンへのリンクを 10 個用意した。表 1 に用意した検索エンジンを示す。各課題の開始時にはブラウザのホームボタンをクリックしてスタートページを表示させた。

ビデオカメラを利用して実験中の画面および被験者の発話プロトコルを記録した。

3.2. 手順

実験は、事前評定・検索・事後インタビューの 3 段階で行った。

表 1: 実験のホームページに用意した検索エンジン

Yahoo Japan...	http://www.yahoo.co.jp/
goo.....	http://www.goo.ne.jp/
Infoseek.....	http://www.infoseek.co.jp/
Infonavigator...	http://infonavi.infoweb.ne.jp/
Lycos.....	http://www.lycos.co.jp/
Altavista Japan...	http://altavista.dec-j.co.jp/
Excite.....	http://www.excite.co.jp/
Titan.....	http://titan.mcnet.ne.jp/
Netplaza...	http://netplaza.biglobe.ne.jp/keyword2.html/
CSJ index ..	http://www.csj.co.jp/csindex/

事前評定: 被験者に表 2 の課題を示し、検索の成績との相関を調べるために(1)課題の答えが WWW 上に存在すると思うか、(2)存在するとしたら、自分はそれを探すことができると思うか、について 5 段階で評定をしてもらった。

検索: 被験者は表 2 の課題の検索を順番に行った。なお本実験での課題提示順序は、すべての被験者で同一である。出題の際、被験者に(1)検索時間は最長で 30 分である、(2)自分で納得できる情報が得られたら検索終了とする、という内容の教示を行った。検索終了時点で、検索の答えの URL と、その答えの信頼性に関する 5 段階評定を回答用紙に記入させた。なお、今回の実験では、被験者の答えが正しいかに関するフィードバックは行わなかった。

実験は適宜休憩をまじえ、3 時間で打ち切った。

事後インタビュー: 事後インタビューでは検索段階で記録したビデオと音声将被験者に提示しながら、検索過程でなぜ検索エンジンやキーワードの変更などを行ったのかなどについて質問を行った。

3.3. 課題

表 2: 実験に使用した検索課題

1.	小淵内閣の閣僚名簿を探してください。
2.	関東地方の最新の週間天気予報を教えてください。
3.	日本国内にある、世界遺産の名前と場所を調べてください。
4.	日本の高速道路の総延長は何 km でしょう。
5.	今年のお年玉つき年賀はがきの当選番号を教えてください。
6.	サザエさんの家の間取りがどのようになっているでしょうか。
7.	トマトの卸値段はいくらでしょうか。関東地方の最近 1 ヶ月以内の値段が分かるものを探してください。
8.	田植えから稲刈りまでの日数は何日くらいでしょう。
9.	世界の高山上位 5 傑を順番に挙げてください。
10.	過去 1 ヶ月間の¥/\$相場の推移がどのようになっているでしょう。
11.	日本人に多い姓上位 10 傑を教えてください。
12.	1 月の関東地方の天気の記録を教えてください。
13.	アメリカの祝日一覧を示してください。
14.	トヨタ自動車の現在の社長は誰でしょう。

___ は以下課題を示す略称として使用する

表 2 に実験で使用した課題を示す。課題の作成にあたっては、WWW 上に存在する情報の多様性を反映させるため(a)百科事典で検索可能であるかどうか、(b)いつの出来事であるか、(c)情報の内容のレベルで正解が一つに定められるかどうか、などの条件を考慮した。百科事典で検索可能な情報の例としては、「高山」課題があり、検索不可能な情報の例としては、「お年玉」課題がある。いつの出来事であるかという観点では、「トマト」課題が過去に関する情報の一例であり、「週間天気」課題は未来に関する情報の一例である。正解が一つに定められる情報の例としては、「閣僚」課題があり、そうでないものの例としては「田植え」課題がある。

課題は、検索しにくい課題が連続しないように提示順序を実験者があらかじめ定め、検索を行わせた。

3.4. 被験者

被験者は、21 歳から 39 歳までの 9 名(男 5 名、女 4 名)とした(表 3)。被験者は、1 ヶ月当たりのインターネット利用時間の自己報告をもとに、高頻度利用群(男 4 名)と低頻度利用群(男 1 名、女 4 名)に分類した。なお、両群共に、ブラウザや検索エンジンの操作に関する基本的な知識を持っていた。

表 3: 被験者のプロフィール

	構成(名)	利用頻度(h/月)
高頻度利用群	男 4	常時接続
低頻度利用群	男 1 女 4	1~30

4. 結果と分析

結果の分析は、時間の制約のため次のように行った。

検索にかかった時間とステップ、検索が完了したかどうか、検索結果のばらつき、検索ノウハウの活用については Delegate のログから得られる情報を利用して全員の被験者の分析を行った。

チャートの作成および「戻る」行動の分類については、被験者の行動を撮影したビデオとプロトコルを分析し、Delegate のログに反映されない情報も利用した。

4.1. 全被験者の課題の成績

被験者が 3 時間以内に検索を終了した課題は、9 問から 12 問であった。今回は、すべての被験者が検索を行った課題 9 問についてのみ解析を行った。

4.1.1 検索にかかった時間とステップ

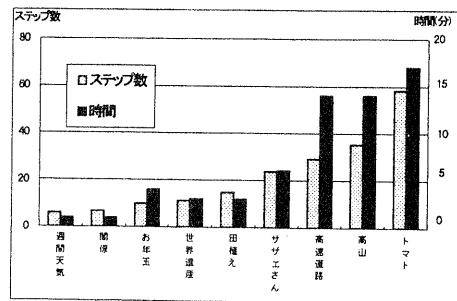


図 1: 検索にかかった時間とステップ数

図 1 は検索にかかった時間とステップ数の関係をグラフ化したものである。時間は、検索開始から最後の画面が表示されるまでの時間、ステップ数は表示した画面の数とした。図 1 の数値は Delegate のログを利用しており、ネットワークの混雑状況やキャッシュの再読み込みなどの情報を考慮していない。そのため大まかな傾向しかつかめないが、検索にかかった時間とステップには相関が見られる。

4.1.2 検索が完了したかどうか

検索の結果は、検索完了(答が見つかった)、検索断念(検索をあきらめた)、検索中止(時間切れ)に分けて整理した。表 4 に示すように課題「トマト」を除いて、ほとんどの被験者は答であるとするサイトを探すことができた。今回の実験では、検索できたかどうかは、ユーザの判断をもとにした。そのため、たとえば「高山」課題について各大陸の最高峰を紹介しているサイトを答であるとするなど、実験者の意図とは違う情報が載っているサイトを答であるとして検索を完了した被験者もいた。

図 1 および表 4 から「高速道路」「トマト」「高山」の課題が比較的難しかったことがわかる。

4.1.3 検索結果のばらつき

次に、課題の難しさとの相関を調べるために、被験者が答であるとしたサイトのばらつきを調べた。

その結果「サザエさん」ではすべての被験者が同じサイトを答であるとした。一方、「高山」や「田植え」のように答のばらつきの大きいものもあった。表 5 に示すように、課題の難しさと答であるとしたサイトの間には目立った関係はなかった。

表 4: 検索が完了したかについて

	検索完了	検索断念	検索中止
閣僚	9		
週間天気	9		
世界遺産	9		
高速道路	8	1	
お年玉	9		
サザエさん	9		
トマト	3	4	2
田植え	9		
高山	8		1

表 5: 被験者が答であるとしたサイトのばらつき

	サザエさん	高速道路	トマト	閣僚	お年玉	週間天気	世界遺産	高山	田植え
サイトA	9	7	2	7	6	5	3	2	2
サイトB		1	1	1	2	1	3	1	1
サイトC			1	1	1	1	1	1	1
サイトD						1	1	1	1
サイトE						1	1	1	1
サイトF								1	1
サイトG								1	1
サイトH									1
あきらめた		1	4						
時間切れ			2					1	

4.1.4 検索ノウハウの活用

ビデオによる記録および Delegate のログから (1)検索エンジンの使い分け、(2)検索式の利用について調べた。検索課題に応じて検索エンジンを使い分けしている被験者のプロトコルおよび事後インタビューを分析した。その結果、一部の被験者は、検索エンジンには得意な分野がそれぞ

れあるというイメージを持っていることが確認された。しかし、その使い分けの基準は明確でなかった。論理演算子(AND, OR, NOT)を用いる検索式については、ほとんどの被験者が活用していなかった。大部分の被験者はスペースの後に検索語を追加し、絞込検索を行っていた。明示的に検索式を使用した被験者は、1名のみであった。

4.2 WWW 検索行動を記述するための枠組み

4.2.1 記述の枠組み

今回、われわれは、WWW 検索行動をモデル化するために「戻る」行動の特徴を調べた。

ユーザの情報検索行動については、野島[7]がオンラインデータベースの一種であるキャブテンシステムの検索行動をモデル化している。しかし、このモデルは被験者が検索をどの段階であきらめるのかということを記述するためのモデルであり、「戻る」という行動を分析するためには適当ではない。「戻る」という行動を分析するためには、どこからどこへ「戻る」のかという情報が必要である。そのため、時間の経過を盛り込んだ記述方法が必要である。

被験者の WWW 検索行動を観察した結果、ほとんどすべての検索課題について「戻る」行動が見られ、ステップ数全体の約 20%を占めていることがわかった。ただし、最も順調に検索が進んだ場合(閣僚)では、検索エンジンの選択、キーワード・ディレクトリ選択、サイトへジャンプ、サイト内検索がこの順序で現れ、前の行動へ戻ることはなかった(図 2)。このような被験者の WWW 検索行動の持つ特徴を明らかにするため、SPROT[6]を参考にして被験者の行動を記述した。

われわれのチャートでは、縦軸に時間の経過を示すため、被験者が表示した画面を時系列に沿って記述した。縦軸には、最も少ないステップ数で検索が進んだときに見られる行動を順調な検索として、検索の段階を記述した。検索の段階は、「検索エンジンの選択」、「キーワード検索」、「ディレクトリの選択」、「サイトへジャンプ」、「サイト内を移動」の 5 種類に分類した。この分類により、検索が順調に進んでいる状態は、チャートにおいては、左上から右下へ直線的に枠が塗りつぶされた状態として描ける。「検索エンジンの選択」は、検索エンジンのトップページを表示させる行動と定義した。「キーワード検索」は、キーボードなどから何らかのキーワードを入力し、検索結果を表示させる行動、「ディレクトリの選択」は、(1)ディレクトリ型検索を行っている場合で、ディレクトリを選択し、検索結果を表示させる行動、(2)ディレクトリ、キーワード両方の検索ができる検索エンジンでキーワード検索を行い、検索結果の中から一致カテゴリを選ぶ行動と定義した。「キーワード検索」および「ディレクトリの選択」については、さらに詳しく、キーワードの追加、ディレクトリの選択行動の内容まで記述した。「サイトへジャンプ」は(1)検索エンジンからサイトへの移動、(2)あるサイトから別のサイトへの移動という行動と定義した。「サイト内を移動」はサイト内のほかのページを見る行動と定義した。図 2 に実際の検索行動の記述例を示す。

このチャートにおいて、「戻る」という行動の例は以下のよ

URL	検索エンジン	キーワード検索					ディレクトリの選択				サイトヘッパ	サイト内を移動	
		山	登山	トエベレス	世界	ブギネックス	ギネス	レサブディ	レサブディ	レサブディ			レサブディ
http://www.yahoo.co.jp/													
http://www.yahoo.co.jp/Reference/													
http://www.yahoo.co.jp/Science/Weights_and_Measures/													
http://www.yahoo.co.jp/Reference/													
http://www.yahoo.co.jp/Reference/Indices/													
http://greenwood.cpm.ehime-u.ac.jp/sato/link/													
http://www.yahoo.co.jp/													
http://www.yahoo.co.jp/Science/													
http://search.yahoo.co.jp/キーワード「山」													
http://search.yahoo.co.jp/キーワード「山」+「登山」													
http://search.yahoo.co.jp/キーワード「山」+「登山」次のページ													
http://www.goo.ne.jp/													
http://www.goo.ne.jp/キーワード「エベレスト」+「世界」													
http://www.php.co.jp/shinkan/ISBN4-569-57233-2.html													
http://www.goo.ne.jp/キーワード「エベレスト」+「世界」													

図2: 被験者の WWW 検索行動の例

うに記述できる。

- ・ブラウザのバックボタンや検索履歴を利用して戻る(図2の矢印A)
- ・キーワードを入力し直す(図2の矢印B)

このチャートを利用すると、以下のことがわかる。

- (1) ある時点において被験者が検索のどの段階にいるのかということ
- (2) 被験者がどのような検索方法を選択しているのかということ
- (3) 被験者のとった検索方法が時間とともにどのように変化しているのかということ

このチャートを用いた記述から、たとえば、被験者が一つの方法に固執するといった状態を取り出し、その状態を分析することで、検索支援の質の向上に役立てることができる可能性がある。

4.2.2 「戻る」行動の分類

どこからどこへ「戻る」のかということを明らかにするため分類を行った(表6)。分類を行う際、「戻る」行動の持つ現在の方法が失敗であったという判断と新たな行動の選択という側面を分けて考えるため、チャート上でキーワード/ディレクトリ選択行動として記述される行動を、「検索結果表示」と「検索方法選択」に更に細かく分類した。「検索結果表示」とは、検索結果の一覧を表示させる行動、「検索方法選択」とは、キーワードあるいはディレクトリを新たに選択し直す行動である。

表6は、例えば、「高山」課題において検索結果表示画面から検索エンジンの選択を行った回数が10回あったということを示している。表6から、「戻る」行動は、「検索結果表示」から「検索方法選択」と、「サイトヘッパ」から「検索結果表示」において多く現れることがわかる。

そこで、「戻る」行動の持つ意味を明らかにすることを試みた。われわれは、戻った後の行動を分析することは、「戻る」行動が生じた理由を推測する有効な手段であると考えた。そこで、戻った後の行動と、どこに戻ったのかということとの対応を明らかにするため分類を行った。結果を表7に示す。分類に際し、チャートから読み取れる検索行動と発話プロトコルを利用し、「戻る」行動の前と後で被験者の行動がどのように変化しているのかを記述した。

表7において、「戻る」行動の後に新たなキーワードの入力やディレクトリの選択が続く行動を「変更」、キーワードの追加を「詳細化」、キーワードの削除を「拡大」に分類した。現在のキーワードやディレクトリの検索を続ける行動を、「維持」に分類した。

表6: 戻る行動の見られた場面の分類

どこから	どこへ	検索結果表示		サイトヘッパ		
		検索エンジン選択	検索方法選択	検索エンジン選択	検索結果表示	検索方法選択
簡潔						
サザエさん		1	1	1	2	
トマト		2	10	3	14	2
田植え			3		9	2
高山	1	10	14	3	17	6

表7: 「戻る」行動の後に続く行動の分類

	変更	詳細化	拡大	維持	行きつ まり
検索エンジン選択	15		1	4	1
検索方法選択	20	16	1		1
検索結果表示				41	1

被験者		トマト	野菜	J A	農協	卸	卸売	卸値	値段	卸値段	市場	青果市場	中央青果市場	中央卸売市場	関東	卸売市場	東京都中央卸売市場
A	時間切れ																
B	あきらめ																
C	あきらめ																
D	できた																

図3:使用したキーワードの広がり例

4.3. 検索範囲の広がり

表7の分類において、検索方法選択の後に続く行動が変化しているものは、「変更」「詳細化」「拡大」に該当する。そこで、後に続く行動の具体的変化を調べるため、検索が難しかった「トマト」課題について被験者がどのようなキーワードを使ったのかを一覧にした。被験者は複数のキーワードを使って検索しており、10個以上のキーワードを使って検索した被験者もいた。われわれはこれらのキーワードの中に相互に似たものがあることに着目した。キーワードをa)概念とその構成要素、b)類義語、c)同義語という点からグループ化を行ったところ、「トマト」課題については、「野菜」、「農協」、「関東」、「卸値」、「市場」という5つに分類できた。多くのキーワードを使って検索している被験者は、キーワードの言い換えを頻繁に行っており、必ずしも多くのカテゴリを検索しているわけではないことが明らかになった。

5. まとめと展望

本研究ではWWW検索行動において「戻る」という行動の持つ意味を明らかにするために、SPROTを参考にしてWWW検索行動を記述した。このチャートを利用して、(1)行動の変化と目標変化との対応、(2)検索範囲の広がりについて検討した。

今回の研究では、分析したデータ数が限られているため、一般的な結論を導き出すことはできなかったが、検索のどの段階へ「戻る」のかによって次に続く行動が違うという結果が得られた。今後、「戻る」行動について、プロトコルを詳細に分析し、被験者がどの程度の確信を持って行動を選択しているのかということに対応づける試みを行う。

謝辞

本研究は、北陸先端科学技術大学院大学とNTTコミュニケーション科学基礎研究所との連携講座の実習として行われた。本研究を行うにあたって、研究の機会を与えてくださった、NTTコミュニケーション科学基礎研究所の石井健一郎氏、萩田紀博氏、吉川厚氏、相川清明氏と、実験実施および結果の解析に際し、貴重なアドバイスをいただいたマルチモーダル対話研究グループの皆様にご感謝いたします。

参考文献

- [1]日本インターネット協会(編):インターネット白書98.インプレス社.1998.
- [2]川上善郎(他):インターネット利用調査報告.
<http://www.ntv.co.jp/bekkoame/index.html>.1996.
- [3]福原知宏:協調フィルタリングに関する研究動向.
http://cairo.aist-nara.ac.jp/~tomohi-/Docs/cf_review/cf_review.html.1998.
- [4]Resnick,P. & Varian,H.R.:Recommender Systems. Communications of the ACM,40(3),56-58.1997.
- [5]Newell,A.:On the analysis of human problem solving protocols. In Jonson-Laird,P.N. & Wason,P.C.(Eds.),Thinking: Readings in Cognitive Science, Cambridge University Press,46-61.1977.
- [6]Itoh,M., Mamine,Y., Goto,S. & Asano,T.: Analyzing Qualitative data with SPROT. In Salvendy,G., Smith,M.J. & Koubek,R.J. (Eds.), HCI International'97,541-544.1997.
- [7]野島久雄, 三宅芳雄: 情報検索における心理要因. テレビジョン学会技術報告, ICS70-7,37-42. 1984.