

検索行動の可視化によるメタ認知促進アプローチからの検索効率向上

和田 洋祐[†] 相場 亮[†]芝浦工業大学大学院 工学研究科 電気電子情報工学専攻[†]

1. はじめに

現在 World Wide Web 上には膨大で多様な Web 文書が存在し、我々は目的に合う情報を探すために Google などの検索システムを用い、日常生活やビジネスの場面で意思決定、問題解決に関わるようになってきた。それにつれてユーザにとって未知の領域を取り扱う場面も増えてきているが未知領域を対象とする検索でユーザは適切なキーワードを思いつけない。またどのような情報が適切かという基準も曖昧なため、ユーザは既有知識から検索を行い、検索結果から新たな知識を獲得し、再度検索を行う試行錯誤的過程を踏むことになる [1]。

検索の結果得られる情報は多量で粒度も様々。さらにリンクによる複雑な構造をしている。そのため、知識獲得を伴う検索行動での認知的負荷は大きく、ユーザは代わり映えのしないキーワードを繰り返し用いたり、手当たり次第に検索したりするので検索が拡散してしまい、目的の情報にたどり着けないことがある [3]。

これらの問題の原因はユーザが自身の検索行動やその過程で得られた知識を把握できず、検索行動が制御できないためであると考えられる。この「自身の認知行動の観察と制御を行うための能力」としてメタ認知が注目されている。本研究では知識獲得を伴う検索行動を行うユーザに対し、知識把握の助けになる情報の提示によってメタ認知を促し、検索の効率化を図ることを考える。

2. 認知活動

未知領域での情報検索はユーザにとって学習を伴う認知活動だと考えることができる。未知領域ではユーザは自身が新しく得た知識を既有知識と関連付け、どのような情報があり、何が不足しているのかという状態を把握し、次の検索を考えていく必要がある。このような機能を果たす人間の能力としてメタ認知がある。

メタ認知は一般に「認知に対する認知」と言われている。そのメタ認知の一つの側面であるメタ認知的活動は認知の過程や状態を観測し、制御する役割を担っている。しかし、メタ認知は認知との並列作業であるために認知的負荷が大きく、また、認知という内界を対象とする観測は非常に困難であることがメタ認知をうまく機能させることができない原因になっている [2]。

3. メタ認知支援の研究

知識獲得を必要とする検索行動の際のメタ認知に関する研究 [3]が行われている。ここではユーザが検索に行き詰ったときに必要なメタ認知に関する考察が行われ、新しく得られた情報から次の検索のリソースを得るためのメタ認知が重要だと指摘されている。学習行動へのメタ認知支援として学習過程をユーザに提示してリフレクションを支援する研究が行われている [4]。リフレクションとは認知状態を可視化することで、メタ認知の観測対象を直接認識可能にし、活動の問題点の発見やその修正を容易にする方法である。可視化方法として、ユーザが学習過程で得た情報同士を自身で関連付けを行う知識マップを提案している。実験では学習リソースを用いて被験者の探索行動の観察を行い、知識マップによるリフレクション支援の有効性が示唆された。

これから、知識獲得を伴う検索行動でもユーザ自身がどのような知識を得たのか把握するためのリフレクションの支援が必要であると考えた。

4. 提案手法

本研究では知識獲得を伴う検索行動に対して、知識獲得の状態を自動的に可視化することでリフレクションの支援を行い、検索の効率化を目指す。

検索行動で見られる過程のうちユーザの知識に変化があったと推測される行動として、クエリの変更とブックマークによるページの保存がある。クエリによって得られた検索結果はクエリに対する周辺情報を含み、話題としての知識構造を作ることができる [5]。しかし、知識獲得を含むプロセスでは、過去に得られた知識と新しい知識の関連を示すことが必要である。

そこでキーワードマップ (以降 KWM) 構成に必要な機能として、(1) 全体の話題を示すのではなくクエリを軸として KWM を構成し、クエリノードの周りに語を展開する。(2) 一度 KWM に現れた要素をすぐには消さず、新しいクエリとの関連を示す。(3) ブックマークで得られた情報を KWM に反映することの三つを考える。

クエリ q_k から得られた検索結果 D_k をもとに KWM G_k を作成する。まず D_k に含まれる検索結果のタイトルとスニペットを単一の文章として形態素解析を行い、単語の出現頻度および共起頻度求める。 q_k に含まれる全キーワード $w_h^k (h = 1, 2, 3, \dots, H)$ を G_k にノードとして加える。 w_h^k と共起する全キーワードとの相互情報量の標準偏差を求め、3 以上のものを関連のある語として KWM に加え、既に KWM に加えられている関連のある語とリンクを結び、新しく加えられた語に対しても同様の操作を行う。これによりクエリと関連のある語とさらにその語と関連のある語が KWM に加わったことになる。次に各ノードとリンクの重要度を求める。

Retrieval visualization for improve efficiency from promoting metacognition approach.

[†]Yosuke Wada, Akira Aiba, Department of Electrical Engineering and Computer Science, Graduate School of Engineering, Shibaura Institute of Technology

$node_i^k = D_k$ が母集団の語 i の出現頻度の標準偏差
 $link_{ij}^k = D_k$ が母集団の語 i, j の相互情報量の標準偏差
 最後に q_k に含まれる語を一つの語と見なし w_0^k を KWM に加え、さらに w_0^k とリンクを結ぶ。次にひとつ前の検索で用いたクエリ q_{k-1} から得られた KWM との合成を行う。合成によって得られた KWM を G_k' とし、 $G_k' = G_k + G_{k-1}'$ とする。ただし G_0' はノードもリンクも持たない空の KWM とする。KWM の合成では同一の語を持つノードとリンクに対して以下の操作を行った。

$$node_i^{k'} = \max (node_i^k, \alpha \cdot node_i^{k'-1})$$

$$link_{ij}^{k'} = \max (link_{ij}^k, \alpha \cdot link_{ij}^{k'-1})$$

α (0~1) は KWM 内での要素の保持率で、値が小さいほど要素の切り替えが激しい。ただし、ブックマークのタグに付加された語とその組合せに対応する要素の保持率は 1 に変更され、KWM から消えないようにする。

5. 実験

5.1. 実験条件

図 1 の環境で実験を行った。被験者は大学生と大学院生の計 8 名。被験者の発話、しぐさ、モニター映像、ブラウザ操作履歴を記録した。提案手法 (B) との比較対象にクエリによる過程を表示しない単純なキーワードマップによるシステム (A) を用意した。この二つと課題の組み合わせからユーザを 4 種に分けて計測を行った。

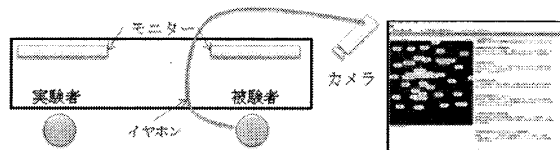


図 1 実験環境と被験者モニターの表示例

5.2. 実験課題

被験者は情報収集を目的とした検索課題を行った。

- ・課題 1
日本の教育が抱える問題点について、どのような問題があるか調べてください。
- ・課題 2
日本と海外のボランティアの違いについて調べるため、ボランティア団体を探してください。

海外のボランティアの定義は日本以外の国に本部があるものとした。また、どちらも最終的にレポートにまとめるという状況設定があることとした。

5.3. 実験手順

被験者は発話の訓練とシステムの利用法について説明を受けた後、課題に取り組んだ。また被験者に対し、各課題の開始前に課題に関する既有知識、終了後に検索過程の説明と新規獲得知識についてインタビューを行った。課題は一つ 25 分行い、5 分毎にキーワードマップを見ながら、現在の検索過程への考察を求めた。実験中は考えていることを発話するように求め、沈黙が 30 秒続く場合には「今何を考えていますか」と質問を出した。

6. 結果・考察

記録できた被験者 6 名のデータの一部を表 1 に示す。

表 1 実験データ

被験者	2	3	4	5	7	8
第一課題組合せ	2A	2B	1B	1A	2B	1B
表示結果件数	210	190	140	200	130	110
選択ページ数	10	7	21	14	15	16
第二課題組合せ	1B	1A	2A	2B	1A	2A
表示結果件数	170	240	180	190	110	150
選択ページ数	17	7	22	16	12	12

表示結果件数は検索結果を合計何件表示したかを示し、選択ページ数は検索結果から何件閲覧したかを示す。

検定に現れるほどの差はなかったが、AB 間での表示結果件数に対する選択ページの割合を比較したところ、システム (B) の方が全被験者共通して大きく、平均して 0.03 大きいことが分かった。これは検索結果 100 件中に 3 件多くユーザの関心を引く結果をもたらしていることになる。また、操作履歴の分析から、ユーザが KWM から新しいキーワードを探す際にクエリノードから離れたところを探索する行動が見られた。これはユーザがクエリノードの周辺に展開されている情報が既知で離れるほど未知である可能性が高いと認知しているからだと考えられる。発話データからは 2 名の被験者が提案システムによるノードの拡張の仕方によって自身の検索範囲の広さに対する考察をしていることが確認された。

7. まとめ

知識獲得を伴う検索過程に対するリフレクション支援による検索効率向上を試みた。ユーザが用いたクエリを軸としてグラフを展開することでユーザの知識変化を抽象的に表現し提示した。一部のデータから効率の向上が確認できた。また発話やシステム操作からも期待したメタ認知を行っていることが確認できた。システムは通常の検索より認知的負荷をかけるため効率化も起こりにくい可能性がある。だが、ユーザがシステムの理解を進めることでより効果が表れると考えられる。

参考文献

- [1] Marcia J Bates, The design of browsing and berrypicking techniques for the online search interface, <http://comminfo.rutgers.edu/~tefko/Courses/612/Articles/Bates.pdf>, 1989
- [2] 平嶋宗, メタ認知の活性化支援, 人工知能学会誌, Vol. 21, No. 1, 2006
- [3] 吉岡敦子, インターネット情報検索における知識構造を促進させるメタ認知の検討日本教育工学会論文誌, Vol. 31, No. 2, p115-123, 2007
- [4] 柏原昭博, 坂本雅直, 長谷川忍, 豊田順一, ハイパー空間における主体的学習プロセスのリフレクション支援, 人工知能学会論文誌, Vol. 18, No. 5B, 2003
- [5] 高間康史, WWW 文書集合族からの時系列的話題情報の抽出・可視化手法の提案, 情報処理学会研究報告, ICS, Vol. 45, p111-116, 2002