

「汎用的な技能」の認知モデル ～検索して編集する認知過程を中心に～

藤田 彬 鈴木 雅之 新井 紀子

国立情報学研究所

E-mail: { a-fujita, m-suzuki, arai }@nii.ac.jp

1 はじめに

近年、日本の高等教育では、学士力(中央教育審議会, 2008)や社会人基礎力(経済産業省, 2007)のような汎用的な技能の獲得が重視されている。例えば、自社の製品のセールスポイントを明確にするためには、他社の製品の情報を収集し、情報を統合することで製品の類似点・相違点などを特定する必要がある。本研究では、このような「ある問題について、必要な情報を情報源から探し出し、収集した情報を編集して、問題を解決する技能」を汎用的な技能の一つとして取り上げる。

こうした問題解決の背景にある認知プロセスは、いまだ未解明であるが、これらのプロセスを明らかにすることで、汎用的な技能を育成するための知見を得ることができる。そこで本研究では、「あるトピックについて文書の適切な箇所を参照して編集するというタスク(以下、検索編集タスク)」を取り上げ、この問題解決の認知プロセスをモデル化し、その妥当性について検討する。

具体的には本研究では、実験参加者に論述問題を課し、検索インタフェイスを介して資料から必要な情報を探し出せる環境下で、解答の作成を求める。実施者は、この解答作業のプロセスの行動ログを収集する。実験参加者が「何をどのように探すか」「一覧できる程度のページの中から該当部分をどのように探すか」「該当部分から求められる観点を満たすように取捨選択ができるか」という三点に絞り、これらの力をどのように総合して検索編集タスクにあたるかを、行動ログとタスクの達成度(評価結果)から分析する。

2 検索及び編集行為の定義と先行研究

本研究では、「設定された目標に対してどのような検索語をキーとすると、情報として必要十分な結果が得られるか(情報を限定できるか)」の推測が伴う検索行為を取り上げる。必要十分な結果が得られるかの推測とは、検索語が「探し出したい文書に必ず含まれること」「関係がない文書に表れないこと」を意識して、検索語を検討する過程をいう。例えば、「内容語に対して、機能語はどのような文書にも表れるから択一性がない」「内容語の中でも、一般名詞でなく固有名詞の方が、文書の特徴づける用語として使われやすい」といった判断が検索語の検討過程として想

定される。

Navarro-Prieto, Scaife & Rogers(1999)は、検索の方策が、1)Top-down: 抽象的内容の文書を概観してから具体性のある文書を絞り出す、2)Bottom-up: 求める情報と一致する内容が書かれた文書が見つけれられるまで、具体的なキーワードを組み合わせた検索語の入力の試行錯誤を繰り返す、3)mixed: 1)と 2)の方策を併用する、というパターンに一般化できることを、実験によって確認した。また Hölischer & Strube (2000)では、設定された課題に対して検索語を導出するプロセスが、検索経験の豊富なユーザと初心者では異なることが示された。

編集行為とは、一定の制約(字数制限、含むべき語など)を守りながら、ある主題について必要十分な事項を含意させるように文章を最適化する行為を指すこととする。資料から主題に関連する箇所(文、節)を抜き出す抽出行為と、抜き出した箇所を制約のもとに圧縮・縮退させる要約行為の2種が含まれる。

Takamura & Okumura(2009)では、文章の要約プロセスを最大被覆問題としてモデル化し、自動的に要約文章を生成する試みがなされているが、検索と編集行為をあわせた一連の活動をモデル化した研究はこれまでにない。

3 認知モデル

本研究では、検索編集タスクの認知プロセスとして、図1のモデルを提案する。このモデルでは、5種の行動が3種の認知活動を反映していることを仮定している。「目標設定」では、問題の解決を目標として問題の解決に必要な情報を集める手立てが思考され、その結果として検索語が決定される。次に「情報の符号化」は、得られた情報の解釈が行われる段階であり、検索結果から適切と思われるペ

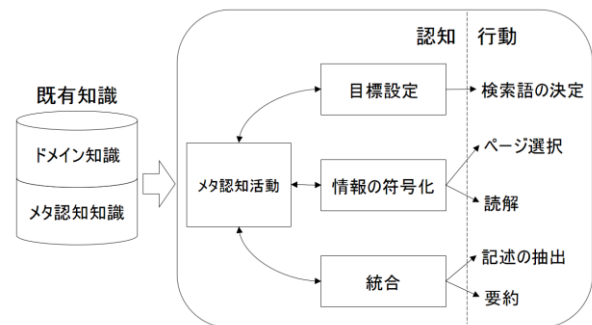


図1 検索編集タスクの認知モデル

ージの選択や、ページに記述された情報の読解が行われる。この目標設定と情報の符号化の2つのフェーズが、検索行為の認知プロセスになる。「統合」では、文字数などの制約条件を考慮しながら、情報の構成が行われることになる。これは、読解した記述から情報を抽出し、それを要約するという行動によって表出される。この統合が、編集行為の認知プロセスである。

これらの認知活動は、メタ認知活動によって支えられる。つまり、各行動の過程では、自身の行動が目標達成にとって適切であるかどうかのモニタリングが行われる。目標達成に不適切と評価すれば、コントロールが行われる。例えば、情報を符号化する際に、問題解決に必要な情報がないと評価すれば、目標が再設定されるだろう。また、自身が産出した文章が、問題に対する解答として不適切・不十分と判断すれば、情報の再統合や、場合によっては目標の再設定が行われるであろう。

これらの認知活動やメタ認知活動は、ドメイン知識(当該課題の領域に関する知識)とメタ認知知識(問題解決方略や課題の性質などに関する知識)の影響を受けると考えられる。例えば、ドメイン知識が豊富であれば、情報間の関連度の強さが分かるため、検索語が特定しやすくなる。また、一般名詞ではなく固有名詞で検索するというのは、検索に関する方略知識の影響を受けたものと考えられる。こうした、何を検索語とすれば良いかという点に限らず、情報の符号化の効率性・正確性や、どの情報を重要視して統合を行うかは、ドメイン知識やメタ認知知識といった既知知識の影響を受けると考えられる。

上記のように仮定したモデルの妥当性を、5節に示す実験により検証する。

4 認知プロセスの分析方法について

本研究では、認知プロセスの妥当性を検証するにあたって、タスク中で入力(更新)される「検索語」「表示ページ」「解答文の編集履歴」を絶対時間とともに記録する機構を用意し、行動過程を自動的に記録する。検索行為が行われるたび、また表示ページを遷移する行為が行われるたびに、検索語、表示ページのそれぞれを記録する。また一定の時間間隔で、その時点で入力されていた解答文を記録する。これら時間軸上に記録された行為(Action)を、隠れ状態である認知活動(Cognition)の説明変数とみなして質的な観点で分析することで、「どのフェーズにどのくらいの時間をかけてどのような認知をしたか」を事後に解明することができる。認知プロセスに直接的にアクセスし、記録・分析する方法としては、問題解決の過程で考えたことや感じたことを口頭で報告させ(think-aloud)、その結果として得られたプロトコル・データを分析する手法であるプロトコル分析(Protocol Analysis)がある(Ericsson & Simon, 1993)。しかし、本研究で用いる課題に対する問題解決のプロセスは言語化が困難であり、ページ選択は必ずしも意識的に行

われるものではないことから、発話プロトコルが実際の認知過程を反映する可能性は低い。したがって、本研究ではより客観的なデータを基に認知プロセスの検証を行うため、発話思考させずに実験を行う。

2節で紹介した先行研究は、サービスとして一般に供される検索エンジンに備わる検索機能を利用して行われたものである。これらの検索エンジンには、検索語が含まれる文書を抽出するという機能に加えて、ユーザの選好性やクエリに対する文書の尤度を考慮した機能が含まれる。例えば、表示ページの順位決定計算の適合性判断に検索者の入力クエリの遷移を考慮する適合性フィードバック(Rocchio, 1971)といった手法が実装されている。本研究では、これらのユーザを補助する機能をバイアスに当たる要素とみなし除外して、認知プロセスを検討する。

また、情報検索のためのインタフェースを提供し、特定の語を含むページを見つけることにかかる時間を統制する。これは、問題に関連する情報を探す際に「対象文書内の関連記述の発見までにかかる時間に、実験参加者の能力とは無関係な差がでること」および「該当箇所にとどりつくまでに運が左右すること」を防ぐためである。

さらに、実験参加者がそれまでに参照したことのない資料を与えることで、問題の答えとなる記述を資料中から探しだすまでにかかる時間を統制する。これは、資料を事前に読んだ経験があるか否かによって、どのような検索語を入力すると、解答に必要な記述を効率的に導き出せるかについての前提知識に差が生じてしまうためである。

5 方法

5.1 実験参加者

男子高校1クラス(36名)と女子高校1クラス(34名)に所属する高校三年生70名である。両校とも偏差値は65程度であり、実験参加者の全員が大学への進学を希望している。

5.2 課題

個人の主張を述べる性質の文章は、その適切性判断(他者による事後評価)に主観が入りやすいことが知られており(Ishioka & Kameda, 2006)、最終的なパフォーマンスの高低について議論することが困難である。一方で、事実について書かれた文章の編集については、その適切性判断に主観が入りにくい。特に、学校教育で扱われる学習教材は内容が統制されており、その教材の内容を要約した文書の適切性判断にも、一般性があり一定の評価基準が設けられている。そこで、文章の評価に評価者の主観が加わらないように、学校教育で扱われる学習内容を題材とし、事実のみを記述する文章の編集をタスクの目標とする。

【問題】: 日本史に関する2種の問題を出題する。両問の内容および出題設定は、日本史教育の専門家により監修されたものであり、実験参加者の意見を記述する性質の問題ではなく、事実を記述する問題である。両問とも解答完了までに15分の制限時間を設ける。

問1は、ある一つのトピックページを探し出し、ある連続した箇所を抜き出すことで、過不足なく答えられる。一方、問2は、ある二つのトピックページを探し出し、それぞれのページ中のある連続した箇所を1か所ずつ抜き出した後、内容を要約することで、過不足なく答えられる。該当箇所を抽出した状態では解答文の文字数が字数制限を超過するため、要約を行う必要がある。

【資料】: 資料には、検定済の日本史教科書(東京書籍『日本史B』(東京書籍株式会社, 2010)を用いる。資料は、歴史上特筆されるイベントや時代を特徴づける社会背景を取り扱う「トピック」という単位の文書にわかれている。1トピック辺りの文書の平均分量は1000字である。

【評価基準】: センター試験の評価基準と学習指導要領に基づいて、専門家が各問の評価基準を策定した。問1の評価基準は2項、問2は7項ある。1項の評価基準を含意したとき、1点が付与される。日常的に記述式解答の採点を行う専門家が評価基準に沿って、答案の評価を行う。1件の答案を異なる2名が評価し、結果が違った際は合議をして評価結果を確定するという形式である。

5.3 手続き

実験参加者には、資料を検索しながら、解答文を編集できるインタフェイスを提供する。このインタフェイスは、検索フォームと解答文入力フォームを有しており、実験参加者が検索編集タスクにあたる間、各種の行動ログを記録する機能をもつ。検索フォームからは、解答文入力フォームへコピー&ペーストが可能となっている。

検索フォームでは、通常の完全一致検索と併せて複数の語が同時に出現するトピックを探索するAND検索ができる。検索を実行すると、まず検索語を含むトピックのページ(以下、トピックページ)へのハイパーリンクとそのトピックページの冒頭部の組み合わせのリスト(以下、ヒットページリスト)が表示される。トピックページの上部には、直前に表示していたページに戻るハイパーリンク、時代の時間軸上での前後のトピックを確認するハイパーリンクがある。「検索行為」および「トピックページの遷移」は、その行為があった時刻と共に入力した検索語、遷移した先のページのURLが行動ログに記録される。またおよそ1秒ごとに解答文入力フォームの中に入力された文字列が行動ログに記録される。

6 結果と考察

6.1 検索編集行為の傾向

はじめに、検索と編集行為の特徴について量的な検討を行う。まず検索行為に関して、同じ検索語で複数回検索している場合には1語とみなして、検索語のカウントを行った。また、AND検索により、複数の検索語を組み合わせている場合は、そのセットを1語の検索語とみなした。その結果、問1の検索語の平均は2.93語($SD = 2.27$; レンジ:0—10個)、問2は3.30個($SD = 1.80$; レンジ:1—8個)であった。このうち、解答として抜き出せる箇所を含むペー

ジがヒットする検索語で検索をした割合は、いずれの問題においても7割程度であった。また、タスク開始後の初回行動は、計140試行(70名の2問分の行動ログ)のうち、133試行が「検索」であった。また、問1は、解答情報を含むページが1つだけであり、この正解ページを表示できなかったのは1名だけであった。一方で問2では、完全正答になるためには、2つのページを参照する必要があり、2ページとも表示できたのは31名、いずれか1ページを表示できたのは38名、いずれも表示できなかったのは1名であった。これらの結果から、ほとんど全ての実験参加者が、解答情報を含むページにいきつくことはできたといえる。

次に編集行為の特徴について検討する。まずテスト得点について、問1の平均点は1.39点($SD = 0.60$)、正答率は70%であり、問2は平均点が3.36点($SD = 1.55$)、正答率が48%であった。ほとんどの実験参加者が、解答となる情報を含むページを参照したにもかかわらず、これだけの正答率であったことから、問題解決に必要な情報を特定するための検索は可能である一方で、テキストの中から解答のために求められる情報を抽出・統合するスキルに問題を抱えている可能性が示唆される。

この点についてより詳細に検討するために、解答となる情報を含むページを最初に参照した直後の行動について、検討を行った。その結果、問1と問2のいずれも、ページから正しい情報を選択して編集を開始した人は全体の約50%であり、ページ内から間違った情報を選択した人が約15%、別のページへジャンプ、別の語で再検索、ヒットページリストに戻った人がそれぞれ約10%であった。これらの結果から、実験参加者は、何をキーワードに検索を行い、どこに必要な情報が載っているかを判断することができても、求められる情報を適切に抽出する能力に問題を抱えている可能性が高いといえる。

ただし、本稿で提示した認知モデルに添えば、目標を再設定し、検索を再度行った後に、解答を含むページを改めて参照し、適切な情報を抽出することも考えられる。そこで以下では、実験参加者の行動ログを質的に分析することで、こうした認知プロセスについて詳細にみていく。

6.2 認知プロセスの検証

表1に、検索とトピックページ遷移の回数が多い実験参加者の行動ログの冒頭を示す。この実験参加者は、表1に示した行動の後も、ヒットしないもしくはヒットするページが多い検索試行を繰り返す。ページ遷移の回数も20回以上に及ぶ。「情報を絞り込む」というメタ認知知識は有しており、適切に目標設定を繰り返してはいるが、情報の符号化の評価が不適切であるため、これら3つの認知フェーズの間を行き来する認知活動を行っていたと考えられる。最終的には、問題と関係のない箇所を抜き出し、解答には評価基準に示された事項が一切含意されなかった。一方、表2に行動ログの冒頭を示した実験参加者は、検索回数が少なく、かつ的確に正解トピックページを選択している。前述の実験参加者と同様に、最初は正解でないペー

ジを選択しながらも、符号化すべき情報がないと適切に判断し、改めて必要な情報が記述されたページを参照し、適切な情報を抽出できている。ここで、検索を改めて行うのではなく、検索結果の中からページ選択を再度行ったのは、目標設定は適切であると判断したためと推察される。しかしながら、解答の文字数が字数制限を上回っており、要約を行った記録がない。したがって、統合の認知活動は完全ではなかったと考えられる。

表 1 検索とページ遷移の回数が多い行動ログの例 (問 1)

経過時間 (m:s)	行動内容
1:04 -3:56	検索試行： 「Query:平安時代の女性」→ヒットなし、 「Query:摂関政治 女性」→ヒットなし、 「Query:摂関政治」→9 トピックヒット(正解トピック含む)、他 10 回の検索。いずれも”正解トピックを含む多数のヒット”か”ヒットなし”
4:43	トピックページ表示：正解でないトピック
5:37	解答追記：「聖武天皇が即位すると、…」(評価基準の含意なし)
6:42 -10:14	ページ遷移：隣接する 15 トピックを参照(2～3 秒の間隔でページ遷移)
…	(中略)
15:00	解答完了：(137 字)「桓武天皇の次に…」

表 2 的確な検索および抽出を行った行動ログの例 (問 2)

経過時間 (m:s)	行動内容
0:25	検索試行：「Query:地頭」→17 トピックヒット(2 つの正解トピック含む)
0:37	トピックページ表示：正解でないトピック
1:07	トピックページ表示：正解トピック①
2:18	解答追記：「頼朝は、後白河法皇に対し、義経の追討を名目として」(評価基準含意)
3:45	検索試行：「Query:地頭 承久の乱」→2 トピックヒット(一つの正解トピック含む)
3:50	トピックページ表示：正解でないトピック
4:24	トピックページ表示：正解トピック②
5:34	解答追記：「承久の乱以降、幕府の支配拡大を背景に、しだいに国衙や荘園領主におさめる官物・年貢を横領したり、国衙の代官や荘園領主の代官を追放したり…」(評価基準含意)
…	(中略)
14:17	解答完了：(397 字)「頼朝が後白河法皇に…」

7 おわりに

本研究では、汎用的な技能の 1 つとして、問題解決に必要な情報を検索・特定し、情報を編集する技能を取り上げ、この問題解決の認知プロセスについて検証を行った。その結果、行動ログの特徴は背景にある認知プロセスによって説明が可能であり、提案した認知モデルの一定の妥当性は確認された。ただし、本研究ではドメイン知識やメタ認知知識の測定を行っておらず、これらの既有知識の個人差が、検索編集行為の個人差とどう関連しているかを実証できていない。また、課題も日本史の論述問題に限定されていることから、今後はこれらの点を改善し、さらなる検証を行う必要がある。

謝辞

本研究に際して、埼玉県立総合教育センターより、研究機会をご提供いただいた。また、学校法人高宮学園代々木ゼミナールより、日本史に関する専門的知識をご提供いただいた。ここに感謝の意を表する。

参考文献

- 中央教育審議会. (2008). 学士課程教育の構築に向けて、文部科学省.
- Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.
- Hölscher, C., & Strube, G. (2000). Web search behavior of Internet experts and newbies, *Proceedings of the 9th international World Wide Web conference on Computer networks*, p.337-346.
- Ishioka, T., & Kameda, M. (2006). Automated Japanese Essay Scoring System based on Articles Written by Experts, *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, , pp.233-240.
- Kammerer, Y., & Gerjets, P. (2013). The role of thinking-aloud instructions and prior domain knowledge in information processing and source evaluation during Web search, *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, pp.716-721.
- 経済産業省. (2007). 社会人基礎力.
- Navarro-Prieto, R., Scaife, M., & Rogers, Y. (1999). Cognitive strategies in Web searching, *Proceedings of the 5th conference in Human Factors and the Web*.
- Rocchio, J. J. (1971). Relevance Feedback in Information Retrieval, *In the SMART Retrieval System: Experiments in Automatic Document Processing*, pp. 313-323. Prentice-Hall Inc.
- Takamura, H., & Okumura, M. (2009). Text summarization model based on maximum coverage problem and its variant, *EACL'09 Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pp.781-789.
- 東京書籍株式会社. (2010). 日本史 B.