

クエリと解の特徴情報を付与するクエリ推薦インタフェース

正元 修平[†] 山本 岳洋^{††,†††} 田中 克己^{††}

[†] 京都大学工学部情報学科 〒606-8501 京都府京都市左京区吉田本町

^{††} 京都大学大学院情報学研究科社会情報学専攻 〒606-8501 京都府京都市左京区吉田本町

^{†††} JST CREST

E-mail: [†]{shogen,tyamamot,tanaka}@dl.kuis.kyoto-u.ac.jp

あらまし 本稿では、Web 検索エンジンのクエリ推薦時に、推薦クエリとともにそのクエリと解の特徴をユーザに提示する手法を提案する。多くの商用検索エンジンでは、ユーザが入力したクエリに応じて動的にクエリを推薦するクエリ自動補完インタフェースが用いられている。しかし、従来のクエリ自動補完インタフェースでは、ユーザは実際に推薦されたクエリを実行して検索結果を閲覧するまで、その推薦クエリを実行したときに求める情報が得られるかが分からない。そこで本研究では、推薦クエリから得られる画像や検索結果の特徴語、推薦クエリに対する推薦クエリに着目し、そうした情報をクエリ推薦時に付与する手法を提案する。これにより、ユーザは推薦クエリを実行する前に、そのクエリが持つ情報の一部を閲覧でき、適切な推薦クエリを選択することができると思われる。また、提案手法を実装し、ユーザ実験を通して提案手法の有用性を検証する。

キーワード クエリ推薦, クエリ自動補完, ユーザインタフェース

1. はじめに

Web の発展に伴い、膨大な情報の中からユーザの求める情報を発見するための手段として、検索エンジンの重要性はますます高まってきている。検索エンジンを利用する際、ユーザは自らの求める情報（情報要求）を表すクエリを、1 語もしくは複数のキーワードの集合として表現する必要がある。しかし、ユーザが入力したクエリは曖昧性を含んでいたり、長さが不十分であったりすることが多く [1]、ユーザにとって自らの情報要求をうまく表現したクエリを作成することは容易ではない。

ユーザのクエリ作成を支援するために、Google^(注1)、Bing^(注2)、Yahoo!^(注3)に代表される多くの商用検索エンジンでは「クエリ自動補完」と呼ばれるインタフェースをユーザに提供している。クエリ自動補完インタフェースでは、ユーザが入力したクエリに応じて、関連するクエリをリスト形式でユーザに推薦する。図 1 に、一般的なクエリ自動補完インタフェースの例を示す。この例では、「京都大学」という入力クエリに応じて「京都大学 図書館」、「京都大学 学園祭」、「京都大学 偏差値」というクエリを推薦・提示している。ユーザは推薦されたクエリを見て、自らの情報要求に最も近い推薦クエリを選択して実行するか、別のクエリを入力して実行するかを選ぶことができる。例えば、ユーザが京都大学の図書館について調べたいときには、「京都大学 図書館」という推薦クエリを選択して実行すれば、求める情報を得ることができる。このように、クエリ自動補完インタ



図 1 一般的なクエリ自動補完インタフェース

フェース（以降、本稿では「クエリ推薦インタフェース」と呼ぶ）はユーザのクエリ作成支援に有用である。

しかし、クエリ推薦インタフェースを利用する際に、以下に示すような問題が生じることがある。京都大学の定期試験について知りたいユーザが、「京都大学 試験」という推薦クエリを見たとする。そして、この推薦クエリを実行すれば求める情報が得られると考え、実行したとする。しかし、実際に「京都大学 試験」というクエリを実行すると、検索結果には京都大学の入学試験についての情報が多く出現し、定期試験についての情報はほとんど出現しないことが分かる。この場合、ユーザは求める情報を得るため、クエリを「京都大学 試験」から別のクエリに修正し、再び検索する必要がある。こうした問題が生じると、ユーザは本来自らの情報要求に適合する情報に到達できなかったり、本来不必要なクエリ修正をしなければならなかったりすることがある。

上記のような問題が生じる原因として、実際に推薦クエリを実行して検索結果を閲覧するまで、そのクエリの意味や、そのクエリを実行したときの検索結果の内容が分からないという点が挙げられる。そのため、実際に推薦クエリを実行するまで、自らの情報要求に適合する情報を得られるかどうかの判断が難しくなっている。ここで、推薦クエリの実行前に、そのクエリ

(注1): <http://www.google.com/>

(注2): <http://www.bing.com/>

(注3): <http://www.yahoo.com/>

を実行したときに求める情報を得られるかどうかの判断に有用な情報をユーザに提示することができれば、ユーザはその情報をもとに、その推薦クエリを実行すべきかどうかの適切な判断を下すことができるだろう。

そこで、本研究では、クエリ推薦インタフェースによって推薦されたクエリから特徴を抽出し、得られた特徴を推薦クエリとともにインタフェースに付与する手法を提案する。ここで、「推薦クエリの特徴」とは、そのクエリを実行したときに自らの情報要求に適合する情報が得られるかを判断する上で有用な情報を、ユーザの理解しやすい形で表したもののことである。

本稿では、推薦クエリに付与する特徴を(1)その推薦クエリを実行した際の検索結果を表す特徴(2)その推薦クエリの意味を表す特徴、(3)その推薦クエリで検索されるトピックの範囲が適切であるかどうかをユーザに気づかせる特徴の3つに分類し、実際にこれらの特徴を取得するための手法を提案する。推薦クエリにこれらの特徴を付与することで、ユーザが推薦クエリを選択すべきか、選択するのであればどの推薦クエリを選択するかを判断することを支援し、検索タスクの効率化を図ることができると思われる。

また、本稿では、提案手法を実際にシステムとして実装し、ユーザ実験を通して提案手法の有用性を検証した。実験では、被験者10名に、既存のクエリ推薦インタフェースと、提案手法により得られたインタフェースを使って実際に検索タスクを行ってもらい、その際の検索行動やアンケートを分析した。

本稿の構成は以下の通りである。2章では、本研究に関連する先行研究について述べる。3章では、本研究が対象とする問題と基本的なアプローチについて説明する。4章では、提案手法とその実装について述べる。5章では、評価実験の方法とその結果について述べる。6章では、提案手法に対する考察について述べる。7章では、本稿のまとめと今後の課題について述べる。

2. 関連研究

2.1 クエリ推薦

クエリ推薦とは、ユーザによって検索エンジンに入力されたクエリに応じて、関連するクエリを提示することでユーザのクエリ作成を支援する技術のことである。多くのクエリ推薦システムでは、入力クエリに対して、検索エンジンのクエリログ[2]やクリックスルーデータ[3]から関連クエリを抽出し、ユーザに提示している。例えば、Caoら[4]は、セッションデータとクリックスルーデータの両者を用いて、現在の検索セッションに適したクエリを推薦する手法を提案している。本研究で提案する手法は、どのようなクエリ推薦を提示するかによらないため、こうした既存のクエリ推薦手法と組み合わせることが可能である。また、Katoら[5]は、クエリ推薦がどのようなときに利用されるのかを分析しており、クエリ推薦は推薦クエリが曖昧でないときによく利用されていると考察している。推薦クエリの曖昧性を解決することは本研究の目的の1つであり、この研究とも関連する。

2.2 クエリ推薦の提示手法

本研究では、クエリ推薦インタフェースによって推薦されたクエリに特徴を付加することで、ユーザがどの推薦クエリを実行すべきかについての判断支援を行っている。これに関連する研究として、Zhaら[6]は、Visual Query Suggestionというシステムを提案している。これは、画像検索を行う上で、クエリ推薦インタフェースにより得られた推薦クエリとともに、クエリの例としてその推薦クエリを実行した際の検索結果画像を数件取り出し、インタフェースに付与するというものである。ユーザはクエリの例を見ることで、文字情報だけではうまく表現できなかった自らの情報要求を表現することができる。本研究は、Web検索エンジンを対象にしておらず、さらに、目的も推薦クエリを選択する判断を支援することで検索タスクの効率化を図るものであるということが、彼らの研究と異なる。また、他にも推薦クエリをグループ化し構造的に提示する手法[7][8]や、階層的に提示する手法[9]なども提案されている。Yahoo!では、クエリ推薦インタフェースにより得られた推薦クエリにマウスカーソルを合わせると、その推薦クエリを実行した際の検索結果の上位数件からタイトル・URL・アイコン(ファビコン)を取り出して得られた検索結果情報が提示される。このシステムは、推薦クエリの実行前に検索結果の一部を提示することで、そのクエリを実行すべきかの判断を支援しているという意味で、本研究の一部と捉えることができる。しかし、このシステムのアプローチは検索結果から数件を切り出して提示するというものであり、検索結果を集約したものを提示している本研究とは、アプローチの点で異なる。

3. 推薦クエリの実行適合性の判断支援

本章では、最初に「クエリの実行適合性」という概念を導入する。その後、クエリの実行適合性の判断時における現状の問題点を述べ、この問題点を解決するための方法について説明する。

3.1 クエリの実行適合性

現在のクエリ推薦インタフェースの問題点を述べるために、まず「クエリの実行適合性」という概念を定義する。クエリの実行適合性とは、ユーザがそのクエリを実行したとき、得られた検索結果中に自らの情報要求に適合する情報が存在するかどうかのことである。例えば、京都大学の学園祭について知りたいユーザがいたとする。ここで「京都大学 図書館」というクエリについて考えると、このクエリを実行して検索結果を閲覧したとしても、検索結果中にユーザの情報要求に適合する情報はほとんどないため、このクエリの実行適合性は低い。それに対して「京都大学 学園祭」というクエリについて考えると、このクエリを実行して検索結果を閲覧すると、ユーザの情報要求に適合する情報が数多く出てくるため、このクエリの実行適合性は高い。

この、クエリの実行適合性という概念を用いて、クエリ推薦インタフェースを利用した際のユーザの一連の検索プロセスを表現したものを図2に示す。図に示すように、ユーザは検索エンジンから提示された推薦クエリを見ながら、各々の推薦クエ

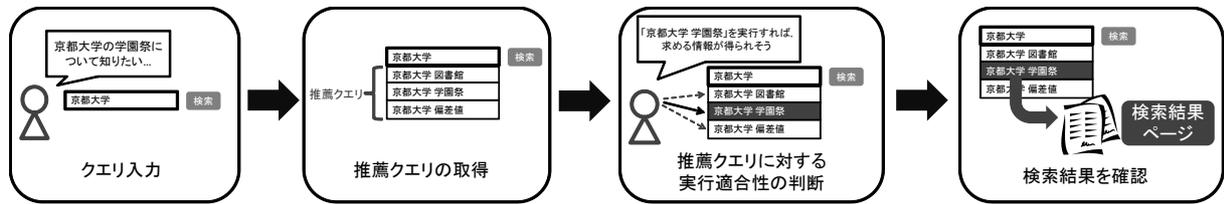


図 2 クエリ推薦インターフェースを用いたときのユーザの一般的な検索プロセス

りについて実行適合性が高いかどうかを判断する．そして，ある推薦クエリについて，その推薦クエリの実行適合性が高いと判断した場合は，そのクエリを実行し検索結果を閲覧していくと考えられる．

3.2 推薦クエリの実行適合性の判断における問題点

3.1 節で述べたクエリの実行適合性を正しく判断できれば，ユーザは求める情報に到達することが可能となる．しかし，この判断はユーザの事前知識や主観に左右されるものであり，また，非常に短い時間で行われることが多い．そのため，以下に示すような問題点が生じると考えられる．

- (1) クエリに曖昧性が含まれていると，ユーザはクエリの実行適合性を誤って判断してしまうことがある．1 章で述べた例がこれに相当する．
- (2) クエリにユーザの知らない言葉が含まれていると，ユーザはそのクエリの実行適合性が高いかどうかを判断することができないことがある．
- (3) 推薦クエリを実行したときの検索結果のトピックの範囲が適切であるかユーザには分からず，トピックの絞り込みを行ったクエリと比べて，このクエリの実行適合性が高いかどうかを判断できないということがある．

(1) (2) (3) のいずれの場合も，ユーザによる推薦クエリの実行適合性の判断が失敗してしまうことによって，本来情報要求に適合するような情報を得られなくなったり，検索タスクの時間的効率が低下したりするといった問題点が生じている．

3.3 推薦クエリの特徴提示による実行適合性の判断支援

本研究では，3.2 節で述べた問題点が生じる共通の原因として，ユーザが推薦クエリを見る段階では，その推薦クエリの実行適合性を判断するのに有用な情報が十分に示されていないことがあると考えた．そこで本稿では，推薦クエリの実行適合性の判断に有用な情報を，ユーザが推薦クエリを見る段階で提示することを提案する．

本研究では，推薦クエリに付与する情報として，以下に示す 3 種類の「特徴」を提案する．

● 解特徴

推薦クエリの解特徴とは，その推薦クエリを実行したときの検索結果の内容を，ユーザが理解しやすい形で表現した特徴である．推薦クエリを実行したときにどのような検索結果があるか分からないユーザが推薦クエリの解特徴を見れば，その推薦クエリの実行適合性を判断しやすくなると考えられる．

● 意味特徴

推薦クエリの意味特徴とは，その推薦クエリの言葉の意味を説明する特徴である．推薦クエリという言葉の意味が分ら

ないユーザが推薦クエリの意味特徴を見れば，その推薦クエリがどのようなトピックを表しているのかが分かり，その推薦クエリの実行適合性を判断しやすくなると考えられる．

● 検索範囲特徴

推薦クエリの検索範囲特徴とは，その推薦クエリを実行することで検索されるトピックの範囲（以降，本稿では「推薦クエリの実行範囲」と呼ぶ）が適切であるかどうか，ユーザに気づかせるための特徴である．推薦クエリによる検索範囲が適切であるか分からないユーザが推薦クエリの実行範囲特徴を見れば，より実行適合性の高いクエリを見つけることができると考えられる．

上記を満たすような情報を各々の推薦クエリから抽出し，クエリ推薦インターフェースに付与することで，ユーザによる推薦クエリの実行適合性の判断を支援することが本研究の目的である．

4. 推薦クエリに付与する特徴の抽出手法

本章では，3.3 節で述べた解特徴，意味特徴，検索範囲特徴の 3 種類の特徴を抽出するための具体的な手法について説明する．以降，本章では，ユーザが入力したあるクエリに対する推薦クエリの集合を $Q = \{q_1, q_2, \dots, q_n\}$ と表す．

4.1 各特徴に共通する指標

推薦クエリに付与する特徴に共通して重要な指標として，以下の 2 点を考えた．

- クエリ推薦インターフェースによって推薦された他のクエリには出現しにくいものである必要がある．

これは，ユーザが推薦クエリのリストを見る上で，提示される各々の推薦クエリの間付加情報に差異がなければ，その情報をもとにどの推薦クエリを実行すべきかを判断することが困難となるからである．

- ユーザが付加情報を理解しやすいような分量・表現形式である必要がある．

これは，ユーザが理解できないような大量の情報を提示したり，表現形式がユーザにとって理解しにくかったりすると，その情報をもとに推薦クエリの実行適合性を判断することが困難となるからである．

これらの点に，各特徴に特有な指標を加えることで，各特徴に必要な指標とする．

4.2 検索結果を用いた解特徴の取得

3.3 節で述べたように，推薦クエリの解特徴に重要な指標は，推薦クエリを実行したときの検索結果を適切に表現していることである．本稿では，解特徴として推薦クエリを実行した際の Web 検索結果から抽出した特徴語を取得する方法を提案する．

各推薦クエリを実行した際の検索結果ページからタイトルとスニペットを抽出・結合して文書とし、得られた文書集合から tf-idf 法を用いて文書中の各単語の解特微量を取得する。そして、解特微量が上位に来るような単語をその推薦クエリの解特微とする。

実際の手法は以下の通りである。

- (1) 推薦クエリ $q_i (\in Q)$ を実行し、検索結果ページの内容を取得する。
- (2) 得られた検索結果ページの内容から、検索結果の上位 α 件についてタイトルとスニペットを抽出し、これらすべてをスペース区切りで連結することで、文書 d_i を得る。
- (3) (1) (2) を、 Q 中のすべての推薦クエリ q_i について行い、文書集合 $D = \{d_1, d_2, \dots, d_n\}$ を得る。
- (4) 文書 d_i を形態素解析することで、 d_i 中に含まれる単語の集合 $W_i = \{w_{i1}, w_{i2}, \dots, w_{im_i}\}$ を得る。
- (5) W_i 中のすべての単語 w_{ij} について、 d_i 中に w_{ij} が出現する回数 $\text{tf}_{\text{ans}}(w_{ij})$ と、全文書の内 w_{ij} が出現する文書数 $\text{df}_{\text{ans}}(w_{ij})$ を求める。
- (6) これらを用いて、文書 d_i 中の単語 w_{ij} の解特微量 $f_{\text{ans}}(w_{ij})$ を次のように定義する。

$$f_{\text{ans}}(w_{ij}) = \frac{\text{tf}_{\text{ans}}(w_{ij})}{\sum_{w_{ij} \in W_i} \text{tf}_{\text{ans}}(w_{ij})} * \log \frac{|D|}{\text{df}_{\text{ans}}(w_{ij})} \quad (1)$$

- (7) $f_{\text{ans}}(w_{ij})$ を W_i 中のすべての単語 w_{ij} について求め、 $f_{\text{ans}}(w_{ij})$ が上位 β 件に来るような w_{ij} を、推薦クエリ q_i の解特微として取得する。
- (8) (4) ~ (7) を D 中のすべての文書 d_i について行う。

4.3 画像検索結果を用いた意味特徴の取得

3.3 節で述べたように、推薦クエリの意味特徴に重要な指標は、その推薦クエリの意味を適切に表現していることである。本稿では、意味特徴として推薦クエリを入力クエリとしたときの画像検索結果から抽出した特徴画像を取得する方法を提案する。各推薦クエリを入力クエリとして画像検索を行い、得られた画像の URL 集合を文書として、得られた文書集合から tf-idf 法を用いて文書中の各画像の意味特徴量を取得する。そして、意味特徴量が上位に来るような画像をその推薦クエリの意味特徴とする。

実際の手法は以下の通りである。

- (1) 推薦クエリ $q_i (\in Q)$ を入力クエリとして画像検索を実行し、画像検索結果の上位 α 件を取得する。ここで、取得した画像検索結果集合を $A_i = \{a_{i1}, a_{i2}, \dots, a_{i\alpha}\}$ とする。
- (2) A_i 中のすべての画像検索結果 a_{ij} からその画像の URL を取得し、URL 集合 $U_i = \{\text{url}(a_{i1}), \text{url}(a_{i2}), \dots, \text{url}(a_{i\alpha})\}$ を得る。ただし、 $\text{url}(a_{ij})$ は画像検索結果 a_{ij} から取得した URL である。
- (3) (1) (2) を、 Q 中のすべての推薦クエリ q_i について行い、集合 $U = \{U_1, U_2, \dots, U_n\}$ を得る。
- (4) $\text{tf}_{\text{sem}}(a_{ij})$ を、 $\text{tf}_{\text{sem}}(a_{ij}) = \gamma(\alpha - \beta \text{rank}(a_{ij}))$ のように定義する。ただし、 $\text{rank}(a_{ij})$ は a_{ij} の画像検索における順位を表し、 β, γ は $\text{rank}(a_{ij})$ による $\text{tf}_{\text{sem}}(a_{ij})$ の値のばらつきを調整するためのパラメータである。

きを調整するためのパラメータである。

- (5) 集合 U の要素 U_i の内、 $\text{url}(a_{ij}) \in U_i$ となるような要素の数を $\text{df}_{\text{sem}}(a_{ij})$ とする。
- (6) 画像検索結果 a_{ij} の意味特徴量 $f_{\text{sem}}(a_{ij})$ を次のように定義する。

$$f_{\text{sem}}(a_{ij}) = \frac{\text{tf}_{\text{sem}}(a_{ij})}{\sum_{a_{ij} \in A_i} \text{tf}_{\text{sem}}(a_{ij})} * \log \frac{|U|}{\text{df}_{\text{sem}}(a_{ij})} \quad (2)$$

- (7) $f_{\text{sem}}(a_{ij})$ を A_i 中のすべての画像検索結果 a_{ij} について求め、 $f_{\text{sem}}(a_{ij})$ が上位 δ 件に来るような a_{ij} を、推薦クエリ q_i の意味特徴として取得する。
- (8) (4) ~ (7) を U 中のすべての要素 U_i について行う。

4.4 推薦クエリに対する推薦クエリを用いた検索範囲特徴の取得

3.3 節で述べたように、推薦クエリの検索範囲特徴に重要な指標は、その情報を見ることで、その推薦クエリの検索範囲を適切に理解できることである。本稿では、検索範囲特徴として推薦クエリを入力クエリとみなしたときの推薦クエリ（以降、本稿では「推薦クエリの推薦クエリ」と呼ぶ）を取得し、そこから特徴語を抽出する方法を提案する。ここで、クエリ修正の種類には 5 種類ある [10] と言われているが、本稿では、5 つの修正タイプの中でも特に頻度の高い、クエリの絞り込みに注目することとする。各推薦クエリについて得られた推薦クエリの推薦クエリから、絞り込みに使われるキーワードを抽出し、それをその推薦クエリの検索範囲特徴とする。

実際の手法は以下の通りである。

- (1) 推薦クエリ $q_i (\in Q)$ を入力クエリとみなして、クエリ推薦インタフェースにより推薦クエリの推薦クエリの集合 $Q_i = \{q_{i1}, q_{i2}, \dots, q_{im_i}\}$ を取得する。
- (2) 推薦クエリ q_i について、これをスペースで区切ることでキーワード集合 $K_i = \{k_i^{(1)}, k_i^{(2)}, \dots, k_i^{(l_i)}\}$ を取得する。
- (3) 推薦クエリ q_i から取得された推薦クエリの推薦クエリ q_{ij} についても、これをスペースで区切ることでキーワード集合 $K_{ij} = \{k_{ij}^{(1)}, k_{ij}^{(2)}, \dots, k_{ij}^{(l_{ij})}\}$ を取得する。
- (4) Q_i 中の推薦クエリ q_{ij} について、 $K_i \subset K_{ij}$ が真であるかどうかを調べ、真であれば、キーワード集合 S_{ij} を $S_{ij} = K_{ij} - K_i$ のように定義する。
- (5) Q_i 中のすべての推薦クエリ q_{ij} について、 $q_{i1}, q_{i2}, \dots, q_{im_i}$ の順に (4) を実行し、 S_{ij} の各要素を取得する。その上位 α 件を検索範囲特徴として取得する。
- (6) (1) ~ (5) を Q 中のすべての推薦クエリ q_i について行う。

5. 評価実験

4 章で述べた特徴を付与したクエリ推薦インタフェースの有用性を検証するため、ユーザ実験を行った。本実験の目的は、4.2 節 ~ 4.4 節で述べた各特徴を付与したクエリ推薦インタフェースが、検索ユーザの効率や検索行為に及ぼす影響を明らかにすることである。

5.1 実験に用いたインタフェース

本実験では、以下に示す 5 つのインタフェースを実装して実



図 3 実験に用いた各インタフェースの動作例

験に用いた。

- (a) 付加情報なしのインタフェース
付加情報なしのインタフェース（以降、本稿では「インタフェース (a)」と呼ぶ）は、本実験におけるベースライン手法の 1 つである、既存のクエリ推薦インタフェースである。
- (b) 検索結果の一部をそのまま提示するインタフェース
検索結果の一部をそのまま提示するインタフェース（以降、本稿では「インタフェース (b)」と呼ぶ）は、本実験における比較対象の 1 つである、2.2 節で述べた Yahoo! を模したインタフェースである。4.2 節で述べたように、本研究では解特徴として特徴的な単語をインタフェースに追加するが、このインタフェースはそのような集約を行わず、検索結果の一部をそのままの形で提示するというものである。推薦クエリにマウスポインタを合わせると、その推薦クエリを実行した際の検索結果上位 5 件のタイトル・URL・スニペットが提示される。
- (c) 検索結果の特徴語を付加したインタフェース
検索結果の特徴語を付加したインタフェース（以降、本稿では「インタフェース (c)」と呼ぶ）は、解特徴として 4.2 節で説明した手法を用いて取得した特徴語を推薦クエリに付加したものである。本実験では、Web 検索結果の上位 10 件から解特徴量を取得し、解特徴量の上位 5 件を解特徴として提示した。
- (d) 意味特徴を付加したインタフェース
意味特徴を付加したインタフェース（以降、本稿では「インタフェース (d)」と呼ぶ）は、意味特徴として 4.3 節で説明した手法を用いて取得した特徴画像を推薦クエリに付加したものである。本実験では、画像検索結果の上位 10 件から意味特徴量を取得し、意味特徴量の上位 5 件を意味特徴として提示した。
- (e) 検索範囲特徴を付加したインタフェース

検索範囲特徴を付加したインタフェース（以降、本稿では「インタフェース (e)」と呼ぶ）は、検索範囲特徴として 4.4 節で説明した手法を用いて取得した特徴語を推薦クエリに付加したものである。本実験では、得られた特徴語の内、上位 5 件を検索範囲特徴として提示した。

これらのインタフェースでは、入力クエリに対する推薦クエリ、Web 検索結果、画像検索結果を、それぞれ Google Suggest API^(注4)、Yahoo!ウェブ検索 API^(注5)、Yahoo!画像検索 API^(注6)を用いて取得した。推薦クエリについては、関連クエリの内、上位 10 件を取得するようにした。また、文書の形態素解析には、日本語形態素解析器である MeCab^(注7)を用いた。「京都大学」を入力クエリとした際の各インタフェースの動作例を図 3 に示す。なお、本実験では、特徴情報の付与がユーザに与える影響のみを分析するため、インタフェース間で特徴抽出にかかる計算時間が等しくなるようにシステムを実装した。

5.2 実験に用いたタスク

様々な条件下での提案手法の有用性を検証するため、本実験では、「Lookup 型タスク」と「Exploratory 型タスク」[11] という二種類の検索タスクを用意した。Lookup 型タスクとは、特定の実事やページを探すような、ユーザの情報要求が具体的な検索タスクである。一方、Exploratory 型タスクは、物事の計画や分析をするようなタスクで、Lookup 型タスクに比べて情報要求が曖昧な検索タスクである。実験では、表 1 に示すように、Lookup 型タスクと Exploratory 型タスクをそれぞれ 5 つずつ、合計 10 個用意した。

5.3 実験手順

本実験の被験者として、20 代の男性 9 人と女性 1 人の計 10 人に参加してもらった。以下に実験の手順を示す。

(注 4) : <http://www.google.com/complete/search>

(注 5) : <http://developer.yahoo.co.jp/webapi/search/websearch/v2/websearch.html>

(注 6) : <http://developer.yahoo.co.jp/webapi/search/imagesearch/v2/imagesearch.html>

(注 7) : <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

表 1 検索タスクの内容と種類．表中の L は Lookup 型のタスク，E は Exploratory 型のタスクであることを示している．

タスク ID	タスク内容	タスクの種類
1	スニーカーブランドの公式ページをできるだけ多く探してください．	L
2	JR 京都駅から JR 東京駅までの交通手段について書かれているページをできるだけ多く探してください．	L
3	ペペロンチーノの作り方について書かれているページをできるだけ多く探してください．	L
4	スペインの名産品についてのページをできるだけ多く探してください．	L
5	Windows パソコンに CD のデータを入れる方法について書かれているページをできるだけ多く探してください．	L
6	あなたはハブとマンダースの関わりについてレポートを執筆することになりました．執筆に有用だと思われる情報について調べてください．	E
7	あなたは北海道に観光旅行に行こうと思いましたが，旅行計画を立てる際に有用だと思われる情報について調べてください．	E
8	あなたは前のスマートフォンを壊してしまい，買い換えようと思いましたが，買い替えに有用だと思われる情報について調べてください．	E
9	あなたは今からダイエットをしようと思いましたが，計画に有用だと思われる情報について調べてください．	E
10	あなたは来年中国に留学することになり，中国語を勉強する必要性に迫られています．有用だと思われる情報について調べてください．	E

- (1) はじめに，被験者に 5.1 節で示した 5 つのインタフェースを使用してもらい，各インタフェースの使い方に慣れてもらった．
- (2) 被験者に，インタフェース (a) ~ (e) のそれぞれについて 2 つずつ，計 10 個の検索タスクを実行してもらった．この際，Lookup 型のタスクでは，主に使用するインタフェースによる検索タスクの時間的効率の違いを評価するため，5 分間検索を実行し，タスクが示すトピックに適合するページをできるだけ多く探してもらった．Exploratory 型のタスクでは，主に使用するインタフェースによる検索行動の違いを評価するため，10 分間通常の検索と同じように検索を実行し，タスクが示すトピックに適合するページを探してもらった．どちらの種類のタスクにおいても，タスク実行中は被験者の検索行動をログデータとして記録した．
- (3) 各検索タスクの終了時と，全検索タスク終了後にアンケートに回答してもらい，被験者の主観に依存するような指標を得た．各アンケートは，5 段階のリッカート尺度にしたがって回答してもらった．また，これらとは別に，各アンケートには自由記述欄を設けた．

なお，それぞれのインタフェースを使用する順序及びそれぞれのタスクを実行する順序については，順序効果が生じないように被験者間でバランスを取った．

5.4 実験結果

5.4.1 検索タスクの効率

はじめに，本実験で得られたログデータから，検索タスクの効率についての分析を行った．表 2 に (a) ~ (e) の各インタフェースについての結果を示す．表中の「1 検索タスクあたりの適合ページ数」は，各検索タスクについて，被験者がトピックに適合していると判定した Web ページの数を求めた後，検索タスクの種類ごとに平均を求めたものである．また，表中の「1 推薦クエリあたりの適合ページ数」は，被験者が推薦クエリを選択して検索を行ったときについて，検索結果中で被験者がトピックに適合していると判定した Web ページの数を求めた後，検索タスクの種類ごとに平均を求めたものである．まず，「1 検索タスクあたりの適合ページ数」を見ると，本実験で用いたどの種類のタスクにおいても，インタフェース (a) は他のインタフェースと比較して高い適合ページ数を示している．こ

表 2 1 検索タスクあたりに得られた適合ページ数の平均と，推薦クエリの実行ごとに得られた適合ページ数の平均のインタフェース間での比較 () 内は標準偏差を示している．

		インタフェース				
		a	b	c	d	e
1 検索タスクあたりの適合ページ数	Lookup 型	10.70(6.44)	11.00(4.20)	10.00(4.92)	9.10(4.39)	8.10(2.39)
	Exploratory 型	12.50(5.70)	11.40(5.78)	10.60(5.57)	11.60(5.85)	12.60(6.92)
	全検索タスク	11.60(6.14)	11.20(5.06)	10.30(5.26)	10.35(5.32)	10.35(5.64)
1 推薦クエリあたりの適合ページ数	Lookup 型	1.85(2.67)	2.96(3.02)	4.19(3.21)	3.05(2.97)	2.77(3.00)
	Exploratory 型	2.41(1.75)	2.03(1.82)	1.95(1.92)	1.72(1.55)	1.83(1.33)
	全検索タスク	2.17(2.21)	2.39(2.41)	2.58(2.56)	2.14(2.20)	2.16(2.11)

表 3 被験者が入力した平均クエリ数，被験者が推薦クエリを閲覧した平均回数，推薦クエリが提示されてから推薦クエリが選択されるまでの平均時間 (秒) のインタフェース間での比較 () 内は標準偏差を示している．

		インタフェース				
		a	b	c	d	e
入力クエリ数	Lookup 型	5.50(4.72)	5.30(4.50)	4.80(4.60)	5.20(5.19)	3.80(2.71)
	Exploratory 型	5.20(1.99)	5.60(3.14)	6.50(2.97)	7.90(4.68)	6.80(2.18)
	全検索タスク	5.35(3.62)	5.45(3.88)	5.65(3.97)	6.55(5.12)	5.30(2.88)
推薦クエリの閲覧数	Lookup 型	5.90(4.39)	5.70(4.24)	6.30(4.17)	4.80(3.25)	4.30(2.57)
	Exploratory 型	6.10(2.47)	7.00(3.41)	7.70(4.15)	7.70(3.69)	7.80(2.32)
	全検索タスク	6.00(3.56)	6.35(3.90)	7.00(4.22)	6.25(3.77)	6.05(3.01)
推薦クエリの閲覧時間 (秒)	Lookup 型	15.72(30.24)	14.04(18.11)	7.79(6.01)	11.96(22.16)	4.61(2.47)
	Exploratory 型	5.26(2.83)	10.69(8.13)	12.00(19.90)	16.43(38.25)	7.91(4.21)
	全検索タスク	9.33(19.66)	12.13(13.46)	10.86(17.39)	15.00(34.01)	6.69(4.00)

れは，他のインタフェースでは，付加された情報を閲覧している時間が影響し，適合ページを探す時間が少なくなっていることが原因と考えられる．一方，「1 推薦クエリあたりの適合ページ数」を見ると，インタフェース (a) と比較してインタフェース (b) ~ (e) は，Lookup 型の検索タスクでは適合ページ数が増加している．特に，解特徴を提示するインタフェース (c) は，Lookup 型の検索において，1 推薦クエリあたりに被験者が得た適合ページ数が多いことが分かる．これは，検索結果の内容をあらかじめ特徴語という形でユーザに提示することで，不要な推薦クエリの実行を抑制することができたためではないかと考えられる．それに対して，Exploratory 型の検索タスクでは適合ページ数が減少している．これは，Exploratory 型の検索タスクのような情報要求が曖昧なタスクでは，付加情報の種類によっては，被験者が付加情報を閲覧することで，どの推薦クエリを選ぶべきかについて，かえって混乱してしまう場合があるということが原因と考えられる．

5.4.2 ユーザの検索行動

次に，提案手法がユーザの検索行為にどのような影響を与えるのかを検証するため，本実験で得られたログデータから，ユーザの検索行動についての分析を行った．表 3 に (a) ~ (e) の各インタフェースについての結果を示す．表中の「入力クエリ数」「推薦クエリの閲覧数」「推薦クエリの閲覧時間 (秒)」は，それぞれ，各検索タスクについて，被験者が入力したクエリ数，被験者が推薦されたクエリを閲覧した回数，推薦クエリが提示されてから被験者が推薦クエリを選択するまでの時間を表している．「入力クエリ数」を見ると，Exploratory 型の検索タスクでは，インタフェース (d) の入力クエリ数が比較的多くなっている．これは，Exploratory 型の検索タスクは，意味の

分からない単語が現れることが多く、意味特徴によって推薦クエリの意味を理解することができたからではないかと考えられる。また、副次的なものとして、推薦クエリに画像を付与することで、推薦クエリに関する興味を誘発したことも原因の1つと考えられる。「推薦クエリの閲覧数」を見ると、Exploratory型の検索タスクについて、インタフェース(a)と比較すると、インタフェース(b)~(e)は推薦クエリを閲覧した回数が多くなっている。これは、Exploratory型の検索タスクはトピックが曖昧で、被験者がクエリを作成するのが難しいため、クエリを作成する手がかりとして特徴を利用したことが原因ではないかと考えられる。「推薦クエリの閲覧時間」を見ると、インタフェース(a)と比較して、インタフェース(b)~(d)は推薦クエリが提示されてから被験者が推薦クエリを選択するまでの時間が長くなっている。これは、被験者が付加情報を注視し、閲覧している時間が長くなっていることが原因と考えられる。また、インタフェース(c)(e)と比較すると、インタフェース(b)(d)では時間の増加が顕著である。これは、検索結果の一部をそのまま提示したり、画像を提示したりする手法では、特徴語を提示する手法に対して、被験者が閲覧する情報量が多くなってしまうことが原因と考えられる。

5.4.3 主観的指標

本実験で得られたアンケートから、ユーザの主観に依存する指標の分析を行った。各検索タスク終了時および実験後アンケートの内容と、その回答を各インタフェースごとにまとめたものを表4に示す。検索タスクのトピックについての知識量に関する質問Q1については、インタフェース間で大きな差はみられなかった。検索タスクの難易度に関する質問Q2については、インタフェース(e)が他のインタフェースと比較して若干高い数値を示しており、インタフェース(e)を用いて検索タスクを実行した被験者は、自らが実行したタスクを難しいものと捉えていた。検索タスクにクエリ推薦インタフェースがどれだけ役に立ったかに関する質問Q3については、インタフェース(b)~(e)がインタフェース(a)と比較して若干高い数値を示しており、インタフェース(b)~(e)を用いて検索タスクを実行した被験者は、クエリ推薦インタフェースを役に立つものだと捉えていた。検索タスクで見つけた情報に対する満足度に関する質問Q4については、インタフェース間で大きな差はみられなかった。付加情報が推薦クエリの選択に役に立ったかに関する質問Q5については、インタフェース(d)が他のインタフェースと比較して低い数値を示しており、インタフェース(d)を用いて検索タスクを実行した被験者は、付加情報が推薦クエリを選択する上で役に立たないものだと捉えていた。実験で使用したインタフェースの使いやすさに関する質問Q6については、インタフェース(b)(c)(e)が、インタフェース(a)と比較して高い数値を示しており、被験者はこれらのインタフェースを使いやすいと捉えていたのに対し、インタフェース(d)は他のインタフェースと比較して低い数値を示しており、被験者はインタフェース(d)を使いにくいと捉えていた。実験で使用したインタフェースによる心理的負担に関する質問Q7については、インタフェース(b)(d)が他のインタフェー

スと比較して高い数値を示しており、被験者はインタフェース(b)(d)に心理的な負担を感じていたことが分かる。インタフェース(b)は、検索結果の一部をそのまま特徴として付加しているため、文字数が多くなってしまい、情報を閲覧する上で被験者の負担につながったのではないかと考えられる。また、インタフェース(d)は、特徴が画像であるため、クエリ推薦インタフェースの画面上でのサイズが大きくなってしまい、被験者が推薦クエリを閲覧する上での負担が大きくなってしまったのではないかと考えられる。インタフェースを実験で使用した際の推薦クエリの実行適合性に関する質問Q8については、インタフェース(b)~(e)がインタフェース(a)と比較して著しく高い数値を示しており、インタフェース(b)~(e)を用いて検索タスクを行った被験者は、推薦クエリの実行適合性を確信を持って判断することができたことが分かる。その中でも、特にインタフェース(b)の値の高さは顕著である。検索結果の一部を特徴として付加しているため、検索結果の情報要求への適合性を判断するのは容易であると考えられる。また、インタフェース(c)(e)についても、インタフェース(b)には劣るが、実行適合性の判断に確信を持たたという結果が出ている。これは、特徴語の形で表された解特徴と検索範囲特徴が、クエリの実行適合性の判断のための手がかりになったことが原因と考えられる。インタフェース(d)については、インタフェース(b)(c)(e)と比較すると低い値になっているが、これは、上で述べたインタフェースの使いにくさなどが原因ではないかと考えられる。

5.5 本実験のまとめ

本実験では、本稿で提案した3つの手法から得られたインタフェース(c)(d)(e)が、推薦クエリの実行適合性の判断支援と検索タスクの時間的効率の向上に有用であることをユーザ実験を通して検証した。

この実験の結果として、解特徴として検索結果の特徴語を加えたインタフェース(c)は、検索タスクの効率や推薦クエリの実行適合性の判断に対する確信度の観点から見て、比較的よい結果が得られており、同じ解特徴として検索結果の一部をそのまま付与したインタフェース(b)と比較しても、付加した特徴が被験者の心理的負担にならないといった利点があった。次に、意味特徴として画像検索結果を加えたインタフェース(d)は、インタフェース(a)と比較すると検索タスクの効率は若干低く、推薦クエリを見るのに心理的な負担が大きいということが分かった。最後に、検索範囲特徴として推薦クエリに対する推薦クエリを加えたインタフェース(e)は、検索タスクの効率についてはインタフェース(a)を下回る結果となったが、実行適合性の判断に対する確信度の点ではインタフェース(c)を上回る結果となった。

6. 考 察

本章では、本稿で提案した手法と5章で述べた実験についての全体的な考察を行う。5章の実験で得られたアンケートの自由記述欄を見ると、「推薦クエリ単体で出てきそうな結果が分かりやすいものだったので、確認程度に使った」(インタフェー

表 4 アンケートの質問内容とその回答の平均値。()内は標準偏差を示している。

質問 ID	質問	インタフェース				
		a	b	c	d	e
各検索タスク終了時	Q1 実行した検索タスクのトピック (話題) について、どのくらい知っていましたか?	2.80(1.36)	2.65(1.31)	2.75(1.37)	2.85(1.31)	2.30(1.35)
	Q2 実行した検索タスクはどのくらい難しかったですか?	2.30(1.05)	2.20(1.08)	2.15(1.15)	2.65(1.16)	3.10(1.18)
	Q3 実行した検索タスクに、クエリ推薦インタフェースは役に立ちましたか?	2.80(1.29)	3.25(0.89)	3.40(1.02)	3.25(1.18)	3.45(0.92)
	Q4 実行した検索タスクで見つけた Web ページに、あなたは満足出来ましたか?	3.80(0.87)	4.00(0.71)	4.10(0.89)	3.85(0.73)	3.80(0.73)
実験後	Q5 実行した検索タスクにおいて、推薦クエリに追加された付加情報は、推薦クエリを選択する上で役に立ちましたか?	-	3.30(1.05)	2.90(1.00)	2.25(1.09)	2.95(0.92)
	Q6 実験で使ったインタフェースはどのくらい使いやすかったですか?	3.10(0.94)	3.80(0.60)	3.70(0.90)	2.70(0.64)	3.80(0.87)
	Q7 実験で使ったインタフェースで推薦クエリを見るのは心理的に負担になりましたか?	1.80(1.17)	2.70(1.00)	1.80(0.75)	3.10(1.22)	1.80(0.60)
	Q8 実験で使ったインタフェースを使ったとき、自分の欲しい情報を得るために推薦クエリを実行すべきかどうかを自信を持って判断できましたか?	2.80(0.75)	4.20(0.75)	3.80(0.40)	3.20(0.75)	3.90(0.54)

ス(c)を用いて実験タスク2を行った被験者の回答)、「付加情報に料理名が出ていたので探しやすかった」(インタフェース(b)を用いて実験タスク3を行った被験者の回答)、「中国語の文法や発音を調べるのに画像はあまり役に立たなかった」(インタフェース(d)を用いて実験タスク10を行った被験者の回答)などの意見があった。これは、検索タスクの種類や、提示される推薦クエリの種類によって、それぞれの提案手法で得られた付加情報が役に立つ場合と役に立たない場合があるという可能性を示している。また、他に多かった意見として、「『ダイエット おすすめ』に対し、『リバウンド』などが付与された。このクエリも有用な情報を探すために必要なものとなるため、それを気づかせてくれた」(インタフェース(c)を用いて実験タスク9を行った被験者の回答)、「『函館 観光』に対し、『史跡』という単語が付加されたので、気になって実行した」(インタフェース(c)を用いて実験タスク4を行った被験者の回答)、「特徴語を直接クリックしたい」(実験後アンケートに対する回答)など、提示された特徴語をクエリ語として入力し、検索を行う被験者が多かった。これは、提案手法が、実行適合性の判断支援だけでなく、クエリ推薦としてユーザのクエリ作成の支援にも寄与していることが理由だと考えられる。

7. まとめ

本稿では、推薦されたクエリから、ユーザの求める情報が得られるかどうかの判断に有用な情報を取得し、これを特徴としてクエリ推薦インタフェースに付与するという手法を提案した。また、特徴を取得するための手法として、推薦クエリを実行したときの検索結果を集約して特徴語を得る方法、推薦クエリを入力クエリとして画像検索を行った際の検索結果から特徴画像を得る方法、推薦クエリに対する推薦クエリから特徴語を得る方法の3つを提案した。また、提案手法をシステムとして実装し、ユーザ実験を行うことで提案手法の有用性を検証した。

今後の課題としては、付加特徴を推薦クエリの種類に応じて変更するという手法を考えていく予定である。6章で述べたように、例えば、実行適合性の判断支援に解特徴が有効なクエリ、意味特徴が有効なクエリ、のように、提示する推薦クエリの種類によって有効な付加情報は異なる可能性がある。クエリに対してこれらの分類を行い、それぞれのクエリに対して最も有効な特徴を見つけ出すことができれば、その特徴を動的にクエリ

推薦インタフェースに付加することで、ユーザの検索タスクをより効率化できると考えられる。

謝辞 本研究の一部は、文科省科研費基盤(A)「ウェブ検索の意図検出と多元的検索意図指標にもとづく検索方式の研究」(24240013, 研究代表者: 田中克己)によるものです。ここに記して謝意を表します。

文 献

- [1] C. L. A. Clarke, M. Kolla, and O. Vechtomova. An effectiveness measure for ambiguous and underspecified queries. In *Proc. of ICTIR: Advances in Information Retrieval Theory*, pages 188–199, 2009.
- [2] Z. Zhang and O. Nasraoui. Mining search engine query logs for query recommendation. In *Proc. of WWW*, pages 1039–1040, 2006.
- [3] Q. Zhao, S. C. H. Hoi, T-Y Liu, S. S. Bhowmick, M. R. Lyu, and W-Y Ma. Time-dependent semantic similarity measure of queries using historical click-through data. In *Proc. of WWW*, pages 543–552, 2006.
- [4] H. Cao, D. Jiang, J. Pei, Q. He, Z. Liao, E. Chen, and H. Li. Context-aware query suggestion by mining click-through and session data. In *Proc. of KDD*, pages 875–883, 2008.
- [5] Makoto P. Kato, Tetsuya Sakai, and Katsumi Tanaka. When do people use query suggestion? a query suggestion log analysis. *Information Retrieval*, to appear, 2013.
- [6] Z-J Zha, L. Yang, T. Mei, M. Wang, and Z. Wang. Visual query suggestion. In *Proc. of MM*, pages 15–24, 2009.
- [7] M. P. Kato, T. Sakai, and K. Tanaka. Structured query suggestion for specialization and parallel movement: effect on search behaviors. In *Proc. of WWW*, pages 389–398, 2012.
- [8] J. Guo, X. Cheng, G. Xu, and H. Shen. A structured approach to query recommendation with social annotation data. In *Proc. of CIKM*, pages 619–628, 2010.
- [9] H. Joho, C. Coverson, M. Sanderson, and M. Beaulieu. Hierarchical presentation of expansion terms. In *Proc. of SAC*, pages 645–649, 2002.
- [10] P. Boldi, F. Bonchi, C. Castillo, and S. Vigna. From "dango" to "japanese cakes": Query reformulation models and patterns. In *Proc. of WI*, pages 183–190, 2009.
- [11] G. Marchionini. Exploratory search: from finding to understanding. *Communication of the ACM*, Vol. 49(No. 4), pages 41–46, 2006.