

Web 検索質問の自動分類と質問応答への応用

藤井 敦[†] 秋葉 友良[‡]

[†] 筑波大学 大学院図書館情報メディア研究科

〒 305-8550 つくば市春日 1-2

[‡] 豊橋技術科学大学 情報工学系

〒 441-8580 豊橋市天伯町雲雀ヶ丘 1-1

E-mail: fujii@slis.tsukuba.ac.jp

Web を検索するユーザの情報要求は多種多様であり、検索質問の種類によって必要とされる検索手法が異なるという知見が得られている。本研究は、検索質問を「調査型（ある事項に関する一般的な調べ物が目的）」と「誘導型（ある事項に関するトップページを探すことが目的）」に自動分類し、その結果に基づいて検索手法を動的に変更する手法を提案する。誘導型の検索質問に使われる語句は、特定のページにリンクするためのアンカーテキストに現れやすい傾向がある。そこで、検索質問中の語句が Web 上のアンカーテキスト群にどのように分布しているかを解析し、検索質問を分類する。NTCIR の Web 検索用テストコレクションを用いた実験の結果、本手法の有効性が確認された。また、検索質問の分類を質問応答に応用する手法を提案し、その効果について議論する。

A Method for Classifying Web Queries and its Application to Question Answering

Atsushi Fujii[†], Tomoyosi Akiba[‡]

[†] Graduate School of Library, Information and Media Studies, University of Tsukuba

[‡] Department of Information and Computer Sciences, Toyohashi University of Technology

E-mail: fujii@slis.tsukuba.ac.jp

There are a number of types of queries on the Web and the ideal retrieval method can vary depending on the query type. We propose a method to classify queries into the “informational” type for which the purpose is to find one or more pages related to a topic or the “navigational” type for which the purpose is to find the representative page of a topic, and use different retrieval methods depending on the query type. Because keywords in navigational queries often appear in anchor texts to link to another pages, we analyze the distribution of a target keyword in the anchor texts on the Web for query classification purposes. We use the NTCIR Web test collection and show the effectiveness of our method experimentally. We also apply our query classification method to question answering and discuss its effects.

1 はじめに

World Wide Webには多種多様な情報が存在する。そこで、Webを検索するユーザの情報要求もまた多種多様である。Broder [4]は、Web上の検索質問をユーザの情報要求に基づいて以下の3種類に分類した。

- informational query (調査型の検索質問)
ある話題についてWeb上の情報を調査するために使用される検索質問である。Web以外の一般的なテキスト検索でも使用される。
- navigational query (誘導型の検索質問)
ある事柄(人物, 組織, 商品, イベントなど)に関するトップページや代表的なページを検索するために使用される検索質問である。
- transactional query (取引型の検索質問)
ソフトウェアのダウンロードやオンラインショッピングなどのように、Webサイトを中継して別の実体に到達するために使用される検索質問である。

Broderは、検索エンジンのユーザに対するアンケート調査と検索エンジンのログ解析によって、各種種類の質問が無視できない一定の割合で存在することを示した。

近年の情報検索研究によって、上記3種類の検索質問に対して、必要とされる検索の手法が根本的に異なることが明らかになっている。TRECやNTCIRでは、調査型と誘導型の検索質問を対象としたWeb検索のテストコレクションが構築された。種々の実験によって、調査型の検索質問にはページの本文を用いた検索が有効であり、誘導型の検索質問にはアンカーテキストやリンク構造を用いた検索が有効であるという知見が得られている [5, 7, 8, 10, 15]。

アンカーテキストとは、あるページから別のページにリンクをはるときに使用される文字列である。アンカーテキストは、あるページに対して第三者がどのように評価しているのかを知る上で重要である。

Liら [13]は、人手で作成した規則によって取引型のページとそれ以外のページを自動的に分類し、取引型の検索質問に対する検索精度を向上させた。

しかし、ユーザは検索質問を入力しても、検索質問の種類は明示しない、あるいはできない。そこで、

様々な検索質問に対して高い検索精度を実現するためには、入力された検索質問の種類を自動的に特定し、異なる検索手法を選択的に使用する必要がある。

本研究は、検索質問を自動的に分類し、その結果に基づいて検索手法を動的に変更する手法を提案する。ただし、取引型を対象としたテストコレクションが存在しないため、調査型と誘導型の検索質問を対象とする。

2章で検索質問の自動分類に関する先行研究について検討し、本研究の位置付けを明確にする。3章で筆者が研究開発しているWeb検索システムの全体像について説明する。4章で検索質問の分類手法について説明し、5章で評価実験について説明する。6章で、検索質問の分類手法を質問応答へ応用する手法を提案し、その効果について議論する。

2 先行研究と本研究の違い

Kangら [10]は、TRECの10ギガバイトWebテストコレクションを用いて、検索質問を調査型と誘導型に自動分類する手法を提案した。しかし、対象のテストコレクションは10ギガバイトという小さな規模である。また、検索質問の分類に焦点が当てられており、分類によって検索精度が向上したかどうかは明らかになっていない。

Leeら [12]は、被験者のWeb検索行動を観察することによって、検索質問を調査型と誘導型に分類するための特徴量を提案した。しかし、検索実験は行っていない。そこで、Leeらが提案した特徴量を用いて検索質問を分類することによって、Web検索の精度がどのように変化するかは明らかになっていない。

以上の背景から、本研究は、NTCIR-3/4で構築された100ギガバイトのテストコレクションを用いて、Web上の検索質問を調査型と誘導型に自動分類することを目的とする。当該テストコレクションは、Kangらが使用したテストコレクションよりも10倍大きく、現実のWeb検索により近い環境での実験を可能とする。さらに、本研究では、検索質問を自動分類することによって、Web検索の精度がどのように変化するかを明らかにする。

本研究で提案する手法は、Baeza-Yatesら [3]が提案した検索質問分類の手法とは異なり、大量の検索ログを必要としない。大学や研究所でWeb検索の研

究を行う場合は、商用の検索サービスとは異なり、大量の検索ログを入手することが困難であることが多い。そのような状況においても本研究で提案する分類手法は適用可能である。

3 Web 検索システムの概要

筆者が研究開発している Web 検索システムの概要を図 1 に示す。図 1 は、大きく分けて、「検索質問分類」、「コンテンツ検索」、「アンカー検索」で構成されている。

システムへの入力は、キーワード集合や自然文による検索質問である。検索質問が自然文の場合は、索引語を抽出して検索に利用する。システムの出力は、検索質問に合致するページの順位付きリストである。ページの順位は、検索質問に対するスコアによって決定される。

検索質問が入力されると、検索質問分類によって「調査型」か「誘導型」のいずれかに分類される。

次に、検索質問の種類に依らずに、コンテンツ検索とアンカー検索を独立に行って 2 つの順位付きリスト（初期リスト）を作成する。コンテンツ検索はページ本文を検索に利用し、アンカー検索はアンカーテキストを検索に利用する。現在、コンテンツ検索には Okapi BM25 [17] を用い、アンカー検索には、筆者が提案したアンカーモデル [9] を用いる。このモデルは、アンカーテキストを用いて、各ページが検索される確率を計算する。さらに、アンカーテキスト集合から同義語を自動抽出し、検索質問の拡張に利用する。

最後に、2 つの初期リストを併合して出力する。ここで、検索質問の種類によって、コンテンツ検索とアンカー検索の結果をそれぞれどの程度重視するかを決定する。具体的には、検索されたページ d に対して、2 つの初期リストにおけるスコアを加重平均し、新しいスコア $S(d)$ を計算する。最終的な順位付きリストにおいて、各ページは $S(d)$ に基づいて降順に配列される。

しかし、コンテンツ検索とアンカー検索で計算されるスコアの意味は異なる。そこで、ページ d の初期スコアにおける順位の逆数を用いて、式 (1) によって $S(d)$ を計算する。

$$S(d) = \frac{\alpha}{R_c(d)} + \frac{1-\alpha}{R_a(d)} \quad (0 \leq \alpha \leq 1) \quad (1)$$

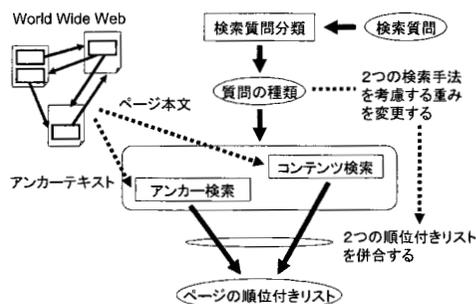


図 1: Web 検索システムの概要

式 (1) において、 $R_c(d)$ と $R_a(d)$ は、それぞれ、コンテンツ検索とアンカー検索で得られた初期リストにおける d の順位である。 α は 0 以上 1 以下の値を取り、検索質問の種類によって変更する必要がある。 α の値は、調査型の質問に対しては 0.5 よりも大きな値に設定し、誘導型の質問に対しては 0.5 よりも小さな値に設定する。

本研究で提案する検索質問の分類手法は、検索質問の種類を決定する際に α の値も決定する。そこで、 α の値を手で決定する必要がない。

4 検索質問の自動分類手法

誘導型の検索質問によって検索されるべきページとは、ある事柄に関するトップページや代表的なページである。そのようなページは、他のページから、特定のアンカーテキストによってリンクされることが多い。例えば、筑波大学のトップページ (<http://www.tsukuba.ac.jp/>) にリンクをはる場合は、アンカーテキストとして「筑波大学」が使われることが多いだろう。また、筑波大学と無関係なページにリンクをはる場合には、「筑波大学」がアンカーテキストとして使われることは稀であろう。

しかし、特定のトップページと対応しない一般的な事柄の場合には、上記のような現象は起こりにくい。例えば、「情報検索」というアンカーテキストでリンクされるページは、検索サービスのページや情

報検索に関する解説のページなど多岐にわたる。

以上の仮説に基づいて、Leeら [12] は、検索質問の文字列と同じアンカーテキストからリンクされているページの出現分布 (anchor-link distribution: ALD) を分析してヒストグラムを作り、ALDの歪度 (skewness) によって検索質問の種類を特定した。歪度が小さい場合には調査型に分類し、そうでなければ誘導型に分類する。

しかし、Leeらは、検索質問の文字列がそのままの形でアンカーテキストとして使用されている場合だけを対象とした。そこで、検索質問と完全一致するアンカーテキストがなければALDを分析することができない。それに対して、本研究は、検索質問がそのままの形でアンカーテキストとして使われていない場合には、検索質問を索引語に分割して個別にALDを分析することで当問題を解消する。

以下、本研究で提案する分類手法について説明する。検索質問 q を構成する索引語の集合を T_q とする。ChaSenを用いて q を形態素解析し、名詞を索引語として使用する。索引語 $t \in T_q$ に対して、 t を含むアンカーテキストによってリンクされているページの集合を D_t とする。ここで、 t とアンカーテキストが同一である必要はない。 D_t から無作為に選んだページが d である確率 $P(d|t)$ の分布を分析し、式(2)によってエントロピーを計算する。

$$H(D_t|T_q) = - \sum_{t \in T_q} P(t) \sum_{d \in D_t} P(d|t) \log P(d|t) \quad (2)$$

q に関するALDの歪度が大きいほど、 $H(D_t|T_q)$ は小さくなり、 q は誘導型に分類されやすくなる。

検索質問は少数のキーワードで構成されることが多いため、検索質問における t の出現頻度は考慮しない。すなわち、 $P(t) = \frac{1}{|T|}$ とする。

t を含むアンカーテキストが存在しない場合は、アンカーテキスト集合から t の同義語を自動抽出し、同義語が存在すれば、 t の代わりに使用する。アンカーテキスト集合からの同義語抽出には筆者の手法 [9] を使用する。

実際には、 $P(d|t)$ を計算する際に、複数のページを一定の階級にまとめる方が分類の精度が向上する。

さらに、 $H(D_t|T_q)$ を $\log |D_t|$ で割って、0以上1以下の範囲に正規化した値を $i(q)$ とする。 $i(q)$ を式(1)の α として利用することで、 $i(q)$ の値が大きい q に対してはコンテンツ検索の結果が重視される。すな

わち、システムの内部で q の種類を具体的に決定する必要はない。

検索質問の種類を具体的に決定する必要がある場合は、 $i(q)$ の値が0.5以上の場合は q を調査型に分類し、それ以外の場合は q を誘導型に分類する。

本手法には2つのパラメタがある。一つは、複数のページをまとめるための階級幅であり、現在は5としている。

また、 t がアンカーテキストに出現しない場合は、本来、ALDの歪度が小さいにも拘らず、全ての $d \in D_t$ に対して $P(d|t)$ が0になるため、 $H(D_t|T_q)$ が小さくなる。そこで、このような場合は、 t を含むアンカーテキストから N 件のページに対して均一にリンクがはられていると見なす。現在は $N = 10000$ としている。

これら2つのパラメタは、対象とするテストコレクションの規模に応じて適宜設定する必要がある。

5 評価実験

5.1 実験データ

NTCIR-3とNTCIR-4で構築されたWeb検索のテストコレクションを使用して提案手法の評価実験を行った。NTCIR-3では調査型の検索質問が作られた [8]。NTCIR-4では調査型と誘導型の検索質問が作られた [7, 15]。NTCIR-3とNTCIR-4の検索対象は共通であり、JPドメインから収集した100ギガバイト (約1千万ページ) のページ集合である。正解のレベルには「適合」と「部分適合」がある。本実験では、「適合」だけを正解として使用した。適合文書が存在する検索質問の件数は、NTCIR-3の調査型が47件、NTCIR-4の調査型が80件、NTCIR-4の誘導型が168件である。すなわち、合計295件の検索質問を用いて、100ギガバイトのページ集合を検索する実験を行った。

5.2 検索質問分類の評価

検索精度を評価する前に、検索質問の分類精度を評価した。 $i(q)$ の閾値を0.5として、 $i(q)$ が0.5以上の場合は q を調査型に分類し、それ以外の場合は q を誘導型に分類した。

比較対象として Kang ら [10] の手法と Lee ら [12] の手法を用いた。Kang らは、検索質問の分類に用いる 4 つの特徴量を提案した。しかし、「検索質問中の索引語がアンカーテキストで使用される頻度」だけを特徴量として使用した。それ以外の特徴量を用いても検索質問の分類精度は 50%未満だった。人手で決定すべきパラメタは、予備実験によって最適化した。

実験の結果、Kang らの手法、Lee らの手法、本手法の分類精度は、それぞれ、75.6%、72.5%、79.3%であり、本手法の分類精度が最も高かった。

5.3 検索精度の評価

検索質問の分類が検索精度に及ぼす影響について評価した。検索精度を評価する尺度として、MAP (Mean Average Precision) と MRR (Mean Reciprocal Rank) を用いた。それぞれ、調査型の検索質問に対する検索と誘導型の検索質問に対する検索を評価する目的に適している。

MAP は精度と再現率の両方を重視した評価尺度であり、テキスト検索の評価によく用いられる。それに対して、MRR は精度を重視した評価尺度であり、誘導型の検索質問に対する検索の評価 [15] や質問応答の評価 [18] に用いられる。

どちらの評価尺度も検索結果の上位から一定数の文書を対象に計算される。調査型の検索では多くの文書が閲覧されることを考慮して、上位 100 件までを評価対象とした。それに対して、誘導型の検索では、上位 10 件までを評価対象とした。

以下 (a)~(f) の検索手法について MAP と MRR を比較した。これらの手法は、式 (1) における α の決定方法だけが異なる。

- (a) 検索質問の分類をせずに、常に $\alpha = 0.5$ とする。
- (b) Kang らの手法で検索質問を分類し、調査型の場合は $\alpha = 0.7$ とし、誘導型の場合には $\alpha = 0.3$ とする。
- (c) Lee らの手法で検索質問を分類し、手法 (b) と同じ方法で α を決定する。
- (d) 本手法で検索質問を分類し、手法 (b) と同じ方法で α を決定する。

(e) 本手法で検索質問を分類し、 $i(q)$ によって α を自動的に決定する。

(f) 検索質問の正しい分類を使用し、手法 (b) と同じ方法で α を決定する。

手法 (b), (c), (d), (f) における α の値は、予備実験によって検索精度が最も高くなるように設定した結果である。

手法 (a) と (f) の MAP や MRR は、それぞれ期待される検索精度の下限と上限である。本研究の提案手法は (e) である。手法 (d) は手法 (e) の変形であり、 α の決定方法を揃えることで手法 (b) や手法 (c) との直接的な比較を行うために使用する。

実験結果を表 1 に示す。MAP と MRR のどちらに対しても、本手法は手法 (b) や手法 (c) の結果を上回った。手法 (e) によって α の値を自動的に決定すると、手法 (d) に比べて MAP は若干向上し、逆に MRR は若干低下した。

表 1: 検索精度の比較

	(a)	(b)	(c)	(d)	(e)	(f)
MAP	25.4	28.1	26.5	30.0	30.3	31.2
MRR	.468	.504	.485	.519	.514	.545

適合文書が 1 位で検索された質問の件数を調査したところ、手法 (b) は 119 件、手法 (c) は 113 件だったのに対して、手法 (c) と (d) では 127 件に増加した。

本手法は、検索質問の分類が常に正しい手法 (f) に比べると MAP と MRR がともに低下するものの、検索質問の分類を行わない手法 (a) に比べて MRR を向上させた。すなわち、検索質問を自動分類することが Web 検索の精度向上に貢献することが分かった。また、本手法は、既存の分類手法よりも Web 検索の精度を向上させることが分かった。

5.4 誤り分析

手法 (e) の分類誤りについて分析を行い原因を特定した。また、検索質問の分類誤りによって、手法 (f) と比較して手法 (e) の MRR がどのように変化したのかを分析した。分析の結果を表 2 に示す。表 2 において、(i) と (ii) は調査型の検索質問に対する誤りの原因であり、(iii)~(vi) は誘導型の検索質問に対

表 2: 検索質問分類の誤り原因と MRR の変化

誤り原因	質問の種類	MRR の変化		
		低下	等価	向上
(i)	調査型	10	3	1
(ii)	調査型	5	3	1
(iii)	誘導型	6	16	5
(iv)	誘導型	1	0	0
(v)	誘導型	0	4	0
(vi)	誘導型	1	0	0

する誤りの原因である。表 2 より、分類誤りのために総じて MRR が低下したことが分かる。

しかし、分類誤りによって MRR が向上した質問もあった。例えば、「京都、寺、神社」という調査型の検索質問である。この質問に対してアンカー検索の検索結果が重視され、その結果、京都の観光に関する代表的なページが検索された。これは、ユーザは調査型を意図して質問を入力したにも拘らず、誘導型の質問として処理したことが良い結果につながった例である。

(i) は、検索質問が特徴的な索引語に分割されたために、個々の索引語に関する ALD の歪度が大きくなったことが原因である。(ii) は、分割された索引語全てを含むアンカーテキストが存在したために、ALD の歪度が大きくなったことが原因である。

(iii) は、(i) の逆であり、検索質問が一般的な索引語に分割されたために、個々の索引語に関する ALD の歪度が小さくなったことが原因である。(iv) は、「遺伝子組み換え食品」という検索質問が「遺伝子組み換え食品ホームページ」や「本当はどうなの？ 遺伝子組み換え食品」のように様々な文脈でアンカーテキストに使用されていたために、ALD の歪度が小さくなったことが原因である。(v) は、検索質問も分割後の索引語もアンカーテキストに存在しないことが原因だった。(vi) は、他の誤り事例に比べると ALD の歪度は大きいものの、 $i(q)$ が 0.5 という閾値より若干大きかったことが原因である。

6 質問応答への応用

6.1 概要

質問応答の研究は、知識ベースと推論機構によって回答を導き出す「人工知能的アプローチ」と、大

量の文書集合から回答となる語句を抽出する「情報検索的アプローチ」に大別することができる。

情報検索的アプローチは、対象の領域に応じて知識ベースを構築する必要がないため、適用範囲が広い。近年、TREC や NTCIR の評価ワークショップで研究されている質問応答システムは情報検索的アプローチに基づいている。例えば、「東京ディズニーランドはどこにありますか?」というような自然言語による質問文を入力として、東京ディズニーランドに関する文書群から回答の候補となる語句を出力する。本研究もこのアプローチに焦点を当て、以降、単に「質問応答」と呼ぶ。

Web を文書集合として用いる質問応答には種々の事例 [1, 2, 6, 11, 14, 16] がある。多少の差異はあるものの、一般化すると、「質問文解析」、「文書検索」、「回答抽出」という一連の処理に分割できる。

具体的には、質問文解析を実行して、質問文で尋ねられている内容（例えば、人物なのか場所なのか）を特定する。次に、質問文から抽出したキーワード等を用いて検索質問を構成して文書検索を実行する。最後に、回答抽出を実行して、検索された Web ページから回答となる語句を抽出する。この段階で、複数の回答候補が得られた場合は、各候補に対して妥当性の検査を行うことがある。

本研究で提案した検索質問の分類手法は、上記の「文書検索」において応用が可能である。すなわち、質問文の内容に基づいて、Web 検索の手法を動的に変更させる点に特長がある。

なお、混乱を避けるために、質問応答への入力を「質問文 (question)」と呼び、文書検索への入力を「検索質問 (query)」と呼ぶことで、両者を区別する。

6.2 検索質問分類の効果

質問文解析では、「～は誰ですか」、「～はいつですか」、「～はどこにありますか」というような質問固有の表現に基づいて、尋ねられている内容が 5W1H 要素のどれかであるかを特定する。例えば、「東京ディズニーランドはどこにありますか?」は、「場所 (Where)」を尋ねている質問文であることが分かる。質問文解析の結果に応じて、回答抽出において抽出対象となる語句の種類が異なる。

それに対して、本研究で提案した検索質問の分類

手法は、「東京ディズニーランド」という名詞を用いる。さらに、分類の結果に応じて文書検索の手法を変更する。このことが質問応答に与える効果について議論する。

「東京ディズニーランド」は、誘導型の検索質問に分類され、文書検索によって東京ディズニーランドのトップページが検索される。そこで、検索されたトップページから住所等を抽出すれば、詳細かつ正確な情報をユーザに提示することができる。

しかし、「イラクの首都はどこにありますか?」という質問文では、「イラク」や「首都」という検索質問は調査型に分類され、イラクに関して記述された種々のページが検索される。それらのページから、具体的な住所ではなく、首都の名前等を回答として抽出する必要がある。

以上、場所に関する2つの質問文から、検索質問の分類によって、「具体的な住所」と「大まかな場所」のどちらを回答すべきかという「詳細度」の制御が可能になることが分かる。また、誘導型の検索質問に分類された場合には、対象となる名詞に関するトップページを検索することによって、回答の信頼性向上が可能になることが分かる。

尋ねられている内容が「人物(Who)」の場合にも、検索質問の分類が質問応答に与える効果がある。例えば、対象の人物が歴史上の人物であって、その人物に関するトップページが存在しなければ、調査型の検索質問に分類される。そこで、対象の人物に関するページ(例えば、Wikipediaの記事)が検索される。しかし、対象の人物が歌手や研究者であり、その人物に関するトップページが存在すれば、誘導型の検索質問に分類される。そこで、質問文の内容に応じて、そのトップページから必要な情報を抽出するか、もしくはそのページそのものを回答として提示する。

既存の質問応答で用いられる質問文解析は、「ユーザが入力した質問文に何が書かれているか」という点に着目して質問の内容を特定する。すなわち、「質問駆動型の分析」を行う。それに対して、本研究で提案した検索質問の分類は、「対象の名詞に関する情報がWeb上でどのように提供されているか」という点に着目して、質問の内容を特定する。すなわち、「文書駆動型の分析」を行う。本研究の特長は、異なる分析手法を統合して、質問応答を高度化する点にある。

6.3 関連研究

既存の質問応答手法では、文書集合としてWebを用いるかどうかにかかわらず、文書検索で用いる検索手法は固定的である。本研究のように、質問文中に含まれる名詞によって、検索手法を動的に変更し、回答の詳細度を制御したり、回答の信頼性を向上させる手法は存在しない。

ただし、GoogleやYahoo!等の検索サービスで提供されるAPIを文書検索に用いる場合は、APIの内部で検索手法が適宜変更されているかもしれない。しかし、質問応答手法を提案する研究者がAPI内部の動作を制御しているわけではない。

Yangら[19]は、一つの質問文に対して複数の回答が要求される「リスト型の質問文」に対して、Webページの種類に基づく質問応答手法を提案した。リスト型の質問文では回答の網羅性が重要であるため、リンク集のような収集的なページから回答を抽出することが好ましい。そこで、文書検索で得られたページ群を「収集的なページかどうか」という観点で分類して、質問応答の網羅性を向上させた。しかし、本研究のように、回答の詳細度や信頼性の問題に取り組んだわけではない。

7 おわりに

Web検索では、検索質問の種類によって必要とされる検索モデルが異なる。本研究は、検索質問を「調査型」か「誘導型」に分類し、その結果に基づいて検索手法を動的に変更する手法を提案した。NTCIRの100ギガバイトWebテストコレクションを用いた評価実験の結果、提案した手法は既存の手法よりも分類精度が高く、かつWeb検索の精度を向上させることが分かった。

本研究で提案したWeb検索の手法は、それ自身が独立したアプリケーションとして意義がある。また同時に、Webを利用する自然言語処理の研究に対しては、基盤(ミドルウェア)としての意義がある。このような例として、本分類手法を質問応答へ応用した場合の効果について議論した。今後の課題は、質問応答への効果を実験によって明らかにすることである。

謝辞

本研究の一部は、文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しいIT基盤技術の研究」(研究課題: WWWのオンラインマニュアル化による知識検索, 課題番号: 19024007)によって実施された。

参考文献

- [1] Eugene Agichtein, Steve Lawrence, and Luis Gravano. Learning search engine specific query transformations for question answering. In *Proceedings of the 10th International World Wide Web Conference*, pp. 169–178, 2001.
- [2] David Azari, Eric Horvitz, Susan Dumais, and Eric Brill. Web-based question answering: A decision-making perspective. In *Proceedings of AAAI-2003*, 2003.
- [3] Ricardo Baeza-Yates, Liliana Calderón-Benavides, and Cristina González-Caro. The intention behind Web queries. In *Proceedings of the 13th International Conference on String Processing and Information Retrieval*, pp. 98–109, 2006.
- [4] Andrei Broder. A taxonomy of web search. *ACM SIGIR Forum*, Vol. 36, No. 2, pp. 3–10, 2002.
- [5] Nick Craswell, David Hawking, and Stephen Robertson. Effective site finding using link anchor information. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 250–257, 2001.
- [6] Susan Dumais, Michele Banko, Eric Brill, Jimmy Lin, and Andrew Ng. Web question answering: Is more always better? In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 291–298, 2002.
- [7] Koji Eguchi, Keizo Oyama, Akiko Aizawa, and Haruko Ishikawa. Overview of the WEB task at the fourth NTCIR workshop. In *Proceedings of the Fourth NTCIR Workshop on Research in Information Access Technologies Information Retrieval, Question Answering and Summarization*, 2004.
- [8] Koji Eguchi, Keizo Oyama, Emi Ishida, Noriko Kando, and Kazuko Kuriyama. Overview of the Web retrieval task at the third NTCIR workshop. In *Proceedings of the Third NTCIR Workshop on Research in Information Retrieval, Automatic Text Summarization and Question Answering*, 2003.
- [9] Atsushi Fujii, Katunobu Itou, Tomoyosi Akiba, and Tetsuya Ishikawa. Exploiting anchor text for the navigational web retrieval at ntcir-5. In *Proceedings of the Fifth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-lingual Information Access*, pp. 455–462, 2005.
- [10] In-Ho Kang and GilChang Kim. Query type classification for Web document retrieval. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 64–71, 2003.
- [11] Cody C. T. Kwok, Oren Etzioni, and Daniel S. Weld. Scaling question answering to the Web. In *Proceedings of the 10th International World Wide Web Conference*, pp. 150–161, 2001.
- [12] Uichin Lee, Zhenyu Liu, and Junghoo Cho. Automatic identification of user goals in Web search. In *Proceedings of the 14th International World Wide Web Conference*, pp. 391–400, 2005.
- [13] Yun Yao Li, Rajasekar Krishnamurthy, Shivakumar Vaithyanathan, and H. V. Jagadish. Getting work done on the Web: Supporting transactional queries. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 557–564, 2006.
- [14] Jimmy Lin, Aaron Fernandes, Boris Katz, Gregory Marton, and Stefanie Tellex. Extracting answers from the Web using knowledge annotation and knowledge mining techniques. In *Proceedings of the 11th Text REtrieval Conference*, 2002.
- [15] Keizo Oyama, Koji Eguchi, Haruko Ishikawa, and Akiko Aizawa. Overview of the NTCIR-4 WEB navigational retrieval task 1. In *Proceedings of the Fourth NTCIR Workshop on Research in Information Access Technologies Information Retrieval, Question Answering and Summarization*, 2004.
- [16] Dragomir Radev, Weiguo Fan, Hong Qi, Harris Wu, and Amardeep Grewal. Probabilistic question answering on the Web. In *Proceedings of the 11th International World Wide Web Conference*, 2002.
- [17] S.E. Robertson and S. Walker. Some simple effective approximations to the 2-poisson model for probabilistic weighted retrieval. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 232–241, 1994.
- [18] Ellen M. Voorhees and Dawn M. Tice. Building a question answering test collection. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 200–207, 2000.
- [19] Hui Yang and Tat-Seng Chua. Web-based list question answering. In *Proceedings of the 20th International Conference on Computational Linguistics*, pp. 1277–1283, 2004.