

# factoid 型質問応答における異なる Web 検索エンジンの組合せの効果

金井 明<sup>†</sup> 佐藤 充<sup>†</sup> 石下 円香<sup>†</sup> 森 辰則<sup>‡</sup>  
<sup>†</sup>横浜国立大学 大学院 環境情報学府 <sup>‡</sup>横浜国立大学 大学院 環境情報研究院  
 E-mail: {a-kanai,mitsuru,ishioroshi,mori}@forest.eis.ynu.ac.jp

## 1 はじめに

質問応答(QA)は、質問に対して関連文書を探して提示するだけでなく、直接質問の答を提示してくれるシステムである。また、近年では、新聞記事集合のような静的かつローカルな文書群ではなく、文書が豊富で日々追加・更新される Web 文書を情報源とした質問応答システム(Web QA)が研究されている。通常、Web 検索エンジンを質問応答用に独自に用意するのは非現実的なので、既存の Web 検索エンジンが利用される。

ここで、我々は複数の異なる検索エンジンが Web QA に使用できることに注目している。各検索エンジンの出力する検索結果が同一であることはなく、それらを組合せることで情報源の多様性を増す事ができ、Web QA の精度向上への何らかの効果が期待される。そこで本論文では、QA において異なる Web 検索エンジンを組合せることによる効果について調査を行なった。

## 2 関連研究

初の Web QA システムの 1 つである START の最近の版では、複数の情報源を活用している [3]。Radev ら [5] は Web QA における確率に基づくアプローチを提案しているが、そこでは 3 つの主要な Web 検索エンジンを組合せて上位 40 文書を得ている。

これらの研究では異なる情報源から得られた文書を使用しているが、文書検索よりも後の段階においては、情報源の違いは考慮していない。これに対して、節 4 で述べる我々の手法では、異なる情報源の間でのデータの冗長性を活用しようとしている。

## 3 基本となる WebQA システム

本研究で使用する Web QA システムは、Web 検索エンジンの出力を利用し、日本語の factoid 型質問に対し、日本語で答を返すシステムである。また、数百もの Web 文書をダウンロードすることは、非常に時間がかかる処理であるので、応答時間の短縮のために Web 検索エンジンによる短い抜粋出力である snippet を Web QA の情報源として用いている。

図 1 にその Web QA システムの構成を示す。質問文解析部は利用者から質問文を受け取り、キーワードのリストや質問文の型などの情報を抽出する。キーワードのリストが検索質問として Web 検索エンジンに入力され、snippet 群が検索される。文照合部は、文書集合から抽出された文集合をパッセージ抽出部から受け取り、それら进行处理する。ここで得られた各文を本論文では**検索文**と呼ぶ。各検索文中の各形態素が一つの解候補として扱われ、それらに対して、次節に述べる方法によりスコアが与えられる。

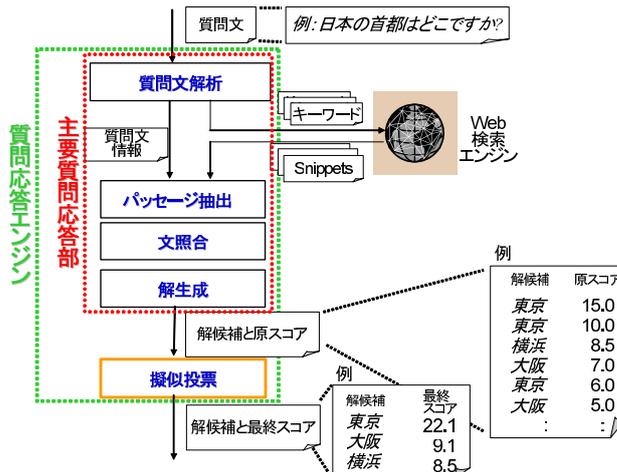


図 1: 基本となる WebQA システム

### 3.1 解候補のスコア付け

基本となる QA システムでは、解候補に対する複合的な照合スコアを採用している。

$$S(AC, L_i, L_q) = Sb(AC, L_i, L_q) + Sk(AC, L_i, L_q) + Sd(AC, L_i, L_q) + St(AC, L_i, L_q) \quad (1)$$

本論文ではこのスコアのことを**原スコア**と呼ぶ。このスコアは、 $i$  番目の検索文  $L_i$  にある解候補  $AC$  に対し、質問文  $L_q$  に関する以下の 4 つの部分スコアを計算し、その線形結合を求めたものである。

1.  $Sb(AC, L_i, L_q)$ : 文字 2-grams の観点で計算した照合スコア
2.  $Sk(AC, L_i, L_q)$ : キーワードの観点で計算した照合スコア
3.  $Sd(AC, L_i, L_q)$ : 解候補とキーワードの間の依存構造の観点で計算した照合スコア
4.  $St(AC, L_i, L_q)$ : 質問文の型の観点で計算した照合スコア

### 3.2 探索制御の枠組みにおける疑似投票手法

既存の多くの QA システムでは解候補に関する大域的な情報を利用している。特に冗長性は最も基本的であり、かつ重要な情報である。例えば、文書中に複数回出現する解候補に対し、そのスコアを増加させるという投票手法がある [1]。

一方で、我々の QA システムのように解候補の探索に基づく枠組においては投票手法をそのまま利用できない。なぜならば、同枠組では上位  $n$  件の解候補が見つかったら検索をそこで終了してしまうので、文書集合全体における解候補についての正確な頻度情報が得られないためである。そこで、投票手法の一つの近似として、以下に述べる疑似投票手法を導入している。まず、上位  $n$  件の解が必要な場合には、**表面表現の異なる解候補が  $n$  個見つかるまで探索を継続していることに注意された**

い。そのため、探索の過程において、すでにスコア計算の終了状態に至っている解候補と全くおなじ表層表現を持つが別の解候補が新たに終了状態に至ることが有り得る。よって、探索の過程において終了状態に至ったすべての解候補を記録することで、解候補の頻度情報を部分的に利用できる。本論文では、解候補  $AC$  に対する疑似投票スコア  $S^v(AC, L_q)$  を次のように定義する。

$$S^v(AC, L_q) = (\log_{10}(\text{freq}(AC, \text{AnsList})) + 1) \cdot \max_{L_i} S(AC, L_i, L_q) \quad (2)$$

ここで、 $\text{AnsList}$  は  $n$ -best 探索において終了状態に至った解候補のリスト、 $\text{freq}(x, L)$  を  $L$  における  $x$  の頻度とする。また、本論文では、この疑似投票スコアを最終スコアと呼ぶことにする。

#### 4 異なる Web 検索エンジンの出力結果を組合せる手法

異なる Web 検索エンジンの出力結果を組合せる手法として以下の 3 手法を検討する。

**手法 A:** 各検索エンジンから得られた文書集合 (snippet の集合) をそのまま併合し、それを 1 つの質問応答エンジンに送る。

**手法 B:** 各検索エンジンから得られた文書集合をそれぞれ個別の主要質問応答部に送り、疑似投票を行なう前に、得られた解候補 (原スコアを有する) の各リストを併合する。その後、式 (2) を使って疑似投票を行なう。次の手法 C とは異なり疑似投票は 1 回だけ行なわれる。

**手法 C:** 各検索エンジンから得られた文書集合をそれぞれ個別の質問応答エンジンに送り、疑似投票まで行なった結果として得られた解候補 (最終スコアを有する) の各リストを併合する。併合結果の中においても同じ表層表現を持つ複数の解候補が得られることがあるので、再び式 (2) を使って疑似投票を行なう。

図 2, 3, 4 に手法 A, B, C をそれぞれ示す。各図において、併合部は、複数の入力から各々データのリストを受けとり、それらを単純に併合してリストを生成する部分プログラムである。

さて、手法 A はベースライン手法であり、質問応答の処理をする前に各検索エンジンの出力結果を併合する。節 2 で述べたように、これと同等の方法が先行研究でも採用されている。

一方、手法 B と C は我々の提案手法であり、質問応答処理を行なう前に文書集合を併合するのではなく、質問応答処理を個別に行なった後に、解候補を併合するという手法である。ここで、手法 B, C が、情報源

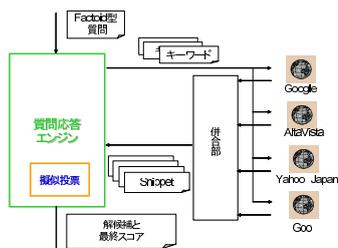


図 2: 複数の Web 検索エンジンを用いる QA システム: 手法 A (ベースライン)

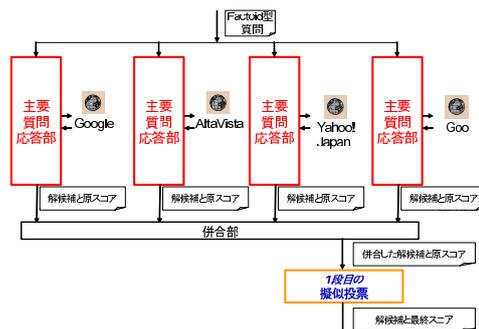


図 3: 複数の Web 検索エンジンを用いる QA システム: 手法 B

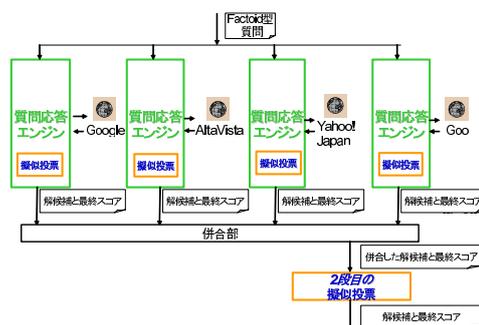


図 4: 複数の Web 検索エンジンを用いる QA システム: 手法 C

の多様性の観点において手法 A と異なるということに注意されたい。手法 A と B は、いずれも、疑似投票の前に、Web 検索エンジンが出力する snippet から、原スコアの観点においてより上位の解候補を抽出し、その後に疑似投票を行なうという過程になっているので、一見すると、両者はほとんど同じ解候補を最終的に出力するように見える。しかし、両者の間には次のような差異が存在する。手法 A においては、由来する検索エンジンによる区別をせず、各解候補を同等に扱い、混合された状態でスコアによる順位付けが行なわれるため、最終的に疑似投票にかけられる解候補が、ごく少数の検索エンジンのみ由来し、情報源の利用に偏りが存在する可能性がある。一方で、手法 B ではすべての検索エンジンについて、それに由来する解候補が必ず同数ずつ疑似投票にかけられるので、情報源の多様性が確保されるようになっている。

手法 C も手法 B と同様の状況にあるが、更に、2 段目の疑似投票処理において、より多くの異なる検索エンジンに由来する解候補に対してより高いスコアが与えられる点が手法 B と異なる。

なお、ここでの議論は、解候補に対する確信度が、由来する情報源の多様性によって、ある程度測定できるという我々の仮説に基づいている。この仮説が成り立つのであれば、手法 A よりも、手法 B, C のほうが精度が高いということが期待される。

#### 5 評価実験

前節で述べた 3 つの組合せ手法を評価、検討するために、以下のような評価実験を行なった。特に、利用する Web 検索エンジンの数の効果を調査するために、質問応答における各種設定を同一にした状況において、i) 各 Web 検索エンジンを単独で使用した場合、ii) Web 検索エンジンを 2 つ組合せた場合、iii) 3 つ組合せた場

合, iv) 4つ組合せた場合の各々で実験を行なった。

評価に用いる質問文集合としては, NTCIR-3 QAC1 [2] の質問文 200 問のうち, 検索 API のタイムアウト等によりいずれかの QA エンジンが一定時間内に回答を返さなかった 8 問を除外した, 計 192 問を評価に用いた。Web 検索エンジンとしては, i) AltaVista, ii) goo, iii) Yahoo! Japan, iv) Google, を使用した。

また, 本実験で設定した各種パラメタの値を表 1 に示す。この設定は, 森 [4] において, 新聞記事を情報源とした時に最も求解精度が良かったものである。なお, 一つのパッセージは隣接する 3 文から構成されている。

表 1: 評価実験における QA エンジンのパラメタ設定

| パラメタ名 | 設定値 | 説明                            |
|-------|-----|-------------------------------|
| $a$   | 10  | 検索すべき解候補の数                    |
| $d$   | 250 | 検索すべき文書 (snippet) の数          |
| $ppd$ | 5   | 各文書 (snippet) から抽出するパッセージの最大数 |
| $p$   | 30  | 求解で考慮すべきパッセージの最大数             |

ここで注意すべきことは, 表 1 の設定をいずれの実験においても共通に採用すると, 利用する Web 検索エンジンの数を変化させた際に, 検索される文書 (snippet) の延べ総数もそれに比例して変化する点である。ここにおいて何をもって公平な比較と考えるかは難しい問題である。もう一つの可能性としては, 最終的にシステム全体で参照する文書 (snippet) の延べ総数が同じになるように, 表 1 のパラメタ  $d$  を調整することが考えられる。しかしながら, 森 [4] によると, 検索する文書数を増やすことが必ずしも求解精度の向上につながらないことが報告されている。これは, 順位の低い文書まで参照することにより, その結果, 原スコアは低いものの出現頻度が高い誤った解候補が出現しやすくなるためである。そこで, 本論文では単純に, 同研究で良いとされた表 1 に示すパラメタを用いることにする。

## 6 実験結果と考察

求解精度の尺度として MRR (正解が最初に現れた順位の逆数の全質問平均) を採用した。その値を図 5 に示す。図 5 によれば, 単独の Web 検索エンジンを使ったシステムより Web 検索エンジンを複数組合せるシステムのほうが精度がよくなる傾向にある。特に, 質問応答処理を行う前に文書集合を併合する手法 A に比べ, 質問応答処理を個別に行なった後に, 解候補を投票処理により併合する手法 B, C のほうが精度がよいことがわかる。一方, 手法 B, C の間の差は顕著ではない。以上より, 文書集合を先に併合してから質問応答処理を行う方法よりも, Web 検索エンジンの出力毎に独立に質問応答処理により解候補を精選した後, それらを投票処理により併合する方法のほうが精度が良いということがわかる。

また, 組合せる Web 検索エンジンの数が増えるに従って, 概ね精度が向上することも観察される。一方で, 図 5 の矢印 1 に示されるように, 最も精度が良かった組合せは, すべての Web 検索エンジンを利用した場合ではなく, goo の出力を用いず, AltaVista, Yahoo! Japan, Google を組合せた場合であることに注意されたい。ここで, 図 5 の矢印 2 に示される箇所に注目すると, goo と Google の組合せにおいて, 他の組合せよりも精度が低い傾向にあることがわかる。この組合せにおいては, Google 単独と同程度以下の精度しか得られていない。その原因を調べるために, 同一質問に対して goo ならびに Google が検索結果として何を出力しているかを調べてみた。質問「スイスの公用語は何語ですか。」(QAC1-1140-1) に対する出力を表 2 に

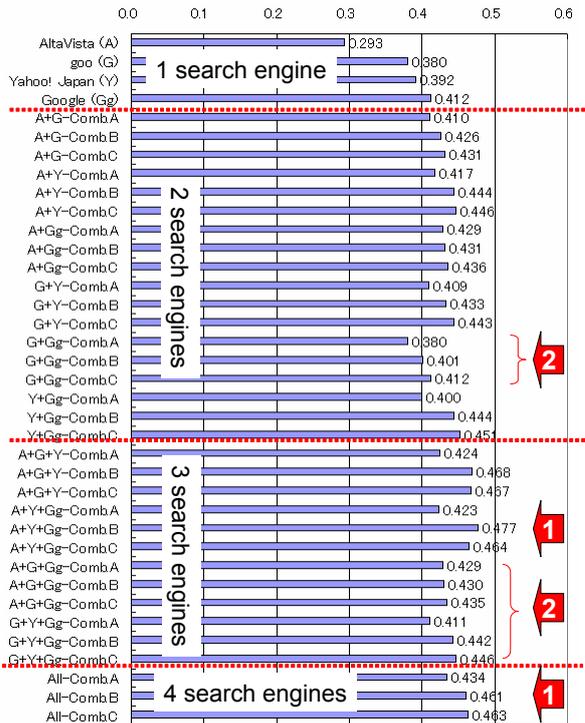


図 5: Web 検索エンジンの各組合せに対する MRR 値を示す。この表において, Google の snippet 20 件に対して

表 2: Google と goo の snippet の比較 (質問が「スイスの公用語は何語ですか。」(QAC1-1140-1) の場合)

| rank | Snippet (Google)         | Snippet (goo)            |
|------|--------------------------|--------------------------|
| 1    | a 16世紀に、スイス連邦は13の自       | ★ スイス連邦の正式名称は4種の         |
| 2    | b 「スイス」200年近くにわたり中立      | a 16世紀に、スイス連邦は13の自       |
| 3    | c 様々な風景が凝縮された国スイス        | b 「スイス」200年近くにわたり中立      |
| 4    | d トロミテ・ラティン語(ドイツ語)       | c 様々な風景が凝縮された国スイス        |
| 5    | e 「国別情報(国際機関)」           | d トロミテ・ラティン語(ドイツ語)       |
| 6    | f スイスは4つ公用語を持つ世界で        | e 「国別情報(国際機関)」           |
| 7    | g 言語、スイスにはドイツ語、フラン       | ★ 参照(クリック) [イタリア語(伊語)    |
| 8    | h スイスでは、3つの公用語を使用        | f スイスは4つ公用語を持つ世界で        |
| 9    | i だが、スイスとの関連で言えば何        | g 言語、スイスにはドイツ語、フラン       |
| 10   | j 彼が教えて英語で書かれた拙稿         | h スイスでは、3つの公用語を使用        |
| 11   | k スイスに住んでいる人はフランス語       | i だが、スイスとの関連で言えば何        |
| 12   | l 言語事情ドイツ語(63.6%)、フ      | j 彼が教えて英語で書かれた拙稿         |
| 13   | m スイスは人口七百万人余りの小国        | k スイスに住んでいる人はフランス語       |
| 14   | n なかなか難しい案件だったよう         | l 言語事情ドイツ語(63.6%)、フ      |
| 15   | o Quickbar   image=250px | m 年生各学年週4時限履修。2002       |
| 16   | p アンドラ公国では公用語、オック語       | n なかなか難しい案件だったよう         |
| 17   | q スイス スイス連邦。欧州中部         | o Quickbar   image=250px |
| 18   | r 時差・時間、一8時間(サマータイ       | p アンドラ公国では公用語、オック語       |
| 19   | s スイスでは地理的又は歴史的な理        | q スイス スイス連邦。欧州中部         |
| 20   | t スイスでは、ドイツ語圏とフランス       | r 時差・時間、一8時間(サマータイ       |

し, 異なるアルファベットを ID として付与してある。同様に goo の snippet に対しても Google の snippet と同一の文字列となっている場合に, 対応するアルファベットを付与してある。星印は Google の snippet に対応物が存在しない goo の snippet である。表 2 によれば Google と goo は同一の snippet を, 20 件中 15 件と数多く出力していることが分かる。このような snippet の同一性は, goo が 2003 年以降, Google と提携し, Google の検索結果を利用しているためだと考えられる。このように数多くの同じ snippet が得られる Web 検索エンジンの組合せにおいては, snippet の表現ならびにそこから抽出される解候補の多様性を担保することができない。例えば, 先の質問において, goo, Google, AltaVista を各々単独で用いた場合に抽出される解候補と, goo と Google との組合せ, ならびに, AltaVista と Google との組合せにおいて抽出される解候補について調べてみよう。図 6 にスコアの降順に各上位 10 件の解候補を示す。

この図に示すとおり, goo 単独と Google 単独の場

| Rank | goo (G) | Score | Rank | Google (Gg) | Score | Rank | AltaVista (A) | Score |
|------|---------|-------|------|-------------|-------|------|---------------|-------|
| 1    | 4つ      | 67.3  | 1    | ロマンシュ語      | 63.4  | 1    | ロマンシュ語        | 67.4  |
| 2    | ロマン語    | 52.8  | 2    | イタリア語       | 60.4  | 2    | フランス語         | 62.3  |
| 3    | ロマンシュ語  | 48.8  | 3    | フランス語       | 60.4  | 3    | 英語            | 61.1  |
| 4    | スイスドイツ語 | 48.4  | 4    | ロマン語        | 52.8  | 4    | ドイツ語          | 47.0  |
| 5    | イタリア語   | 46.4  | 5    | 7つ          | 51.7  | 5    | イタリア語         | 47.0  |
| 6    | フランス語   | 46.4  | 6    | 4つ          | 48.7  | 6    | 国             | 46.4  |
| 7    | 憲法      | 45.9  | 7    | スイスドイツ語     | 48.4  | 7    | 言葉            | 46.3  |
| 8    | 独語      | 45.2  | 8    | ドイツ語        | 46.4  | 8    | ユ語            | 45.6  |
| 9    | 仏語      | 45.2  | 9    | 憲法          | 45.9  | 9    | 三言語           | 45.4  |
| 10   | 伊語      | 45.2  | 10   | 独語          | 45.2  | 10   | 四言語           | 44.0  |

| Goo+Google |                   | AltaVista+Google |                   |         |      |
|------------|-------------------|------------------|-------------------|---------|------|
| Rank       | G+Gg-Comb.B Score | Rank             | A+Gg-Comb.B Score |         |      |
| 1          | 4つ                | 76.5             | 1                 | ロマンシュ語  | 82.8 |
| 2          | ロマンシュ語            | 72.0             | 2                 | フランス語   | 76.7 |
| 3          | ロマン語              | 68.7             | 3                 | イタリア語   | 69.4 |
| 4          | フランス語             | 68.6             | 4                 | ドイツ語    | 61.2 |
| 5          | イタリア語             | 68.6             | 5                 | 英語      | 61.1 |
| 6          | スイスドイツ語           | 62.9             | 6                 | ロマン語    | 52.8 |
| 7          | 憲法                | 59.7             | 7                 | 7つ      | 51.7 |
| 8          | 独語                | 58.8             | 8                 | 4つ      | 48.7 |
| 9          | 仏語                | 58.8             | 9                 | スイスドイツ語 | 48.4 |
| 10         | 伊語                | 58.8             | 10                | 国       | 46.4 |

Correct answer

図 6: 各組合せにおける解候補のスコア付けの比較 (質問が「スイスの公用語は何語ですか。」(QAC1-1140-1)の場合)

合の解候補を比較すると、正解、不正解に関わらず同じ文字列が多く出現していることが分かる。一方で、Google 単独と AltaVista 単独の場合を比較すると、正解である解候補は共通して登場しているもの、不正解の解候補はどちらか一方に出現するのみである。このような状況にあるため、goo と Google を組合せた時には、投票処理によって運悪く不正解の解候補のスコアが上昇することもあり、結果として RR 値が相対的に下がることがある。図 6 左下に示す例では不正解である「4つ」という解候補が 1 位となってしまう<sup>1</sup>。一方で、Google と AltaVista の組合せにおいては、図 6 右下の例のように、正解である解候補のスコアが投票処理によって適切に上昇し、結果として上位 4 件がいずれも正解となっている。

次に各手法間において RR の代表値に統計的有意差があるかどうかを調べるために、ウィルコクソンの符号付順位和検定(両側検定)を行なった。RR の代表値に統計的な有意差が見られた手法間の比較の代表例を表 3 に示す。表中の「p 値」はウィルコクソンの符号付順位和検定統計量の p 値である。また、表中で「手法 X < 手法 Y」と記されている時には、手法 Y による RR の代表値のほうが手法 X による RR の代表値よりも大きく、なおかつ、その差が統計的に有意であることを示しており、手法 Y の方が手法 X より精度が良いことを表す。表 3 の (1) に示す、「Web 検索エンジンの組合せ数に関する比較」の結果を見ると、組合せる検索エンジンの数を増やすほど精度が向上することが統計検定によって確認がなされている。紙面の都合で割愛したが、この表に示した以外の組合せについても、各 Web 検索エンジンを単独で使用した場合と組合せて使用した場合の間で、統計的有意差が見られるものが数多くあった。次に表 3 の (2) に示す「組合せ手法に関する比較」の結果を見ると、手法 B、C と手法 A との間で統計的有意差が見られ、提案した組合せ手法 B ならびに C の有効性が確認された。また、表 3 の (3) に示す「goo と Google の組合せに関する比較」の結果を見ると、AltaVista、Google、goo の組合せと AltaVista、Google、Yahoo の組合せの間に統計的有意差が見られ、検索結果の snippet が似通っている

<sup>1</sup> 質問文中にある表現「何語」は「なにご」(何の言語の意味)と解釈すべきであるが、局所的にみると「なんご」(いくつの単語の意味)とも解釈できるので、今回評価に利用したシステムでは、「数」も質問の型としている。

Google と goo を組合せることが精度を劣化させることがわかる。

表 3: 統計検定により RR の代表値に差が確認された手法間の比較の例とその p 値

| RR の代表値に差が確認された手法間の比較                 | p 値                    |
|---------------------------------------|------------------------|
| (1) Web 検索エンジンの組合せ数に関する比較             |                        |
| AltaVista (A) < A + Y + Gg (手法 B)     | $2.3 \times 10^{-8**}$ |
| Yahoo (Y) < A + Y + Gg (手法 B)         | 0.001**                |
| goo (G) < ALL (手法 C)                  | $7.4 \times 10^{-4**}$ |
| Google (Gg) < A + Y + Gg (手法 B)       | 0.011*                 |
| Y + Gg (手法 B) < A + Y + Gg (手法 B)     | 0.018*                 |
| A + Gg (手法 B) < A + Y + Gg (手法 B)     | 0.018*                 |
| G + Gg (手法 A) < A + G + Gg (手法 A)     | 0.020*                 |
| G + Gg (手法 B) < G + Y + Gg (手法 B)     | 0.015*                 |
| G + Gg (手法 B) < A + G + Gg (手法 B)     | 0.033*                 |
| A + Y (手法 B) < A + G + Y (手法 B)       | 0.010*                 |
| A + G + Gg (手法 B) < ALL (手法 B)        | 0.023*                 |
| A + G + Gg (手法 C) < ALL (手法 C)        | 0.049*                 |
| (2) 組合せ手法に関する比較                       |                        |
| A + Y + Gg (手法 A) < A + Y + Gg (手法 B) | 0.040*                 |
| Y + Gg (手法 A) < Y + Gg (手法 C)         | 0.032*                 |
| A + G + Y (手法 A) < A + G + Y (手法 B)   | 0.030*                 |
| A + G + Y (手法 A) < A + G + Y (手法 C)   | 0.045*                 |
| (3) goo と Google の組合せに関する比較           |                        |
| A + Gg + G (手法 B) < A + Gg + Y (手法 B) | 0.009**                |

表中 \*\* は 1% 水準、\* は 5% 水準で有意差があることを示している。

## 7 おわりに

本論文では、Web 質問応答において、異なる Web 検索エンジンを組合せることによる効果について調査した。評価実験より、質問応答処理の前に検索結果を併合する手法よりも、QA エンジンによる個別の質問応答処理の後に、抽出された解候補を組合せるほうが効果があることが示された。

今後は、更に精度の良い質問応答の実現のため、a) Web 検索エンジンの出力のより効果的な組合せ手法、b) snippet における表現の多様性を積極的に引き出すため、Web 検索エンジンに入力する検索要求を生成する手法を複数用意して組合せる方法、などを検討したい。

## 謝辞

本研究の一部は文科省科研費特定領域「情報爆発 IT 基盤」(課題番号 19024033)によるものである。

## 参考文献

- [1] Charles L.A. Clarke, Gordon V. Cormack, and Thomas R. Lynam. Exploiting redundancy in question answering. In *Proceedings of SIGIR '01: the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 358-365, 2001.
- [2] Jun'ichi Fukumoto, Tsuneaki Kato, and Fumito Masui. Question Answering Challenge (QAC-1) — Question answering evaluation at NTCIR Workshop 3 —. In *Working Notes of the Third NTCIR Workshop meeting - Part IV: Question Answering Challenge (QAC1)*, pp. 1-6, 2002.
- [3] B. Katz, M. Bilotti, S. Felshin, A. Fernandes, W. Hildebrandt, R. Katzir, J. Lin, D. Loreto, G. Marton, F. Mora, and O. Uzuner. Answering multiple questions on a topic from heterogeneous resources. In *Proceedings of TREC 2004*, 2004.
- [4] Tatsunori Mori. Japanese question-answering system using A\* search and its improvement. *ACM Transactions on Asian Language Information Processing (TALIP)*, Vol. 4, No. 3, pp. 280-304, 2005.
- [5] Dragomir R. Radev, Weiguo Fan, Hong Qi, Harris Wu, and Amardeep Grewal. Probabilistic question answering on the web. *Journal of the American Society for Information Science and Technology*, Vol. 56, No. 3, March 2005.