

## How-to 型質問応答の実現に向けた質問回答文書の特徴分析

麻野間 直樹<sup>†</sup> 古瀬 蔵<sup>†</sup> 片岡 良治<sup>†</sup>

<sup>†</sup> 日本電信電話株式会社 NTTサイバーソリューション研究所

〒239-0847 神奈川県横須賀市光の丘 1-1

E-mail: <sup>†</sup> {asanoma.naoki, furuse.osamu, kataoka.ryoji}@lab.ntt.co.jp

あらまし 我々は従来の Web 検索よりも効率的に方法の回答らしい文書を探し出す、方法説明に特化した検索の実現を目指している。方法のような説明的な文書のスタイルは多様であるため、方法質問に対して満足いく回答かどうかを測定するために、説明的確さと詳しさの観点で、回答合致度および説明詳細度の二つの評価尺度を導入した。方法質問によって検索された文書に対して、評価尺度の値の高い文書にはどのような特徴があるか、また文書の特徴から評価尺度が予測可能かを検証した。その結果、箇条書きの有無と評価尺度との間に正の相関があることがわかった。さらに回答合致度を予測することが可能であることを回帰分析により示すことができた。回答合致度の低い文書を、誤って高いと予測した事例を分析すると、予測精度向上のためには、より深い言語解析処理が必要であることが確認された。

キーワード 方法説明検索, How-to 型質問応答, 回答らしさ, 文書スタイル, 箇条書き

## Feature Analysis of Explanatory Documents for How-to Type Question Answering

Naoki ASANOMA<sup>†</sup> Osamu FURUSE<sup>†</sup> and Ryoji KATAOKA<sup>†</sup>

<sup>†</sup> NTT Cyber Solutions Laboratories, NTT Corporation

1-1 Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847 Japan

E-mail: <sup>†</sup> {asanoma.naoki, furuse.osamu, kataoka.ryoji}@lab.ntt.co.jp

**Abstract** We are investigating a retrieval mechanism that provides adequate explanations to how-to type questions more efficiently than conventional web searches. Because explanatory documents have various document styles, we defined two measures for evaluating the probable likelihood of answer to how-to type questions. We examined the features of documents with high values of each measure, and the feasibility to estimate the measured values based on the document features. The experimental results showed a positive correlation between the presence of itemization and each measure. Regression analysis showed the possibilities for estimating the accordance value as one of two measures. Although the analysis of the cases in which the documents with low accordance values was assigned a high estimation value, deep natural language analysis is necessary to improve the estimation accuracy.

**Keyword** Explanatory procedure retrieval, How-to type question answering, Answer likelihood, Document style, Itemization

### 1. はじめに

何かの方法を知りたいと思った時、インターネット上からどのようにしたら回答を探せるだろうか。

「OKWeb (教えて!goo)」や「はてな」などの Q&A サイトでは、誰かの質問に他の誰かが答えて得られた

Q&A 集が検索可能となっており、その膨大な Q&A 蓄積データから自分の知りたい疑問を解決する回答を探し出すことが可能である。

このような Q&A サイトから回答を見つけられない、あるいは Q&A サイト自体になじみの薄いユーザは、

通常の Web 検索を試さざるをえない。現状でも、「方法」「手順」などの方法特有の検索語句をつけて質問キーワードを Web 検索に入力すれば、ある程度所望の回答を得られる場合はある。

しかし、その検索結果に付与されている概要、もしくは原文書を一一つ読んで、役に立ちそうなページかをユーザが判断する必要があり、満足いく回答を得るのに非常に時間がかかる場合が多いのが現実である。このように方法を検索するという目的において、Web 検索は必ずしも使い勝手がよいわけではない。そこで方法の回答取得に特化した Web 検索が必要となってくる。

方法に特化した検索は、方法質問に対する質問応答技術の一要素として見ることができる。質問応答技術 (Question Answering, 以下、QA) の研究は近年も盛んに行われており、TREC, NTCIR などの情報検索の会議においては QA が 1 タスク [1] として定着するほどとなっている。しかし、これまでの QA の研究 ([2][3] など) の多くは、人物・場所のような固有名詞 1 語、または数単語程度で答えられる形式の質問が対象となっており、比較的長い回答を求める「方法」などの質問応答技術はこれから取り組む課題の一つになっている。

ことばの定義、理由、方法などの説明的な答えをユーザに提示する QA の研究は、最近徐々に検討が進んできている。例えば、理由の回答を特定の言語現象を鍵として探し出す方法 [4] や、掲示板に現れる問い合わせ文章を対象として問題状況の類型化をする試み [5] などが行われている。また、蓄積した FAQ データに含まれる回答を再利用する方法検索の試みとしては、Tomuro らの FAQFinder [6] が例にあげられる。これはあらかじめ用意された FAQ の質問あるいは回答部分と、入力された質問文と照合し、最もよく照合する Q&A から回答を導き出すことを行う。

我々は、WWW を知識源として用いて、方法の質問に対してユーザが満足する回答を、従来の Web 検索よりも効率的に探し出せるような、方法説明に特化した検索機能を検討している。具体的には通常の Web 検索をベースとして、検索結果内の個々の文書に対して、「質問で尋ねている方法かどうか」について回答としての適切性の評価尺度を自動的に計算し、検索結果にその情報を付与することで、どの文書が満足度の高い回答かをユーザが判断できるようにする。

本報告では、このような方法説明検索の実現に向けて、方法に関する検索結果の文書を分析し、質問で尋ねている方法かどうかの回答らしさの評価尺度をどのように与えるべきか、与えた評価尺度によって、満足度の高い説明回答を効率的に検索できるかどうかの検

証を行う。

## 2. 評価尺度

ユーザ満足度が高くかつ効率的な方法検索を実現するにあたり、検索された文書が質問に対する方法説明を与えているか、与えているとしたらその回答は適切かどうかを判断する評価尺度を設定する必要がある。

「質問で尋ねている方法かどうか」の直接的な判断基準として、質問に合致した回答説明を含んでいるかを捉えることとし、この観点での評価尺度をまず設定する。

別の観点としては、方法の回答説明が詳しく書かれているかどうかを評価することもできる。

以降、方法に関する文書を分析する際には、次に示す二つの指標を評価尺度として用いることとする。

**回答合致度**：説明部分は質問に対して的確に回答しているか、外的れか

**説明詳細度**：説明部分は詳細に説明されているか、簡単な説明か

ある文書を評価するとき、これらの評価尺度について Yes/No の 2 値を与えることは可能である。しかしながら実際には、方法のような説明的な文書には、「かすってはいるが、聞いていることとあまり関係がない」、あるいは「問題解決にはなるが説明量が期待しているより非常に少ない」といった場合があり、質問に対する回答として適切かどうかの判断がゆれることが多い。したがって本稿の評価ではこれらの評価尺度を 0~1 の連続値とし、幅を持たせることにした。

## 3. 文書収集と評価の付与

方法についての質問に対する回答文書がどのような説明をしているかを把握するため、方法を尋ねる質問を検索エンジンに実際に入力し、その検索結果から辿ることのできる全ての文書を調査した。

最初に、何かの方法・手順を問うような質問を 30 個設定した。質問はそのキーワード列を見ただけで、何の方法を調べようとしているか明確にわかるものにした。図 1 は設定した質問 (クエリ) の一覧である。

次に、検索サイト goo<sup>1</sup> の「ウェブ検索」に図 1 の質問クエリを入力し検索結果を得る。それからそれぞれの質問クエリに対する検索結果の上位 90 件の HTML 形式の文書のスナップショット (現存文書、もしくは“キャッシュ”) を取得し保存した。このようにして取得した計 2,700 の文書 (以下、この文書の集合を HTML 文書と呼ぶ) を評価・分析対象とする。

さらに保存した HTML 文書に対して、前節で述べた

<sup>1</sup> <http://www.goo.ne.jp/>

グラタン レシピ  
 梅干 漬け方  
 鍋 ご飯 炊き方  
 ウール セーター 洗い方  
 フローリング 傷 補修 方法  
 アリ 家の中 駆除 方法  
 エアコン 掃除 方法  
 年賀状 パソコン 作り方  
 ポップアップ 広告 消す 方法  
 Excel 重複 データ チェック 方法  
 東京から 北海道 交通手段  
 確定申告 手順  
 シャッキリ 止め方  
 ジャガイモ 育て方  
 パスポート 申請方法  
 腹筋 鍛え方  
 犬 トイレ しつけ方  
 バラ 剪定 方法  
 ミルクティー 淹れ方  
 花粉症 予防対策  
 廃油から 石鹸 作り方  
 たけのこ アク抜き 方法  
 正しい 目薬 点眼 方法  
 ワイン 家庭で 保存 方法  
 スライム 作り方  
 有限会社 設立 手順  
 公正証書 遺言 作成 方法  
 内容証明郵便 出し方  
 鼻うがい 方法  
 包丁 研ぎ方

図 1 : 30 の質問クエリ一覧

二つの評価尺度による 3 段階評価 (1, 0.5, 0)<sup>2</sup>を与える。個人差を考慮して、同じ一つの文書に対して 3 人の評価者がこの評価を付与した。

また HTML 文書の文書構造の属性情報も付与した。具体的には HTML 文書をブラウザ上で閲覧し、見た目として箇条書きや表の形式がページ中に存在しているかをすべての文書に対して調査した。この際に、HTML のタグ (table, ul, li など) を使っているかどうかは問わず、画像や記号文字によって構成される箇条書き・表でも、見た目としてその形態になっていればよいことにした。

収集した 2,700 の HTML 文書のサイズの傾向としては、単語数の平均が 1,913、最大値 49,326、最小値 90 で、文書サイズにはかなりのばらつきがあった。HTML 文書のスタイルの傾向も、ブログを含む個人の日記、趣味でまとめたノウハウ集から、メーリングリストやメールマガジンのコピー、企業が運営するコラム、製品の紹介ページまで様々である。

#### 4. 評価尺度の傾向分析

2 節で定義した各評価尺度がどういった性質を持っているかをここで検証する。

まずそれぞれの評価尺度がどのような文書の性質

<sup>2</sup> 評価者に対する事後アンケートによると、3 段階評価で適切だった、との意見が得られている

を測るものかを理解するために、どちらか一方の尺度が高い値で他方は低い値となるような事例を分析してみる。以下にその一例を示す。

**回答合致度だけ高い文書：**リンク集のスタイルの文書で、回答に合致するページのリンクがその中の一つとして含まれている (例：色々な漬物のレシピへのリンクのみ羅列されている中に、目的の梅干の漬け方へのリンクがある)

**詳細説明度だけ高い文書：**質問からは外れた方法説明がこと細かに記述されている (例：腹筋の鍛え方の質問に対し、背筋の鍛え方の詳細な説明)

このように全く違うスタイルの文書が現れることから、それぞれの評価尺度は異なる観点で文書の回答らしさを測定していることがうかがえる。

またこれらの回答らしさの評価尺度と、goo ウェブ検索の検索順位とのふるまいの違いを見るために、前節で収集した文書を用いて、各評価尺度の値で降順に並べたときの 1 位文書が、実際の goo ウェブ検索の順位では最高何位になっているかを調べた。その結果、回答合致度では 30 のうち 13、説明詳細度では 22 の質問クエリにおいて、評価尺度での 1 位文書が Web 検索では 1 位になっていなかった。

これより、二つの評価尺度は、通常の Web 検索とも異なる基準で方法説明の回答らしさを測定していることがわかる。

ところで前節で収集した文書には、質問を解決する上で参考になるか、参考にならないかを示す有用度 (役立ち度) の評価も 3 段階で付与されている。満足度にもつながるこの有用度と二つの評価尺度との相関を調べることで、これらの尺度の妥当性を検証できる。調査の結果、有用度との相関係数は、回答合致度とは 0.93、説明詳細度とは 0.78 であった。どちらの評価尺度も高い正の相関を示し、妥当な指標であることが示された。言い換えればこれらの評価尺度の高い方法説明文書は、ほとんどの場合問題の解決に役立つことを示している。

以上のことから、方法検索を実現した際には、それぞれの評価尺度の観点で、方法説明に関する有用な情報をユーザに提供可能であると考えている。

#### 5. 方法説明文書の分析

3 節で収集した HTML 文書を用いて、方法説明文書にはどのような特徴があるのかを分析する。具体的には、HTML 文書から得られる各種の属性情報と、二つの評価尺度との相関関係を調べる。本分析では、各評価尺度は 3 人の評価者が与えたスコアの平均値をその代表値として用いることにする。

##### 5.1. 言語外情報との関係

最初の分析として、HTML 文書から得られる言語外

情報の特徴と、二つの評価尺度との相関を調べた。言語外情報は、箇条書きおよび表の有無（存在すると判定されていれば1、なければ0）、文書の検索順位、検索順位の逆数を示す Reciprocal Rank (RR) を特徴として選ぶ。

箇条書きや表を文書の特徴として選んだのは、箇条書きのような項目の順序を整理する文書構造が方法の説明によく現れるだろうとの仮説によるものであり、実際に予備検討の段階でもその傾向は確認していた。

表1は、これらの特徴と二つの評価尺度との相関係数を示している。

表1：言語外情報と評価尺度との相関

	箇条書き	表	検索順位	RR
回答合致度	0.502	0.171	-0.326	0.203
説明詳細度	0.520	0.177	-0.256	0.167

この結果から、まず箇条書きの有無は、二つの評価尺度のいずれとも正の相関が認められ、仮説は確かめられた。さらに検索順位と二つの評価尺度とは弱い負の相関が認められ、検索順位が高いほど質問に合致する文書が含まれている傾向があるという直感にあっている。逆に、表の有無とRRはごく弱い相関しか認められなかった。

以上より、箇条書きの文書構造は方法説明文書の特徴付ける要素であることが示された。

## 5.2 言語情報との関係

次に、HTML文書内のHTMLタグ以外のテキストを形態素解析した上で、その単語列から得られる言語情報と二つの評価尺度との相関を調べる。

以下に挙げる四つの統計量を、方法説明を特徴づけるものと仮定し、分析と数値化を行った。

**総単語数**：HTML文書に含まれる全単語の数。

**キーワード数**：FAQFinder[6]の文書判別実験では、「手順(方法)」に関するキーワードをあらかじめピックアップし、判別時の特徴としてそのキーワードの頻度を用いられていた。本分析でも同様の特徴を扱う。

まずHTML文書での高頻度語を集計し、その上位単語から方法説明特有と思われるキーワードを18個ピックアップし、方法説明の「キーワード」として設定した。ここで用いたキーワードを以下に示す。

方, 方法, 対策, 手, しかた, レシピ, 法, 手順, 手続き, 手続, 仕方, 方式, 手段, やり方, 策, 方針, 手法, 工法

キーワード数はこれら全キーワードにおける文書

中の出現頻度となる。

**クエリワード数**：HTML文書は、方法を問う質問による検索結果なので、各文書にはそれを得るもととなった質問クエリが対応している。

クエリワード数は、文書内において、その質問クエリに含まれる各単語（以下では、クエリ単語と呼ぶ）の出現頻度を合計した値である。クエリ単語が何度も出る文書は回答を含んでいる、との予想からこれを測定することにした。

**クエリカバー率**：さらに、クエリ単語がHTML文書のタイトルや見出しに出てくるような文書は、さらに重要である、とも予想できる。

ここでは同一のHTMLタグで囲まれている文字列中にどれだけ多くのクエリ単語が含まれているかを測定するために、クエリ単語の出現頻度を用いるのではなく、全クエリ単語のうちタグ中に出現したクエリ単語の数の比率をクエリカバー率として計算することにした。複数箇所の見出しでクエリ単語が出現する場合は、この比率は文書内での最大値を用いる。つまり、

$$(\text{クエリカバー率}) = \frac{\max((\text{同じ見出しタグ内で出現するクエリ単語数}) / (\text{クエリ単語数}))}{}$$

となる。なおタイトル・見出しとしてのタグは、今回"title","h1"~"h5"のみを対象とした。

これらの特徴と二つの評価尺度との相関係数を計算した結果を表2に示す。

表2：言語情報と評価尺度との相関

	総単語数	キーワード数	クエリワード数	クエリカバー率
回答合致度	-0.155	-0.063	0.025	0.360
説明詳細度	-0.000	0.033	0.128	0.254

まずクエリカバー率は、各評価尺度と弱い正の相関があることがわかった。つまり、より多くのクエリ単語がタイトルや見出しに現れるような文書は、方法の回答を説明しているという傾向があり、これは直感にも合致する。

一方、総単語数、キーワード数、クエリワード数のような単語の頻度を用いる特徴については、二つの評価尺度と非常に弱い相関か、もしくはまったく相関がないことが判明した。

文書全体を対象として頻度をカウントしたことや、方法特有のキーワードを分析なしで選んだことなどを改善すれば、二つの評価尺度との相関がより高くなることも期待できるが、そもそも頻度のような単純な特徴では役に立たないということも考えられる。

### 5.3. 評価尺度の予測実験

ここまでで定義した言語外情報と言語情報の特徴を用いて、二つの評価尺度である回答合致度と説明詳細度を予測する回帰分析の実験を行った。

予測実験は、評価尺度が付与された 2,700 文書を実測(正解)データとし、10-fold の交差検定を行って評価した。二つの評価尺度の予測精度は、回帰式による予測値と実測値との重相関係数で測ることができる。

実験では、箇条書きの有無、表の有無、検索順位の三つのみを特徴として用いる場合 A と、それに単語数、キーワード数、クエリワード数の特徴を加えた場合 B、さらにクエリカバー率の特徴を加えた場合 C の 3 通りにて評価を行う。

表 3 は、予測値と実測値との重相関係数(セル内上)および平均絶対誤差(セル内下)の値をそれぞれの実験条件で求めたものである。

表 3：予測値と実測値との重相関係数と平均絶対誤差

	A	B	C
回答合致度	0.571 (0.247)	0.576 (0.245)	0.594 (0.241)
説明詳細度	0.561 (0.212)	0.573 (0.209)	0.577 (0.209)

実験結果は、全体的にどの場合でも 0.56~0.6 程度の相関を示しており、回帰式から評価尺度がある程度予測可能であることを示した。ベースラインとしての箇条書きと評価尺度との相関係数(表 1) 0.5~0.52 と比較しても、予測精度は上がっていると言える。

また平均して 0.2~0.25 程度の実測値との誤差があるという結果が出たが、もともと実測値が(0, 0.5, 1)の 3 値なので、たとえば予測値も 3 値に正規化してしまえば見た目の誤差は減少すると予想される。

逆に、条件 A に対して言語情報の属性を加えた場合 B、C でも、飛躍的な精度向上が見られなかった。これらのことから、二つの評価尺度を予測するには、それぞれの特徴の抽出方法を見直すことと、別の観点での特徴を加えることが必要であると考えられる。

次に、実際にこの予測結果を検索画面に反映した場合、見た目上どのような印象になるか調べた。

一つのクエリに対して得られた検索結果 90 件に対して、それぞれの評価尺度の予測値が降順になるよう再ランキングしたときを想定する。比較対象として、goo ウェブ検索による通常の検索順位のランキングを用いる。それぞれの場合において得られる上位 10 件<sup>3</sup> の実際の各評価尺度の和をとり増減を調べた。

<sup>3</sup> goo 検索ではデフォルトで上位 10 件が出力されるので、これにあわせた

表 4 の値は 30 のクエリについて、再ランキング前の評価尺度の値の和(前)より再ランキング後の評価尺度の値の和(後)が向上したクエリ数と低下したクエリ数、および再ランキング前後で高い値の評価尺度を示したクエリ数を表している。

表 4：再ランキングによる変化

	推移		評価尺度の和 7~10	
	向上	低下	(前)	(後)
回答合致度	21	8	10	22
説明詳細度	25	4	1	11

30 クエリ中、7 割以上のクエリで検索結果としての評価尺度が向上し、1、2 割のクエリで低下することが示された。さらに 7~10 の高い値の評価尺度を示すクエリは再ランキング後には大幅に増えることがわかった。

つまり、回帰予測の重相関係数はそれほど高い値ではなくても、検索画面として再ランキング結果を示すことによって、良質なサービスを提供することができることがわかった。

## 6. 誤り分析

5.3 節の実験における回答合致度の予測を例にとり、残差の絶対値が大きかった事例を誤り分析の対象として検証した。

### 6.1. 予測が外れた例

回帰分析によって回答合致度が 1 に近いと予測した文書が、実際は回答合致度が 0 か非常に小さかった(つまり質問とはほとんど関係なかった)いくつかのケースを抽出した。

誤り原因としてもっとも目立ったのは、別の手順の説明の一部分に、クエリ単語の一部あるいは全部が出現してしまっているケースである。表 5 に、クエリ、誤って予測された文書の実際の内容、実測値、および予測値を示して、事例の一部を列挙する。

表 5：予測が大きく外れた例

クエリ	文書の内容	実測値	予測値
パスポート申請方法	ビザの「申請方法」に「パスポート」が必要との記述	0	0.96
犬 トイレしつけ方	猫の「トイレのしつけ方」の説明に	0	0.92
たけのこ アク抜き方法	栗きんとんのレシビ中に栗の「アク抜き」の方法を説明	0	0.78

改善策として、「犬」と「トイレ」、「たけのこ」と「アク抜き」のような重要なクエリ単語間の関係を認識し

た上で、その関係を崩すような「猫」や「栗」のような証拠を捉えたらペナルティを与えるとといった、深い言語処理が必要である。

予測が当たるか、外れるかは、方法説明の検索結果の印象に直接影響するので、このような予測外れの誤りを減らすように改良する（言い換えれば適合率を上げる）ことは最重要課題である。

## 6.2. 予測できなかった例

一方、回帰分析によって回答合致度が低いと予測した文書が、実際は回答合致度が1か、1に近い値だったケースについて検証してみる。

これらのケースの傾向として、箇条書きなどの文章構造はなく、ブログやコラムのような記事として淡々と文章で方法説明がなされているボタンが多く見受けられた。表6にそのような実例の一部を挙げる。

表6：予測できなかった例

クエリ	文書の内容	実測値	予測値
しゃっくり止め方	ブロッグー記事に対するコメント欄に多くのしゃっくりの止め方が含まれる	1	0.2
犬 トイレしつけ方	犬のトイレのしつけ方が体験談的に文章で書かれている	1	0.23

注目すべき点として、予想できなかった例の中には、クエリ単語が HTML タグの "b" や "strong", あるいは "table" タグによる装飾で強調されていることがあった。このような本稿の分析で対象としていたタイトルや見出し以外に、クエリワードの強調部分を広く認識できるようになれば、回帰分析の予測精度が向上することが期待できる。

また他の傾向として、本稿では特別扱いとしなかった写真や図などの画像を併用しながら、順を追って文章で方法の説明がされているケースもあった。図解つきの方法説明が役に立つことは直感的にも理解しやすいので、この特徴も考慮に入れるべきであろう。

## 7. まとめと課題

方法説明に特化した検索に実現に向けて、質問に対しての回答らしさの尺度である2種類の評価尺度を導入し、この評価尺度と文書内の言語外情報や言語情報の特徴との相関分析、および二つの評価尺度の回帰予測実験を行った。これによって以下のような知見が得られた。

- ・二つの評価尺度は回答らしさの妥当な指標である
- ・箇条書きの有無やクエリカバー率（見出し内のクエリ単語の出現率の最大値）といった特徴は、評価尺度の予測に重要な役割を果たす

・二つの評価尺度の回帰分析はある程度の予測能力を示し、この回帰分析によって実用に耐えうる方法説明検索が実現可能である

さらに回答合致度の回帰予測実験で、誤り事例からいくつかの課題点を抽出した。例えば、検索結果から方法らしい文書を厳選するためには、より深い言語処理を行うなどの新しい枠組みが必要であることがわかった。

我々が目指す方法検索を実現する上では、人手で抽出していた特徴や、より詳しい文書構造の自動抽出が、大きな課題となってくる。具体的には以下のような課題の解決が必要と考えている。

**箇条書きの自動特定：**人手によっていた箇条書き構造の抽出を自動化する必要がある。HTML タグによる箇条書きは当然として、テキスト情報のみによる箇条項目の自動特定まで可能になることが望ましい。

**クエリワードの強調部分の自動特定：**HTML タグを駆使したり、記号や罫線といったテキスト文字のみで強調したりして、見た目としての語句の強調部分を自動特定できるとよい。

**説明範囲の特定：**正確な特徴量の測定や予測能力向上のため、文書中で質問に回答している説明部分の範囲が自動特定できるとよい。

今後は文書構造の自動抽出方法の検討と、深い言語解析から得られる有効な特徴量の検討を行い、方法説明に特化した検索サービスの実現を目指す予定である。

## 文 献

- [1] J. Fukumoto, T. Kato, and F. Masui, "Question Answering Challenge for Five ranked answers and List answers - Overview of NTCIR4 QAC2 Subtask 1 and 2 -," Working Notes of the 4th NTCIR Workshop Meeting (NTCIR4), pp.283-290, Tokyo, Japan, June 2004.
- [2] J. Lin, and B. Katz, "Question Answering from the Web Using Knowledge Annotation and Knowledge Mining Techniques," Proc. of 12th Intl. Conf. on Information and Knowledge Management (CIKM 2003), pp.116-123, New Orleans, Louisiana, Nov.2003.
- [3] H. Isozaki, "NTT's Question Answering System for NTCIR QAC2," Working Notes of the 4th NTCIR Workshop Meeting (NTCIR4), pp. 326-332, Tokyo, Japan, June 2004.
- [4] 森本 格行, 福本 淳一, "Why 型質問に対する回答抽出," 第10回言文処年次大会, pp.293-296, March 2004.
- [5] 林良彦, "問題解決知識を求める自然言語メッセージの分析," 第11回言文処年次大会, pp.576-579, March 2005.
- [6] N. Tomuro and S. L. Lytinen, Retrieval Models and Q and A Learning With FAQ Files, in New Directions in Question Answering, ed. M. T. Maybury, Chapter 14, AAAI Press / The MIT Press, Menlo Park, Cambridge and London, 2004.