

Webからの上位概念抽出による比較可能概念の発見

岡田 崇[†] 湯本 高行^{††} 新居 学^{††} 高橋 豊^{††}

[†] 兵庫県立大学工学部

〒 671-2280 兵庫県姫路市書写 2167

^{††} 兵庫県立大学院工学研究科

〒 671-2280 兵庫県姫路市書写 2167

E-mail: †eo05p023@steng.u-hyogo.ac.jp, ††{yumoto,nii,takahasi}@eng.u-hyogo.ac.jp

あらまし 本論文では Web から、ユーザによって与えられた単語と比較可能な概念を発見する手法を提案する。これまで、多くの同位語獲得手法が提案されてきたが、本論文では HTML 文書のナビゲーションメニューに注目し、共通の上位概念をもつ語を同位語として取得する手法を提案する。ナビゲーションメニューとはパンくずリストや簡条書きされたリンクのことで、利用者がサイト内で現在の位置を見失わないようにし、サイト内を容易にナビゲートするものである。このような HTML タグにより与えられる HTML 文書の構造を利用し、同位語を発見する。提案手法ではパンくずリストから上位概念及びそのリンク先の Web ページを収集し、そのページ内に含まれる簡条書きから上位概念についての下位概念を発見することで同位語を取得する。

キーワード パンくずリスト, ナビゲーションメニュー, 系列パターンマイニング

Discovering Comparable Concept by Extracting Superordinate Concept from Web

Takashi OKADA[†], Takayuki YUMOTO^{††}, Manabu NII^{††}, and Yutaka TAKAHASHI^{††}

[†] School of Engineering, University of Hyogo

2167 Syosya, Himeji, Hyogo 671-2280, Japan

^{††} Graduate School of Engineering, University of Hyogo

2167 Syosya, Himeji, Hyogo 671-2280, Japan

E-mail: †eo05p023@steng.u-hyogo.ac.jp, ††{yumoto,nii,takahasi}@eng.u-hyogo.ac.jp

Abstract In this paper, we propose a method for discovering comparable concept of a given keyword from Web. In our method, we extract coordinate terms, which have common superordinate concept, from a navigation menu. A navigation menu consists of a topic path and a itemization of categories, etc. It helps users to navigate through the Web site easily. We extract superordinate concepts and the corresponding Web pages from topic paths. From itemization in these Web pages, we find subordinate concepts about the superordinate concept.

Key words topic path, navigation menu, sequential pattern mining

1. はじめに

近年、商品やサービスを購入するためにインターネットを利用して情報を収集する人が増加している。これにはインターネット上で商品を購入することができるショッピングサイトの増加やネット普及率の増加などが考えられる。ネットショッピングには時間・場所を問わずに商品を購入できるといったメリットがあり、利用者数も年々増加傾向にある。インターネットを利用して商品の情報を調べるための手段の一つとして検索エンジンを利用することがあげられる。検索エンジンではキー

ワードを入力することによってインターネットで公開されている情報を検索することができる。ユーザのニーズに合った商品が見つければその商品を購入することになるが、ユーザは可能な限り多くの商品の中から最適な商品を選択したいと考える。しかし、入力したキーワードと比較すべき商品を検索することはユーザ自身がある程度知識を持っていないと難しい。なぜならユーザ自身が知らない商品については具体的なキーワードを入力することができないからである。

そこで本研究では、ユーザが入力したキーワードと比較可能な商品を提示することで、より良い商品の発見を支援する。例

例えば、ディスプレイを購入したいが「液晶ディスプレイ」しか知らないような場合、このユーザに「プラズマディスプレイ」、「CRT ディスプレイ」といった同位語を提示して比較対象を思い起こさせるサービスが考えられる。本論文では HTML 文書より、ユーザによって与えられた単語の同位語を発見する手法を提案する。従来より、多くの同位語獲得手法が提案されてきたが、本論文では HTML 文書のナビゲーションメニューから共通の上位概念をもつ語を同位語として取得する。ナビゲーションメニューとは、3 節で述べるパンくずリストや箇条書きされたリンクのことで、利用者がサイト内で現在の位置を見失わないようにし、ナビゲーションを助けるために使用される。また、上位概念から同位語を発見することでどのような観点から比較の対象になり得るか、根拠として上位概念を用いる。上位概念から見た同位語の例を図 1 に示す。ユーザが入力したキーワードを液晶ディスプレイとし、上位概念を出力装置とした場合、出力装置という観点から液晶ディスプレイとプリンタは同位語になり得るが、上位概念をディスプレイとしたとき、液晶ディスプレイとプリンタは同位語ではないと考えられる。この様に上位概念の取り方次第で同位語が異なることに注目し、比較可能概念を発見する。

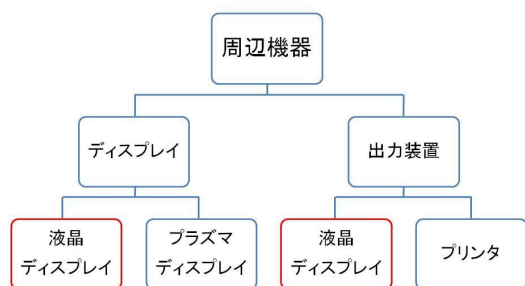


図 1 上位概念から見た同位語の例

2. 関連研究

大島らは、同位語が並列助詞「や」で接続されることが多いことに着目し、同位語発見手法を提案している [1]。クエリ Q と並列助詞「や」で接続されている語、つまり「 Q や」の直後と「や Q 」の直前に出現する語を取得し、「や」の両側に出現する語を同位語とみなす。同位語とは共通する上位概念をもつような語のことである。同位語は共通した属性を持つ場合が多いため、本研究では同位語から比較可能概念を発見する。

新里らは、HTML タグにより与えられる HTML 文書の構造、大量のテキストから獲得した名詞と動詞の係り受け関係、従来より情報検索などで使われてきた df , idf などの統計量、ヒューリスティックなルールの 4 つの要素を組み合わせ、HTML 文書より単語間の上位下位関係を自動獲得する手法を提案している [2]。HTML タグにより与えられる HTML 文書の構造とは、箇条書きやリストボックス、テーブルのセルなどの繰り返し要素のことである。本研究ではこのような繰り返し要素の他に上

位概念のページへのリンクを用いて同位語を発見する手法を提案する。この手法を用いることによって製品の型番などの特定の単語についての同位語を発見することができると考えられる。

HTML 文書内の変則的な表現をしている箇条書きを発見するために系列パターンマイニングを利用する。系列パターンマイニングとは、頻出する部分系列パターンを高速に抽出する方法の総称である。系列パターンマイニングツールの 1 つとして PrefixSpan [3] があり、容易に部分系列パターンを抽出することができる。PrefixSpan は Pei らによって提案された系列パターンマイニングのアルゴリズムである [4]。PrefixSpan は Prefix Projection という射影方法とその射影によって生成される射影データベースを用い、深さ優先でマイニングを行う。Prefix Projection とは、射影元のシーケンスから射影対象のシーケンスより後ろに存在するアイテムからなるシーケンスのみを抽出する射影である。本手法では HTML を木構造ではなくテキストストリームと見なすことで系列パターンマイニングを適用する。

3. 手 法

本研究では、ユーザの入力としてクエリを 1 つ利用し、それと比較可能な概念の発見を目的とする。クエリの入力から比較可能概念の発見までの処理の流れを以下に示す。また、処理の概念図を図 2 に示す。

(1) Web 上からのページの収集

まず、検索エンジンを用いてパンくずリスト中にクエリが出現する Web ページを収集する。パンくずリストとは「家電 > ディスプレイ > 液晶テレビ」のような階層構造の上位ページへのリンクのリストである。図 3 にパンくずリストの例を示す。

(2) パンくずリストからの上位概念の発見

(1) で収集した Web ページからパンくずリストを抽出し、上位概念とそのリンク先のページを取得する。この操作は図 2 中の (2) の矢印に相当する。抽出した各語には適切な上位概念を選ぶ基準となるスコアを与える。

(3) 系列パターンマイニングによる同位語候補の発見

(2) で取得した上位概念のリンク先のページの HTML のタグから、箇条書きやリストのような繰り返しパターンを発見し、同位語候補を取得する。図 3 に箇条書きの例を示す。この操作は図 2 中の (3) の矢印に相当する。これは、HTML 文書中に現れる箇条書きやリストなどの繰り返し要素に含まれる語は意味的に類似していると考えられるからである。HTML 文書の HTML のタグから箇条書き部分を発見するために系列パターンマイニングを用いる。

(4) ノイズの排除

(3) で収集した語の中には同位語とするには適切でない語が多く存在する。そのような語をノイズとして排除する。

(5) 比較可能概念の決定

収集した同位語候補から、出現するサイト数が閾値 θ_s 以上であるものを比較可能概念とする。

閾値 θ_s はサイト集合を $S = (s_1, s_2, \dots, s_n)$ としたとき、(1) 式で与えられる。

$$\theta_s = \left\lceil \frac{|S| + 1}{2} \right\rceil \quad (1)$$

3.1 Web 上からのページの収集

検索エンジンを用いてパンくずリスト中にクエリが出現する Web ページを収集する。サイト集合を $S = (s_1, s_2, \dots, s_n)$ とし、各サイト s の検索エンジンを使用し、クエリ q を用いて検索を行う。検索結果として表示される Web ページ p_1 にはクエリに関連する商品が掲載されている Web ページ ($p_2, \dots, p_n$) へのナビゲーションメニューが含まれている。このような検索結果で表示される Web ページ p_1 や商品が掲載されている Web ページ ($p_2, \dots, p_n$) にはパンくずリストが出現すると考えられるので、このような Web ページ集合 $P = (p_1, p_2, \dots, p_n)$ を収集する。

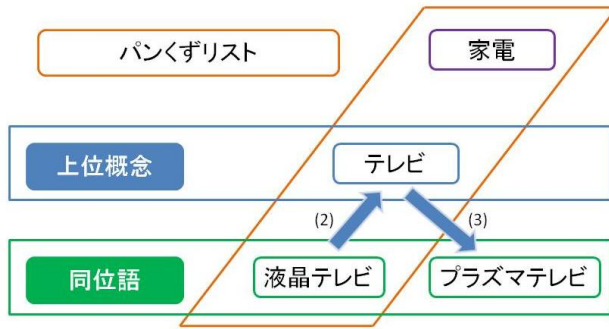


図2 処理の概念図
パンくずリスト



図3 ナビゲーションメニューの例

3.2 パンくずリストからの上位概念の発見

3.1 節で収集した Web ページからパンくずリストを抽出し、クエリの上位概念とそのリンク先のページを取得する。まず、Web ページからパンくずリストを抽出する。パンくずリストを抽出する際は、Web ページの HTML からタグを取り除き、クエリを含むパンくずリストを抽出する。次に、得られたパンくずリストから上位概念とそのリンク先のページを取得する。パンくずリストの上位ページへのリンクのテキストを上位概念とする。例えば、クエリが A 、パンくずリストが「 $C > B > A$ 」であれば、 B と C は A の上位概念である。このように、入力された語より階層が上の概念とそのリンク先の Web ページを収集する。また、抽出した各語には適切な上位概念を選ぶ基準と

なるスコアを与える。本提案手法では、適切な上位概念を決める方法として、対象となる語に隣接する語から上位概念を決定する方法とスコアによって上位概念を決定する方法を提案する。

3.2.1 対象となる語に隣接する語からの上位概念の決定

パンくずリスト ($B > A$) 内の語 B と A の階層間距離 $d(B, A)$ は (2) 式を定義する。

$$d(B, A) = 1 \quad (2)$$

また、パンくずリスト ($B > \dots > x > \dots > A$) 内のある語 x とその下位概念 A との階層間距離 $d(x, A)$ 、上位概念 B とある語 x との階層間距離 $d(B, x)$ とすると、 A と B との階層間距離 $d(B, A)$ は (3) 式を定義する。

$$d(B, A) = d(x, A) + d(B, x) \quad (3)$$

例えば対象となる語が A 、パンくずリストが「 $C > B > A$ 」であれば、上位概念 B と A の階層間距離 $d(B, A)$ は 1 であり、上位概念 C と A との階層間距離 $d(C, A)$ は 2 である。対象となる語 t と上位概念 w の階層間距離 $d(t, w)$ が 1 である上位概念 w を t に隣接する上位概念とする。例えばユーザに与えられた語が A 、パンくずリストが「 $C > B > A$ 」であれば、 B が隣接する上位概念である。このように t に隣接する上位概念を w 、サイト集合を $S = (s_1, s_2, \dots, s_n)$ としたとき、スコア $score(w)$ は (4) 式で与えられる。

$$score(w) = \sum_{s \in S} exist(w, s) \quad (4)$$

$$exist(w, s) = \begin{cases} 1 & \text{if } (w \in s) \\ 0 & \text{if } (w \notin s) \end{cases} \quad (5)$$

$exist(w, s)$ はサイト s について t の上位概念 w が存在する ($w \in s$) ならば 1、そうでなければ ($w \notin s$) 0 の値をとる。(4) 式で与えられるスコア $score(w)$ が閾値 θ_{next} 以上ならば、その上位概念 w を適切な上位概念とする。

$$score(w) \geq \theta_{next} \quad (6)$$

閾値 θ_{next} はサイト集合を S としたとき (7) 式で与えられる。

$$\theta_{next} = \left\lceil \frac{|S| + 1}{2} \right\rceil \quad (7)$$

3.2.2 スコアによる上位概念の決定

パンくずリスト内での対象となる語 t と上位概念の階層が近いほど関連が深いと考え、スコア $score(w)$ はサイト集合を $S = (s_1, s_2, \dots, s_n)$ 、上位概念を w としたとき (8) 式で与えられる。

$$score(w) = \sum_{s \in S} \frac{1}{Num(t, s)} \sum_{t \in T_s} r^{d-1} f(w, t) \quad (8)$$

d : 階層間距離

r : 対数減衰率

$$f(t, s) = \begin{cases} 1 & \text{if}(w \in t) \\ 0 & \text{if}(w \notin t) \end{cases} \quad (9)$$

ここで階層間距離 d は (2), (3) で定義したものである。また, 対数減衰率 r は $r < 1$ の条件を満たすものとする。 $f(t, w)$ は サイト s についてのパンくずリスト集合 $T_s = (t_1, t_2, \dots, t_m)$ の内, パンくずリスト t に上位概念 w が存在する ($w \in t$) ならば 1, そうでなければ ($w \notin t$) 0 の値をとる。 サイト s について, 上位概念 w が存在するパンくずリスト t の数 $Num(t, s)$ で正規化する。 (8) 式で与えられるスコア $score(w)$ が閾値 θ_{next} 以上ならば, その上位概念 w を適切な上位概念とする。 閾値 θ_{next} をサイト集合を S としたとき (7) 式で与える。

3.3 系列パターンマイニングによる同位語候補の発見

3.2 節で取得した上位概念のリンク先のページから同位語候補を発見する。 パンくずリストから得た上位概念のリンク先のページの HTML 文書から, 入力した語が含まれる繰り返し要素を発見することで入力した語の同位語候補を得る。 HTML 文書中に現れる変則的な表現をしている箇条書きやリストなどの繰り返し要素に含まれる語は意味的に類似していると考えられる。 HTML 文書から変則的な表現をしている箇条書き部分を発見するために系列パターンマイニングを用いる。 例えば, 図 4 のように li タグで箇条書きされた部分を発見する方法の手順を以下に示す。

(1) 中間コードの生成

繰り返しパターンを発見するための前処理として必要な中間コードの生成について述べる。 HTML のタグは基本的にタグ名, 属性名, 属性値, 要素値で構成され, 以下のように記述される。

<タグ名 属性=属性値> 要素値 </タグ名>

また, 画像を表示するなどの用途で用いられる要素値がない空要素は, 以下のように記述される。

<タグ名 属性=属性値>

タグが同じ順序で出現する部分を繰り返し要素として発見する。 このとき, 不要な属性および属性値を除去し, 要素値を一旦同じ文字列に置き換える。 繰り返し要素を発見するために HTML のタグを置き換えたものを中間コードと呼ぶ。 中間コードの例を以下に示す。

[小文字のタグ名_要素値を置き換えた文字列_大文字のタグ名]
[小文字のタグ名]

具体的には, 図 4 の HTML のタグを中間コードに書き換えると図 5 のようになる。 このとき, 要素値を置き換えた文字列は “Text” とする。 このように置き換えることで繰り返し要素を発見することができる。

(2) 繰り返しパターンの発見

系列パターンマイニングツール PrefixSpan を用いて繰り返しパターンを発見する。 要素値を含むタグの開始から終了までの文字列のみを繰り返しパターンとし, PrefixSpan を用いて中間コードから繰り返し要素を発見する。 PrefixSpan に入力するデータは, タグの開始を “[小文字のタグ名_”, タグの終了を “_大文字のタグ名]”, 要素値を “Text”, 空要素を “[小文字の

タグ名]” と表記し, 4 種類の要素に分割したものとする。 例えば, 図 5 の中間コードから PrefixSpan に入力されるデータは (“[ul_”, “[li_”, “Text” , ..., “_UL]”) となる。 また, 図 5 より得られる繰り返しパターンは, タグの開始 “[li_”, 要素値 “Text”, タグの終了 “_LI]” の要素からなる “[li_Text_LI]” となる。

(3) 同位語候補の発見

繰り返しパターンから同位語候補を発見する手順について述べる。 中間コードに書き換えた際に要素値を一旦別の文字列に置き換えたので, 元の文字列に戻す必要がある。 まず, 繰り返しパターンが出現する箇所を発見し要素値に対応する文字列を特定する。 例えば, 図 5 の “Text” 部分に対応する語を図 4 から抽出すると

- プラズマテレビ
- 液晶テレビ
- CRT ディスプレイ

が得られ, これらは同位語候補となる。

```
<ul>
  <li> プラズマテレビ </li>
  <li> 液晶テレビ </li>
  <li> CRT ディスプレイ </li>
</ul>
```

図 4 HTML 文書内の箇条書き部分

```
[ul_
  [li_Text_LI]
  [li_Text_LI]
  [li_Text_LI]
 UL]
```

図 5 図 4 の中間コード

3.4 ノイズの排除

3.3 節で収集した同位語候補からノイズを排除する。 ストップワードとして, どのようなクエリを与えても同位語にならないと判断できる語を削除する。 なお, 本稿では表 1 に示す語をストップワードとして用いる。

表 1 ストップワード

その他, ホーム, トップ, カテゴリ, ランキング
一覧, 詳細, 検索, 商品

4. 実 験

3. 節で提案した手法を用いて, 実験を行った。 入力として「液晶」, 「プラズマ」, 「一眼レフカメラ」, 「ピアノ」を与え, サイト集合 $S = (Amazon.co.jp [5], 価格.com [6], 楽天市場 [7], yourguide [8], bidders [9])$ に対してそれぞれの語について比較可能概念を求めた。

4.1 パンくずリストからの上位概念の発見実験

4.1.1 対象となる語に隣接する語からの上位概念の決定
適切な上位概念を決定するために, パンくずリストから対象となる語に隣接する語を求め, (4) 式により与えられるスコア

表 2 4.1.1 についての上位概念のスコア計算結果

(a)「液晶」の上位概念の候補

上位概念	スコア $score(w)$
テレビ	3
アクセサリ	2
ディスプレイ	1
携帯電話・IP フォンアクセサリ	1
携帯電話・IP フォン	1
モニタ	1
家電	1
パソコン	1
本体周辺機器 (DS Lite 用)	1
ディスプレイ・モニター (その他)	1
周辺機器	1
入力機器・アクセサリ	1
ディスプレイ・モニタ	1
デジカメアクセサリ	1

(b)「プラズマ」の上位概念の候補

上位概念	スコア $score(w)$
テレビ	4
物理学	1
家電	1
プラズマテレビ	1

(c)「一眼レフカメラ」の上位概念の候補

上位概念	スコア $score(w)$
カメラ	2
ニコン	1
カメラ・ビデオ	1
本体	1
カメラ・光学機器	1
デジタルカメラ	1

(d)「ピアノ」の上位概念の候補

上位概念	スコア $score(w)$
楽器	4
鍵盤楽器	2
ピアノ・電子ピアノ・キーボード	1
ジャズ	1
ポピュラー	1
キッズ・ベビーブランド	1
クラシック	1
ヤマハ	1
ピアノ	1
楽器玩具	1
ピアノ・鍵盤楽器	1

$score(w)$ を算出した。それぞれの上位概念のスコア $score(w)$ を計算し、表 2 にまとめる。表 2 より、「液晶」の上位概念は「テレビ」、「プラズマ」の上位概念は「テレビ」、「ピアノ」の上位概念は「楽器」が適切であるという結果が得られ、「一眼レフカメラ」については適切な上位概念は得られなかった。また、「一眼レフカメラ」の上位概念として「ニコン」、「ピアノ」の上位概念として「ヤマハ」などのように、パンくずリストにメー

カーなどの属性が出現することがわかった。

4.1.2 スコアによる上位概念の決定

適切な上位概念を決定するために、パンくずリストから対象となる語の上位概念を全て求め、(8) 式により与えられるスコア $score(w)$ を算出した。それぞれの場合について計算した上位概念のスコア $score(w)$ の上位 6 件を表 3 にまとめる。表 3

表 3 4.1.2 についての上位概念のスコア計算結果

(a)「液晶」の上位概念候補

上位概念	スコア $score(w)$
テレビ	3.0
周辺機器	1.3
モニタ	1.0
ディスプレイ	1.0
ディスプレイ・モニタ	1.0
携帯電話	0.5

(b)「プラズマ」の上位概念候補

上位概念	スコア $score(w)$
テレビ	3.9
家電・AV・カメラ	1.2
物理学	1.0
家電	1.0
テレビ・レコーダー	0.5
科学・テクノロジー	0.5

(c)「一眼レフカメラ」の上位概念候補

上位概念	スコア $score(w)$
カメラ	2.0
家電・AV・カメラ	1.0
カメラ・ビデオ	1.0
デジタルカメラ	1.0
カメラ・光学機器	1.0
ニコン	1.0

(d)「ピアノ」の上位概念候補

上位概念	スコア $score(w)$
楽器	3.3
鍵盤楽器	2.0
ピアノ・鍵盤楽器	1.0
クラシック	1.0
ジャズ	1.0
ヤマハ	1.0

より、「液晶」の上位概念は「テレビ」、「プラズマ」の上位概念は「テレビ」、「ピアノ」の上位概念は「楽器」が適切であるという結果が得られ、「一眼レフカメラ」については適切な上位概念は得られなかった。

4.1.3 考 察

実験で得られたパンくずリストには「一眼レフカメラ」の上位概念として「ニコン」、「ピアノ」の上位概念として「ヤマハ」などのように、属性が出現した。このような語は適切な上位概念では無いため、取り除く必要がある。また、このようなパンくずリスト内に出現する属性を利用することで比較可能概念の発見に役立てることが今後の課題であると考えられる。「一眼

レフカメラ」の上位概念が得られなかった理由として、上位概念であると考えられる「カメラ」は「カメラ・～」といった複数の語からなる上位概念が多く、語の組み合わせが異なるためサイト毎のスコアは高いが、サイト集合としてのスコアは低くなってしまったと考えられる。

4.2 系列パターンマイニングによる同位語候補発見実験

PrefixSpan を用いて適切な上位概念とそのリンク先の Web ページから同位語候補を発見する実験を行った。2つ以上のサイトで出現した語を表4にまとめる。ただし、表1のストップワードはノイズとして排除した。

表4 系列パターンマイニングによる同位語候補発見実験の結果

入力	上位概念	同位語
液晶	テレビ	ポータブルテレビ プラズマテレビ 有機 EL テレビ
プラズマ	テレビ	ポータブルテレビ 液晶テレビ 有機 EL テレビ
ピアノ	楽器	ベース 音楽ソフトウェア 弦楽器 マイク 打楽器 アコースティックギター DJ 楽器 楽譜 管楽器 和楽器 オルガン

5. まとめと今後の課題

本論文ではパンくずリストや箇条書きなどのナビゲーションメニューを用い、HTML の構造を利用した比較可能概念発見手法の提案及び実験を行った結果について述べた。本手法はパンくずリストから上位概念及びそのリンク先の Web ページを収集し、そのページ内に含まれる箇条書きなどの変則的な繰り返しパターンから上位概念についての下位概念を発見することで同位語候補を取得し、その同位語候補から比較可能概念となる語を発見するといった手法である。実験として、提案手法により「液晶」、「プラズマ」、「一眼レフカメラ」、「ピアノ」についての比較可能概念を求めた。実験ではパンくずリスト内の対象となる語に隣接する語から上位概念を決定する方法とスコアから上位概念を決定する方法の2通りの方法から比較可能概念を求めた。複数の語からなる上位概念の扱いやパンくずリストに出現する属性などの語を比較可能概念の発見に役立てることが今後の課題であると考えられる。

謝辞 本研究の一部は、平成20年度科研費基盤研究(B)(2)「ユーザの潜在的意図を用いたレス・コンシャス情報検索基盤の構築」(課題番号:20300039)によるものです。ここに記して謝意を表すものとします。

文 献

- [1] 大島 裕明, 山口 雅史, 小山 聡, 田中 克己. Web 検索とクエリログを併用した同位語発見手法. 日本データベース学会 Letters, Vol.5, No.4, pp.37-40, 2007.
- [2] 新里圭司, 鳥澤健太郎. HTML 文書からの単語間の上位下位関係の自動獲得. 自然言語処理 Vol.1 No.1., pp.95-102, 2003.
- [3] PrefixSpan. <http://prefixspan-rel.sourceforge.jp/>.
- [4] J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal and M-C. Hsu. PrefixSpan: Mining Sequential Patterns Efficiently by PrefixProjected Pattern Growth. In. Proc. 2001 Int. Conf. Data Engineering (ICDE'01), pages 215-224, Heidelberg, Germany, April 2001.
- [5] Amazon.co.jp. <http://www.amazon.co.jp/>.
- [6] 価格.com. <http://kakaku.com/>.
- [7] 楽天市場. <http://www.bidders.co.jp/>.
- [8] yourguide. <http://shopping.yourguide.co.jp/>.
- [9] bidders. <http://www.bidders.co.jp/>.