

「係り受け句」による検索語の展開とそれに基づく文献選択

藤井佐和子 阪口哲男 杉本重雄 田畑孝一

図書館情報大学

〒 305 つくば市春日 1-2

Tel: 0298-52-0511 Fax: 0298-52-4326

E-mail: {sawako, saka, sugimoto, tabata}@ulis.ac.jp

概要

本論文では、全文検索システムを利用するエンドユーザに対して、検索要求の明確化を支援するアプローチを提案する。その手法は、文献集合(日本語)からエンドユーザが指定した検索語とその係り受け関係から検索語と関連性が高いと考えられる語群とを抽出し、生成した「係り受け句」の一覧の提示による文献選択を試みるものである。「係り受け句」は検索語が単独で持つ概念をより多様に、より深化させた概念を表現する句である。また、「係り受け句」は原文の表現をほぼそのまま反映しているので、各「係り受け句」はそれが抽出された文献とその抽出個所をほぼ特定することになる。従って、「係り受け句」一覧はエンドユーザに対して、その一覧のブラウジングによる検索要求の明確化の手がかりと、その一覧から興味をひかれた「係り受け句」の選択による文献選択とを同時に提供できる場となっている。このような文献選択方法では、複雑な検索式の作成が不要なため、検索要求が不明確なあるいは検索システムに不慣れたエンドユーザでもより容易な文献選択が可能になると考えられる。

本手法の有効性を検証するために、作成したプロトタイプシステムによる実験とアンケートによる評価を32人に依頼した。選択した「係り受け句」が抽出された文献個所を被験者に一読してもらったところ、その72%は「係り受け句」を見たときに期待した、またはそれ以上の知見が得られるものであったと評価された。

キーワード

係り受け、ブラウジング、検索語、全文検索、句

Query Term Expansion by "Dependency Phrase" and its application to Document Retrieval Support

Sawako Fujii, Tetsuo Sakaguchi, Shigeo Sugimoto, Koichi Tabata
University of Library and Information Science
Address: 1-2 Kasuga, Tsukuba, Ibaraki, 305, JAPAN
Phone: +81-298-52-0511 Fax: +81-298-52-4326
E-mail: {sawako, saka, sugimoto, tabata}@ulis.ac.jp

Abstract

In this paper, we propose a new method of the document selection support by the query term expansion that was named "Dependency Phrase". Usually, a user retrieves document by using a query term, and has to choose a document fulfilling his/her request. But it is difficult to find a relevant document from the query results, and a second query is often necessary. This method allows a query term expansion by extracting informations based on Dependency relationship. Each "Dependency Phrase" extracted from the document is displayed, and to choose a "Dependency Phrase" almost equals to choose a document itself. Browsing a list of "Dependency Phrase" related to the query term helps the end-user to choose a retrieval document, even if his/her request is not clearly formulated at first.

In this paper, we present the evaluation of this method. The evaluation was made by way of a questionnaire, answered by 32 persons who actually used the prototype system. It was found that 72% of the retrieved documents were profitable/relevant because of the use of the "Dependency Phrase".

Keywords

Dependency, Browsing, Query term, full-text Search, Phrase

1-1 はじめに

近年のネットワークの普及・一般化により、検索システムの主体も検索の専門家からエンドユーザへと移行している。こうした背景のもと、エンドユーザの検索行動に対して、検索要求が不明確なまま検索を始める傾向 [2] や検索での試行錯誤の多さ [4] が指摘されている。通常、ユーザは検索要求を言語化した検索語を入力し、その検索結果から自分の要求に合う文献を選び出さなければならない。しかし、一般的に検索結果の表示情報だけからの選択が困難であることや、検索結果集合が大きすぎるため更に絞込みが必要となることが多いという問題がある。検索語の追加による絞込みも、自らの検索要求が曖昧であったり、発想の限界という壁に阻まれた場合は有効な検索結果を得ることは難しい。この検索要求の未分化や発想の限界は、エンドユーザ自身による検索における特徴的な問題であり、自己解決が困難であるため何らかの外部支援を行うことが有効だと考えられる。

そこで、本論文では、全文検索システムを利用するエンドユーザに対して、検索要求の明確化を支援するアプローチを提案する。

1-2 研究の背景

本手法の特徴の一つは、発想支援の考え方を導入した点である。発想支援の考え方を検索で生かそうとする研究には、選択されたキーワードの履歴からユーザの興味ある分野を推論し、その興味に関連するキーワード群と文献群を提示することで文献選択を支援する研究 [1] や、ユーザの興味を類似概念群として理解し、それを知識ベースの構造化に採り入れることで、ユーザの興味を満足させる情報を提示する研究 [2] がある。一方、本研究ではユーザが指定した検索語の関連記述を原文献から抽出し、提示するという形の外部刺激により、ユーザの検索語に関する概念の展開 (発想) と検索要求の明確化を支援するものである。

また、本研究の先行研究として、文献集合 (英語) から出現頻度の高い語を中心に語からその語を含む句へ、更に句からその句を含む文へとブラウジングによる選択で順次拡大して行くことで文献集合の理解を試みる研究がある [3]。

本研究は、文献 [3] を参考として、日本語文献を対象とした名詞句のブラウジングによる文献選択を試みた。原文献から検索語の関連/派生事項を名詞句として抽出するために日本語の場合に重要だとされる係り受け情報を利用した。

2-1 「係り受け句」とは

2-1-1 「係り受け句」の有効性

本研究では、ユーザが予め指定した検索語を含む一文を文献から抽出し、更にその文から名詞句を抽出する。そのためには、検索語を含む一文から、検索語の関連/派生事項だけをその関連性に依拠してまとめ、漏れなく抽出する手法が重要となる。

単純に一文から検索語を中心にその前後の数語をそのまま検索語関連の重要な記述部分とみなして抽出する方法がある。しかし、その方法は日本語のように語順の自由度が高く、省略など多い言語では、この手法の効果はあまり期待出来ないとされている [7]。

一方、日本語には、文末以外の全ての語が自分より後に出現する他の語に係るという2項関係 (係り受け関係) の積み重ねでその構造を解釈する係り受け文法がある。この係り受け情報を用いた検索についての研究は既にいくつか行われている [6]。実際、新美によると日本語の近接度による関連性は3語以上離れた場合、急速に不確実なものとなるため、係り受け情報が有効な手段となる [5]。そこで、この係り受け情報

をもとに一文中で検索語と関連性の高い語を認識する方法を本研究では採用した。このようにして、抽出した検索語関連の記述を更に名詞句相当の形式に加工したものを本論文では「係り受け句」と呼ぶ。

プロトタイプシステム上では「係り受け句」の約 40 %が原文中で非接続の文節から、また、約 24 %が 3 語以上離れた文節から構成されていた。

2-1-2 「係り受け句」の定義

文はその文を構成する文節間の係り受け関係をもとに文節をノード・係り受け関係をアークとするツリー構造で表現できる。このツリーから検索語を含む句を係り受け情報に基づいて、以下の手順で抽出する。(本文中で使用する品詞名、品詞区分あるいは文節の単位は、後述する形態素解析ツール JUMAN と日本語構文解析システム KNP で使用されているものに準拠している)

(1) 検索語を含む文節からツリーをリーフとルートの方角にたどり、それぞれ 4 文節以内に名詞節があれば、検索語を含む文節から最も検索語に近い名詞節までの全文節を抽出する。

(2) (1) の名詞節が形式名詞もしくは副詞的名詞の場合には、独立した名詞節とはみなさないで、それに係る動詞節を含めて 1 つの名詞節とみなし、抽出する。

(3) (1) の名詞節が時相名詞もしくは数詞の場合には、その次に検索語に近い名詞節が 4 文節以内にあれば、その名詞節までの全文節を抽出する。

(4) 4 文節以内に、名詞節が無く、サ変動詞節がある場合、検索語から最も近いサ変動詞節までの全文節を抽出する。

(5) 上記の (1)~(4) のいずれにも該当しない場合には、検索語を含む文節のみを抽出する。

(6) 抽出対象となる句内の名詞節に係る形容詞節がある場合には、その形容詞節も含めて抽出する。

抽出された句を以下の手順で加工した句を「係り受け句」と呼ぶ。

(1) 句の最後の文節から名詞性接尾辞以外の付属語を削除する。

(2) 句の最後の文節が形式名詞あるいは副詞的名詞の場合には、その直前にある用言の活用形を原形にする。但し、その用言がサ変動詞の場合は形式名詞節・副詞的名詞節を削除する。

(3) 句の最後の文節がサ変動詞の場合、サ変動詞“する”を削除する。

(4) 句内の読点や句点や括弧を削除する。

従って、この例文からは<デジタル化することで出現する図書館はデジタルライブラリー><ネットワーク上に出現する図書館はデジタルライブラリー><新しい図書館はデジタルライブラリー>の 3 つの「係り受け句」が最終的に抽出されることになる。

2-2 検索語の展開と文献選択

2-2-1 「係り受け句」一覧の提示の効果

本手法では、全文献に存在するユーザが指定した検索語の総数よりも抽出される「係り受け句」の種類の方が多くなる。ユーザが検索語に持っていた固定概念の枠を超える検索語の展開(検索要求の明確化)をするために、この多種多様な「係り受け句」一覧の提示が有効な支援となることが期待される。

例文：情報や資料をデジタル化することでネットワーク上に出現する新しい図書館はデジタルライブラリーと呼ばれる。

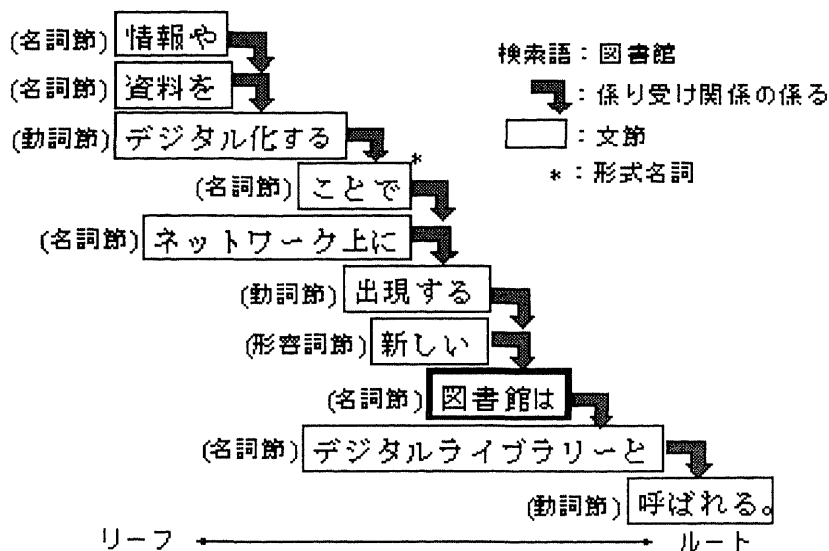


図1：KNPによる係り受け解析結果としてのツリー構造の例

2-2-2 「係り受け句」一覧からのパラグラフ表示

本手法の文献選択では、「係り受け句」を選択し、その「係り受け句」が抽出されたパラグラフを参照した上で、必要に応じて文献を選択する方法を採っている。このパラグラフだけの表示を用意したのは、読むべき場所と量を限定するためである。そうすることで必要とする情報が在るのか無いのか、ひいては最終的に文献を読むべきかどうかの判断を素早く且つ的確に行えると考えたからである。

3 プロトタイプシステムの概要

3-1 「係り受け句」リストの作成

3-1-1 原文献

検索対象とする原文献は、図書館学と情報学を中心とする全124文献とした。

3-1-2 係り受け解析

(1) 形態素解析ツールと係り受け解析ツールについて本研究では京都大学長尾研究室で開発された日本語形態素解析システムJUMANとJUMANで得られた結果をもとに解析を行う日本語構文解析システムKNPを形態素および構文解析ツールとして用いた。

(2) KNP原文献への前処理

(3) 文献内の検索対象範囲

(4) 文字列一致による検索語を含む文の抽出

- (5) JUMAN による形態素解析
- (6) 係り受け解析の精度を保つための前処理
- (7) KNP を用いた係り受け解析

この際、係り受けの可能性行列、文節間の類似行列なども表示するオプション (detail) を指定した。(注：本来、JUMAN と KNP は、一括処理が可能である。)

3-1-3 「係り受け句」の抽出・加工

- (1) 定義に準じて句となりうる全パターンを抽出
- (2) 同一文内の類似表現の削除

同一文から抽出されたパターンの内、他のパターンの部分集合となるものを省いた。この処理によって、検索語が全文献中の出現頻度が2番目に高い「図書館」の場合、抽出された句の数(3,853)を約87%(3,378)に削減できた。

- (3) 名詞句的形式に整形

3-2 文献選択とその画面

3-2-1 検索語選択のための初期画面

プロトタイプシステムはブラウザ上での利用とアンケート調査のための実験を前提として作成した。そのため自由な入力を受け付ける画面ではなく、予め用意した検索語を選択する画面を初期画面としている。ここで、検索語を選択するとその検索語を含む「係り受け句」一覧の画面が提示される。

3-2-2 K W I C形式で表示された「係り受け句」一覧

「係り受け句」選択画面は画面1のように、大小2つの表示部から構成されている。メインの表示部には「係り受け句」一覧が表示される。この一覧は、通常のK W I C形式のソートとはやや異なり、最初のソートのキーを検索語より句の後に出現する付属語を無視し、最初の自立語とした。このソート方法によって、見た目の一覧性だけでなく、内容的な一覧性も向上すると思われる。

また、ソートの第一のキーとなる自立語を同じくする「係り受け句」が5つ以上ある場合には、メインの画面上では検索語と自立語だけを表示し、別途詳細を表示することにした。この処理によって検索語：図書館の場合、抽出された「係り受け句」数(3,378)が選択画面のリスト上で約41%(1,399)に圧縮できた。

また、画面下の細長い表示部は、「係り受け句」が抽出元の一文のみを表示するために設けた。「係り受け句」は原文のある一文から抽出されるものなので、それが一文中でどのような構成要素であったかを知るために必要だと考えたためである。

3-2-3 選択された文献のパラグラフ

ある「係り受け句」をブラウザ上で選択すると、それが抽出されたもとのパラグラフだけでなく、その文献のタイトルや著者名といった書誌事項をも表示した別のブラウザが立ち上がる。(画面2参照) 更に、ユーザの利便性を考慮し、該当「係り受け句」に含まれる付属語以外の単語を赤い字で、またその前後数行は字を大きく表示している。この表示画面には、原文献へのリンクも用意した。

Netscape: 指定検索語: 「<全文>/」

ファイル 編集 表示 ジャンプ ブックマーク オプション ディレクトリ ウィンドウ ヘルプ

Back Forward Home Reload Open Print Find

場所: file:///usr/users/student/sawako/public_html/zenFrame.html

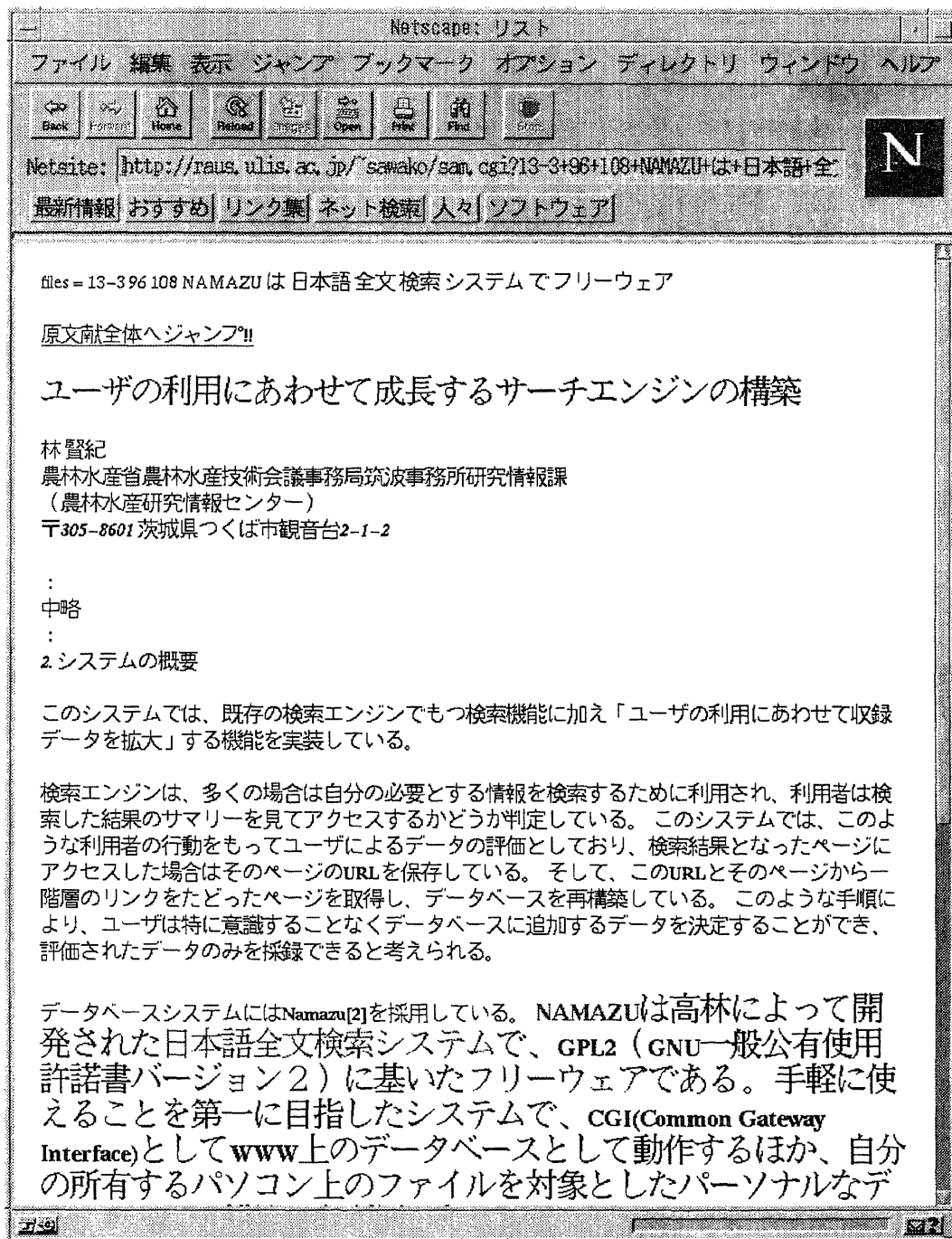
最新情報 おすすめ リンク集 ネット検索 人々 ソフトウェア

指定検索語: 「全文」のリスト

ID 1	周辺の句	一文	全文
ID 2	周辺の句	一文	
ID 3	周辺の句	一文	本稿は全文
ID 4	周辺の句	一文	イメージデータとして蓄積した学術文献全文
ID 5	周辺の句	一文	検索に加えて蓄積した学術文献全文
ID 6	周辺の句	一文	機関では蓄積した学術文献全文
			xxxxx全文 xxDBxxxxx
ID 19	周辺の句	一文	全文 FULL TEXT データベース化
ID 20	周辺の句	一文	無修正全文 OCRによる方式
ID 21	周辺の句	一文	人手を要しない無修正全文 OCRによる方式
ID 22	周辺の句	一文	本文全文 PLAIN TEXT
			xxxxx全文 xxDBxxxxx
ID 30	周辺の句	一文	論文中の全文 にあるか検討
			xxxxx全文 xx検索xxxxx
ID 32	周辺の句	一文	全文 をしている関係
ID 33	周辺の句	一文	収録対象文献の全文 としての特性
ID 34	周辺の句	一文	著作物本文の全文 になっているのでより巧緻
ID 35	周辺の句	一文	全文 及びイメージで提供

指定検索語: 「全文」の原文

画面1: 「係り受け句」の選択画面



画面 2 : 選択した「係り受け句」が抽出された原文献のパラグラフの表示画面

4 プロトタイプシステムを用いた評価実験 (被験者とアンケート)

本論文で提案した「係り受け句」による文献選択の有効性を確かめるため、アンケート調査を行った。この評価法を用いたのは、発想支援的な面の評価のためには通常の再現率や適合率といった評価尺度より、アンケートによる評価の方が適切だと考えたためである。

原文献は情報学分野の論文なので、この分野の専門的知識をある程度有する大学生・大学院生を中心とした32名を被験者とした。検索語は予め決めてある33個の内から、興味のある語を5つ以上選択してもらうように依頼した。検索語の「係り受け句」リストから選択する句の数はユーザの任意とした。(アンケートの質問項目については付録を参照のこと)

5 アンケート結果の分析とその考察

5-1 全般的な回答状況

まず、各被験者によって選択された検索語の総数は149語であったので、用意した33語から一人平均4.66語を選択したことになる。選択されなかった検索語は2語だけであった。また、無効は検索語の選択段階で2語(全149語中)・句の選択段階で9句(全595句中)あった。

5-2 結果の分析とその考察

まず、質問<「係り受け句」を選択した際に、期待した内容が実際に読んだ原文献のパラグラフで得られたかどうか>に対する回答では、グラフ1で示されたような結果が得られた。

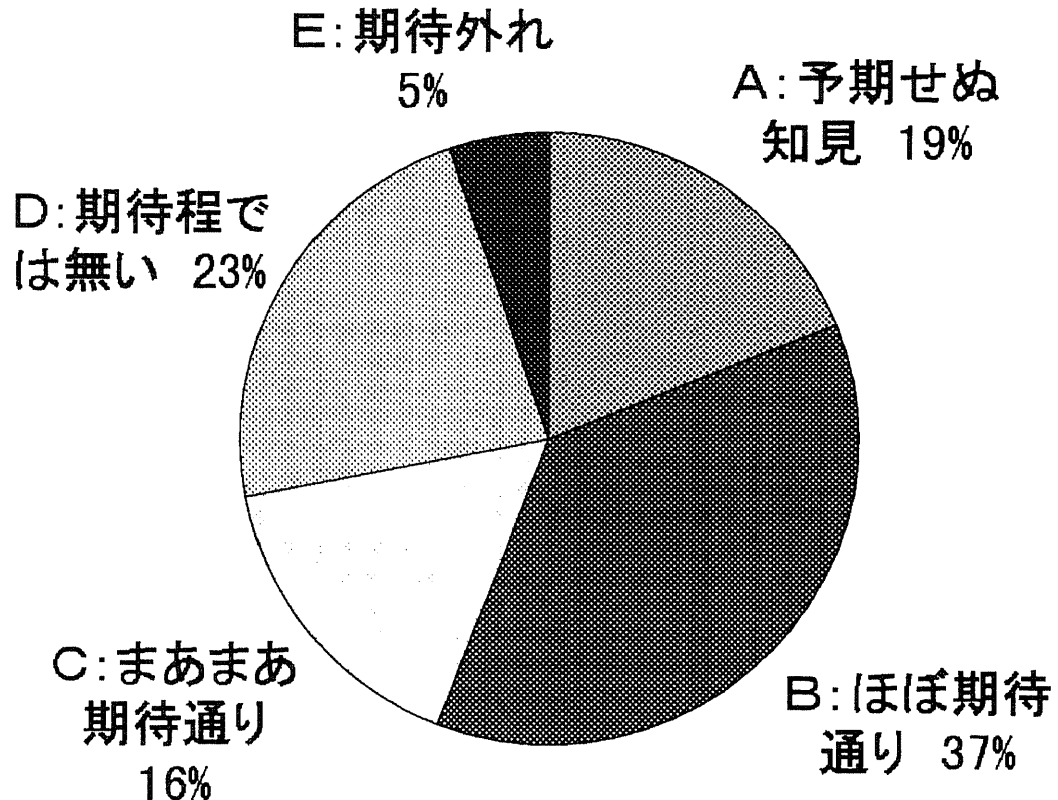
「係り受け句」の抽出元のパラグラフを読んだ被験者は、それらのパラグラフの内72%(A,B,C)が「係り受け句」を選択した際に期待した情報に合った、あるいは、期待以上の知見が得られるものであったと評価している。これは、「係り受け句」の多くが抽出元のパラグラフの情報を適切に表現できていたことを意味している。従って、「係り受け句」による文献選択は有効な手法だといえることになる。特に、Aの<予期せぬ知見が得られた>が全体の16%にもおよび、「係り受け句」が新規な情報発見の一手段となりうる可能性が示される結果を得た。

次に、先の質問に対する被験者のパラグラフへの予備知識の程度を示す結果は以下の表1のようになった。検索結果のパラグラフの内容をdの<全く知らない>と被験者が回答した199の「係り受け句」についても、やはり71%(A,B,C)がプラスの評価を被験者から得ているので、検索対象分野について専門的知識が不足していても、「係り受け句」は有効な手法となりうることが示された。

次に、「係り受け句」が抽出元の一文中の構成要素として、どの程度重要だと被験者が思ったかについての結果は以下の表2のようになった。グラフ3を見ると抽出元の一文中ではc<単なる1要素に過ぎなかった>と被験者が判断した「係り受け句」でも、その内54%(A,B,C)は期待した、あるいはそれ以上の情報を含む文献のパラグラフを選択する手段となっていたことがわかる。従って、一文を読むことより、「係り受け句」を見る方が検索要求にあったパラグラフを選択するためのより有効な手段であるといえる。

5-3 自由記述のまとめ

被験者のシステム全体に対する自由記述の内、いくつかを以下に挙げる。



グラフ1:「係り受け句」一覧で選択したパラグラフから得られた知見への評価

- ユニークな検索方法で、外れても思わぬ発見があり楽しい。「コンピュータの情報検索では、ALL or Nothing でブラウジングができず serendipity (思わぬ発見) がない」という限界を超える先駆的な試みだと思う。
- マシンによる検索語の切り出しは、語に対して固定概念に捕われずに、新たなイメージを持たせることができると思った。
- 句から段落を見るのは便利だし、時間の省略にもなる。
- 同じ文献が出てくるのが何度かありました。一度出てきたらそれで良いと思います。
- 助詞に少し手を加えて、自然な表現にするとより良いと思った。

6 まとめと今後の課題

6-1 まとめ

これまで述べてきたように本手法は、プロトタイプシステムでの実験評価では文献選択の手法として高い評価を得たといえる。プロトタイプシステム上の欠点の指摘や不満の声もあったがその多くは改善可能なものであり、手法自体は有効なものだと考えている。また、今回の実験ではある1つの検索語は平均で32.45文献に出現し、また平均で99.00パラグラフ中に出現していた。これだけの数のパラグラフを一つ一つ参照することは困難であるが、「係り受け句」を用いることで比較的労力が少なく十分な情報が得られることになる。この点でも検索に「係り受け句」を用いることの有効性が示されていると考える。

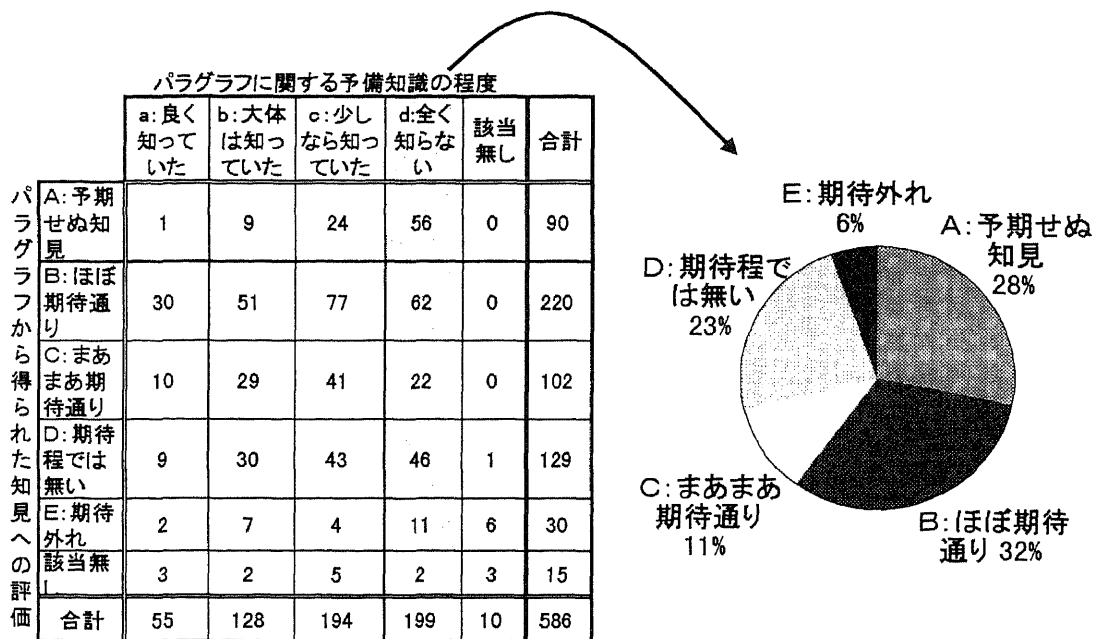


表 1 : 質問<パラグラフから得られた知見への評価>と質問<パラグラフへの予備知識の程度>の集計結果
 グラフ 2 : d<全く知らない>パラグラフから得られた知見への評価分布

6-2 今後の課題

今後このプロトタイプシステムの実用化には以下のような改良が必要ではないと思われる。

- 一度読んだ文献を画面上認識できるようにする
- 「係り受け句」選択画面上のリストの大きさを更に圧縮する
- 入力検索語として複合語を受け付けるために、その形態素解析を行う機能を付加する
- 類義語問題：シソーラス的機能を補完する

参考文献

- [1] 大澤幸生, 須川敦史, 谷内田正彦. ユーザの変化する興味を理解し表現する文献検索支援システム Index Navigator. 人工知能学会誌. Vol.13, No.3, p.461-469(1998)
- [2] 高野敦子, 平井誠, 北橋忠宏. 情報検索の自然言語インターフェースにおける対話管理の枠組み. 自然言語処理. 124-14, p.103-110(1998)
- [3] Craig G. Nevill-Manning, Ian H. Witten and Gordon W. Paynter "Browsing in Digital Libraries: A Phrase-based Approach", 2nd ACM International Conference on Digital Libraries, July 1997, pp.230-246

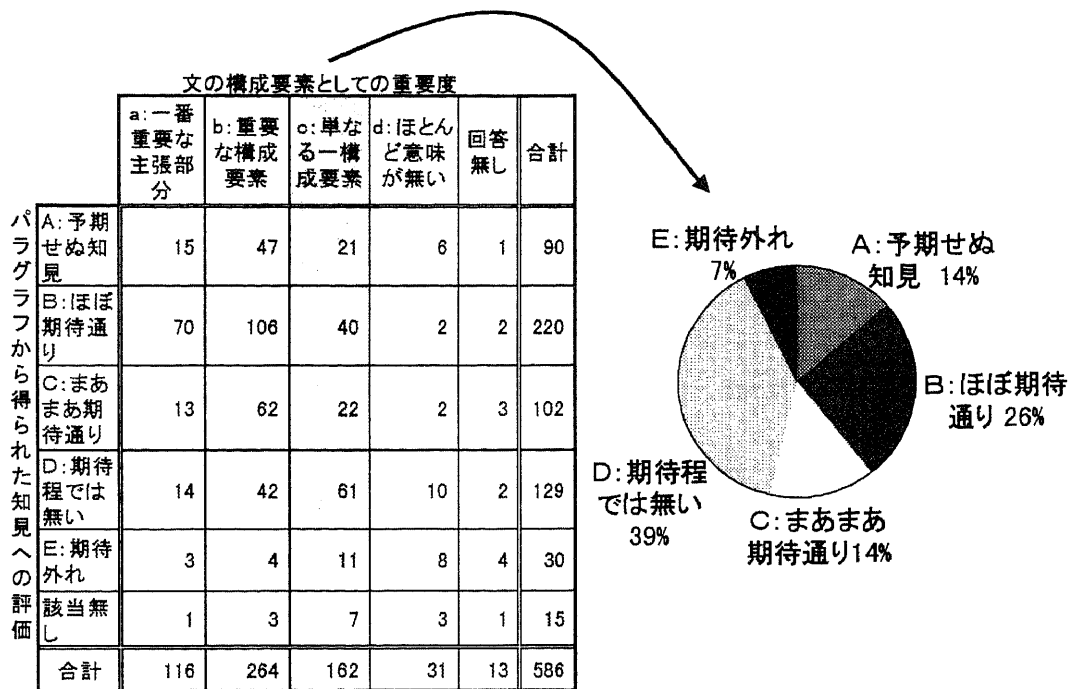


表2：質問<パラグラフから得られた知見への評価>と質問<「係り受け句」の抽出元の一文中での重要度>の集計結果

グラフ3：c<一文中で単なる構成要素>に過ぎない「係り受け句」の抽出元のパラグラフから得られた知見への評価分布

- [4] 大山敬三. インターネットに適應した全文検索データベース検索システムの構成. 学術情報センター紀要. 第7号(1996) <http://www.nacsis.ac.jp/rd/bulletin/no7/cont-j.html>
- [5] 新美和彦, 兵藤安昭, 池田尚志. 係り受け情報を用いた全文検索とその評価. デジタル図書館. No.11(1997) <http://www.ulis.ac.jp/DLworkshop/old-programs.html> 「デジタル図書館」ワークショップ
- [6] 池田和幸, 高須淳宏, 安達淳. 単語間の係り受け情報を用いた文献検索手法. 学術情報センター紀要. 第9号, p. 143-159(1997)
- [7] 長尾真編. 自然言語処理. 第2版. 岩波書店. 1996. (岩波講座：ソフトウェア科学, 15巻)
- [8] 松本裕治, 黒橋禎夫, 山地治, 妙木裕, 長尾真. 日本語形態素解析システム JUMAN 使用説明書 version 3.3. 京都大学工学部. 奈良先端技術大学院大学
- [9] 黒橋禎夫. 日本語構文解析システム KNP 使用説明書 version 2.0 b4. 京都大学工学部

付録：アンケートの質問用紙

これは、情報検索に関するアンケートです。検索対象となる原文献集合は、デジタル図書館を中心とする情報学に関連した論文を収集したものです。最初に枠の上の

"Your name: No. "に名前の記入をお願いします。

では、以下の質問について、お答えください。

質問(1) 以下の30個の検索語候補の中から興味があるものを5つ以上出来るだけたくさん選んでください。以降の質問は、選んだ検索語1つ1つに対して個別に解答していただくことになるので、検索語1つに対して、この用紙1枚を使用してください。では、選んだ検索語を1つだけ○を付けてください。

1. 著作権
2. マルチメディア
3. オンライン
4. 多言語
5. 目録
6. 画像
7. シソーラス
8. サイト
9. OPAC
10. インターネット
11. タグ
12. 蔵書
13. オブジェクト
14. 情報検索
15. ブラウザ
16. リンク
17. ソフトウェア
18. 索引
19. HTML
20. 出版
21. デジタル化
22. 分類
23. メタデータ
24. 学習
25. SGML
26. Web
27. 検索語
28. 課金
29. 辞書
30. 主題

*これ以降の質問(2)から質問(7)までの一連の質問に解答している途中で、新たな検索対象となりうる(選んだ検索語よりも興味を惹かれる)表現を見つけた、あるいは思い付いたとき、その表現と解答中のID番号、質問番号を(7)に記入してください。もし、そういったことがなかった場合には、答える必要はありません。

質問(2) この”句”のリストから、興味を惹かれたものを好きなだけ選んで、リストの一番左にあるID番号を記入してください。”周辺の句”の参照は利用してもしなくても構いません。

(2) 選択した”句”のID: _____

**ここで、一度、最後の質問(7)までを読んでみてください。その方が、全体の流れが分かり、解答がしやすいと思います。

質問(3) 画面の各句の一番左のIDが表示された列のリンクを開いてください。新しいブラウザで表示されたのがその句が存在した段落(原文)です。それを読んでから以下の質問にお答えください。

今読んだ”段落(原文)”の内容は、選択した各”句”を見た時に期待した情報と、どの程度一致していましたか? 選択した各”句”のID番号を該当する項目の横に記入してください。

(A) 予期せぬ知見が得られた ---- 該当”句”のID: _____

(B) ほぼ期待通りの内容だった ---- 該当”句”のID: _____

(C) まあまあ期待に沿った内容だった ---- 該当”句”のID: _____

(D) 期待した程の内容はなかった ---- 該当"句"のID: _____

(E) 全く期待はずれだった ---- 該当"句"のID: _____

もし良ければ、(E)を選んだ、理由を教えてください。例)話しの方向性が期待とは違う

理由: _____

(F) 選んだ句が読むべき段落を持たなかった ---- 該当"句"のID: _____

質問(4) 目を通した段落(原文)の内容は、既知っているものでしたか?

質問(3)が(F)を選択した以外の"句"について、そのID番号を該当する項目の横に記入してください。

(A) よく知っていた ---- 該当"句"のID: _____

(B) 大体は知っていた ---- 該当"句"のID: _____

(C) 少しなら知っていた ---- 該当"句"のID: _____

(D) 全く知らなかった ---- 該当"句"のID: _____

質問(5) 原文献の他の場所も読みたいと感じましたか、あるいは読む価値がありそうだと
思いましたか? 選択した各"句"のID番号を該当する項目の横に記入してください。

「はい」の場合:

(A) 検索語とは直接関係ないが面白そうだから ---- 該当"句"のID: _____

(B) 必要な情報は得られたが、更に興味がわいたから ---- 該当"句"のID: _____

(C) その段落だけでは不十分だったから ---- 該当"句"のID: _____

(D) その他: _____

「いいえ」の場合:

(A) その段落だけで必要な情報が得られたから ---- 該当"句"のID: _____

(B) 期待はずれの内容だったから ---- 該当"句"のID: _____

(C) 読みたくないから ---- 該当"句"のID: _____

(D) その他: _____

質問(6) "句"リストを表示した画面の3番目の列にあるリンクをクリックすると同じ画面の一番下に、先程選んだ"句"が含まれていた1文が表示されます。

その文を読んでください。"句"は"文"の中でどの程度、重要なもの(主題的)だと思いますか? 選択した各"句"のID番号を該当する項目の下に記入してください。

(A) "句"は"文"の1番重要な主張部分だった

---- 該当"句"のID:-----

(B) "句"は"文"の中で重要な構成要素だった

---- 該当"句"のID:-----

(C) "句"は"文"の中で単なる1構成要素に過ぎなかった

---- 該当"句"のID:-----

(D) "句"は"文"の中でほとんど意味のない部分だった

---- 該当"句"のID:-----

質問(7) ID番号____ を解答中、 句 or 一文 or 段落(原文) の表示画面から

見つけた or 思い付いた 検索語よりも興味がある表現

:-----

<画面の説明>

画面に表示されているのは、質問(1)で○を付けた検索語を含む句のリストです。
最初に画面の説明をします。

- ・画面の1番左の列は、句の固有IDです。クリックするとその句を含む段落を中心に原文が別のブラウザ画面に表示されます。
- ・2番目の列をクリックすると原文の同じ節や段落に含まれている周辺の句が参照できます。
- ・3番目の列をクリックすると同じ画面の下の細長い列にその句を含む1文が表示されます。
- ・4番目の列が検索語を含む句です。検索語の後方の語が前方の語に順にソートされています。

注) 5個以上同じ語が検索語の後にある場合、"xx 検索語 xxx"と表示されています。
そこをクリックすると別のブラウザ画面に詳細が表示されます。

注) 複数のIDが付与されている句もあります。異なる文から偶然同じ句が抽出された事を、意味しています。従って、リンクを辿って表示される「段落」や「周辺の句」の内容も異なります。

以下の質問は、最後に答えてください。

質問(8) 所属は以下のどれですか？

A. 学部生：_____年 B. 大学院生 C. 研究生 D. その他：_____

質問(9) 専攻分野は何ですか？

A. 図書館学 B. 情報学 C. それ以外：_____

質問(10) 以下から1つ選んで、○をしてください。

句を選択するとき、「周辺の句」(参照)を

利用した・利用しなかった・両方の場合があった

質問(11) 上の(10)で利用した・両方の場合があったを選んだ方への質問です。

「周辺の句」は 役にたった・あまり役に立たなかった・どちらともいえない

記名をお願いします。：_____

最後のお願い：この「句」から、その句が存在する段落(文献全体ではなく)を選択する方法やこのシステム全体の構成について、何かコメントを頂ければ幸いです。

コメント：_____

以上で、終了です。長い時間お付き合い下さりましてありがとうございます。

ご協力を心から感謝します。