

特集：電子文書処理

研究論文

## NTCIR-3 WEB: Web 検索のための評価ワークショップ

## NTCIR-3 WEB: An Evaluation Workshop for Web Retrieval

江口 浩二

国立情報学研究所

Koji EGUCHI

National Institute of Informatics

大山 敬三

国立情報学研究所

Keizo OYAMA

National Institute of Informatics

石田 栄美

国立情報学研究所

Emi ISHIDA

National Institute of Informatics

神門 典子

国立情報学研究所

Noriko KANDO

National Institute of Informatics

栗山 和子

白百合女子大学

Kazuko KURIYAMA

Shirayuri College

### 要旨

著者らは、2001年から2002年に開催された第3回 NTCIR ワークショップにおいて Web 検索タスク（以下、NTCIR-3 WEB）を実施した。その主な目的は、共通のデータセットを用いて Web サーチエンジンの検索有効性を評価するとともに、Web 検索に関する再利用可能なテストコレクションを構築することであった。上記の目的のもと、検索対象の文書データとして JP ドメインから収集した 100GB 及び 10GB の Web ページデータを使用し、多様な入力形式による検索実行結果に対して、複数のユーザモデル、複数の文書モデルを仮定した評価を行った。ここに入力形式として、語、文、文書を設定した。また、ユーザモデルとして、網羅的に適合文書を求めるモデルと検索結果上位における精度を重視したモデルを想定した。文書モデルとしては、ページを単位としたモデル、リンクで結合された文書群を単位としたモデルを仮定した。本論文では、NTCIR-3 WEB において構築されたテストコレクションの概要、提案された評価手法、及び評価結果について述べる。評価の結果から、簡潔なクエリを用いた検索結果のランキング上位において特にリンク解析が有効であることなどが示唆された。

### ABSTRACT

The authors conducted the Web Retrieval Task ('NTCIR-3 WEB') from 2001 to 2002 at the Third NTCIR Workshop. In the NTCIR-3 WEB, they attempted to assess the retrieval effectiveness of Web search engine systems using a common data set, and to build re-usable test collections that are suitable for evaluating Web information retrieval systems. With these objectives, they evaluated on searches using various types of user input, user models and document models. As the document data sets, they constructed 100-gigabyte and 10-gigabyte document collections that were gathered from the '.jp' domain. The user input was given as query term(s), sentence, and document(s). They assumed two user models where comprehensive relevant documents are required, and where precision of the top-ranked results is emphasized. They also assumed several document models, such as a document as an individual page, and a document as a page set connected by hyperlinks. This paper describes an overview of the test collections constructed in the NTCIR-3 WEB, the proposed evaluation methods, and the evaluation results. The evaluation results suggested that the link-based techniques can perform more effectively when short queries are input.

[キーワード]

評価手法, テストコレクション, Web 情報検索, サーチエンジン

[Keywords]

Evaluation Methods, Test Collections, Web Information Retrieval, Search Engines

## 1 はじめに

本論文では、2001年8月から2002年10月にかけて開催された第3回 NTCIR ワークショップにおいて実施された Web 検索タスク（以下、**NTCIR-3 WEB**）について論ずる<sup>[1-3]</sup>。これは、共通のデータセットを用いて Web サーチエンジンの検索有効性を評価するとともに、Web 検索に関する再利用可能な実験評価用テストコレクション（以下、**Web テストコレクション**）を構築することを主目的とした評価ワークショップである<sup>[4]</sup>。ここで、Web テストコレクションは次の三つにより構成される。すなわち、(i) 検索の対象となる**文書セット**、(ii) 情報ニーズを記述した**検索課題**、(iii) 各検索課題に関する**適合判定の結果**である。これら三つの構成要素を用いて、複数の検索システムの有効性を比較評価するという点では、従来のテキスト検索システムに関する評価型ワークショップと同様であるものの、NTCIR-3 WEB では上記それぞれについて、Web 検索に特徴的な観点を採り入れた。

以上の目的のもと、NTCIR-3 WEB は次の手順で実施された。

- (1) **オーガナイザ**（本論文の著者ら）は、JP ドメインから独自に収集した 100GB 及び 10GB の Web ページデータをもとに、検索の対象となる**文書セット**を構築した。
- (2) **参加者**は、オーガナイザが提案した数種のタスクのうちの一つを選択して、検索課題をもとにして独自のシステムに対応したクエリを生成し、上記の文書データのいずれかに対して検索を行ない、その**実行結果**をオーガナイザに報告した。オーガナイザが設計したタスクは以下の通りである。
  - (A) **サーベイ検索タスク**
    - (A1) **トピック検索タスク**
    - (A2) **類書検索タスク**
  - (B) **ターゲット検索タスク**
  - (C) **自由課題タスク**
    - (C1) **分類出力タスク**
    - (C2) **音声入力タスク**

- (3) **オーガナイザ**は、**判定者**に検索実行結果に対する適合判定を指示し、その判定結果をもとにしてシステムごとの最終的な評価値を算出した。以上の評価手法には、Web 検索における特徴的な観点として、リンク構造及びページ内容の重複を考慮した。

NTCIR-3 WEB は、以下の日程で実施された。

**2001年8月1日** 参加募集開始

**2002年1月15日** 文書データセットへのアクセス許可

**2002年2月8日** ドライラン用検索課題の配布

**2002年2月18日** ドライラン実行結果の提出締切

**2002年4月15日** ドライランの適合判定結果の配布

**2002年4月25日** フォーマルラン用検索課題の配布

**2002年5月13日** フォーマルラン実行結果の提出締切

**2002年8月6日** フォーマルランの適合判定結果及び評価結果の配布

**2002年10月8-10日** 成果報告会及びラウンドテーブル会議

ドライランは参加者とオーガナイザが NTCIR-3 WEB の手順に関して経験を得ることを目的として、少数の検索課題を用いて実施した。フォーマルランにおいて正式な評価を行なった。

以下、2章では NTCIR-3 WEB におけるタスク設計について述べる。3章では文書セット、検索課題及び適合判定方法について述べる。4章では評価尺度について述べる。5章では参加状況、評価結果及び考察を述べる。

## 2 タスク設計

Web 検索システムは従来のテキスト検索システムとは異なる点が多いが、NTCIR-3 WEB ではタグとリンクによる構造を備えた大規模 Web 文書データの検索に関するタスクを想定して、評価手法を設計した。タスクの構成は 1章の(2)において示した(A)~(C)である。これらのそれぞれにつき、(I)100GB あるいは (II)10GB の文書データを対象に実施された。

2.1 及び 2.2 において、それぞれサーベイ検索タスク及びターゲット検索タスクの目的及び手順を述べる。これらは後述の通り、前提とするユーザモデルが互いに異なる。また、2.3 においては、自由課題タスクについて紹介する。

## 2.1 サーベイ検索タスク

サーベイ検索タスクは、ユーザが網羅的に適合文書を探るようなユーザモデルを仮定している。クエリの表現形式に応じて、サーベイ検索タスクを (i) 語や文を用いて検索するトピック検索タスク、(ii) 文書を用いて検索する類書検索タスクに分けて実施した。それぞれの概要を 2.1.1 及び 2.1.2 に述べる。

### 2.1.1 トピック検索タスク

トピック検索タスクは、新聞記事や学術文書を対象とした Ad Hoc 型検索<sup>[5-7]</sup>に相当するものであり、テストコレクションの再利用性を鑑みて実施した。ただし、4 章で後述する通り、評価尺度等について Web 検索に特徴的な種々の要素を導入した。

参加グループは、検索課題ごとに実行した検索結果の上位 1,000 件に対応する文書 ID のランク付きリスト（以下、実行結果リスト）を、オーガナイザに提出した。人間が介在して検索課題を参照しつつ実行する検索（以下、対話型システム）とそれ以外（以下、自動システム）のいずれのシステムも許可された。また、トピック検索タスクの参加グループは、少なくとも二件の実行結果リストを提出することが期待されており、一方が検索課題を構成する TITLE のみを用いた検索、他方が DESC のみを用いた検索であった。ただし、検索課題のその他の部分を用いた検索実行結果リストを、最大二件まで提出することが許可された。検索課題の詳細については 3.2.1 に示す。

参加グループは、使用する自動システムあるいは対話型システムの別と、検索課題のどのフィールドを用いたかを報告することとした。なぜなら、システムを評価する際には、自動システム・対話型システムの種別、及び、使用された検索課題のフィールドに従って、別々に有効性が比較されるのが適切だからである。

また、参加グループは根拠パッセージを提出することが期待された。根拠パッセージとは、個々の検索された文書に含まれる、検索システムが適合性を算出した根拠を与える記述を指す。根拠パッセージは補完的な評価において有用であると考えたが、結果としてはいずれの参加グループからも根拠パッセージの提出がなかった。

### 2.1.2 類書検索タスク

類書検索タスクの目的は、(i) ある検索語による検索結果から一件の適合文書を選択し、それをを用いた類似検索を評価することと、(ii) ある検索語による検索結果から少数の適合文書を選択し、それをを用いた類似検索を評価することとである。多くの訓練文書を想定していないのは現実の Web の環境を勘案したことによる。

類書検索タスクでは、提出が必須である実行結果と提出が任意である実行結果を以下の要領で設定した。

**提出が必須の実行結果** 検索課題の RDOC タグで特定された先頭の文書を使用しなければならない。TITLE タグを使用してもよい。RDOC タグで特定された二番目と三番目の文書を使用してはならない。

**提出が任意の実行結果** 検索課題の RDOC タグで特定された先頭の文書を使用しなければならない。TITLE タグを使用してもよい。RDOC タグで特定された二番目と三番目の文書を使用してもよい。

なお、RDOC を始めとした検索課題の形式に関する詳細は 3.2.1 で述べる通りである。当該必須条件は上記の目的 (i) に対応し、任意の条件については目的 (ii) に対応する。

実行結果リストの提出要領はトピック検索タスクと同様であった。なお、3.4 で述べる通り、適合判定は検索課題の全記述によって与えられる判定基準に従って実施され、RDOC で特定された適合文書の内容によって与えられる基準に従うものではないとした。

## 2.2 ターゲット検索タスク

ターゲット検索タスクは、事実を確認しうる記述や疑問に対する答えを含む、少数の文書を求めるようなユーザモデルを仮定しており、この場合、検索結果上位における精度が強調される

ここでは、実行結果リストの提出に関わる条件は、ほぼ 2.1 に述べたサーベイ検索タスクと同様であるが、実行結果リストは検索課題ごとに 20 件のみでよい。4 章に述べる通り、検索課題ごとに実行された検索結果の上位 10 件程度を用いて、独自の観点に基づく評価尺度が適用された。

ところで、TREC High Precision Track<sup>[8,9]</sup>は、判定者が 10 件（ないし 15 件）の最も適合すると判断した文書を、検索課題ごとに 5 分間以内で発見するというものであった。ターゲット検索タスクの設計は、これとは異なるものの、検索結果ランキングの上位における精度が重要とされるという点で類似する部分がある。

また、これまでの TREC Web Track<sup>[6,10,11]</sup>では、数

百件もの検索結果文書を用いた精度・再現率に基づく評価尺度が使用されるとともに、検索結果ランキングの上位における精度についても同一のタスク設計において用いられた。これに対し、NTCIR-3 WEB ではサーベイ検索タスクとターゲット検索タスクを別々に実施した。これにより、ターゲット検索タスクの検索課題数をサーベイ検索タスクよりも多く設定することが可能となった。ターゲット検索タスクのように検索結果ランキング上位にのみ着目する場合の評価は一般に、十分な数の検索課題を使用しないと評価が安定しないといわれている<sup>[12]</sup>が、これを重視する場合、ターゲット検索タスクを独立して実施することは妥当な措置であると思われる。

### 2.3 自由課題タスク

NTCIR-3 WEB が提供した文書セットを用いることを前提に、参加者が自由に提案したタスク設計のうち、複数のグループの参加が見込まれるものについては、自由課題タスクとして、NTCIR-3 WEB におけるタスクの一つとして取り上げることとした。結果として、(i) 分類出力タスクと (ii) 音声入力タスクが採用された。(i) は検索結果の分類に基づく提示方法を評価することが目的であり<sup>[2]</sup>、(ii) は音声で与えられたクエリを用いて Web のテキスト文書を検索する手法を評価するのが目的であった<sup>[13]</sup>。結果としては、(i) の実行結果を提出した参加グループはなく、(ii) については二グループにとどまった。以上の自由課題タスクの詳細については別稿<sup>[2,13]</sup>に委ねて、本論文では割愛する。

## 3 Web テストコレクション

Web テストコレクションは、(i) 文書セット、(ii) 情報ニーズを記述した検索課題、(iii) 各検索課題に対する適合判定結果リストにより構成される。これら構成要素のそれぞれは、現実の Web の環境に適合したものとなることを目指した。上記 (i)~(iii) のそれぞれを 3.1, 3.2 及び 3.4 に示す。また、3.3 には、適合判定に先だって実施されるプーリングについて述べる。

### 3.1 文書セット

検索処理の対象となる文書セットは、Web テストコレクションのために明示的に特定されるのが相応しい。Web を対象とするのは、NTCIR ワークショップにおいて初めての試みであるため、以下の方法を採用した。

- クローラが収集した Web ページの一部を抽出し、検索対象文書データとしての URL 集合を定義する。
- 検索対象文書データを提供する。

上記の方法で得られた文書データは、従来のテストコレクションと同様と見なすことができるため、適合文書集合の特定方法やシステムの評価方法について、既存の手法を拡張することで対処される。また、それにより、再利用可能な Web テストコレクションが構築できると期待される。

以下、3.1.1 においてページ収集戦略を、3.1.2 において Web 文書の定義を述べる。なお、参加グループは文書データ NW100G-01 及び NW10G-01 を、国立情報学研究所に設置されたオープンラボラトリにおいてのみ処理できるとする制限が課せられた (3.1.3)。

#### 3.1.1 ページ収集

クローラ (crawler) は再帰的に、Web ページを取得し蓄積するとともに、各ページにおける未収集のリンクを追跡するようなソフトウェアを指す。文書データを構築するため、2001 年 8 月 29 日から同年 11 月 12 日の間、以下の条件でページ収集を行なった。

**Web サイト:** JP ドメイン上の HTTP サーバ<sup>1</sup>。

**ポート:** 任意。

**ファイル形式:** HTML またはプレーンテキスト・ファイル。

まず、クローラ<sup>2</sup> はページの収集元となる Web サイトを発見し、次にそれらから Web ページを収集することを試みた。具体的には、前記の Web サイト、ポート、ファイル形式に関する条件の下で、次に示すページ収集戦略が適用された。

#### Web サイト発見

- (1) 起点を特定する。国立情報学研究所のトップページ (<http://www.nii.ac.jp/>) を起点とした。
- (2) 起点のページもしくは発見されたサイトのルートページからリンクを抽出し、20 ページだけ収集することを試みる。
- (3) 収集されたページからリンクを検出する。
- (4) 新たに発見されたサイトを抽出する。IP アドレスでなくホスト名でサイトを同定する。
- (5) 別名サイト及び起動していないサイトを除外する。
- (6) サイト発見率が一定値を下回るまで、前記 (2) ~ (5) を並列かつ再帰的に実行する。

<sup>1</sup> ページ収集後に、収集されたページからリンクを抽出し、それらの URL に対応するページでインターネット上にその存在が確認できたものによって文書セットを拡張した (3.1.2) が、それらには JP ドメイン以外のページも含まれる。

<sup>2</sup> クローラとして、Open Text 社の「Livelink Spider」を使用した。

### Web ページ収集

- (1) 発見されたサイトのそれぞれに対して、ルートページの URL を URL リストに追加する。
- (2) URL リストの先頭におけるページの収集を試みる。
- (3) 重複ページあるいは前記のページ収集条件に照らして不適切なページであればそれを棄却し、上記 (2) へ戻る。
- (4) 収集されたページから、同一サイトのページへ張られたリンクのみを抽出する。
- (5) 新たに発見されたページを URL リストの末尾に追加する。
- (6) 収集されたページの数<sup>3</sup>が 2,000 件に達するか、URL リストが空になるまで、前記 (2)~(5) を繰り返す。

ここで、ホスト名とそのポート番号の間に「:」記号を挿入し、その末尾に「/」記号を付与することで、ルートページを特定した。なお、ページ収集では、次に示すようなページを除外した。

- 文書データセットの作成に障害となるページ。画像やアーカイブファイルのような非テキストのページであるにも関わらず、HTTP ヘッダにおける「Content-type」属性が誤って「text/html」あるいは「text/plain」と記されている場合を含む。
- ループパスを引き起こすページ。
- 動的に生成されるページ。ただし、それらのうち初出の 10 件は除去しなかった。

#### 3.1.2 文書セット及び文書データの定義

我々は、ページ収集で得られた Web ページ群から部分集合を抽出し、参加者に提供するための文書データを構成した。また、文書データから JP ドメインに限定せずにリンク情報を抽出し、それが示すページの確認を確認して、文書データを拡張するような URL リストを作成した。その結果、次の二種類の文書セットを定義した。

(1) 提供文書データすなわち 100GB の **NW100G-01** と 10GB の **NW10G-01**、及び、(2) 参照先文書データである。なお、10GB の NW10G-01 は、100GB の NW100G-01 の部分集合である。また、上記 (2) の参照先文書データとしては、NW100G-01 と NW10G-01 のそれぞれに対して、それに含まれる文書から参照されたページ群を用

<sup>3</sup>この制限は経験に基づいて決定された。最終的には、Web ページ収集のあと、サイトごとのページ数を 100GB データに対して 1,300 件以下に、10GB データに対して 20 件以下になるように再調整した。

いた<sup>4</sup>。これらはリンク解析に用いることができ、結果としてその一部が実行結果に含まれることもありうる。

NW100G-01 と NW10G-01 の統計的特徴を表 1、2、3 及び 4 に示す。また、NW100G-01 及び NW10G-01 のそれぞれに対して、以下のファイルをも提供した。

- ページデータ及びメタデータ。メタデータは URL、ページ収集日時、HTTP ヘッダなどから成る。一サイトにつき一ファイルで構成した。
- 収集元サイトのリスト。
- 別名サイトのリスト。
- 収集されたページのリスト。
- 重複ページのリスト。
- 参照先ページのリストすなわち参照文書データに含まれるページのリスト。
- 提供文書データ中のページから提供文書データ中のページへ、もしくは参照先ページへと張られたリンクのリスト。

個々の Web ページとそのメタデータをもって **Web 文書** を定義する。なお、本論文では特に明示しない限りは、Web 文書を単に「文書」と表記する。Web 文書は、以下のタグで示された部分からなる。Web 文書の例を図 1 に示す。

- <NW:DOC> は Web 文書の境界を示す。
  - <NW:META> は次の項目を含むメタデータを示す。
    - <NW:DOCID> は文書 ID を示す。
    - <NW:DATE> はページ収集の日時を示す。
    - <NW:CTYPE> は「Content-Type」を示す。
    - <NW:URL> は URL 文字列を示す<sup>5</sup>。
    - <NW:HTTPH> は HTTP ヘッダを示す。
  - <NW:DATA> はページデータを示す。ページのバイト長を示す <NW:DSIZE> タグで開始され、それに続いてページ内容の原データが記述される。
- 同一のサイトから収集された Web 文書は一つのファイルに格納した。また、提供文書データとして次の三種を参加者に提供した。
- 収集したままの原データ。
  - 日本語文字コードを EUC コードに変換したもの。

<sup>4</sup>ただし、インターネット上にその存在が確認できたものに限った。

<sup>5</sup><NW:URL> にはただ一つの URL を示すものとし、サイト内における同一内容のページ群は重複ページのリストとして別途に提供した。

```

<NW:DOC>
<NW:META>
<NW:DOCID>NW010616091</NW:DOCID>
<NW:DATE>Mon, 05 Nov 2001 09:46:11 GMT</NW:DATE>
<NW:CTYPE>text/html</NW:CTYPE>
<NW:URL>http://www.nii.ac.jp/</NW:URL>
<NW:HTTTPH>HTTP/1.1 200 OK
Date: Tue, 06 Nov 2001 02:24:19 GMT
Server: Apache/1.3.14 (Unix)
Last-Modified: Mon, 05 Nov 2001 09:46:11 GMT
ETag: "ae663-4dce-3be65fe3"
Accept-Ranges: bytes
Content-Length: 19918
Connection: close
Content-Type: text/html
</NW:HTTTPH>
</NW:META>
<NW:DATA><NW:DSIZE>19852</NW:DSIZE>  <IDOCTYPE
HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
<HTML lang="ja">
<HEAD>
<META HTTP-EQUIV="Content-Type" CONTENT="text/html;
charset=ISO-2022-JP">
<TITLE>NII -The National Institute of Informatics-</TITLE>
:
</HTML>
</NW:DATA>
</NW:DOC>
    
```

図 1: Web 文書の例

- 日本語文字コードを EUC コードに変換し、一切のタグ及びコメント部分を除去したもの。ただし、META タグにおいて記述されたキーワードについては残し、行頭に識別子を付した。

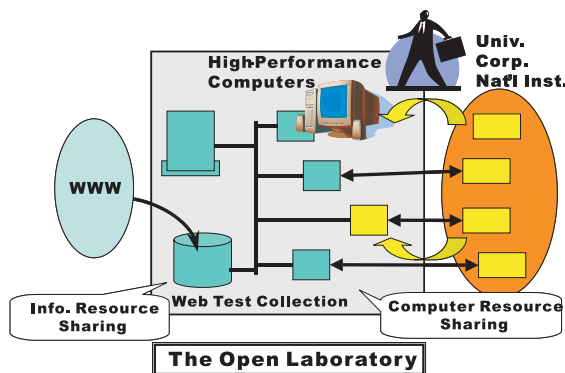


図 2: オープンラボラトリーの枠組

### 3.1.3 オープンラボラトリ

参加者は国立情報学研究所に設置されたオープンラボラトリにおいてのみ、文書データに対するインデクシングなどの処理を許可された<sup>6</sup> (図 2)。参加者はイ

<sup>6</sup>文書データの処理のため、オープンラボラトリのホストコンピュータに対してリモートアクセスを行なうことは許可され

ンデックスファイルなどの加工済みデータを持ちだし、個々の研究室において実験を行なうことができた。これにより、参加者は十分な計算機資源を所有していても、NTCIR-3 WEB に参加することができた。

## 3.2 検索課題

### 3.2.1 検索課題の形式

オーガナイザは、情報ニーズが所定の形式で記述された検索課題を提供した。検索課題の形式は過去の NTCIR ワークショップ<sup>[7]</sup>を参考にしたが、TITLE の定義、新たに導入した RDOC 及び USER タグ、そして NARR タグの形式については独自に設計した。使用することが必須となるタグ部分と、使用することが許可されるタグ部分はタスクごとに指定されている (2 章)。検索課題は下記のタグで示された部分により構成されている。なお、図 3 に検索課題の例を示す。

- <TOPIC> は検索課題の境界を示す。
- <NUM> は検索課題 ID を示す。
- <TITLE> はサーチエンジンへ投入されるクエリを模擬して、以下のように定義した。すなわち、検索課題作成者が、三つの戦略、(a) 同義語・類義語を並べる場合、(b) 異なる観点の語を並べる場合、(c) 戦略 (a) と戦略 (b) の複合のうち、適切と思われる戦略を選択し、それにそってサーチエンジンに入力するであろう三語以内の語をリストアップするものとする。ただし、重要な語から順に左から並べるものとする。なお、図 3 の例における <TITLE CASE="c" RELAT="2-3"> は戦略 (c) においてタイトル中の 2 番目の語と 3 番目の語が同義あるいは類義関係にあることを示している。仮に参加グループがブール型検索を用いる場合、戦略 (a) に対しては OR 演算、戦略 (b) に対しては AND 演算、戦略 (c) に対しては OR 演算と AND 演算の組合せとみなすこともできる。
- <DESC> は情報ニーズの最も基本的な記述であり、一文で表現される。
- <NARR> は、検索の背景・目的、語の定義、及び、適合判定基準の補足を、数段落で記述したものである。これらはそれぞれ <BACK>, <TERM>, <RELE> タグで示される。ただし、これら三つが常に記述されるとは限らない。
- <CONC> は同義語や類義語、関連語からなるリストを示す。これらは検索課題作成者により定義される。
- <RDOC> は三件以下の適合文書 ID を示す。これらは

```

<TOPIC>
<NUM>0004</NUM>
<TITLE CASE="c" RELAT="2-3"> コンピューターウイルス, 予
防, 対策 </TITLE>
<DESC> コンピューターウイルスの予防方法や対策法について
説明している文章を探したい </DESC>
<NARR><BACK> インターネット利用が爆発的に普及する中で
コンピュータウイルスは日常的な問題にまで近づいてきて
いる。そこでどのような予防法をとり、もし感染してしま
った際にはどのような対処をすればよいのか知っておきたい。
</BACK><RELE> 適合文書は、コンピュータウイルスへの予
防・対策についての情報を提供するもの。被害届出やウイル
スの種類についてのみ述べているものは適合としない。特定
のウイルスについてのみ情報を提供するページは部分的適合
とする。 </RELE></NARR>
<CONC> コンピューターウイルス, ワーム, 情報セキュリティ, 不正
アクセス, 予防, 対策, 感染 </CONC>
<RDOC>NW003214039, NW013338047, NW013315769
</RDOC>
<USER> 大学院修士 1 年, 男性, 検索歴 5 年 </USER>
</TOPIC>

```

(a) 検索課題の例

```

<TOPIC>
<NUM>0004</NUM>
<TITLE CASE="c" RELAT="2-3">computer virus, preventive,
countermeasure</TITLE>
<DESC>I want to find sentences that explain preventives or counter
measures against computer viruses.</DESC>
<NARR><BACK>Because the use of the Internet has spread ex
plosively, computer viruses have become a serious problem in
our daily lives. I want to know what kind of preventives are
required, and what kind of countermeasures I should take when
my computer becomes infected.</BACK> <RELE>Relevant docu
ments must provide some information on preventives or counter
measures against computer viruses. Documents that describe only
the victim's reports or the types of computer viruses are regarded
as not relevant. Pages that provide some information on a particu
lar virus are regarded as partially relevant.</RELE></NARR>
<CONC>computer virus, worm, information security, illegal ac
cess, preventive, countermeasure, infection</CONC>
<RDOC>NW003214039, NW013338047, NW013315769
</RDOC>
<USER>1st year of Master Course, Male, 5 years of search
experience</USER>
</TOPIC>

```

(b) 検索課題例の英語翻訳版

図 3: NTCIR-3 WEB (ドライラン) における検索課題の例とその英語翻訳版

類書検索タスク<sup>[2]</sup>で使用される。

- <USER> には、検索課題作成者の属性として、職業、性別、検索歴を記す。

なお、すべての検索課題は日本語で記述された。

### 3.2.2 検索課題作成の戦略

我々は、検索課題を作成するにあたって、以下の戦略を適用した。

- すべての検索課題は、検索システムや適合判定を伴わずに作成した。
- 時間または時間的な変化に過度に依存する検索課題を棄却した。実際の Web に対するユーザの要求を勘案すると、このような課題は重要であるものの、適合判定の手順が複雑になることを避けるため、当該措置をとった。仮に「中田英寿選手の今後の試合予定を知りたい」なる検索課題を想定すると、これは「今後」という概念が時間に強く依存するので除外される。
- 検索課題作成者は、次の制約のもとで DESC を記述した。(1) TITLE で特定された語そのものが DESC に出現しなくとも、その概念や意味が、DESC に含まれる。(2) DESC の示す対象範囲と NARR の示す対象範囲の間に大きな差異が存在しない。これらの制約を課した理由は、検索システムが主に TITLE あるいは DESC を用いて検索を実行するのに対して、判定者は主に NARR で指示された対象範囲に基づいて適合判定を行なうからである。

- RDOC を記述するため、検索課題作成者は、オーガナイザの検索システムによる実行結果 20 件から最も適合する文書を三件以下だけ選択した。これは、検索課題を参加グループに配布するのに先だって行なわれた。

前記の通り、すべての検索課題は、検索システムや適合判定を伴わずに作成した。従って、作成された検索課題の一部は適合文書が十分な数に満たないなどの理由で、検索有効性の比較評価に使用するのに適さなかった。このため、これら不適切な検索課題は除外した。結果として、47 の検索課題をフォーマルランの評価に用いた。これらの検索課題セットはサーベイ検索タスクとターゲット検索タスクのいずれにおいても同じく使用した<sup>7</sup>。

### 3.3 プーリング

プーリングとは、各実行結果リストの上位から十分な数の文書を取り出し、併合するようなサンプリングを指す<sup>[5-7]</sup>。プーリングの結果、プールと呼ばれる文書セットの部分集合を得て、文書セット中の全適合文書を見積るのに使用した。

プーリングでは、各実行結果リストの上位 100 文書を取り出した。また、プールに含まれる文書群に対して、メタサーチエンジンの戦略を用いたランキングを行なった。ランキング戦略には、Borda Count 投票アルゴリズム<sup>[14]</sup>を使用した。

<sup>7</sup> ターゲット検索タスクには 96 の検索課題を用いる方向で検討を進めたが、予期せぬ問題により、結果として 47 の検索課題を用いることとなった。

なお、適合文書の網羅性を改善するための追加のマニュアル検索<sup>[7,15]</sup>は行なわなかった。ただし、プールにおける適合文書の網羅性を改善することを目的として、オーガナイザが参加者と同じ条件で検索システムを実行し、その実行結果を参加者と同様に扱った。

### 3.4 適合判定

プールは各参加者が提出した実行結果リストの上位に含まれる文書で構成されるため、大部分の適合文書が含まれていると仮定した。判定者がプールにおける各文書の適合性を判定する際には、判定用システムを用い(3.4.2)、複数の文書モデルに基づいて(3.4.1)、多値適合レベルを判定した(3.4.3)。

ところで、プール中の文書によっては日本語や英語以外の言語による記述を含みうるが、判定者は日本語あるいは英語で記述された情報のみに基づいて、適合性を判定した。

#### 3.4.1 文書モデル

Web ページの表現形式は多様であり、Web における情報の妥当な基本単位は、リンクで結合されたページ群である場合、個々のページである場合、あるいはページ中のパッセージである場合がありうる。

TREC Web Track ではページを情報の単位と仮定された<sup>[6,10,11]</sup>。この仮定のもとでは、複数の Authority 的ページ<sup>[16]</sup>へのリンクを備えた Hub 的ページは、そのページ自体に適合情報を含まない限りにおいては不適合とみなされる。しかしながら、Web の環境においてはしばしば、上記仮定に基づく Hub 的ページの方が適合ページよりもユーザにとって有用である。

NTCIR-3 WEB では、適合判定に際して次の三つの文書モデルを仮定した。

**1 クリック距離文書モデル** 判定者が所与のページの適合性を判定する際に、そのページの内容だけでなく、それからリンクが張られたページ群のうちプールに含まれるものについてのみ、参考にすることを許可する。これは、適合文書の大部分がプールに含まれることを仮定している。

**ページ単位文書モデル** 判定者が所与のページの適合性を判定する際、そのページ内容すべてに基づく。この仮定は従来と同様である<sup>[5-7]</sup>。

**パッセージ単位文書モデル** 判定者が、ページ中で適合性の根拠を与えるようなパッセージを特定する。

#### 3.4.2 判定用システム

NTCIR-3 WEB で用いた判定用システムは HTTP サーバ上で動作するものであり、CGI を経由して利用することができた。3.3 で述べた通り、判定対象となるプール中の全文書はメタサーチエンジンの戦略により適合度の順にランキングされた。判定対象の個々の文書はプレーンテキストに変換され、そのリンク先ページ群のうちプールに含まれるもののリストが併記された。判定者がある文書の適合性を判定するときには、基本的に、その変換済みテキストと、リンク先文書のテキストを閲覧するように指示した。ただし、判定者は同内容で変換されていないページをも参照することができることとした。

#### 3.4.3 多値適合レベル

判定者は、プール中の個々の文書に対して多値適合レベル、すなわち高適合、適合、部分的適合、不適合のいずれかに判定した。なお、判定者には、各適合レベルに対応する文書の数を意識することのないよう指示した。すなわち、高適合の文書数が非常に少ないとしても、あるがままに判定するよう要求した。本論文では、特に適合のレベルを明記しない限りにおいては、高適合、適合あるいは部分的適合と判定された文書をもって「適合文書」と表記する。

## 4 評価尺度

各参加グループのサーチエンジン・システムによる実行結果リストを評価するため、サーベイ検索タスクでは上位 1,000 件以下、ターゲット検索タスクでは上位 20 件以下に着目した。

評価尺度としては、サーベイ検索タスクに対しては、(i) 精度 (precision) と再現率 (recall) に基づく尺度、及び、(ii) DCG (Discounted Cumulative Gain) を使用した。ターゲット検索タスクに対しては、(i)、(ii) に加えて、(iii) WRR (Weighted Reciprocal Rank) を使用した。上記の評価尺度はページを基本単位として設計されたものであるが、適合判定の前提となる文書モデルが 1 クリック距離文書モデルかページ単位文書モデルのいずれであるかによって、通常は適合文書集合が異なり、最終的に実行結果リストに対する評価結果も異なる。

### 4.1 精度と再現率

精度と再現率に基づく評価尺度として、サーベイ検索タスクにおいては、非補間平均精度 (non-interpolated average precision)、及び、R 精度 (R-precision)<sup>[17]</sup> すなわち検索結果の上位  $|R|$  件における精度を求めた。ここに  $|R|$  は各検索課題に対する適合文書数を示す。また、

再現率 11 点それぞれに対する精度を求めた。上記の評価尺度は、情報検索に関する従来の評価ワークショップで用いられてきたものと同様である<sup>8[5-7]</sup>。

一方、ターゲット検索タスクにおいては、検索結果の上位 10 件における精度を求めた。

ところで、上記のような精度と再現率に基づく評価尺度は、しばしば多値適合レベルを二値の適合レベルに縮退させることを要求する。従って、以下の二種類の適合レベルを仮定した。

**適合レベル 1** ある文書が多値適合レベルにおいて高適合あるいは適合であるとき、二値レベルにおいては適合であるとみなし、そうでなければ不適合であるとみなす。

**適合レベル 2** ある文書が多値適合レベルにおいて高適合、適合あるいは部分的適合であるとき、二値レベルにおいては適合であるとみなし、そうでなければ不適合であるとみなす。

適合レベル 1 及び 2 のそれぞれに従って、実行結果リストごとに、前記の各評価尺度の全検索課題に関する平均値を求めた。

## 4.2 DCG

NTCIR-3 WEB では、多値適合レベルに適した評価尺度の一つである DCG (Discounted Cumulative Gain)<sup>[18,19]</sup>を用いた。DCG は次式で表される。

$$\text{dcg}(i) = \begin{cases} g(1) & \text{if } i = 1 \\ \text{dcg}(i-1) + g(i)/\log_b(i) & \text{otherwise,} \end{cases} \quad (1)$$

$$g(i) = \begin{cases} h & \text{if } d(i) \in H \\ a & \text{if } d(i) \in A \\ b & \text{if } d(i) \in B \end{cases} \quad (2)$$

ここに  $d(i)$  は上位  $i$  ランクの文書を示し、 $H$ 、 $A$  及び  $B$  はそれぞれ高適合、適合、部分的適合と判定された文書集合を示す。式 (2) で定義される利得の度合いとしては、二つの適合レベルのそれぞれについて次のように定めた。

**適合レベル 1**  $(h, a, b) = (3, 2, 0)$ ,

**適合レベル 2**  $(h, a, b) = (3, 2, 1)$ 。

式 (1) における対数関数の底は  $b = 2$  とした。

サーベイ検索タスクにおいては実行結果リストの上位 1,000 ランクまで、ターゲット検索タスクにおいては

<sup>8</sup>当該の評価尺度は、TREC において開発された「trec\_eval」なる評価ツールを用いて計算できる。なお、trec\_eval は次の URL から入手可能である。

([ftp://ftp.cs.cornell.edu/pub/smart/trec\\_eval.v3beta.shar](ftp://ftp.cs.cornell.edu/pub/smart/trec_eval.v3beta.shar))。

上位 10 ランクまでの DCG を求め、ランクごとに全検索課題にわたる平均値を計算した。

## 4.3 WRR

MRR (Mean Reciprocal Rank)<sup>[20]</sup>は、しばしば質問応答システムの評価に用いられ、各質問に対する実行結果リストにおける初出の答えのランクの逆数を、全質問にわたって平均した値で定義される。

我々は、MRR の発想を多値適合レベルに対応するよう拡張した WRR (Weighted Reciprocal Rank) を提案し、ターゲット検索タスクの評価に適用した。WRR は次式で定義される  $wrr(m)$  の全検索課題にわたる平均値として求められる。

$$wrr(m) = \max(r(m)), \quad (3)$$

$$r(m) = \begin{cases} \delta_h / (i - 1/\beta_h) & \text{if } (d(i) \in H \wedge 1 \leq i \leq m) \\ \delta_a / (i - 1/\beta_a) & \text{if } (d(i) \in A \wedge 1 \leq i \leq m) \\ \delta_b / (i - 1/\beta_b) & \text{if } (d(i) \in B \wedge 1 \leq i \leq m) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

ここに  $m$  は実行結果リストにおいて着目するランクの最大値を示す。また、重み係数は、 $\delta_h \in \{1, 0\}$ 、 $\delta_a \in \{1, 0\}$ 、 $\delta_b \in \{1, 0\}$ 、及び、 $\beta_b \geq \beta_a \geq \beta_h > 1$  のそれぞれを満たすものとする。なお、式 (4) において  $\beta_x$  (ただし  $x \in \{h, a, b\}$ ) の値が十分大きいとき、 $(-1/\beta_x)$  の項を省略することができる。

式 (4) における  $m$  の値が 10 の場合における WRR を計算した。このとき、4.1 に示した二つの適合レベルを想定して、 $\delta_x$  を下記のように設定した。

**適合レベル 1**  $(\delta_h, \delta_a, \delta_b) = (1, 1, 0)$ ,

**適合レベル 2**  $(\delta_h, \delta_a, \delta_b) = (1, 1, 1)$ 。

また、本論文では簡単のため、いずれの適合レベルにおいても  $(\beta_h, \beta_a, \beta_b) = (\infty, \infty, \infty)$  と仮定した。

さらに、実行結果リストにおいて着目するランクの最大値が 10 の場合に、適合文書が一つも見つからなかった検索課題のパーセンテージを、二つの適合レベルごとに求めた。

## 4.4 ページ内容の重複を考慮した評価手法

多くのユーザは、サーチエンジンの検索結果において重複したページが出現することを望まない。そこで、重複ページ群が特定できることを前提として、以下のようページ内容の重複を考慮した評価手法を提案する。

- 各実行結果リストにおいて、重複文書群のうち初出の文書については、特別な措置をとらずに他の文書と同様に扱う。

- 重複文書群で二番目以降に出現する文書については、たとえそれが適合と判定された文書であっても、不適合（もしくは部分的適合）として扱う。

結果として、重複文書を含む実行結果リストはペナルティが課せられることになる。

上記のページ重複を考慮した評価手法は、精度・再現率に基づく評価尺度（4.1）や DCG（4.2）と組み合わせることを前提に設計した。ところで、当該評価手法の前提となる重複文書群を特定するため、判定者はプールに含まれる個々の文書に対して適合性だけでなく重複性についても可能な範囲で判定した。また、ページ内容が完全に一致する文書群を機械的に検出することで、前述の人間により判定された重複文書群を補完した。なお、上記の評価手法は、主要な部分が同一と見なされる類似文書群や、ハイパーリンクで何らかの関連性をもって接続された関連文書群を用いることも可能である。さらに、これらを 1 クリック距離文書モデルに基づく適合判定結果と組み合わせることで、Hub 的ページや Authority 的ページを特定することによる Topic Distillation 手法<sup>[16]</sup>の妥当な評価手法になりうると考える。

## 5 評価結果

### 5.1 参加状況

六つの参加グループが不備なく実行結果リストを提出した<sup>9</sup>。また、オーガナイザが、プールにおける適合文書の網羅性を改善することを目的として、参加グループと同じ要領で検索システムの実行結果リストを提出した。参加機関は以下の通りであった（参加機関名の五十音順）。

- 会津大学
- NEC
- 大阪教育大学
- 東京大学・リコー
- 図書館情報大学・産業科学総合研究所
- 奈良先端科学技術大学院大学・通信総合研究所

個々の参加グループは、多様な研究目的のもと NTCIR-3 WEB に従事した。以下に、参加グループ ID とその参加グループが実施した実験の概要を示す（参加グループ ID のアルファベット順）。なお、参加機関と参加グループ ID の対応づけは企業等の参加者に配慮して明示しないこととする。

#### GRACE 確率モデルに基づく疑似適合フィードバック

<sup>9</sup>七つの参加グループが実行結果リストを提出したものの、一つの参加グループは提出期日までに、全検索課題のほぼ半数についてしか実行結果を提出しなかったため、評価の対象に含めることができなかった。

ク手法、及び、Kleinberg らの HITS アルゴリズム<sup>[16]</sup>に基づくリンク解析を用いた再ランキング手法に関する実験が行なわれた。

**K3100** ページのテキスト情報だけでなく、そのページもしくはサイトを参照しているアンカーテキストを用いた検索手法に関する実験が行なわれた。

**NAICR** スコア正規化を伴わない分散情報検索手法に関する実験が行なわれた。検索機能には確率モデルに基づく OKAPI アルゴリズムが使用された。

**OASIS** 分散型情報検索手法に関する実験が行なわれた。検索機能にはベクトル空間モデルに基づく手法が用いられた。

**OKSAT** *N*-gram に基づくインデクシング手法に関する実験が行なわれた。検索機能には確率モデルに基づく手法が用いられた。

**ORGREF** オーガナイザがプールを拡張するため、近接演算や TF-IDF 重み付け手法によるランキングの機能を備えたプール演算型検索システムを実行した。

**UAIFI** 音声認識機能及びテキスト検索機能が統合された音声入力型検索の実験が行なわれた。テキスト検索機能は確率モデルに基づく手法が用いられた。テキスト検索の実行結果はプールに加えたが、音声入力型検索の実行結果は加えなかった。

なお、実行結果の提出状況は表 5 に示す通りである<sup>10</sup>。

### 5.2 実験の条件

実行結果リストを評価する際には、 $\{PL_1, PL_2\} \times \{DM_1, DM_2\} \times \{RL_1, RL_2\}$  のそれぞれの組合せを想定した。これらの各項目について以下に説明する。

#### プーリング手法 (3.3)

( $PL_1$ ) 全実行結果によるプーリング (100GB 及び 10GB の文書データに対する実行結果に基づくプーリング)

( $PL_2$ ) 10GB の文書データに係る実行結果のみによるプーリング

#### 文書モデル (3.4.1)

( $DM_1$ ) 1 クリック距離文書モデル

( $DM_2$ ) ページ単位文書モデル

#### 適合レベル (4 章)

( $RL_1$ ) 適合レベル 1

<sup>10</sup>個々の検索システムの詳細については、以下の論文集を参照されたい。(http://research.nii.ac.jp/ntcir/publication1-en.html)。

### ( $RL_2$ ) 適合レベル 2

ただし、100GB の文書データに対する実行結果に対しては  $PL_1$ 、10GB の文書データに対する実行結果に対しては  $PL_2$  のプーリング手法に基づく適合判定結果リストを用いて評価した。例えば、ある実行結果リストが 100GB の文書データに対して検索実行したものである場合、 $\{PL_1\} \times \{DM_1, DM_2\} \times \{RL_1, RL_2\}$  の四通りの条件で評価値を算出することにした。

## 5.3 評価結果と考察

5.1 に概要を示した実験システムによる実行結果リストの各々に対し、4 章で述べた各種評価尺度を用いて、5.2 に示した条件のもとで有効性を評価した。サーベイ検索タスク及びターゲット検索タスクの評価結果をそれぞれ表 6, 7 に示す。これらの表では、タスク、自動システム・対話型システムの区別<sup>11</sup>及び検索課題中で使用されたフィールドごとに、サーベイ検索タスクについては非補間平均精度順、ターゲット検索タスクについては上位 10 件における精度順に、実行結果 ID が並べられている。ただし、いずれもページ単位文書モデル ( $DM_2$ ) かつ適合レベル 1 ( $RL_1$ ) に基づいた評価値を 47 の検索課題にわたって平均したものを基準とした。

また、100GB の大規模文書データに対する実行結果に関する再現率・精度グラフ及び DCG のグラフを図 4, 5, 6 及び 7 に示す。また、10GB の小規模文書データに対する実行結果に関するグラフを図 8, 9, 10 及び 11 に示す。これらのグラフにおいては、すべての実行結果は自動システムによるものである。対話型システムの実行結果も提出されたが、表 5 に示した通り、少数であったため当該グラフの掲載を割愛する。各グラフにおいて、凡例は実行結果 ID を示しており、サーベイ検索タスクについては非補間平均精度順、ターゲット検索タスクについては上位 10 件における精度順に並べられている。なお、グラフで用いられた実行結果 ID は、参加グループごとで最も性能の高かったものである。

さらに、トピック検索タスク（サーベイ検索タスクの一部）及びターゲット検索タスクの実行結果を用いて、1 クリック距離文書モデル ( $DM_1$ ) 及びページ単位文書モデル ( $DM_2$ ) のそれぞれに基づき、評価尺度ごとのシステム順位（性能が高いものから順）を求めた。それぞれの結果を表 8 及び 9 に示す。ただし、これらの表において示された実行結果 ID は、100GB の大規模文書データに対するもののみであり、適合レベル 1 ( $RL_1$ )

<sup>11</sup> 自動システム及び対話型システムの定義については、2.1.1 において説明した通りである。

に従った結果に限定した。また、対話型システム (2.1) の実行結果リストは除外した。

表 8 及び 9 において下線を付した実行結果 ID は、文書内容に基づいた検索手法とリンク解析を組み合わせたものであり、それ以外はリンク解析を用いていない。そこで、ターゲット検索タスクに関する表 9 において、検索課題の TITLE のみ用いた実行結果と DESC のみを用いたもののうち、リンク解析を用いた実行結果の順位分布を観察する。このとき、TITLE のように短いクエリを用いた検索実行の方が、検索結果上位においては、リンク解析が有効であることが示唆される<sup>12</sup>。さらに、表 8 及び 9 における TITLE を用いた実行結果（上段）について、リンク解析を用いた実行結果の順位分布を観察すると、サーベイ検索タスクよりもターゲット検索タスクの方が、リンク解析が効果的であることが示唆される。

## 6 おわりに

ページ間のリンク関係を考慮した Web 検索システムの評価手法を提案した。NTCIR-3 WEB における評価結果から、以下の事実を示唆する結果を確認した。すなわち、(i) 短いクエリを用いた検索結果のランキング上位において特にリンク解析が有効であることと、(ii) 検索結果を網羅的に調査するようなユーザモデルよりも検索結果上位にのみ着目するユーザモデルにおいてリンク解析が効果的であることである。評価結果に関する詳細かつ定量的な分析は今後の課題である。本論文ではまた、ページ内容の重複を考慮した評価手法を提案した。その適用結果分析、及び、内容的な類似文書群やリンクによる関連文書群を用いた評価に関して、今後さらに検討を行う。

## 謝辞

本研究は、文部科学省科学研究費補助金特定領域研究「情報学」（課題番号 13224087）及び若手研究（課題番号 14780339）による。NTCIR-3 Web 検索タスクのすべての参加者の協力に感謝する。また、Web 検索タスク・アドバイザー委員及び国立情報学研究所・安達教授からは有益な助言を得た。

## 文献

- [1] Eguchi, K.; Oyama, K.; Ishida, E.; Kuriyama, K.; Kando, N., "Evaluation design of Web Retrieval Task

<sup>12</sup> 実行結果 GRACE-LA1-1 及び GRACE-LB-1 が、リンク解析を伴わないもの上位に順位付けられているが、これらは「GRACE」で始まる同グループによる他の実行結果とはシステムのパラメータが異なるため、例外的な検索実行と言える。

- in the Third NTCIR Workshop,” *The 11th International World Wide Web Conference (WWW2002)*, No. poster-22, Honolulu, Hawaii, USA (2002).
- [2] Eguchi, K.; Oyama, K.; Ishida, E.; Kando, N.; Kuriyama, K., “Overview of the Web Retrieval Task at the Third NTCIR Workshop,” *Proceedings of the 3rd NTCIR Workshop on Research in Information Retrieval, Automatic Text Summarization and Question Answering*, Tokyo, Japan (2003).
- [3] “NTCIR Web Task,” (<http://research.nii.ac.jp/ntcir/web/>).
- [4] (特集)「NTCIR: 情報アクセスに関わるテキスト処理技術の評価ワークショップ」, 人工知能学会誌, Vol. 17, No. 3 (2002).
- [5] Voorhees, E. M.; Harman, D., “Overview of the Sixth Text REtrieval Conference (TREC-6),” *Proceedings of the 6th Text REtrieval Conference (TREC-6)*, NIST Special Publication 500-240, pp. 1–24 (1997).
- [6] Hawking, D.; Voorhees, E.; Craswell, N.; Bailey, P., “Overview of the TREC-8 Web Track,” *Proceedings of the 8th Text REtrieval Conference (TREC-8)*, NIST Special Publication 500-246, pp. 131–149 (1999).
- [7] Kando, N.; Kuriyama, K.; Nozue, T.; Eguchi, K.; Kato, H.; Hidaka, S., “Overview of IR tasks at the First NTCIR Workshop,” *Proceedings of the 1st NTCIR Workshop on Research in Japanese Text Retrieval and Term Recognition*, Tokyo, Japan, pp. 11–22 (1999).
- [8] Buckley, C., “TREC 6 High-Precision Track,” *Proceedings of the 6th Text REtrieval Conference (TREC-6)*, NIST Special Publication 500-240, pp. 69–71 (1996).
- [9] Buckley, C., “The TREC 7 High Precision Track,” *Proceedings of the 7th Text REtrieval Conference (TREC-7)*, NIST Special Publication 500-242, pp. 57–63 (1997).
- [10] Hawking, D., “Overview of the TREC-9 Web Track,” *Proceedings of the 9th Text REtrieval Conference (TREC-9)*, NIST Special Publication 500-249, pp. 97–112 (2000).
- [11] Hawking, D.; Craswell, N., “Overview of the TREC-2001 Web Track,” *Proceedings of the 10th Text REtrieval Conference (TREC-2001)*, NIST Special Publication 500-250, pp. 61–68 (2001).
- [12] Buckley, C.; Voorhees, E., “Evaluating evaluation measure stability,” *Proceedings of the 23rd Annual International ACM SIGIR Conference*, Athens, Greece, pp. 33–40 (2000).
- [13] Fujii, A.; Itou, K., “Evaluating speech-driven IR in the NTCIR-3 Web Retrieval Task,” *Proceedings of the 3rd NTCIR Workshop on Research in Information Retrieval, Automatic Text Summarization and Question Answering*, Tokyo, Japan (2003).
- [14] Aslam, J. A.; Montague, M., “Models for metasearch,” *Proceedings of the 24th Annual International ACM SIGIR Conference*, New Orleans, Louisiana, USA, pp. 276–284 (2001).
- [15] Kuriyama, K.; Kando, N.; Nozue, T.; Eguchi, K., “Pooling for a large-scale test collection: An analysis of the search results from the First NTCIR Workshop,” *Information Retrieval*, Vol. 5, No. 1, pp. 41–59 (2002).
- [16] Kleinberg, J. M., “Authoritative sources in a hyperlinked environment,” *Proceedings of the 9th ACM SIAM Symposium on Discrete Algorithms*, San Francisco, California, USA (1998).
- [17] Baeza-Yates, R.(ed.), *Modern Information Retrieval*, Addison-Wesley (1999).
- [18] Järvelin, K.; Kekäläinen, J., “IR evaluation methods for retrieving highly relevant documents,” *Proceedings of the 23rd Annual International ACM SIGIR Conference*, Athens, Greece, pp. 41–48 (2000).
- [19] Voorhees, E. M., “Evaluation by highly relevant documents,” *Proceedings of the 24th Annual International ACM SIGIR Conference*, New Orleans, Louisiana, pp. 74–82 (2001).
- [20] Voorhees, E. M., “The TREC-8 Question Answering Track report,” *Proceedings of the 8th Text REtrieval Conference (TREC-8)*, NIST Special Publication 500-246, pp. 77–82 (1999).

表 1: NW100G-01 に関する基礎的な統計

# of crawled sites	97,561 (# of aliased sites: 3,285 is not included)
maximum # of pages within a site	1,300
# of crawled pages ( <i>i.e.</i> , # of pages included in the document data for providing)	11,038,720 (# of aliased sites: 419,709 is not included)
# of pages for searching (of which existence are confirmed)	15,364,404
# of links (connected from the crawled pages to the pages for searching)	64,365,554

表 2: NW100G-01 のリンクに関する統計

		# of links	# of pages	# of sites
1.	# of links connected from the crawled pages (only text files)	78,175,556		
2.	# of links connected to the pages for searching (of which existence are confirmed)	64,365,554		
3.	# of links connected to the pages not for searching (of which existence could not be confirmed)	13,810,002 (3./1.=0.176)		
2.	# of links that connected to the pages for searching (of which existence are confirmed), and # of their destination pages	64,365,554	15,182,651	
2-1-1.	# of links that are closed in the crawled pages, and # of their destination pages	53,928,019	10,857,715	
2-1-2.	# of links that are not closed in the crawled pages, and # of their destination pages	10,437,535	4,324,936	
2-2-1.	# of links that connected to the pages within the same sites, and # of their destination pages	56,673,429	14,218,861	
2-2-1-1.	In 2-2-1., # of links that connected to the crawled pages, and # of their destination pages	49,960,354	10,800,231	
2-2-1-2.	In 2-2-1., # of links that connected outside the crawled pages, and # of their destination pages	6,713,075	3,418,630	
2-2-2.	# of links that connected to the pages on another crawled site, and # of their destination pages	5,563,383	729,754	
2-2-2-1.	In 2-2-2., # of links that connected to the crawled pages, and # of their destination pages	3,967,665	344,487	
2-2-2-2.	In 2-2-2., # of links that connected outside the crawled pages, and # of their destination pages	1,595,718	385,267	
2-2-3.	# of links that connected outside the crawled sites, and # of their destination pages and sites	2,128,742	600,437	237,432
3.	# of links connected to the pages not for searching (of which existence could not be confirmed), and # of their destination pages	13,810,002 (3./1.=0.176)		
3-1-1.	# of links connected to the pages within the same sites, and # of their destination pages	8,525,716	5,863,863	
3-1-2.	# of links and pages that connected to the pages on another crawled site, and # of their destination pages	1,789,643	687,553	
3-1-3.	# of links, pages and sites that connected outside the crawled sites, and # of their destination pages and sites	3,494,643	1,047,306	217,554

表 3: NW10G-01 に関する基礎的な統計

# of crawled sites	97,561 (# of aliased sites: 3,285 is not included)
maximum # of pages within a site	20
# of crawled pages ( <i>i.e.</i> , # of pages included in the document data for providing)	1,445,466 (# of aliased sites: 141,574 is not included)
# of pages for searching (of which existence are confirmed)	4,849,714
# of links (connected from the crawled pages to the pages for searching)	9,885,538

表 4: NW10G-01 のリンクに関する統計

		# of links	# of pages	# of sites
1.	# of links connected from the crawled pages (only text files)	11,642,167		
2.	# of links connected to the pages for searching (of which existence are confirmed)	9,885,538		
3.	# of links connected to the pages not for searching (of which existence could not be confirmed)	1,756,629 (3./1.=0.150)		
2.	# of links that connected to the pages for searching (of which existence are confirmed), and # of their destination pages	9,885,538	4,810,115	
2-1-1.	# of links that are closed in the crawled pages, and # of their destination pages	4,978,298	1,405,928	
2-1-2.	# of links that are not closed in the crawled pages, and # of their destination pages	4,907,240	3,404,187	
2-2-1.	# of links that connected to the pages within the same sites, and # of their destination pages	8,427,690	4,461,635	
2-2-1-1.	In 2-2-1., # of links that connected to the crawled pages, and # of their destination pages	4,303,577	1,349,118	
2-2-1-2.	In 2-2-1., # of links that connected outside the crawled pages, and # of their destination pages	4,124,113	3,112,517	
2-2-2.	# of links that connected to the pages on another crawled site, and # of their destination pages	1,084,263	193,122	
2-2-2-1.	In 2-2-2., # of links that connected to the crawled pages, and # of their destination pages	674,721	122,070	
2-2-2-2.	In 2-2-2., # of links that connected outside the crawled pages, and # of their destination pages	409,542	71,052	
2-2-3.	# of links that connected outside the crawled sites, and # of their destination pages and sites	373,585	155,358	73,916
3.	# of links connected to the pages not for searching (of which existence could not be confirmed), and # of their destination pages	1,756,629 (3./1.=0.150)		
3-1-1.	# of links connected to the pages within the same sites, and # of their destination pages	822,318	597,855	
3-1-2.	# of links and pages that connected to the pages on another crawled site, and # of their destination pages	934,311	442,940	
3-1-3.	# of links, pages and sites that connected outside the crawled sites, and # of their destination pages and sites	620,663	262,735	80,169

表 5: サーベイ検索タスク及びターゲット検索タスクに関する実行結果の提出状況

Task	QMethod	TopicPart(#runs)	LinkInfo(#runs)	Task	QMethod	TopicPart(#runs)	LinkInfo(#runs)
I-A1	automatic	T (14)	cont (11) link&cont (3)	II-A1	automatic	T (14)	cont (12) link&cont (2)
	automatic	D (10)	cont (7) link&cont (3)		automatic	D (12)	cont (10) link&cont (2)
	interactive	TD (1)	cont (1) link&cont (0)		interactive	TD (1)	cont (1) link&cont (0)
I-A2	automatic	TR[1] (6)	cont (6) link&cont (0)	II-A2	automatic	TR[1] (6)	cont (6) link&cont (0)
	automatic	T (1)	cont (1) link&cont (0)		automatic	T (1)	cont (1) link&cont (0)
	automatic	D (1)	cont (1) link&cont (0)		automatic	D (1)	cont (1) link&cont (0)
I-B	automatic	T (13)	cont (10) link&cont (3)	II-B	automatic	T (13)	cont (11) link&cont (2)
	automatic	D (9)	cont (6) link&cont (3)		automatic	D (9)	cont (7) link&cont (2)

**Task** はタスクの種類を示す。「I」及び「II」はそれぞれ 100GB、10GB の文書データを用いたことを示し、「A1」、「A2」、「B」はそれぞれトピック検索タスク、類書検索タスク及びターゲット検索タスクを示す。

**QMethod** は自動システム、対話型システムのいずれであることを示す。

**TopicPart** は検索課題においてクエリ生成の際に使用した部分を示す。「T」、「D」、「R[n]」はそれぞれ、TITLE、DESC 及び RDOC で特定された  $n$  番目の文書を示す。

**LinkInfo** はシステムが Web ページのリンク情報を使用したかどうかを示す。「link&cont」はリンク及びページ内容の両方を使用したことを示し、「cont」はページ内容のみを使用したことを示す。

表 6: サーベイ検索タスクの評価結果

Task	QMethod	Topic	RunID	LinkInfo	$DM_2 \& RL_1$				$DM_2 \& RL_2$				$DM_1 \& RL_1$			
					aprec	rprec	dcg(100)	dcg(1K)	aprec	rprec	dcg(100)	dcg(1K)	aprec	rprec	dcg(100)	dcg(1K)
I-A1	automatic	T	GRACE-LA1-1	cont	0.1506	0.1707	7.3478	13.0492	0.2109	0.2345	9.2695	16.4611	0.1551	0.1986	9.3798	17.3195
	automatic	T	GRACE-LA1-2	link&cont	0.1489	0.1739	7.2873	12.8568	0.2061	0.2361	9.1492	16.1874	0.1548	0.2019	9.3503	17.1396
	automatic	T	OKSAT-F-04	cont	0.1151	0.1469	6.3920	11.0055	0.1520	0.1972	7.8993	13.7063	0.1104	0.1536	7.3605	13.7549
	automatic	T	ORGREF-LA1-6	cont	0.1111	0.1571	6.0655	10.6693	0.1447	0.1939	7.5225	13.4621	0.1111	0.1603	7.0813	13.5672
	automatic	T	K3100-05	link&cont	0.0951	0.1449	5.8189	10.3197	0.1300	0.1821	7.2762	13.0700	0.1013	0.1560	7.3532	13.8044
	automatic	T	K3100-06	link&cont	0.0937	0.1369	5.7608	10.2257	0.1281	0.1747	7.1691	12.9111	0.1002	0.1486	7.2143	13.5953
	automatic	T	NAICR-I-A1-4	cont	0.0875	0.1298	5.3408	10.0245	0.1135	0.1751	6.8675	12.5391	0.0904	0.1443	6.7687	13.3463
	automatic	T	ORGREF-LA1-5	cont	0.0833	0.1176	4.7709	9.5219	0.1135	0.1623	6.2605	12.0449	0.0868	0.1280	5.7715	12.3219
	automatic	T	UAIFI3	cont	0.0815	0.1204	5.1242	10.1186	0.1346	0.1913	6.9208	13.1294	0.0981	0.1477	7.0491	14.0841
	automatic	T	UAIFI4	cont	0.0808	0.1128	4.9741	10.2023	0.1280	0.1837	6.6714	13.0869	0.0938	0.1432	6.7929	13.9697
	automatic	T	ORGREF-LA1-3	cont	0.0654	0.1039	4.5514	8.0706	0.1009	0.1560	6.0680	10.4609	0.0685	0.1164	5.8615	10.6385
	automatic	T	ORGREF-LA1-1	cont	0.0596	0.0998	4.2086	7.7335	0.0879	0.1339	5.2586	9.7006	0.0588	0.1030	4.7826	9.8060
	automatic	T	ORGREF-LA1-4	cont	0.0526	0.0846	4.1272	7.2806	0.0827	0.1350	5.5365	9.4269	0.0554	0.0983	5.2200	9.5099
	automatic	T	ORGREF-LA1-2	cont	0.0449	0.0728	3.4951	7.2076	0.0653	0.1096	4.3731	8.9193	0.0455	0.0792	4.0205	9.0462
	automatic	D	GRACE-LA1-4	link&cont	0.1555	0.1856	8.0395	13.7795	0.2088	0.2310	9.7968	16.7512	0.1488	0.1918	9.5314	16.9439
	automatic	D	GRACE-LA1-3	cont	0.1548	0.1799	8.0356	13.7948	0.2082	0.2302	9.7899	16.7633	0.1479	0.1906	9.5193	16.9506
	automatic	D	OKSAT-F-06	cont	0.1236	0.1473	6.4498	11.4801	0.1559	0.1902	7.7792	14.2475	0.1206	0.1607	7.4291	14.3643
	automatic	D	NAICR-I-A1-3	cont	0.0907	0.1363	5.7470	10.7955	0.1148	0.1798	7.1050	13.1694	0.0935	0.1529	7.0400	13.9398
	automatic	D	K3100-07	link&cont	0.0897	0.1164	5.2714	8.7714	0.1162	0.1524	6.4217	11.0198	0.0928	0.1325	6.6369	11.4676
	automatic	D	UAIFI1	cont	0.0855	0.1247	5.6399	10.5625	0.1257	0.1809	7.1328	13.1817	0.0982	0.1547	7.5055	14.1661
automatic	D	UAIFI2	cont	0.0843	0.1256	5.5013	10.5842	0.1184	0.1772	6.9297	13.1267	0.0928	0.1461	7.2712	14.1109	
automatic	D	NAICR-I-A1-2	cont	0.0756	0.0990	3.9179	7.6978	0.0945	0.1257	4.7994	9.4511	0.0845	0.1102	5.2435	10.2021	
automatic	D	K3100-08	link&cont	0.0737	0.0991	4.5122	8.9811	0.0980	0.1368	5.5802	11.2223	0.0756	0.1150	5.6447	11.3948	
automatic	D	NAICR-I-A1-1	cont	0.0736	0.0929	3.4572	6.9776	0.0882	0.1222	4.2856	8.5957	0.0841	0.1083	4.9067	9.3417	
interactive	TD	OKSAT-F-02	cont	0.1238	0.1535	6.9065	12.6003	0.1674	0.2032	8.3054	15.3811	0.1191	0.1682	8.0375	15.5999	
I-A2	automatic	TR[1]	GRACE-LA2-3	cont	0.1977	0.2229	10.0645	15.7602	0.2546	0.2770	11.9122	19.1600	0.1998	0.2464	12.3606	20.3841
	automatic	TR[1]	GRACE-LA2-2	cont	0.1966	0.2207	9.9764	15.4430	0.2553	0.2808	11.9402	18.8438	0.1990	0.2291	12.2912	20.1235
	automatic	TR[1]	GRACE-LA2-1	cont	0.1913	0.2092	9.7183	14.9258	0.2430	0.2762	11.5838	18.1920	0.1959	0.2317	11.9774	19.7194
	automatic	TR[1]	GRACE-LA2-4	cont	0.1769	0.2077	9.0702	15.0199	0.2357	0.2570	10.9365	18.4044	0.1811	0.2275	11.2150	19.3453
	automatic	TR[1]	NAICR-I-A2-3	cont	0.0873	0.1073	5.5473	9.6440	0.1000	0.1415	6.3849	11.3242	0.0886	0.1281	7.2916	12.5037
	automatic	TR[1]	NAICR-I-A2-4	cont	0.0860	0.1183	4.9069	8.3797	0.1059	0.1352	5.7773	10.0328	0.0943	0.1344	7.0367	11.4648
	automatic	T	NAICR-I-A2-2	cont	0.0823	0.1214	5.3067	9.9524	0.1102	0.1718	6.7786	12.4229	0.0898	0.1445	6.7268	13.3339
automatic	D	NAICR-I-A2-1	cont	0.0908	0.1356	5.7532	10.7333	0.1168	0.1788	7.1618	13.1620	0.0935	0.1537	7.0990	13.8678	
II-A1	automatic	T	GRACE-SA1-2	cont	0.2164	0.1966	4.1269	5.0773	0.2433	0.2429	5.1583	6.3925	0.1775	0.1935	5.8389	7.7713
	automatic	T	GRACE-SA1-1	cont	0.2056	0.1978	3.7720	4.8112	0.2266	0.2372	4.7265	6.0589	0.1768	0.1912	5.6464	7.6963
	automatic	T	OKSAT-F-03	cont	0.1749	0.1702	3.0193	3.8194	0.1936	0.2069	3.8146	4.9079	0.1271	0.1496	4.2058	5.8870
	automatic	T	ORGREF-SA1-6	cont	0.1747	0.1769	3.0717	3.8409	0.1873	0.1918	3.8668	4.8973	0.1478	0.1682	4.6581	6.1227
	automatic	T	UAIFI11	cont	0.1430	0.1396	2.5043	3.4155	0.1491	0.1663	3.2681	4.4459	0.1201	0.1468	4.2076	5.9526
	automatic	T	UAIFI12	cont	0.1342	0.1281	2.4010	3.3096	0.1407	0.1566	3.1988	4.3501	0.1125	0.1302	4.0756	5.8087
	automatic	T	K3100-01	link&cont	0.1333	0.1341	2.7853	3.5351	0.1641	0.1709	3.6344	4.6416	0.1145	0.1364	4.3749	5.7201
	automatic	T	K3100-02	link&cont	0.1297	0.1335	2.7752	3.5293	0.1613	0.1702	3.6193	4.6315	0.1137	0.1359	4.3485	5.7052
	automatic	T	ORGREF-SA1-1	cont	0.1243	0.1465	2.3034	3.1777	0.1298	0.1553	2.8820	4.0060	0.0951	0.1272	3.2905	4.8968
	automatic	T	ORGREF-SA1-5	cont	0.1214	0.1158	2.6300	3.2386	0.1304	0.1358	3.2642	4.1666	0.1217	0.1510	4.2670	5.5502
	automatic	T	ORGREF-SA1-3	cont	0.1101	0.1109	1.8230	2.4912	0.1041	0.1310	2.3043	3.2148	0.0885	0.1131	2.9778	4.3754
	automatic	T	ORGREF-SA1-2	cont	0.1101	0.1360	2.1614	2.9670	0.1102	0.1408	2.6588	3.7180	0.0863	0.1203	3.0921	4.6049
	automatic	T	NAICR-II-A1-4	cont	0.1028	0.0971	2.5783	3.3424	0.1268	0.1577	3.2079	4.1906	0.0965	0.1331	3.8590	5.2079
	automatic	T	ORGREF-SA1-4	cont	0.0803	0.0892	1.5944	2.2843	0.0882	0.1204	2.0438	2.9751	0.0731	0.0897	2.6996	4.0951
	automatic	D	GRACE-SA1-4	cont	0.2264	0.2368	3.8688	4.7892	0.2458	0.2535	4.7648	5.9638	0.1694	0.2031	5.3314	7.2810
	automatic	D	GRACE-SA1-3	cont	0.1970	0.2168	3.7560	4.7709	0.2241	0.2435	4.6328	5.9133	0.1604	0.1915	5.3131	7.4203
	automatic	D	OKSAT-F-05	cont	0.1859	0.1935	3.2372	4.0581	0.2228	0.2180	4.1085	5.2259	0.1397	0.1591	4.4914	6.2484
	automatic	D	UAIFI9	cont	0.1408	0.1368	2.7048	3.7329	0.1577	0.1605	3.4496	4.7394	0.1259	0.1503	4.4458	6.3179
	automatic	D	UAIFI10	cont	0.1330	0.1247	2.6003	3.6667	0.1529	0.1574	3.3869	4.6950	0.1161	0.1426	4.2757	6.1605
	automatic	D	NAICR-II-A1-2	cont	0.1286	0.1084	2.0882	3.2505	0.1460	0.1597	2.5680	4.0171	0.0971	0.1287	2.9105	4.8595
automatic	D	K3100-03	link&cont	0.1244	0.1070	2.3855	2.8715	0.1588	0.1582	3.1519	3.7956	0.1120	0.1357	3.6493	4.4917	
automatic	D	NAICR-II-A1-1	cont	0.1234	0.1016	2.1688	3.3191	0.1404	0.1514	2.6938	4.1096	0.0871	0.1148	2.9350	4.9196	
automatic	D	K3100-04	link&cont	0.1213	0.1152	2.6217	3.2188	0.1530	0.1615	3.3607	4.2460	0.0993	0.1216	3.7586	4.9564	
automatic	D	NAICR-II-A1-3	cont	0.1183	0.1170	2.8585	4.0558	0.1356	0.1629	3.3936	4.9340	0.1044	0.1310	4.3235	6.3616	
automatic	D	OASIS12	cont	0.0406	0.0366	1.2292	2.0637	0.0481	0.0479	1.4979	2.5557	0.0296	0.0389	1.6223	3.0087	
automatic	D	OASIS11	cont	0.0385	0.0366	1.1471	1.9719	0.0453	0.0435	1.3921	2.4401	0.0275	0.0378	1.4870	2.8542	
interactive	TD	OKSAT-F-01	cont	0.2236	0.2305	3.9013	4.7711	0.2436	0.2513	4.8161	5.9357	0.1635	0.1896	5.2637	7.0661	
II-A2	automatic	TR[1]	GRACE-SA2-2	cont	0.2810	0.2676	4.8953	5.8038	0.3310	0.3173	5.8673	7.0763	0.2264	0.2396	6.8790	8.8193
	automatic	TR[1]	GRACE-SA2-3	cont	0.2781	0.2497	5.0227	5.8906	0.3410	0.3201	6.0741	7.2131	0.2284	0.2502	6.9720	8.9581
	automatic	TR[1]	GRACE-SA2-1	cont	0.2637	0.2453	4.5736	5.5302	0.3153	0.3092	5.5584	6.7654	0.2133	0.2441	6.6967	8.5307
	automatic	TR[1]	GRACE-SA2-4	cont	0.2361	0.2395	4.4749	5.2703	0.2696	0.2825	5.4708	6.5560	0.2017	0.2301	6.2571	8.1387
	automatic	TR[1]	NAICR-II-A2-2	cont	0.1899	0.18										

表 7: ターゲット検索タスクの評価結果

Task	QMethod	Topic	RunID	LinkInfo	$DM_2 \& RL_1$				$DM_2 \& RL_2$				$DM_1 \& RL_1$			
					prec(10)	deg(10)	wrr(10)	%nf(10)	prec(10)	deg(10)	wrr(10)	%nf(10)	prec(10)	deg(10)	wrr(10)	%nf(10)
I-B	automatic	T	GRACE-LB-1	cont	0.2213	2.4940	0.3618	0.2979	0.3511	3.2212	0.4767	0.2979	0.2532	2.9598	0.4060	0.2553
	automatic	T	GRACE-LB-2	link&cont	0.2106	2.3901	0.3581	0.2766	0.3404	3.0686	0.4627	0.2553	0.2447	2.8802	0.4030	0.2553
	automatic	T	K3100-14	link&cont	0.2106	2.3196	0.3544	0.2553	0.2830	2.7305	0.4553	0.1915	0.2447	2.7441	0.3691	0.2553
	automatic	T	K3100-13	link&cont	0.2085	2.3072	0.3796	0.1915	0.2936	2.7736	0.4822	0.1277	0.2426	2.7400	0.3983	0.1915
	automatic	T	ORGFREF-LB-6	cont	0.1915	2.2709	0.3346	0.3191	0.2745	2.7288	0.4301	0.2766	0.2170	2.4918	0.3422	0.3191
	automatic	T	UAIF18	cont	0.1468	1.6883	0.2751	0.3404	0.2702	2.3850	0.4141	0.3191	0.1979	2.3787	0.3766	0.2553
	automatic	T	UAIF17	cont	0.1426	1.5409	0.2338	0.3404	0.2830	2.3470	0.4058	0.3191	0.1872	2.1654	0.3506	0.2766
	automatic	T	NAICR-I-B-1	cont	0.1383	1.7538	0.2992	0.3830	0.2085	2.0991	0.3370	0.2979	0.1787	2.2595	0.3446	0.3404
	automatic	T	ORGFREF-LB-5	cont	0.1340	1.6262	0.2943	0.4255	0.2362	2.1592	0.4036	0.3191	0.1681	1.9934	0.3319	0.3830
	automatic	T	ORGFREF-LB-3	cont	0.1213	1.4486	0.3016	0.4043	0.2319	2.1360	0.4534	0.3191	0.1638	1.9411	0.3563	0.3617
	automatic	T	ORGFREF-LB-4	cont	0.1106	1.2752	0.2646	0.4681	0.2021	1.8616	0.4053	0.3617	0.1447	1.6978	0.3368	0.3830
	automatic	T	ORGFREF-LB-1	cont	0.1064	1.2370	0.1986	0.6383	0.1596	1.4746	0.2311	0.5319	0.1128	1.2778	0.2008	0.6170
	automatic	T	ORGFREF-LB-2	cont	0.0809	0.9163	0.1148	0.7021	0.1106	1.0343	0.1369	0.5957	0.0851	0.9510	0.1169	0.6809
	automatic	D	GRACE-LB-4	link&cont	0.2340	2.9319	0.4214	0.3191	0.3340	3.5002	0.5020	0.2979	0.2681	3.3213	0.4634	0.2553
	automatic	D	GRACE-LB-3	cont	0.2340	2.9707	0.4204	0.3617	0.3298	3.5222	0.5023	0.3191	0.2681	3.3602	0.4589	0.2979
	automatic	D	UAIF15	cont	0.1851	1.9955	0.2543	0.3617	0.2894	2.5765	0.3832	0.2553	0.2319	2.6208	0.3446	0.2553
	automatic	D	UAIF16	cont	0.1830	2.0094	0.3008	0.3830	0.2745	2.5031	0.3910	0.2979	0.2319	2.6879	0.3808	0.2340
	automatic	D	NAICR-I-B-2	cont	0.1596	1.8877	0.3538	0.3191	0.2489	2.3694	0.4682	0.2340	0.1872	2.2507	0.4134	0.2766
	automatic	D	K3100-15	link&cont	0.1574	1.9228	0.3542	0.3830	0.2298	2.3331	0.4892	0.2128	0.1787	2.2030	0.3742	0.3191
	automatic	D	NAICR-I-B-3	cont	0.1319	1.5588	0.2267	0.5745	0.1787	1.8170	0.2776	0.5532	0.1723	2.0386	0.2574	0.5532
automatic	D	NAICR-I-B-4	cont	0.1319	1.5209	0.2213	0.5745	0.1809	1.7859	0.2848	0.5106	0.1681	2.0232	0.2678	0.5319	
automatic	D	K3100-16	link&cont	0.1191	1.4750	0.2881	0.4468	0.1809	1.8565	0.4211	0.3191	0.1426	1.7639	0.3086	0.3830	
II-B	automatic	T	GRACE-SB-2	cont	0.1745	2.3978	0.4222	0.3617	0.2638	2.9225	0.5197	0.2553	0.2255	2.9720	0.4517	0.3191
	automatic	T	GRACE-SB-1	cont	0.1681	2.1770	0.4297	0.2553	0.2511	2.6727	0.5454	0.1702	0.2191	2.8935	0.5146	0.2128
	automatic	T	ORGFREF-SB-6	cont	0.1277	1.8553	0.4331	0.2979	0.1936	2.2583	0.5110	0.2553	0.1617	2.3035	0.4998	0.2766
	automatic	T	K3100-09	link&cont	0.1170	1.6852	0.3987	0.2766	0.1809	2.0590	0.4764	0.1915	0.1681	2.2144	0.4223	0.2553
	automatic	T	K3100-10	link&cont	0.1170	1.6844	0.3880	0.2766	0.1787	2.0512	0.4658	0.1915	0.1681	2.2087	0.4117	0.2553
	automatic	T	ORGFREF-SB-5	cont	0.1149	1.3497	0.2431	0.4043	0.1638	1.6300	0.3142	0.3191	0.1681	2.0671	0.3433	0.3191
	automatic	T	NAICR-II-B-4	cont	0.1021	1.4110	0.3314	0.3404	0.1596	1.7312	0.4254	0.2766	0.1383	1.7645	0.3463	0.3404
	automatic	T	ORGFREF-SB-1	cont	0.0979	1.4558	0.3333	0.3617	0.1489	1.7820	0.4081	0.2766	0.1234	1.7580	0.3821	0.2979
	automatic	T	UAIF15	cont	0.0915	1.1879	0.2889	0.4468	0.1532	1.5306	0.3538	0.3830	0.1319	1.7779	0.3931	0.3830
	automatic	T	UAIF16	cont	0.0851	1.0937	0.2587	0.4681	0.1532	1.4818	0.3445	0.3617	0.1298	1.7609	0.3805	0.3830
	automatic	T	ORGFREF-SB-2	cont	0.0851	1.3603	0.3063	0.4255	0.1319	1.6324	0.3855	0.3191	0.1128	1.6736	0.3606	0.3191
	automatic	T	ORGFREF-SB-3	cont	0.0745	0.9305	0.2474	0.5106	0.1191	1.1893	0.3033	0.3830	0.1191	1.5963	0.3751	0.4468
	automatic	T	ORGFREF-SB-4	cont	0.0617	0.8083	0.1975	0.5745	0.1106	1.0859	0.2770	0.4043	0.1106	1.4979	0.3206	0.4681
	automatic	D	GRACE-SB-4	cont	0.1702	2.1963	0.3970	0.2979	0.2596	2.7014	0.4878	0.2340	0.2213	2.7438	0.4339	0.2553
	automatic	D	GRACE-SB-3	cont	0.1574	2.1215	0.3714	0.2766	0.2340	2.5717	0.4553	0.2128	0.1979	2.6903	0.4606	0.2340
	automatic	D	K3100-11	link&cont	0.1191	1.6266	0.3197	0.3830	0.1809	2.0473	0.4690	0.2553	0.1766	2.2604	0.3662	0.3191
	automatic	D	K3100-12	link&cont	0.1170	1.5794	0.3327	0.4043	0.1702	1.9287	0.4495	0.3191	0.1617	2.1115	0.3757	0.3617
	automatic	D	UAIF13	cont	0.1064	1.3941	0.3003	0.4255	0.1681	1.7433	0.3480	0.3617	0.1617	2.1268	0.4221	0.3617
	automatic	D	NAICR-II-B-1	cont	0.0957	1.2957	0.2628	0.5319	0.1447	1.5883	0.3227	0.4894	0.1213	1.5585	0.2828	0.5106
	automatic	D	UAIF14	cont	0.0915	1.2614	0.2719	0.4894	0.1511	1.6203	0.3435	0.4043	0.1426	1.9744	0.4166	0.3830
automatic	D	NAICR-II-B-3	cont	0.0787	1.0807	0.2045	0.5957	0.1340	1.3784	0.2604	0.5319	0.1128	1.4262	0.2543	0.5319	
automatic	D	NAICR-II-B-2	cont	0.0745	1.0700	0.2172	0.5532	0.1234	1.3504	0.2810	0.5106	0.1043	1.3520	0.2446	0.5319	

RunID は実行結果 ID を示す。また、prec(10), deg(10), wrr(10), %nf(10) はそれぞれ実行結果リストの上位 10 件における精度、DCG, WRR, 及び、適合文書が一つも見つからなかった検索課題のパーセンテージを示す。

表 8: 大規模サーベイ検索タスクに関する評価尺度ごとのシステム順位比較

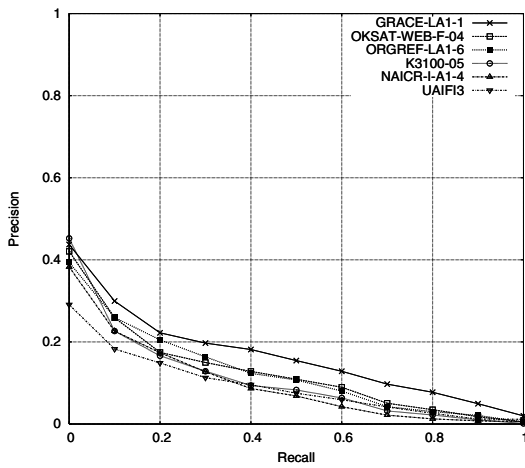
Topic Part	$DM_1 \times RL_1$				$DM_2 \times RL_1$			
	aprec	rprec	dcg(100)	dcg(1K)	aprec	rprec	dcg(100)	dcg(1K)
T	GRACE-LA1-1	<b>GRACE-LA1-2</b>	GRACE-LA1-1	GRACE-LA1-1	GRACE-LA1-1	<b>GRACE-LA1-2</b>	GRACE-LA1-1	GRACE-LA1-1
T	<b>GRACE-LA1-2</b>	GRACE-LA1-1	<b>GRACE-LA1-2</b>	<b>GRACE-LA1-2</b>	<b>GRACE-LA1-2</b>	GRACE-LA1-1	<b>GRACE-LA1-2</b>	<b>GRACE-LA1-2</b>
T	ORGREF-LA1-6	ORGREF-LA1-6	OXSAT-F-04	UAIF3	OXSAT-F-04	ORGREF-LA1-6	OXSAT-F-04	OXSAT-F-04
T	OXSAT-F-04	<b>K3100-05</b>	<b>K3100-05</b>	UAIF4	ORGREF-LA1-6	OXSAT-F-04	ORGREF-LA1-6	ORGREF-LA1-6
T	<b>K3100-05</b>	OXSAT-F-04	<b>K3100-06</b>	<b>K3100-05</b>	<b>K3100-05</b>	<b>K3100-05</b>	<b>K3100-05</b>	<b>K3100-05</b>
T	<b>K3100-06</b>	<b>K3100-06</b>	ORGREF-LA1-6	OXSAT-F-04	<b>K3100-06</b>	<b>K3100-06</b>	<b>K3100-06</b>	<b>K3100-06</b>
T	UAIF3	UAIF3	UAIF3	<b>K3100-06</b>	NAICR-I-A1-4	NAICR-I-A1-4	NAICR-I-A1-4	UAIF4
T	UAIF4	NAICR-I-A1-4	UAIF4	ORGREF-LA1-6	ORGREF-LA1-5	UAIF3	UAIF3	UAIF3
T	NAICR-I-A1-4	UAIF4	NAICR-I-A1-4	NAICR-I-A1-4	UAIF3	ORGREF-LA1-5	UAIF4	NAICR-I-A1-4
T	ORGREF-LA1-5	ORGREF-LA1-5	ORGREF-LA1-3	ORGREF-LA1-5	UAIF4	UAIF4	ORGREF-LA1-5	ORGREF-LA1-5
T	ORGREF-LA1-3	ORGREF-LA1-3	ORGREF-LA1-5	ORGREF-LA1-3	ORGREF-LA1-3	ORGREF-LA1-3	ORGREF-LA1-3	ORGREF-LA1-3
T	ORGREF-LA1-1	ORGREF-LA1-1	ORGREF-LA1-4	ORGREF-LA1-1	ORGREF-LA1-1	ORGREF-LA1-1	ORGREF-LA1-1	ORGREF-LA1-1
T	ORGREF-LA1-4	ORGREF-LA1-4	ORGREF-LA1-1	ORGREF-LA1-4	ORGREF-LA1-4	ORGREF-LA1-4	ORGREF-LA1-4	ORGREF-LA1-4
T	ORGREF-LA1-2	ORGREF-LA1-2	ORGREF-LA1-2	ORGREF-LA1-2	ORGREF-LA1-2	ORGREF-LA1-2	ORGREF-LA1-2	ORGREF-LA1-2
D	<b>GRACE-LA1-4</b>	<b>GRACE-LA1-4</b>	<b>GRACE-LA1-4</b>	GRACE-LA1-3	<b>GRACE-LA1-4</b>	<b>GRACE-LA1-4</b>	<b>GRACE-LA1-4</b>	GRACE-LA1-3
D	GRACE-LA1-3	GRACE-LA1-3	GRACE-LA1-3	<b>GRACE-LA1-4</b>	GRACE-LA1-3	GRACE-LA1-3	<b>GRACE-LA1-4</b>	<b>GRACE-LA1-4</b>
D	OXSAT-F-06	OXSAT-F-06	UAIF1	OXSAT-F-06	OXSAT-F-06	OXSAT-F-06	OXSAT-F-06	OXSAT-F-06
D	UAIF1	UAIF1	OXSAT-F-06	UAIF1	NAICR-I-A1-3	NAICR-I-A1-3	NAICR-I-A1-3	NAICR-I-A1-3
D	NAICR-I-A1-3	NAICR-I-A1-3	UAIF2	UAIF2	<b>K3100-07</b>	UAIF2	UAIF1	UAIF2
D	UAIF2	UAIF2	NAICR-I-A1-3	NAICR-I-A1-3	UAIF1	UAIF1	UAIF2	UAIF1
D	<b>K3100-07</b>	<b>K3100-07</b>	<b>K3100-07</b>	<b>K3100-07</b>	UAIF2	<b>K3100-07</b>	<b>K3100-07</b>	<b>K3100-08</b>
D	NAICR-I-A1-2	<b>K3100-08</b>	<b>K3100-08</b>	<b>K3100-08</b>	NAICR-I-A1-2	<b>K3100-08</b>	<b>K3100-08</b>	<b>K3100-07</b>
D	NAICR-I-A1-1	NAICR-I-A1-2	NAICR-I-A1-2	NAICR-I-A1-2	<b>K3100-08</b>	NAICR-I-A1-2	NAICR-I-A1-2	NAICR-I-A1-2
D	<b>K3100-08</b>	NAICR-I-A1-1	NAICR-I-A1-1	NAICR-I-A1-1	NAICR-I-A1-1	NAICR-I-A1-1	NAICR-I-A1-1	NAICR-I-A1-1

aprec, rprec は、それぞれ非補間平均精度, R 精度を示す。また、dcg(100), dcg(1K) は、それぞれ実行結果リストの上位 100, 1,000 件における DCG を示す。なお、各記号列は実行結果 ID を示す。

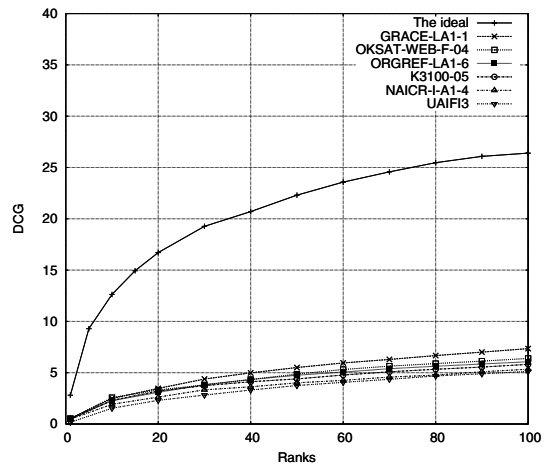
表 9: 大規模ターゲット検索タスクに関する評価尺度ごとのシステム順位比較

Topic Part	$DM_1 \times RL_1$				$DM_2 \times RL_1$			
	prec(10)	dcg(10)	wrr(10)	%nf(10)	prec(10)	dcg(10)	wrr(10)	%nf(10)
T	GRACE-LB-1	GRACE-LB-1	GRACE-LB-1	<b>K3100-13</b>	GRACE-LB-1	GRACE-LB-1	<b>K3100-13</b>	<b>K3100-13</b>
T	<b>K3100-14</b>	<b>GRACE-LB-2</b>	<b>GRACE-LB-2</b>	GRACE-LB-1	<b>GRACE-LB-2</b>	<b>GRACE-LB-2</b>	GRACE-LB-1	<b>K3100-14</b>
T	<b>GRACE-LB-2</b>	<b>K3100-14</b>	<b>K3100-13</b>	<b>GRACE-LB-2</b>	<b>K3100-14</b>	<b>K3100-14</b>	<b>GRACE-LB-2</b>	<b>GRACE-LB-2</b>
T	<b>K3100-13</b>	<b>K3100-13</b>	UAIF8	UAIF8	<b>K3100-13</b>	<b>K3100-13</b>	<b>K3100-14</b>	GRACE-LB-1
T	ORGREF-LB-6	ORGREF-LB-6	<b>K3100-14</b>	<b>K3100-14</b>	ORGREF-LB-6	ORGREF-LB-6	ORGREF-LB-6	ORGREF-LB-6
T	UAIF8	UAIF8	ORGREF-LB-3	UAIF7	UAIF8	NAICR-I-B-1	ORGREF-LB-3	UAIF8
T	UAIF7	NAICR-I-B-1	UAIF7	ORGREF-LB-6	UAIF7	UAIF8	NAICR-I-B-1	UAIF7
T	NAICR-I-B-1	UAIF7	NAICR-I-B-1	NAICR-I-B-1	NAICR-I-B-1	ORGREF-LB-5	ORGREF-LB-5	NAICR-I-B-1
T	ORGREF-LB-5	ORGREF-LB-5	ORGREF-LB-6	ORGREF-LB-3	ORGREF-LB-5	UAIF7	UAIF8	ORGREF-LB-3
T	ORGREF-LB-3	ORGREF-LB-3	ORGREF-LB-4	ORGREF-LB-4	ORGREF-LB-3	ORGREF-LB-3	ORGREF-LB-4	ORGREF-LB-5
T	ORGREF-LB-4	ORGREF-LB-4	ORGREF-LB-5	ORGREF-LB-5	ORGREF-LB-4	ORGREF-LB-4	UAIF7	ORGREF-LB-4
T	ORGREF-LB-1	ORGREF-LB-1	ORGREF-LB-1	ORGREF-LB-1	ORGREF-LB-1	ORGREF-LB-1	ORGREF-LB-1	ORGREF-LB-1
T	ORGREF-LB-2	ORGREF-LB-2	ORGREF-LB-2	ORGREF-LB-2	ORGREF-LB-2	ORGREF-LB-2	ORGREF-LB-2	ORGREF-LB-2
D	<b>GRACE-LB-4</b>	GRACE-LB-3	<b>GRACE-LB-4</b>	UAIF6	<b>GRACE-LB-4</b>	GRACE-LB-3	<b>GRACE-LB-4</b>	<b>GRACE-LB-4</b>
D	GRACE-LB-3	<b>GRACE-LB-4</b>	GRACE-LB-3	<b>GRACE-LB-4</b>	GRACE-LB-3	<b>GRACE-LB-4</b>	GRACE-LB-3	NAICR-I-B-2
D	UAIF5	UAIF6	NAICR-I-B-2	UAIF5	UAIF5	UAIF6	<b>K3100-15</b>	GRACE-LB-3
D	UAIF6	UAIF5	UAIF6	NAICR-I-B-2	UAIF6	UAIF5	NAICR-I-B-2	UAIF5
D	NAICR-I-B-2	NAICR-I-B-2	<b>K3100-15</b>	GRACE-LB-3	NAICR-I-B-2	<b>K3100-15</b>	UAIF6	<b>K3100-15</b>
D	<b>K3100-15</b>	<b>K3100-15</b>	UAIF5	<b>K3100-15</b>	<b>K3100-15</b>	NAICR-I-B-2	<b>K3100-16</b>	UAIF6
D	NAICR-I-B-3	NAICR-I-B-3	<b>K3100-16</b>	<b>K3100-16</b>	NAICR-I-B-3	NAICR-I-B-3	UAIF5	<b>K3100-16</b>
D	NAICR-I-B-4	NAICR-I-B-4	NAICR-I-B-4	NAICR-I-B-4	NAICR-I-B-4	NAICR-I-B-4	NAICR-I-B-3	NAICR-I-B-3
D	<b>K3100-16</b>	<b>K3100-16</b>	NAICR-I-B-3	NAICR-I-B-3	<b>K3100-16</b>	<b>K3100-16</b>	NAICR-I-B-4	NAICR-I-B-4

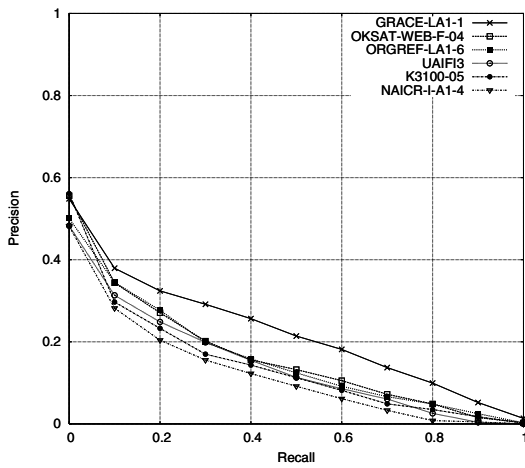
prec(10), dcg(10), wrr(10), %nf(10) は、それぞれ実行結果リストの上位 10 件における精度, DCG, WRR, 及び、適合文書が一つも見つからなかった検索課題のパーセンテージを示す。



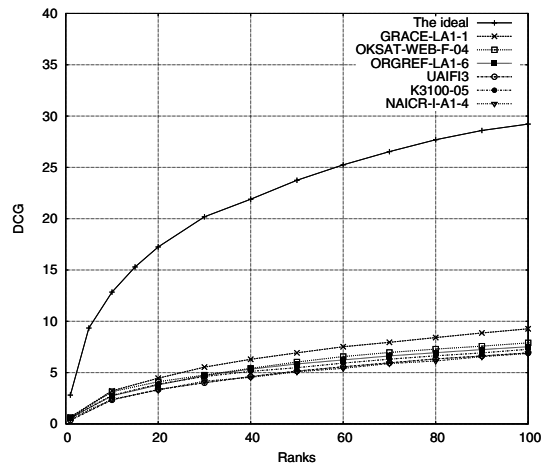
Recall-precision curves for the I-A1 'automatic' and 'TITLE-only' runs without considering links (rigid relevance level)



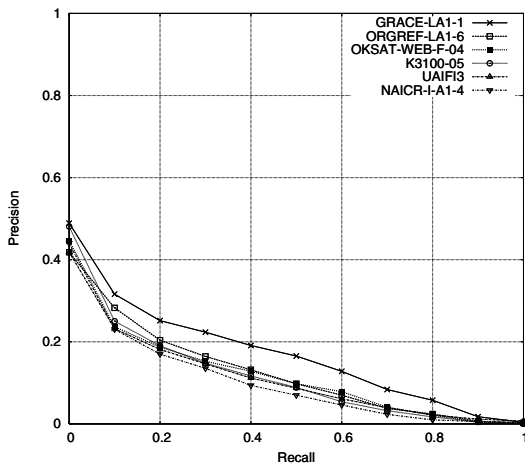
DCG curves for the I-A1 'automatic' and 'TITLE-only' runs without considering links (rigid relevance level)



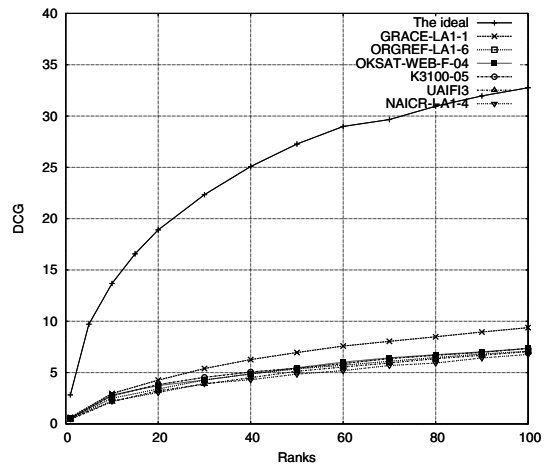
Recall-precision curves for the I-A1 'automatic' and 'TITLE-only' runs without considering links (relaxed relevance level)



DCG curves for the I-A1 'automatic' and 'TITLE-only' runs without considering links (relaxed relevance level)

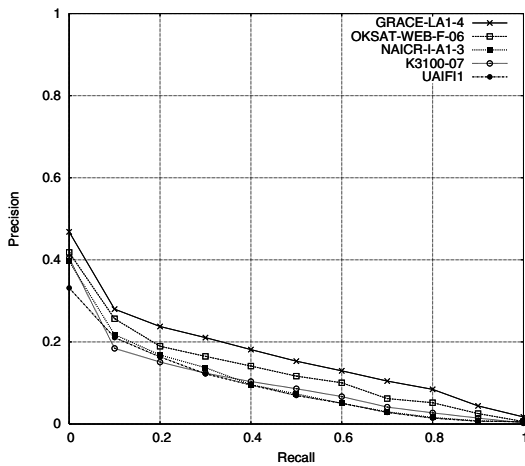


Recall-precision curves for the I-A1 'automatic' and 'TITLE-only' runs with considering links (rigid relevance level)

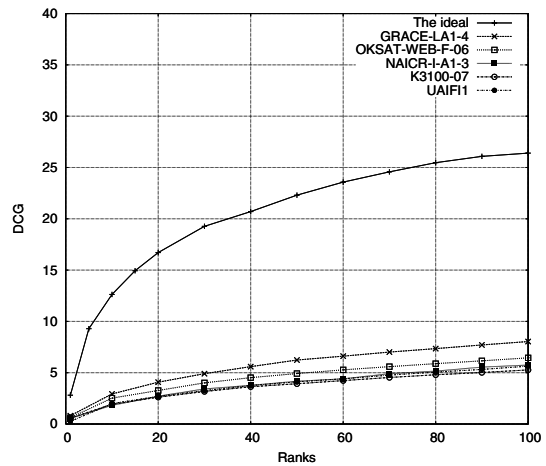


DCG curves for the I-A1 'automatic' and 'TITLE-only' runs with considering links (rigid relevance level)

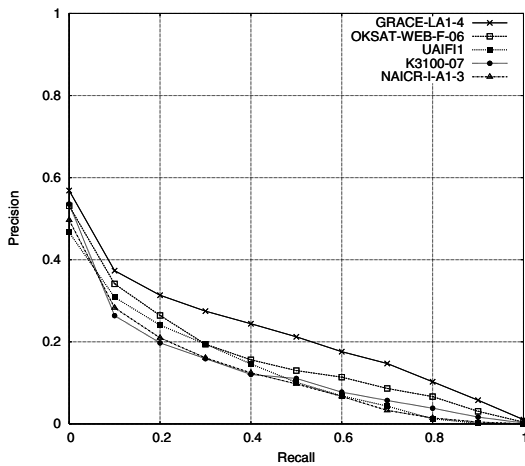
図4: 大規模トピック検索タスク (I-A1) において TITLE のみを使用した自動システムの実行結果に関する再現率・精度グラフと DCG のグラフ



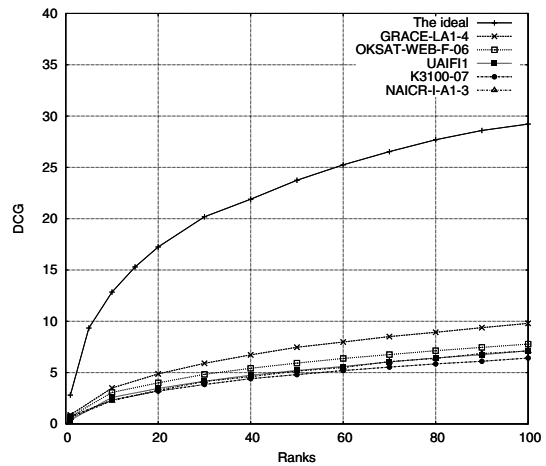
Recall-precision curves for the I-A1 'automatic' and 'DESC-only' runs without considering links (rigid relevance level)



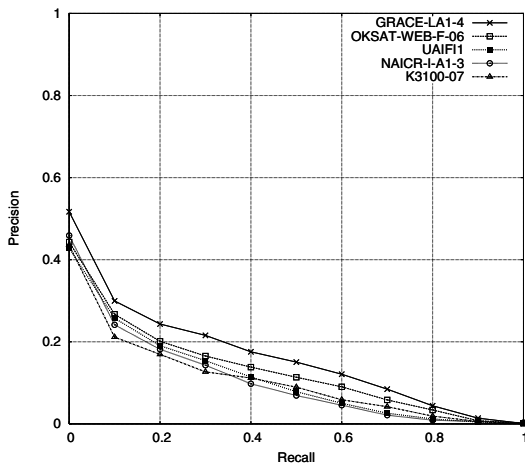
DCG curves for the I-A1 'automatic' and 'DESC-only' runs without considering links (rigid relevance level)



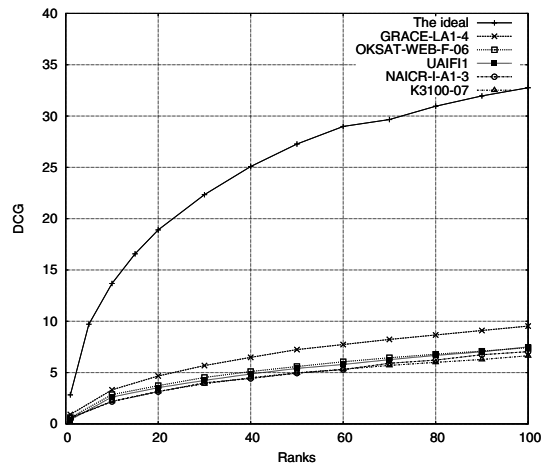
Recall-precision curves for the I-A1 'automatic' and 'DESC-only' runs without considering links (relaxed relevance level)



DCG curves for the I-A1 'automatic' and 'DESC-only' runs without considering links (relaxed relevance level)

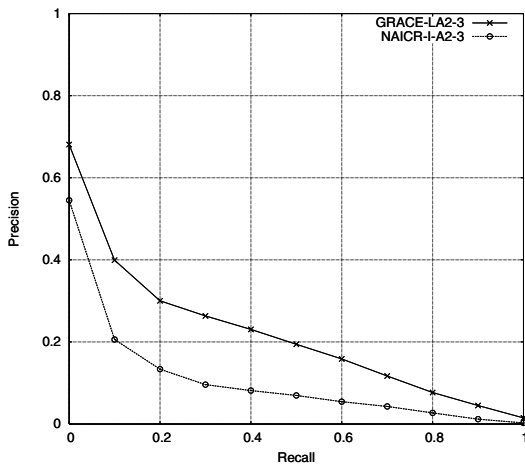


Recall-precision curves for the I-A1 'automatic' and 'DESC-only' runs with considering links (rigid relevance level)

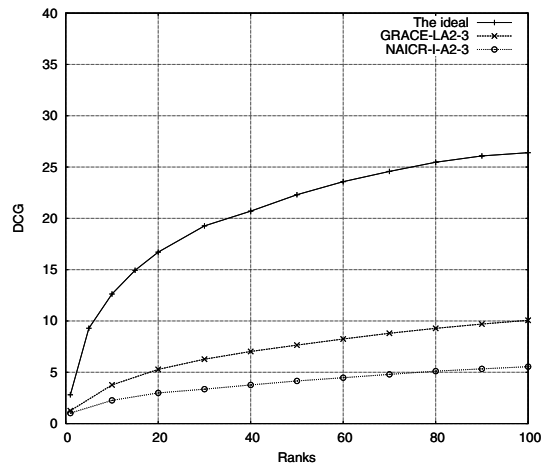


DCG curves for the I-A1 'automatic' and 'DESC-only' runs with considering links (rigid relevance level)

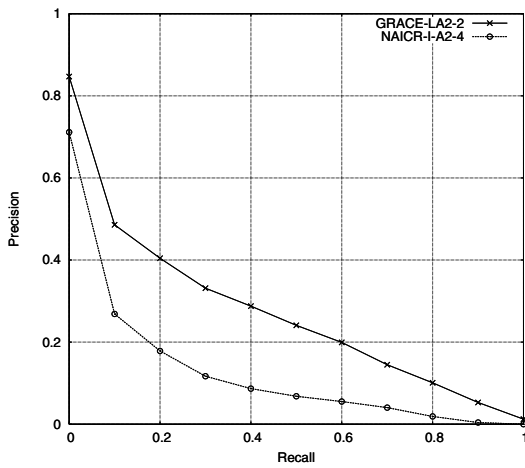
図 5: 大規模トピック検索タスク (I-A1) において DESC のみを使用した自動システムの実行結果に関する再現率・精度グラフと DCG のグラフ



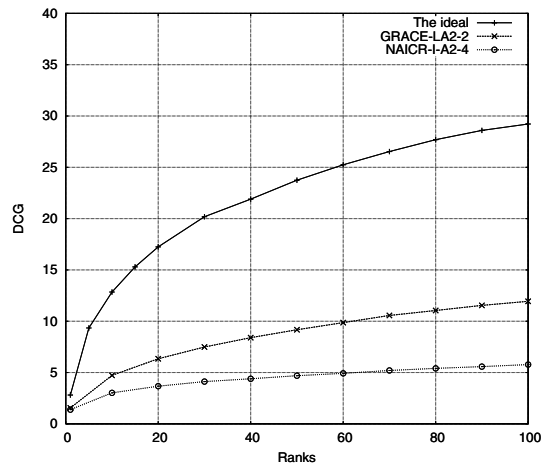
Recall-precision curves for the I-A2 'automatic' and 'TITLE-and-RDOC[1]' runs without considering links (rigid relevance level)



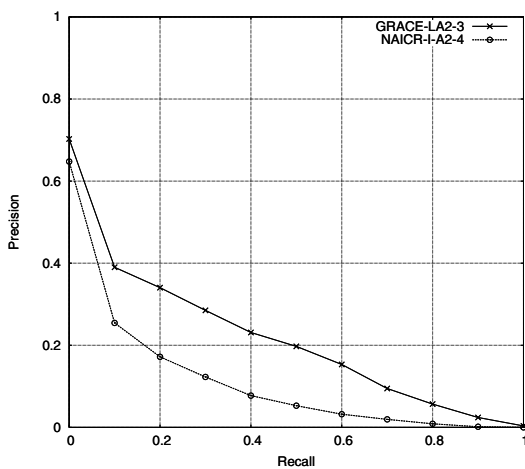
DCG curves for the I-A2 'automatic' and 'TITLE-and-RDOC[1]' runs without considering links (rigid relevance level)



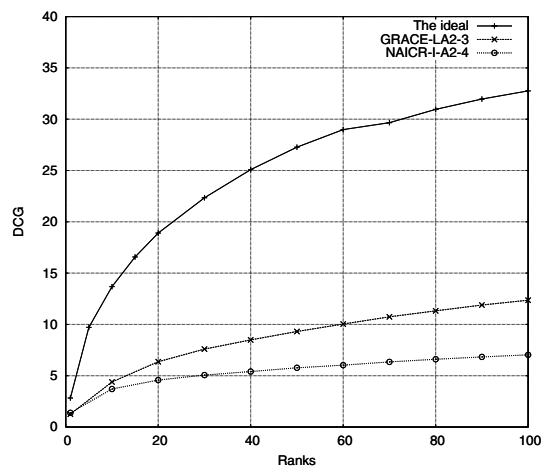
Recall-precision curves for the I-A2 'automatic' and 'TITLE-and-RDOC[1]' runs without considering links (relaxed relevance level)



DCG curves for the I-A2 'automatic' and 'TITLE-and-RDOC[1]' runs without considering links (relaxed relevance level)

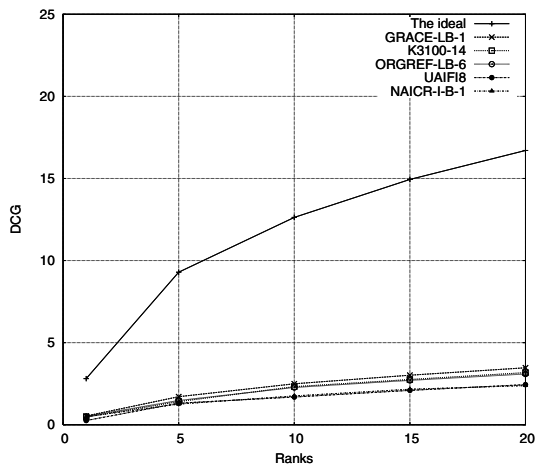


Recall-precision curves for the I-A2 'automatic' and 'TITLE-and-RDOC[1]' runs with considering links (rigid relevance level)

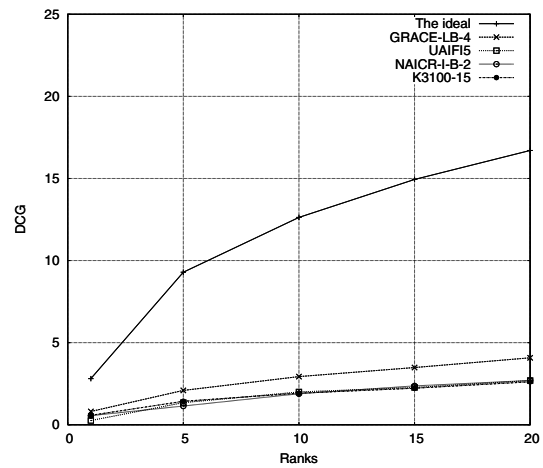


DCG curves for the I-A2 'automatic' and 'TITLE-and-RDOC[1]' runs with considering links (rigid relevance level)

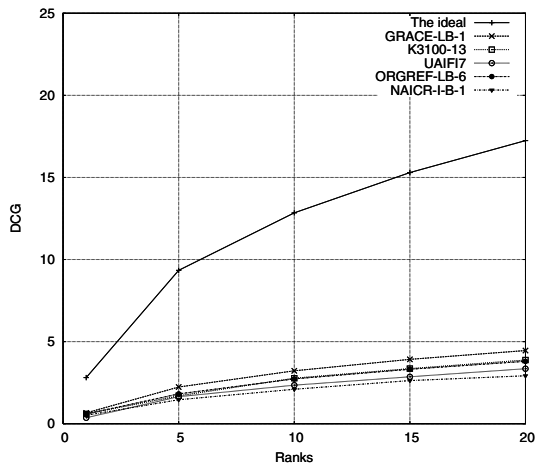
図 6: 大規模類書検索タスク (I-A2) における自動システムの実行結果に関する再現率・精度グラフと DCG のグラフ



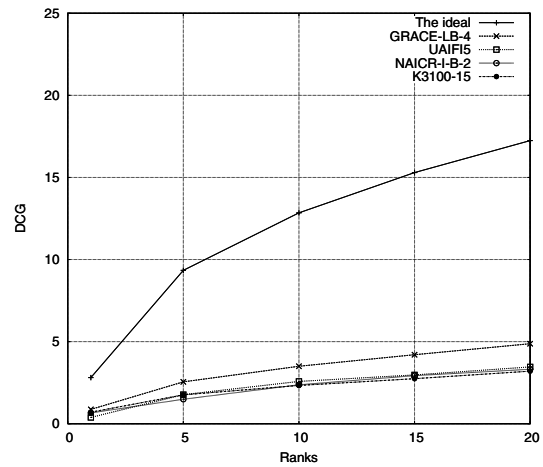
DCG curves for the I-B 'automatic' and 'TITLE-only' runs without considering links (rigid relevance level)



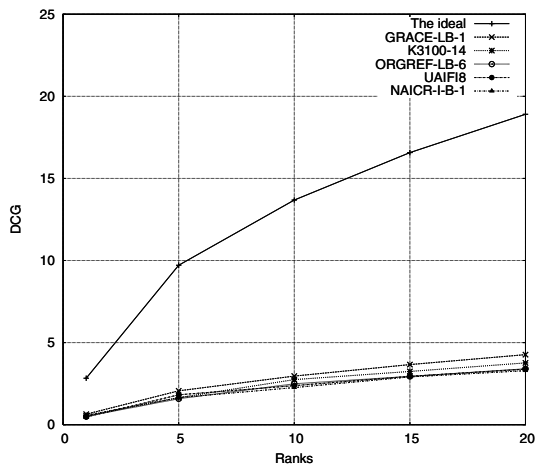
DCG curves for the I-B 'automatic' and 'DESC-only' runs without considering links (rigid relevance level)



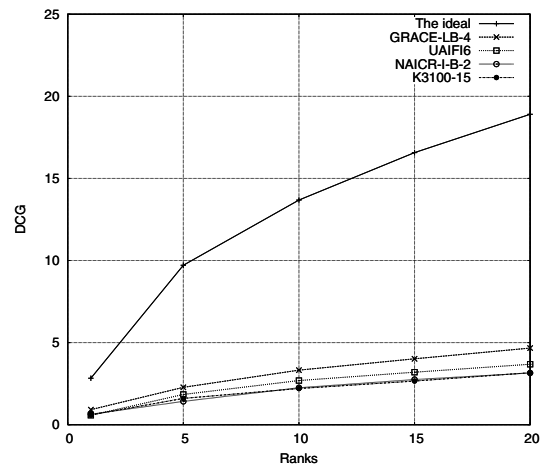
DCG curves for the I-B 'automatic' and 'TITLE-only' runs without considering links (relaxed relevance level)



DCG curves for the I-B 'automatic' and 'DESC-only' runs without considering links (relaxed relevance level)

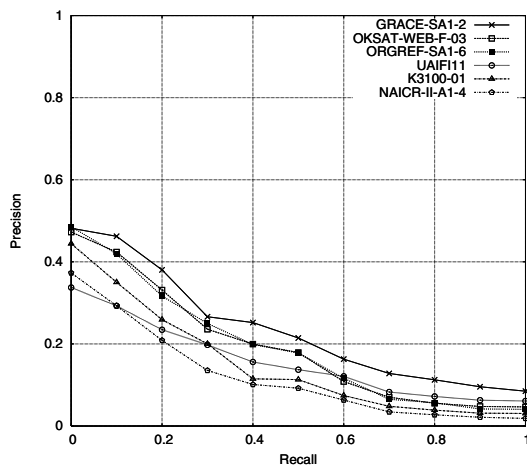


DCG curves for the I-B 'automatic' and 'TITLE-only' runs with considering links (rigid relevance level)

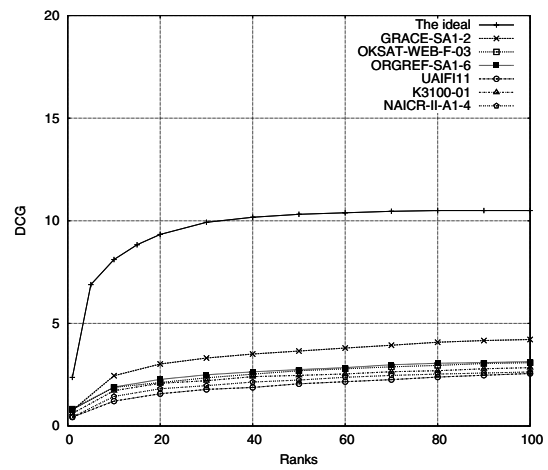


DCG curves for the I-B 'automatic' and 'DESC-only' runs with considering links (rigid relevance level)

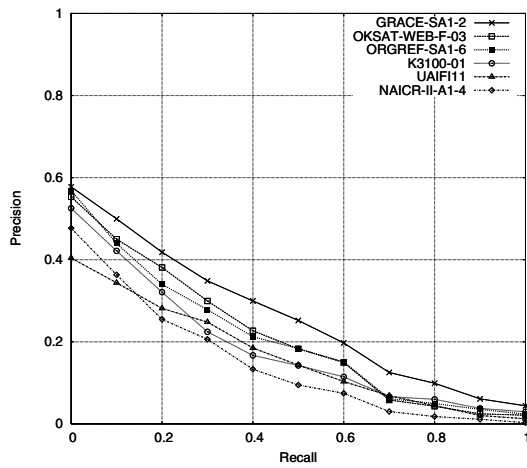
図 7: 大規模ターゲット検索タスク (I-B) における自動システムの実行結果に関する DCG のグラフ



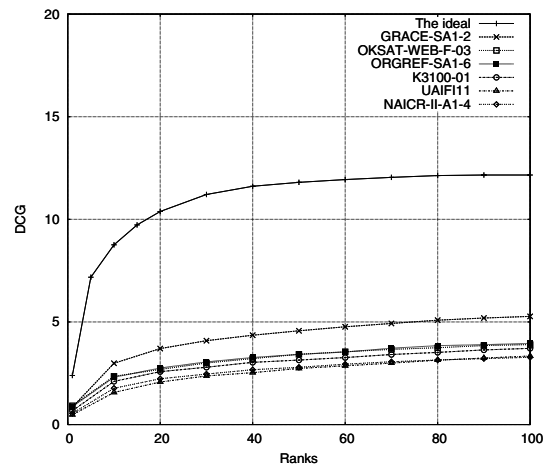
Recall-precision curves for the II-A1 'automatic' and 'TITLE-only' runs without considering links (rigid relevance level)



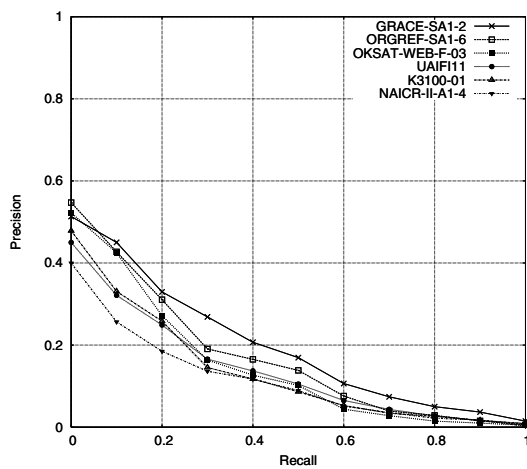
DCG curves for the II-A1 'automatic' and 'TITLE-only' runs without considering links (rigid relevance level)



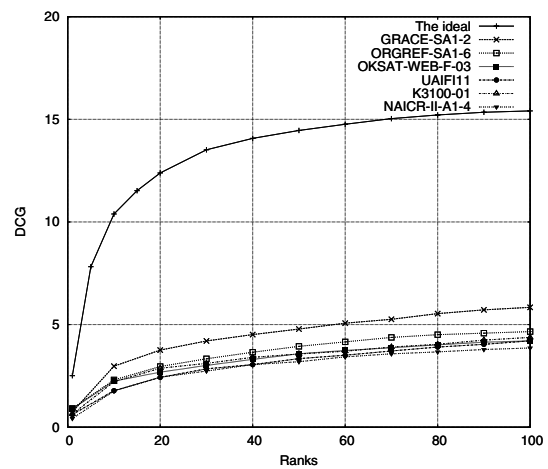
Recall-precision curves for the II-A1 'automatic' and 'TITLE-only' runs without considering links (relaxed relevance level)



DCG curves for the II-A1 'automatic' and 'TITLE-only' runs without considering links (relaxed relevance level)

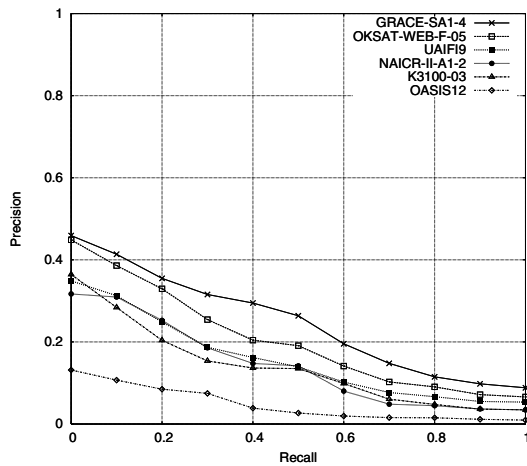


Recall-precision curves for the II-A1 'automatic' and 'TITLE-only' runs with considering links (rigid relevance level)

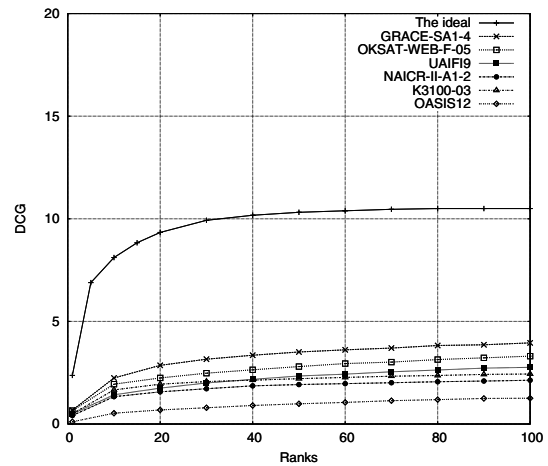


DCG curves for the II-A1 'automatic' and 'TITLE-only' runs with considering links (rigid relevance level)

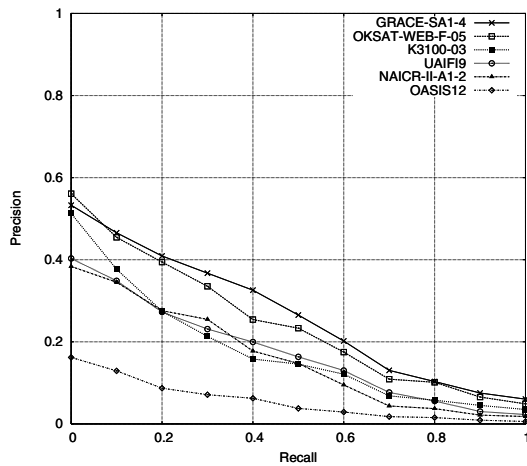
図 8: 小規模トピック検索タスク (II-A1) において TITLE のみを使用した自動システムの実行結果に関する再現率・精度グラフと DCG のグラフ



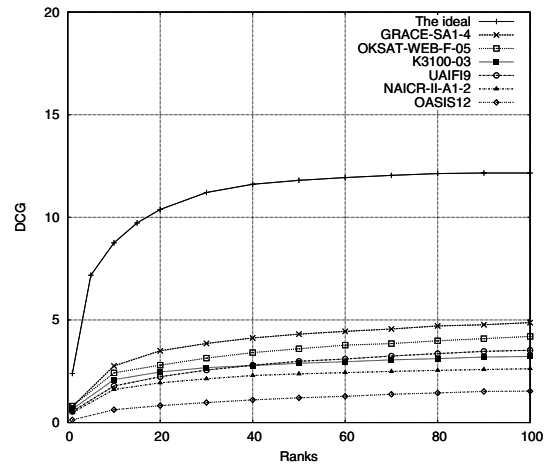
Recall-precision curves for the II-A1 'automatic' and 'DESC-only' runs without considering links (rigid relevance level)



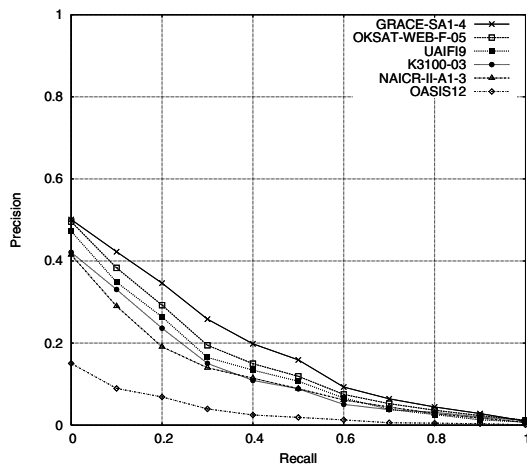
DCG curves for the II-A1 'automatic' and 'DESC-only' runs without considering links (rigid relevance level)



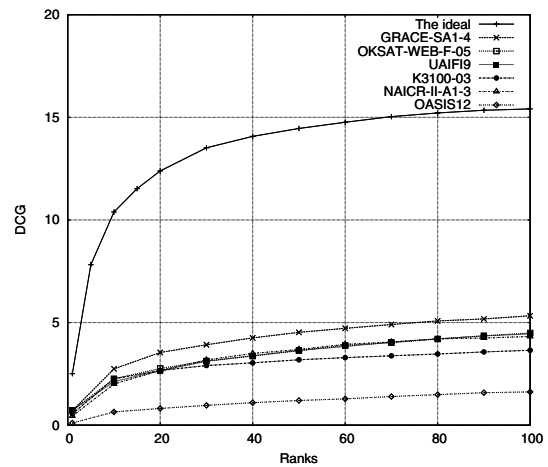
Recall-precision curves for the II-A1 'automatic' and 'DESC-only' runs without considering links (relaxed relevance level)



DCG curves for the II-A1 'automatic' and 'DESC-only' runs without considering links (relaxed relevance level)

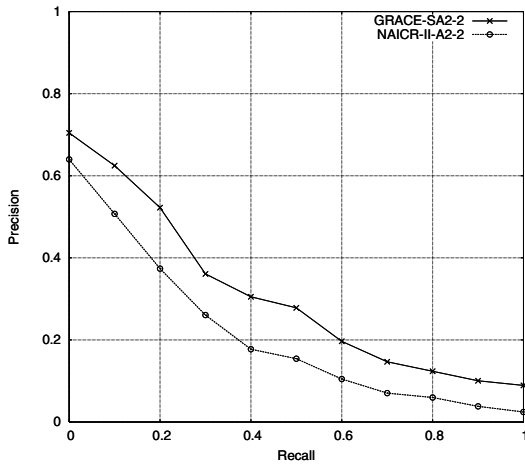


Recall-precision curves for the II-A1 'automatic' and 'DESC-only' runs with considering links (rigid relevance level)

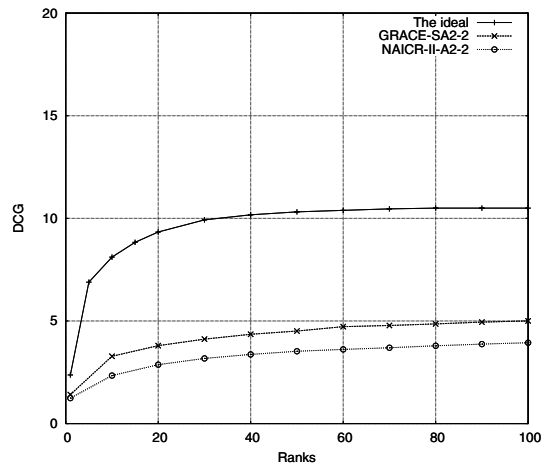


DCG curves for the II-A1 'automatic' and 'DESC-only' runs with considering links (rigid relevance level)

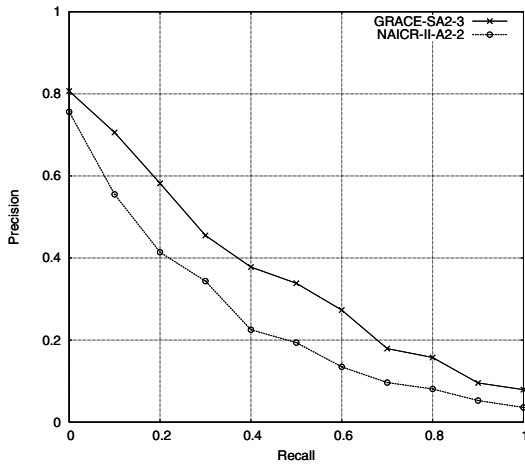
図9: 小規模トピック検索タスク (II-A1) において DESC のみを使用した自動システムの実行結果に関する再現率・精度グラフと DCG のグラフ



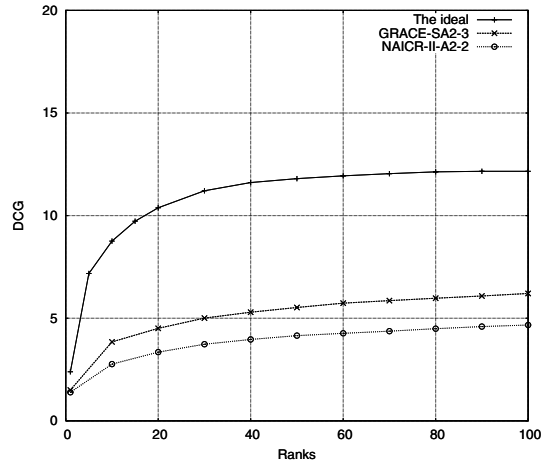
Recall-precision curves for the II-A2 'automatic' and 'TITLE-and-RDOC[1]' runs without considering links (rigid relevance level)



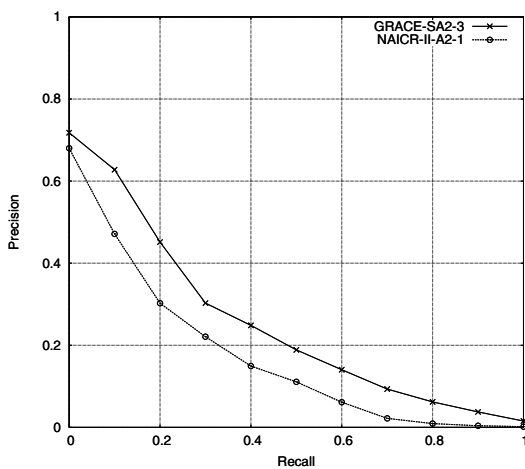
DCG curves for the II-A2 'automatic' and 'TITLE-and-RDOC[1]' runs without considering links (rigid relevance level)



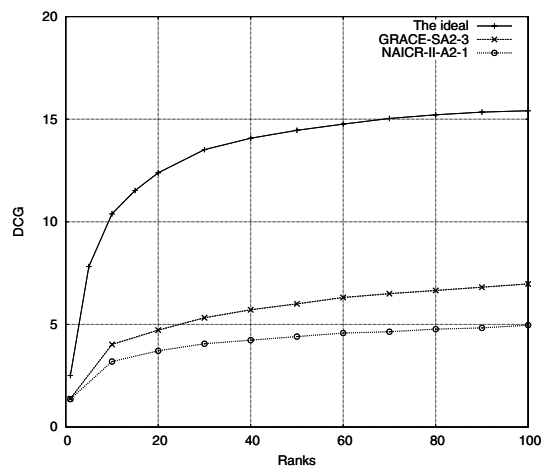
Recall-precision curves for the II-A2 'automatic' and 'TITLE-and-RDOC[1]' runs without considering links (relaxed relevance level)



DCG curves for the II-A2 'automatic' and 'TITLE-and-RDOC[1]' runs without considering links (relaxed relevance level)

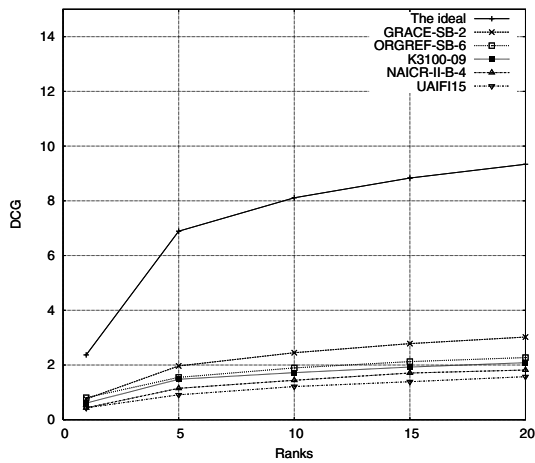


Recall-precision curves for the II-A2 'automatic' and 'TITLE-and-RDOC[1]' runs with considering links (rigid relevance level)

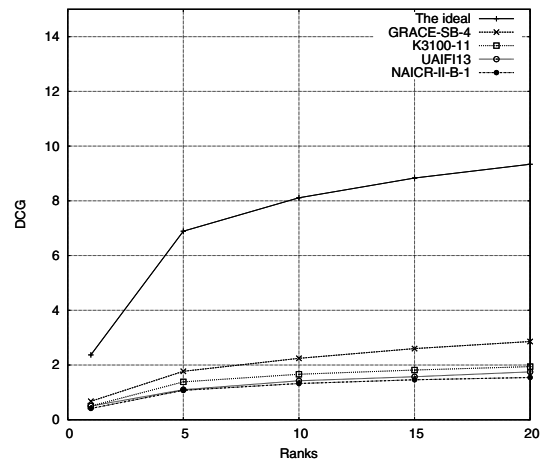


DCG curves for the II-A2 'automatic' and 'TITLE-and-RDOC[1]' runs with considering links (rigid relevance level)

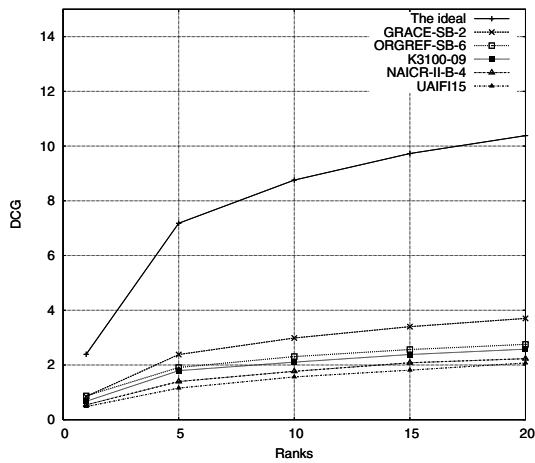
図 10: 小規模類書検索タスク (II-A2) における自動システムの実行結果に関する再現率・精度グラフと DCG のグラフ



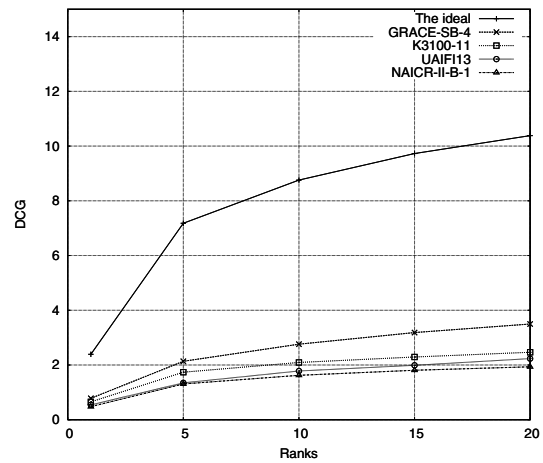
DCG curves for the II-B 'automatic' and 'TITLE-only' runs without considering links (rigid relevance level)



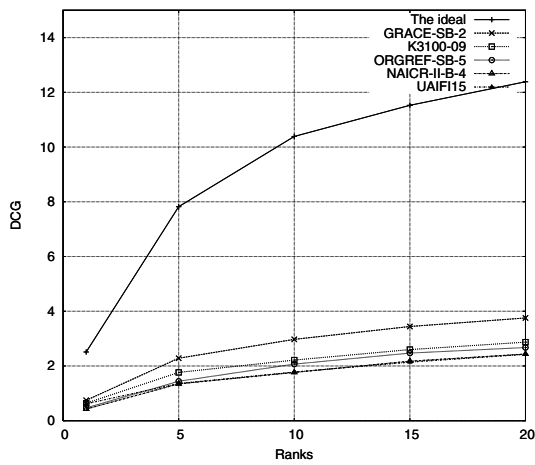
DCG curves for the II-B 'automatic' and 'DESC-only' runs without considering links (rigid relevance level)



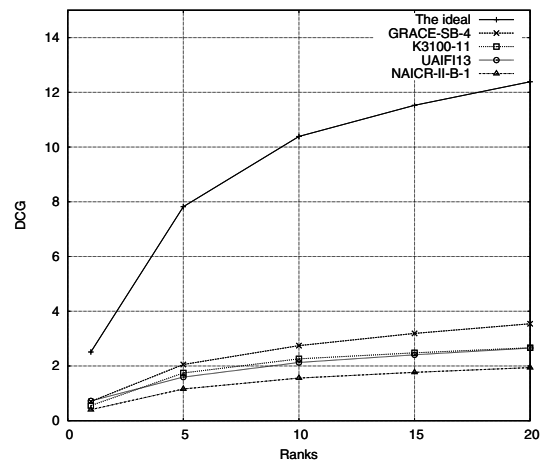
DCG curves for the II-B 'automatic' and 'TITLE-only' runs without considering links (relaxed relevance level)



DCG curves for the II-B 'automatic' and 'DESC-only' runs without considering links (relaxed relevance level)



DCG curves for the II-B 'automatic' and 'TITLE-only' runs with considering links (rigid relevance level)



DCG curves for the II-B 'automatic' and 'DESC-only' runs with considering links (rigid relevance level)

図 11: 小規模ターゲット検索タスク (II-B) における自動システムの実行結果に関する DCG のグラフ