Overview of IR Tasks at the First NTCIR Workshop

Noriko Kando, Kazuko Kuriyama, Toshihiko Nozue, Koji Eguchi, Hiroyuki Kato, Souichiro Hidaka
Research and Development Department
National Center for Science Information Systems (NACSIS) , Japan
URL: http://www.rd.nacsis.ac.jp/~ntcadm/

1. Introduction

This paper serves as an introduction to the first NTCIR Workshop [1] and research using the NACSIS Test Collection 1 (NTCIR-1), which is described in detail in the rest of the volume. The NTCIR Workshop is the first evaluation workshop designed to enhance research in Japanese text retrieval. Since this report is written based on the search results and system description forms which are submitted on March 4, 1999, it should be updated based on the research results presented at the Workshop meeting.

The NTCIR Workshop has the following goals:

(1) to encourage research in information retrieval, cross-lingual information retrieval and related areas by providing a large-scale Japanese test collection and a common evaluation setting that allows cross-system comparisons

(2) to provide a forum for research groups interested in comparing results and exchanging ideas or opinions in an informal atmosphere

(3) to investigate the effective method to construct large scale test collections and IR laboratory type testing

The test collection used in this Workshop is called "NACSIS Test Collection 1" or "NTCIR-1" and consists of more than 330,000 documents, with more than half presented as English-Japanese pairs. Although there is a Japanese test collection called BMIR-J2 consisting of 5,080 newspaper articles [6], and its contribution to IR research in Japan is tremendous, enhancement of the Japanese test collection in the aspects of both the variety of text types and the scale is needed. We place emphasis on cross-lingual retrieval since it is critical in the Internet environment and for Japanese scientific information retrieval [2].

Each participant has conducted one or more of the following tasks:

The Ad Hoc Information Retrieval task - to investigate the retrieval performance of systems that search a static set of documents using new search topics

The Cross-Lingual Information Retrieval task - an ad hoc task in which the documents are in English and the topics are in Japanese

The Automatic Term Recognition and Role Analysis task - (1) to extract terms from titles and abstracts of the documents, and (2) to identify the terms representing the "object", "method" and "main operation" of the main topic of the documents.

Thirty-one groups, including participants from six countries, had enrolled to participate in the first NTCIR Workshop. Among them, 28 groups enrolled in IR tasks (23 in the Ad Hoc task and 16 in the cross-lingual task), and nine in term recognition tasks.

Regarding IR tasks, the search results of 117 runs were submitted from 23 groups. There were 48 runs for the Ad Hoc Retrieval Task from 17 groups and 69 runs for the Cross-Lingual Retrieval task from 10 groups. Nine groups are from Japanese companies (six in the Ad Hoc task and four in the cross-lingual, one did both), 11 are from Japanese universities or national research institutes, and four Non-Japanese groups. Two groups from the United State, one group from Australia, one group from Taiwan, nineteen groups are from Japan and some of them have non-Japanese members or collaboration with research groups outside of Japan. Two groups worked without any Japanese language expertise.,

Table 1. Participants for IR Tasks

	Ad Hoc IR	CrossLingua	sum
#groups enrolled	23	16	28
#groups submitted result	17	10	23
Submitted Runs	48	69	117

The rest of this report concentrates on the IR tasks, i.e., the Ad Hoc Retrieval and Cross-Lingual Retrieval tasks. The overview for the Term Recognition task was prepared separately [4f]. In the next section, we describe the tasks performed in the Workshop. Section 3 shows the test collection (NTCIR-1) used in the Workshop. Section 4 describes the methods of evaluation and Section 5 analyzes the search results submitted by participants. The final section lists issues to be discussed.

2. Task Descriptions for IR Tasks

2.1 The Procedures

Schedule

Nov. 1, '98: Delivered documents, 30 training topics, and relevance judgments for training topics

Dec. 2, '98 : Pretest, then updated relevant judgments for training topics

Feb. 8, '99 : Delivered test topics

March 4, '99 : Results submission

June 12, '99 : Relevance judgments for tests topics were delivered to active participants

Aug.30-Sept.1, '99 : Workshop, Tokyo

Oct. 1, '99 : Test Collection 1 (NTCIR-1) will be available for research purpose

[1] This project is supported by "Research for the Future" Program JSPS-RFTF96P00602 of the Japan Society for the Promotion of Science

From November 1, 1998, the document data, 30 search topics (0001--0030) and their relevance judgments started to be delivered to each IR tasks participant to train their systems. Among them, 21 topics are used as cross-lingual topics. The topics were written in Japanese. Fifty-three new test topics (0031--0083) were distributed to participants on February 8, and the search results for these new topics were submitted from each participant by March 4 as official test runs. The test topics are common for both IR tasks. After relevance assessments, it was revealed that 13 topics of the 53 contained less than five relevant documents per topic in cross-lingual retrieval. These 13 topics were discarded and the remaining 39 topics consist the "Official Topic Set for Cross-Language retrieval task".

Some of the participating groups enrolled for the IR tasks could not submit search results by March 4 with various reasons. Three of them wished to continue as active participants, so we set a second due date and accepted the search results submitted by that date. These additional submissions were included in the evaluation results but were not included in the document pool for relevance assessment by human analysts.

A participant could submit the results of more than one run. Both automatic and manual query constructions were allowed. In the case of automatic construction, the participants were required to submit at least one set of results of searches using <DESCRIPTION> fields of the topics only as the mandatory runs. For optional automatic runs and manual runs, any fields of the topics could be used. In addition, each participant had to complete and submit a form describing the detailed features of their system.

Human analysts assessed the relevance of retrieved documents for each topic. Based on the relevance assessments, interpolated recall and precision at 11 points, average precision (non-interpolated) over all relevant documents, and precision at 5, 10, 15, 20, 30, 100 documents were calculated using TREC's evaluation program, which is available from the ftp site of Cornell University.

3. The Test Collection

The test collection used in the Workshop, "Test Collection 1" or "NTCIR-1", consists of documents, topics, and relevance assessments for each search topic.

3.1 Documents

The documents are author abstracts of conference papers which were presented at academic meetings hosted by 65 Japanese academic societies [5]. Subject domain, length and format of the documents are diversified according to each society. Wide range of subject domains from pure sciences, technology and engineering, social sciences and humanities are included although about half of the documents are from electronic engineering and computer sciences.

The Collection contains three document collections: the JE Collection; the J Collection; and the E Collection. The JE Collection contains 339,483 documents, of which more than half are English-Japanese paired. Both Japanese part and English part are prepared by the author(s) of the document. The J and E Collections are constructed by extracting the Japanese or English parts, respectively, of the documents in the JE Collection.

3.1.1 A sample of the document record

Documents are SGML tagged plain text. A record may contain document ID, title, a list of author(s), name and date of the conference, abstract, keyword(s) which were assigned by the author(s) of the document, name of the host society. (See Fig.1)

<REC>
<ACCN>gakkai-00000111144</ACCN>
<TITL TYPE="kanji">電子原稿・電子出版・電子図書館-「SGML実験誌」の作成実験を通して</TITL>
<TITE TYPE="alpha">Electronic manuscripts, electronic publishing and electronic library </TITE>
<AUPK TYPE="kanji">根岸 正光</AUPK>
<AUPE TYPE="alpha">Negishi,Masamitsu</AUPE>
<CONF TYPE="kanji">研究発表会(情報学基礎)</CONF>
<CNFE TYPE="alpha">The Special Interest Group Notes of IPSJ</CNFE>
<CNFD>1991. 11. 19</CNFD>
<ABST TYPE="kanji"><ABST.P>電子出版というキーワードを中心に、文献の執筆、編集、印刷、流通の過程の電子化について、その現状を整理して今後の動向を検討する。とくに、電子出版に関する国際規格であるSGML(Standard Generalized Markup Language)に対するわが国での動きに注目し、学術情報センターにおける「SGML実験誌」およびその全文CD-ROM版の作成実験を通じて得られた知見を報告する。また電子図書館について、その諸形態を展望する。出版文化に依拠するこの種の社会システムの場合、技術的な問題というのは、その技術の社会的な受容・浸透の問題であり、この観点から標準化の重要性を論じる。</ABST.P></ABST>
<ABSE TYPE="alpha"><ABSE.P>Current situation on electronic processing in preparation, editing, printing and distribution of documents is summarized and its future trend is discussed, with focus on the concept: "Electronic publishing. "Movements in the country concerning an international standard on electronic publishing, SGML (Standard Generalized Markup Language), are assumed to be important, and the results from an experiment at NACSIS to publish "SGML Experimental Journal" and to make its full-text CD-ROM version are reported. Various forms of "Electronic library" are also investigated. The author puts emphasis on standardization, as technological problems for those social systems based on cultural settings of publication of the country, are the problems of acceptance and penetration of the technology in the society.</ABSE.P></ABSE>
<KYWD TYPE="kanji">電子出版 // 電子図書館 // 電子原稿 // SGML // 学術情報センター // 全文データベース</KYWD>
<KYWE TYPE="alpha">Electronic publishing // Electronic library // Electronic manuscripts // SGML // NACSIS // Full text databases</KYWE>
<SOCN TYPE="kanji">情報処理学会</SOCN>
<SOCE TYPE="alpha">Information Processing Society of Japan</SOCE>
</REC>

Fig. 1. A sample of the document record

Table 2. Average Length of Document Fields

Total number of documents: 339,483

	title	author	abstract	keyword	record
tokens	36.5		315.2	31.1	382.8
bytes	122		821.6	107.3	1050.9
number		3.4		8.3	

Since one of the purposes of the original database is to provide an alert information about papers presented in Japanese academic conferences as soon as possible, documents are put in the database without any revision or modification by professional abstractors or

editors. Some of them are refereed, and others are pre- or non-refereed. Both as part of the philosophy of leaving the data as close to the original as possible, and because it is impossible to check all the data manually, there are "errors" in the data. These range from errors in the original data or other typographical errors, to errors in the reformatting done at NACSIS and at Test Collection Project Group. The error-checking has concentrated on allowing readability of the data rather than on correcting content.

3.1.2 Document Collections and Tasks

In the Workshop, as shown in Fig. 2, the JE Collection is used in the Ad Hoc task since in Japanese operational IR environments, especially for retrieval of scholarly/scientific/technical documents and Web documents, retrieving both Japanese and English documents at once is quite natural. The E Collection is used in the Cross-Lingual task. The J Collection is used in monolingual retrieval, which will be the baseline for comparing the search effectiveness with the results in the Cross-Lingual task.

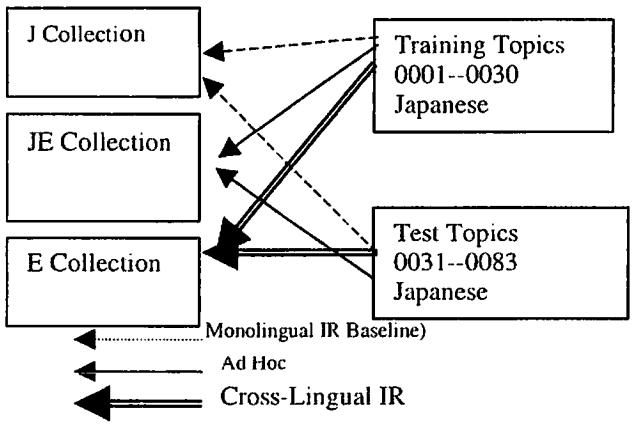

Fig 2. Relation of the Document Collections and Tasks

In Japan, (1) most of the documents published in Japan are in Japanese, but some are in English or in both English and Japanese; (2) a Japanese document quite often contains an English abstract (often an English abstract without a Japanese Abstract); (3) a concept may be represented in any of the four forms: Japanese term; transliteration of an English term using Japanese Katakana characters; English or another language term with its original spelling using roman letters; or an acronym. To overcome the word mismatches caused by this, cross-linguistic techniques are needed, even for Japanese text retrieval. Fortunately, Japanese newspaper texts are well controlled by the editorial policies and the style manuals of each press company, and English terms never appear in the text with their original spelling. On the other hand, in the ordinary texts, usage of English terms in Japanese sentences depends on the preference of the individual writers. In scientific, technical or scholarly documents, in particular, technical terms and new concepts are often represented as English or other language terms with their original spelling using roman letters, and cross-lingual techniques are more important in the retrieval of these documents.

However, the Ad Hoc Retrieval task using the JE Collection, which contains both Japanese documents and English documents, is complicated and a few groups submitted search results using the J Collection only. We evaluated all the search results submitted as Ad Hoc runs, regardless of whether they referred to the JE Collection or the J Collection, since there is a very strong correlation at the level of $p < 0.001$ between the search results using the JE Collection and the J Collection.

3.2 Topics

A topic is a formatted description of a user's information need. We defined the topics as statements of "user need" rather than "queries", which are the strings actually submitted to the system, since we would like to allow both manual and automatic query construction from the topics.

The query format is similar to that used in TREC-1 and 2 and contains SGML-like tags. A topic consists of a title of the topic, a description, a detailed narrative, a list of concepts and field(s). The title is a very short description of the topic and can be used as a very short query that resembles those often submitted by end-users of Internet search engines. Each narrative may contain a detailed explanation of the topic, term definitions, background knowledge, the purpose of the search, criteria for judgment of relevance, and so on. (See Fig.3)

```
<TOPIC q=0005>
<TITLE>
特徴次元リダクション
</TITLE>
<DESCRIPTION>
クラスタリングにおける特徴次元リダクション
</DESCRIPTION>
<NARRATIVE>
オブジェクトのクラスタリングを行なうとき、オブジェクトを特徴ベクトルで表現することが望まれる。アプリケーションによっては、オブジェクトの次元は数千、数万となることがある。このような場合、事前に次元を落とすことが必要になる。正解文書は、特徴次元リダクションの方法について、理論面から、または実験によって、提案、比較などを行なっているもの。画像処理などの実験の操作の一部として特徴次元リダクションを用いているだけでは要求を満たさない。
</NARRATIVE>
<CONCEPTS>
特徴選択, 主成分分析, 情報の粒度, 幾何クラスタリング
</CONCEPTS>
<FIELD>
1.電子・情報・制御
</FIELD>
</TOPIC>
```

Fig. 3 A sample Topic

3.2.1 Topic Preparation

Topics are collected from users with permission to use them as part of a test collection. Some were collected from researchers in several fields by interview or by using the topic collection form[2], some are collected from reference counters of research libraries, and others were created by the analysts based on their research interest or needs. Analysts are mainly graduate students with such backgrounds as computer sciences, pharmacology, biochemistry, and social sciences such as education, linguistics and so on.

[2] Accessible via: http://www.rd.nacsis.ac.jp/~ntcadm/index-e.html

The collection contains 30 training topics and 53 test topics. Among them 21 training topics and 39 test topics are usable for cross-lingual retrieval. All the topics are written in Japanese. English and Korean versions will be available.

After collecting the topics, each topic was examined for its clearness and difficulty by the analysts and project members in NACSIS. The criteria are as follows;

(1) Statements of "user need" rather than "queries"

(2) <Description> contain every concepts that are needed to describe the topic

(3) Not too easy:

(3-1) Simple word matching of query terms can not retrieve every relevant docs.

(3-2) A document containing query terms can be non-relevant.

(4) Five or more relevant documents in top 100 documents retrieved by the retrieval system we used in NACSIS

Sentences were modified when they were too restricted or ambiguous. Function category of each topic was analyzed and assigned based on the function category proposed by BMIR [6]. The category indicates the required level of techniques and knowledge to conduct search of the topic. We tried to balance the topic length, number of relevant documents, and "difficulty"; however, we found that estimating "difficulty" is difficult.

Table 3. Average Length of Fields in aTopic

	title		desc		narrative	
	char	word+phrs	char	word+phrs	char	word+phrs
training	6.8	3.0	22.9	6.2	159.3	24.3
test	9.6	4.3	33.1	8.9	310.0	43.7
all	8.6	3.8	29.4	8.0	255.6	36.7

3.3 Relevance Judgments (Right Answers)

The relevance judgments were done by pooling methods. A certain number of top ranked documents are collected from each submitted run and create a pool of possibly relevant documents. Human analysts assess the relevance of each document in the pool against the topic. The relevance assessment is done using three grades: relevant, partially relevant, non-relevant. Two analysts assess the relevance of a topic separately, and then the primary analyst of the topic, who created the topic, makes the final judgment.

Relevant judgment files contain not only relevance of each document in the pool but also contain extracted phrases or passages showing the reason why the analyst assessed the document as "relevant". Since a narrative of topics may contain some description related to user's situation or purpose of the search, situational-oriented relevance judgments were conducted as well as topic-oriented relevance judgments which are more common in the ordinary IR systems laboratory testing. However, only topic-oriented judgments are used in the formal evaluation of this Workshop. The relevance judgments are often called "right answers" in the Japanese IR community.

Table 4. Number of Relevant Documents par Topic

	ave	max	min
training	66.3	317	5
test	44.2	584	6
total	52.0	584	5

3.4. Linguistic Analysis

Part of the J collection contains detailed hand tagged part-of-speech (POS) tags [7]. Because of the absence of explicit boundaries between words in Japanese sentences, we set the three levels of lexical boundaries (i.e., word boundaries, strong and weak morpheme boundaries), and assigned detailed POS tags based on the boundaries and types of origin, so that the collection can be used to examine suitable term segmentations of Japanese texts for retrieval purpose. The Tagged Corpus was not used for official evaluation of the IR tasks but was distributed to the IR tasks participants on the request

3.5 Evaluation of the Test Collection

We have evaluated the Test Collection from these aspects so that the Collection is usable as a reliable tool for IR system testing;

(A) as a tool for IR system evaluation.
(B) topic by topic analysis
(C) effectiveness of the pooling

The results of these studies have been reported and published various occasion [8-12]. The brief summary of the studies are shown here.

Evaluation of the Test Collection as a Tool for IR system evaluation

Regaring (A), " Evaluation of the Test Collection as a Tool for IR system evaluation", we have investigated on the following issues based on both the results of the pretest and official test in the aspect whether these conditions affect the system evaluation or not;

(1) exhaustivity of the relevance judgments

(2) number of documents pooled

(3) consistency of the relevance assessments done by human assessers.

Topic by Topic Analysis

Regarding (B) "Topic by Topic Analysis", we have also investigated the analysis of the characteristics of each topic from the following aspects;

(1) number of relevant documents

(2) function category described in section 3.2

(3) length of the topic, length of each field in a topic (words, phrases, characters)

(4) term frequency of each term in each topic

(5) inverse document frequency of each term in each topic

(6) median of the average precision of all submitted runs for each topic

(7) distribution of the average precision that each runs obtained.

Effectiveness of the Pooling

Regarding (C) "Effectiveness of the Pooling", we have investigated the effectiveness of the pooling to collect possibly relevant document in both effectively and efficiently. Mark Rorvig's study

on visualization of the collection [13] supported that the NTCIR Collection collected relevant documents effectively through their analyses using visualization technique that had been applied to TREC collections.

3.5.1 Coverage of the Relevance Judgments and Strategy for Pooling: Analysis of Pooling in Pre-test

Some of the issues listed above are investigated using the search results submitted on December 2 as a pre-test. We asked for the top 1000 documents for each training topic to be submitted (1) to obtain feedback for the training topics and the relevance assessments for them, and (2) to confirm the procedure of the test. By adding additional relevant documents found in the pre-test to the initial ones, the relevance assessments for the training topics were revised.

Using the results of the pre-test, we evaluated (1) coverage of the initial pooling done in NACSIS, (2) effectiveness of pooling, and (3) the reliability of the test collection. Assessments were performed by investigating the effects that variations of pooling methods and coverage and variation in relevance assessments had on the evaluation of search effectiveness. The results are as follows:

1. The initial pooling worked well and covered 97% of the total relevant documents.
2. Interactive searches were effective for some particular topics and found 17.5% of unique relevant documents.
3. In terms of exhaustiveness, pooling the top 100 documents from each run worked well for topics with less than 50 relevant documents. For the topics with more than 100 relevant documents, although the top 100 pooling covered only 51.9% of the total relevant documents, the coverage reached 90% if it was combined with interactive searches.
4. The top 100 pooling method has the effect of reducing the size of the document pool to 27% of its possible size.
5. Regarding test topics, based on the analyses above we decided to use the top 100 pooling.

3.5.2 As a Tool for System Evaluation

We found very high similarity among the system rankings produced using different sets of relevance assessments, regardless of the different coverage and pooling methods and regardless of inconsistency among relevance assessments.

The consistency between the final judgment and the primary analysts of each topic and the one between the final judgment and the secondary analysts are, 41.8%, 60.6% for training topics and 75.1%, 66.3% for test topics respectively. Regardless of the inconsistency of the relevance assessments, we found strong correlation among the system ranking produced using the relevant judgements by the primary analysts, the one by the secondary analysts, and the final judgments[8,9,12].

Regarding the pooling methods, we had compared the relevance judgments based on the document pool created by collecting top X documents from one run per system, top X documents from every runs regardless of the number of runs submitted by a system. As a result, regardless of the different coverage of pooled methods, we have found strong correlation among the system rankings produced based on the different pooling for both training topics and test topics [9,12].

As Gerard Salton has pointed out and confirmed by Ellen Voorhees [14], when certain number of topics are used for system evaluation, the results will be solid as a whole regardless of the inconsistency of the relevance judgments or difference of the pooling of each topic. However, further investigation of the

4. Evaluation

Based on the relevant assessments, the interpolated recall and precision at 11 points, the average precision (non-interpolated) over all relevant documents and precision at 5, 10, 15, 20, 30, 100 documents will be calculated using the TREC's evaluation program, which is available from the ftp site of Cornell University.

Since this is our first experience in organizing such an evaluation workshop, it is important to follow a standard method so as to be accepted by the IR community, and to assist in the standardization of the evaluation of IR systems. With the kind understanding and cooperation from Drs Donna Harman and Ellen Voorhees on NIST, USA, we have used the TREC's evaluation program and part of the scripts prepared for the search results handling at TREC.

5. Retrieval Results

This section reports overview of the retrieval results from the aspects of system effectiveness and analyses some of the similarities and differences of the approaches that each participating group took. Since one of the main purpose of the NTCIR Workshop is to enhance the research in Japanese text retrieval and cross-lingual retrieval, to examine various approaches using NTCIR Collection has been encouraged. For the official test runs, new 53 topics (topics 0031--0083) were delivered. Forty-eight Ad Hoc runs and 69 cross-lingual runs were submitted. Relevance judgements were done by human analysts and delivered to the active participating groups by the beginning of June, 1999. For the further detail of each approach, please consult each system paper in this proceedings.

5.1 Ad Hoc Runs

Retrieval results were submitted from seventeen participating groups listed below and three runs from NTCIR organizing group to test the effectiveness of the internal pooling procedure. 'CatA' indicates the Category A participants, that are expected to report the detail of their system and 'CatB' indicates the Category B participants, that have to submitted a paper.

LIST of Ad Hoc Task Active Participants

AIRCS (NTT Communication Science Laboratories) CatB, *

Aoe Laboratory of Tokushima University (Department of Information Science & Intelligent Systems, University of Tokushima) CatB, JALAB

Central Research Laboratory, Hitachi, Ltd. CatA, *

Fuji Xerox ITDC (Fuji Xerox Co., Ltd.) CatB, *

Graduate School of Engineering, University of Tokyo (Graduate School of Engineering, University of Tokyo) CatA, R2D2

Human Media Res. Labs., NEC Corp. (Human Media Research Laboratories, NEC Corporation) CatB, *

Communications Research Laboratory, Ministry of Posts and Telecommunications (Communications Research Laboratory, Ministry of Posts and Telecommunications) CatA, CRL

JUSTSYSTEM CORPORATION (JUSTSYSTEM Corporation) CatB, JCSB

Kameda-laboratory (School of Information Technology, Faculty of Engineering, Tokyo University of Technology) CatA, KLAB

Matsushita Electric Industrial Co., Ltd. (Multimedia Systems Research Laboratory, Matsushita Electric Industrial Co., Ltd.) CatB, NTE15

RMIT Computer Science & CSIRO Maths and Information Science, CatA, RMIT

Shirai Lab., Waseda Univ. (School of Science and Engineering, Waseda University) CatA, WSLab

The Structured Index Team, (NACSIS B) CatA, STIX

Toyohashi University of Technology Software System Laboratory (Toyohashi University of Technology) CatA, sstut

UC Berkeley Text Retrieval Research Group (School of Information Management and Systems, University of California at Berkeley),CatA, BKYTR (Run ID: BK??????)

University of Tsukuba (University of Tsukuba) CatA, UTS

Gotoh Lab., Kanagawa Univ. CatB, TGL[3]

NTCIR IR organizer, IK

The number of submitted runs by the categories of query length and query methods is as follows;

5.1 Ad hoc all

	VeryShort	Short		Long	sum
		w/o title	w/ title		
w/o concept	0	28	0	6	34
w/ concept	2	1	1	11	14
sum	2	29		17	48

ad hoc auto

	VeryShort	Short		Long	sum
		w/o title	w/ title		
w/o concept	0	26	0	4	30
w/ concept	2	0	1	7	10
sum	2	27		11	40

ad hoc inter

	VeryShort	Short		Long	sum
		w/o title	w/ title		
w/o concept	0	2	0	2	34
w/ concept	0	0	0	4	4
sum	0	29		6	35

[3] No system paper included in the Proceedings

Table 5.Number of submitted Ad Hoc Runs

'Long' includes any runs using "narrative" of the topics. 'Short' includes any runs without "narrative", and 'Very Short' is the runs constructing queries without "narrative" nor "description".

(1)Ad Hoc All Runs

The R/P graph of the Ad Hoc All runs are shown in Fig 4. and Table 6.

Jscb -- group uses strong NLP oriented techniques for both Indexing and query processing, normalizing index terms, utilize phrase based on NLP techniques, and relevance feedback for both automatic and interactive runs.

BKJJBIFU -- Berkeley uses rather simple bigram, just discarding HIRAGANA (Phonetics, mostly used for functional words). Berkeley's word based approach was the longest match with dictionary and seems also not so deeply depends on NLP.

The most interesting things found of this evaluation is that, the two systems that took completely different approaches obtain very high scores. JSCB uses NLP techniques very well and BKJJBIFU puts the focus on the statistic approach of weighting algorithm. Traditionally, Japanese IR community have tended to pay too much attention to the methods of segmenting texts into tokens rather than retrieval models or algorithms themselves.

(2) Ad Hoc Short Queries Runs

The R/P graph of Ad Hoc Short Query Runs are shown in Fig. 5 and Table 7.

The rest of 11 point R/P graphs and related tables of top ranked runs of each category of topic length and query methods are shown as an appendix of this report. In the followings we will analyses some of the similarity and differences of the approaches so that this will help to envisage research using the collection in some extent.

(3) Topic Length and Query methods

- Long Query > Short Query, but some runs are vice versa
- Interactive > Automatic
- w/Concept+title+Description > Description or Title only.

As TREC's previous experiments has revealed, longer queries tended to obtain higher results. When comparing the search effectiveness obtained by a pair of runs conducted by the same system, for example, jscb1 and jscb2, BKJJBIDS and BKJJBIDS, several runs by CRL group, KCTRG1 and KCTRG2, and UTS1 and UTS2, longer queries obtain higher search effectiveness. The search effectiveness of interactive systems may vary. The level of human invention and interaction, numbers and characteristics of human subjects were also various. To investigate search effectiveness of interactive systems and cross-system comparison, explicit design of the testing is inevitable and it will be one of the issues we should follow.

Search w/Concept are expected to be more effective than a search w/o Concepts based on the previous TREC experiences. Searches using longer queries obtained higher search effectiveness at NTCIR Workshop but since there is no system submitted both search results with concepts and without concept with the same condition, direct comparison between them were not available.

Fig 4. Ad Hoc - All Runs (Relevant) top 12 runs

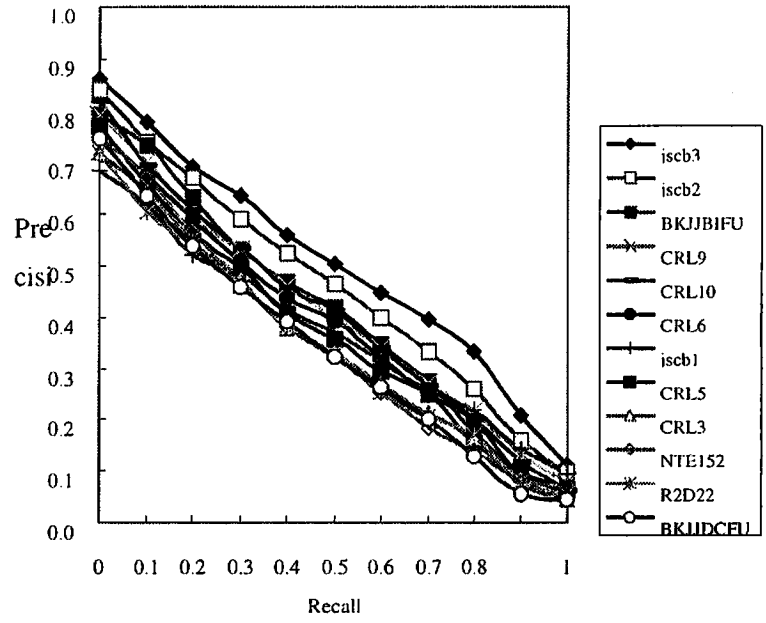

Table 6. Ad Hoc - All Runs (Relevant) top 12 runs

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb3	interact	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	RF(topXwords,reweighting)
jscb2	auto	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
BKJJBIFU	auto	long	w/concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
CRL9	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL10	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL6	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
jscb1	auto	short	w/o concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
CRL5	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL3	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
NTE152	auto	short+title	w/concept	word	longest match(EDR+internal)	inverted	Q=word	vector space	tf/idf+docLength+cooccurence	QE based on term similarity
R2D22	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	document vector expansion based on document interrelationship
BKJJDCFU	auto	long	w/concept	word+phrase	longest match(chasen+edict)	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	

Fig 5. Ad Hoc - Short Queries (Relevant) top 12 runs

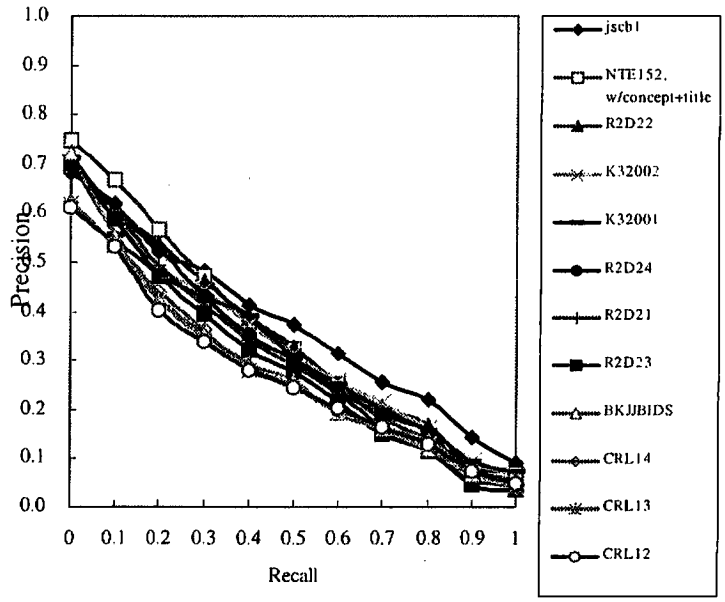

Table 7. Ad Hoc -- Short Queries (Relevant) top 12 runs.

RunID	method	topic	length	IndexUnit	indexing	indexSpec	queryUnit	Model	ranking	queryExpansion
jscb1	auto	short	w/o concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	
NTE152	auto	short+title	w/concept	word	longest match(EDR+internal)	inverted	Q=word	vector space	tf/idf+docLength+cooccurence	AutomaticRF
R2D22	auto	short	w/o concept	word	morphological+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	QE based on term similarity
K32002	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	document vector expansion based on document interrelationship
K32001	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
R2D24	auto	short	w/o concept	word	morphological+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf+proximity	
R2D21	auto	short	w/o concept	word	morphological+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	document vector expansion based on document interrelationship
R2D23	auto	short	w/o concept	word	morphological+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf+proximity	document vector expansion based on document interrelationship
BKJJBIDS	auto	short	w/o concept	bi-gram	discard HIRAGANA	inverted	Q=word	probabilistic	tf/idf+docLength+Qlength+collectionLength	document vector expansion based on document interrelationship
CRL14	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probabilistic	tf/idf+Qidf	
CRL13	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probabilistic	tf/idf+Qidf	
CRL12	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probabilistic	tf/idf+Qidf	

4) Indexing and Index Structure

Index Units for Japanese Text

Uni-gram: rmit, WLab, 1K

Bi-gram+:BKJJBIDS, BKJJBIFU, K3200, UTS, sstut

Word: (Morphological) anno,dove, KCTRG, r2d2, STIX, tgl

Word: (Longest match) NTE15

Word+Phrase: (Morphological) jscb

Word+Phrase:(Longest match) BKJJDCFU

Indexing methods used by the participating groups were roughly devided into four categories; uni-gram, bi-gram or its extention, word based, and word+phrase. When n-gram was used with the index structure of suffix array (or B-tree or Trie structure) or combination with positional information, more flexible searches , such as index by character and search by word+phrase, are available. Japanese text can be segmented into words by morphological analysis and selecting terms based on POS, or by longest matching with dictionary. The best run among all runs used NLP technique for segmentation, and the second best group for all runs used simple bi-gram, and second best group among all runs used simple bi-gram. Only one group (UC Berkeley) tested both bi-gram and word based approach (longest matching). No direct comparison between n-gram and NLP based word segmentation, nor n-gram with suffix array.

(5) Retrieval model and Weighting

Out of 17 groups, 13 groups used vector space model, and the two use Probabilistic model. As for the weighting, most of groups uses tf/idf and its variation. Document length is used by BK, NTE15, K3200, UTS, Dove. In addition to this, Query length and Collection length are used by BK, word co-occurrence in a document by NTE15, string length by UTS.

(6) Query Expansion

Query expansion was used by jscb, NTE15, and CRL4 of runs by CRL. In most of those cases, query expansion seems to work well and provides higher search effectiveness for both interactive search and automatic search.

5.2 Cross-Lingual Runs

Cross-lingual Retrieval task is an ad hoc retrieval task in which the documents are in English and the topics are in Japanese. Sixty-nine retrieval result sets are submitted from ten participating groups and NTCIR IR organizer. The organizer did to assess the effectiveness of the system that used for the initial pooling for this collection. Here is the list of the participants who submitted the search resutls;

LIST of Cross-Lingual Task Active Participants

Communications Research Laboratory, Ministry of Posts and Telecommunications, CatA, CRL

Matsushita Electric Industrial Co., Ltd. (Multimedia Systems Research Laboratory, Matsushita Electric Industrial Co., Ltd.) CatB, NTE15

Natural Language Processing Laboratory, Department of Computer, Science and Information Engineering, National Taiwan University , CatB, TSTAR

NEC C&C MEDIA RESEARCH LABORATORIES, CatB, SONIA

Toshiba R&D Center, CatA, TSB

Toyohashi University of Technology Software System Laboratory, CatA, sstut

UC Berkeley Text Retrieval Research Group (School of Information Managemant and Systems, University of California at Berkeley), CatA, BKYTR (Run ID: BK??????)

University of Library and Information Science, CatA, ULIS

University of Maryland (College of Library and Information Services, University of Maryland), CatA, UMD

FUJITSU ONE (FUJITSU LABORATORIES LTD.), CatB,FLAB[3]

NTCIR IR organizer group, 1K*

[3] no sysstem paper is included in this volume

Table 8.Number of submitted Cross-Lingual Runs

CrossLingual All

	VeryShort	Short		Long	concept only	sum
		w/o title	w/ title			
w/o concept	0	42	0	2		44
w/ concept	5	5	6	8	1	25
w/ Econce	5	5	5	4	1	
sum	5	53		10	1	69

(1) Cross-Lingual All Runs

Fig6. Cross-Lingual - All Runs (Relevant) top 12 runs

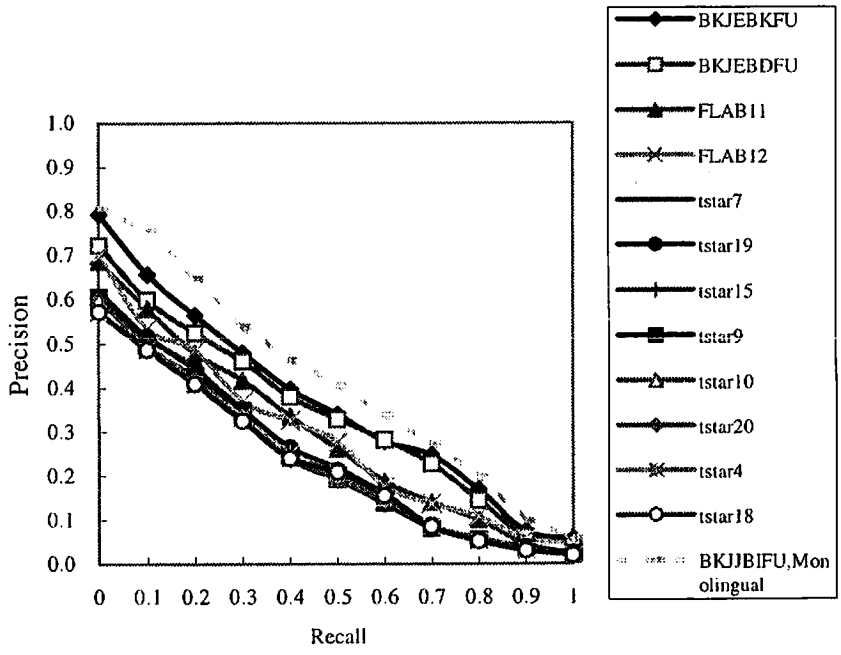

Table 9. Cross-Lingual All Runs (Relevant) Top 12 Runs

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBKFU	auto	long	w/concept	w/Econ	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
BKJEBDFU	auto	long	w/concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
FLAB11	interact	long	w/concept	w/oEcon	-	-	-	
FLAB12	interact	long	w/concept	w/oEcon	-	-	-	
tstar7	auto	very short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar19	auto	very short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar15	auto	very short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar9	auto	very short	w/concept	w/Econ	dictionary	afterTransl(LOB)	vector space	tf/idf+docLength
tstar10	auto	short+title	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar20	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar4	auto	short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar18	auto	short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength

Topic length

- Long Query > Short Query, but some runs are vice versa
- w/Concept > w/o Concepts, but the search using <Concept> only worked poor
- Short query w/title is also effective.
- Retrieval model: Vector space, probabilistic model, extension of probabilistic model, etc.

Regarding topic length and query methods, like Ad Hoc task, a search using longer topic tended to obtain better results, however, some runs are vice versa. Some of the runs could not utilize the narrative of the search topics to improve the search effectiveness.

TSTAR group tested heavily using <Title> fields with <Description> or <Concepts>. In their testing combination of <Title> and <Concepts> produced their best results.

There were only four interactive runs; two interactive runs with short queries without concepts and two interactive runs with long query with concept but without English concepts were submitted.

(2) Cross-Lingual Short Queries without Concepts

Fig 7. Cross-Lingual - Short w/o Concepts (Relevant) top 12 runs

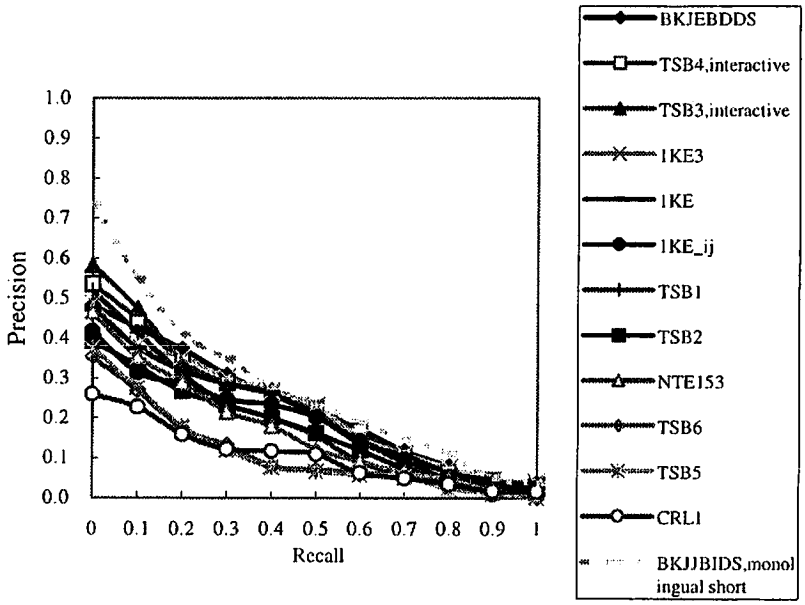

Table 10. Cross-Lingual Short Queries without Concepts Runs (Relevant) Top 12 Runs

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBDDS	auto	short	w/o concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
TSB4	interact	short	w/o concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength
TSB3	interact	short	w/o concept	w/oEcon	MT	local feedback	probablistic	tf/idf+docLength
IKE3	auto	short	w/o concept	w/oEcon	dictionary	QE wcluster	-	tf/idf+docLength
IKE	auto	short	w/o concept	w/oEcon	dictionary	QE wcluster	-	tf/idf+docLength
IKE_ij	auto	short	w/o concept	w/oEcon	dictionary	QE wcluster	-	tf/idf+docLength
TSB1	auto	short	w/o concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength
TSB2	auto	short	w/o concept	w/oEcon	MT	local feedback	probablistic	tf/idf+docLength
NTE153	auto	short	w/o concept	w/oEcon	corpus	noQE	vector space	tf/idf+docLength+proximity
TSB6	auto	short	w/o concept	w/oEcon	MT	noQE	vector space	tf/idf+proximity
TSB5	auto	short	w/o concept	w/oEcon	MT	noQE	vector space	tf/idf+proximity
CRL1	auto	very short	w/concept	w/oEcon	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength

(3) Approaches

Regarding the overall approach for Cross-Lingual retrieval, every runs took the approach of "Query Translation". The methods are categorized into the followings three;

(a) Dictionary Based: BK (except BKJEMTFU), CRL, sstut, TSTAR, ULIS, 1K

(b) Corpus Based : NTE15

(c) Machine Translation: BKJEMTFU (GiSELLE by ISI/USC), TBS (TBS1-2:ASTRANSAC, TBS5-6:Honyaku Professial v4.0), UMD (DQT with edict dictionary), SONIA (original)

(4) Dictionaries

Dictionaries used for query translation are as follows;

- Internal prepared (SONIA) ca.460K entries
- Internal prepared (BK) 373,477 entries, 23MB, using ntc1-je0.
- Internal prepared (CRL2) using ntc1-je0
- Internal prepared (sstut) 579,115 entries, 17MB, combine commercially available 5 bilingual dictionaries of subject domains of architecture, electronic & information, computer science, medicine, science & technology.
- Internal prepared , (1K-ij) 20,636 entries, combine commercially available 4 dictionaries on information processing and computer sciences
- Internal prepared (1KE, 1KE3) 37,170 entries using ca.10% of ntc1-je0.
- Not specified (TSTAR) 3MB
- EDR (CRL1) ?
- EDR (ULIS) technical terms: 24K entries, 0.50MB
- EDR (ULIS) General terms: ca. 260K entries, 7.4MB
- Edict (UMD) 64,433 entries, 2.6MB

(5) Technical terms

One of the difficult part of the document collection and topics of NTCIR-1 is technical terms. These are sometimes too specific or technical, or new to look up in the ordinary dictionaries. Moreover these technical terms are often expressed as English terms with original English spelling of Roman alphabet in Japanese text. In Japanese texts, a concept can be represented in four forms; Japanese terms (mainly Chinese characters), transliterated form of English terms using Katakata's, English acronyms, and English terms and term mismatch is caused by the variation. This problem is quite popular when we handle ordinary Japanese documents other than particular types of documents that are controlled by strict editorial policies not to use English terms like newspaper articles. These are also problematic to tune MT systems for the purpose.

In the other hand, quasi-paired documents of the own language and English like ntc1-je0 documents are not so difficult to find in the non-English speaking countries. These documents can be found various places such as, on the Web, scholarly documents, commercial documents describing the company products, or government documents, and so on. Using these documents to prepare bilingual or multilingual lexical resources that are usable for cross-lingual information access will be one of the practical approaches for the problem.

Using ntc1-je0 (JE Collection) to train systems or to extract knowledge are allowed in this NTCIR Workshop. However, using the bilingual lexical resource created from ntc1-je0 to translate queries is a kind of closed testing in the aspect of lexical resource preparing. It is wonderful to develop the appropriate methods to utilize the corpus to create bilingual lexical resource and to investigate the highest ceiling of the search effectiveness of cross-lingual retrieval using this collection. However, to obtain more solid evidence, these results should be tested against a new document set in the future. The Workshop organizer should provide the appropriate document sets for the purpose in the near future.

(6) Query Expansion & Disambiguation

Query expansion & word send disambiguation are conducted several groups. For example, TSTAR group looked at post-translate expansion and disambiguation using various kind of English monolingual corpora including ntc1-e0, TREC disk 4&5, LOB corpus. On the other hand, SONIA group looked at pre-translate expansion and disambiguation using ntc1-j0 or other parallel corpus (not specified). TSB2 and 4 incorporated automatic local feedback. ULIS, sstut, 1K groups took more natural query expansion, i.e. translating into more than one target language terms. It means translation and expansion at a time.

Though these line of investigation is promising, effects of these approaches are not so explicit in the search results submitted on March 4, more investigation and updating the results are expected.

(7) Model & Weighting Scheme

Vector space model is used by NTE15, TSTAR, ULIS, and TSB5 and TSB6. Probabilistic model is used by BKRTR, CRL, TSB1-4, UMD. TSB group tested both but MT systems used for TSB1-4 and the one for TSB5-6 are different, therefore direct comparison between these two models are not available.

Regarding weighting scheme, most runs use tf/idf or its extention. Idf is used by CRL, tf/idf is used by SONIA, sstut, ULIS, tf/idf+proximity is used by TSB5&6, tf/idf+document length is used by TSB1-4,TSTAR, and UMD. Tf/idf+document length+ proximity by NTE15, and tf/idf+document length+query length+collection length is used by BKRTR.

6. Summary and Future Directions

Through the overview of the Workshop, we have found various approaches and investigation have been tested using the NTCIR Collection. The results of the research have already reported at several international conferences. The lists of the publication on NTCIR and research using NTCIR-1 is listed at http://www.rd.nacsis.ac.jp/~ntcadm/paper1-en.html and will be enhanced.

Through the overview of the submitted runs of the first Workshop, the following things were found.

- Searches using longer topics obtain higher search

effectiveness than the ones using shorter topics

- <Concepts> and <titles> are effective to improve the search effectiveness, but the search using <concepts> only did not work well.
- Bigram, Word based approach with morphological and NLP analysis, and Suffix array (or B-tree or Trie) are effective for Japanese text retrieval.
- For Cross-lingual retrieval, both statistical approach and NLP approach are effective.
- Query expansion seems effective for both Ad Hoc and Cross Lingual retrieval but not so solid evidences were achieved in this Workshop. Further investigation and update are expected.
- Transliteration of Katakana terms improved the search effectiveness of Cross-lingual retrieval. Also the approach can be tested on Ad Hoc retrieval.

For the further study, we need to consider the following issues for further enhancement.

(1) Schedule: is the training period sufficient?

Regarding the next NTCIR Workshop, the call for participation is planed to start from April, 2000, documents and training topics will be distribute in May, test topics will be distributed in September or October, 2000, and the Workshop meeting will be in March, 2001.

(2) Evaluation of Cross-Lingual Retrieval

We used the evaluation method that search English documents by Japanese queries and comparing the system effectiveness against the monolingual searches of Japanese documents by Japanese queries. We understood that this is different from TREC's Cross-Language track. Futher consideration and discussion is needed.

(3) International collaboration for the cross-lingual evaluation.

(4) Evaluation method for the difficulty of search topics.

(5) Enhance the variation of text types. Copyright issues

(6) Is the support regarding Japanese text processing sufficient for non-Japanese participants?

(7) Further discussion of evaluation scheme and subtasks.

Especially evaluation of interactive system, using real Web documents including hyperlinks, post retrieval processing such as automatic abstracting, pinpointing the answers in the retrieved documents and so on.

ACKNOWLEDGMENTS

We thank all the participants for their contributions and the analysts who worked very hard with surprisingly excellent concentration. Special thanks are due to Donna Harman, Ellen Voorhees, Ross Wilkinson, Sung H. Myaeng, and Mun Kew Leong for their substantial advice and continuous support.

REFERENCES

[1] NTCIR Project. http://www.rd.nacsis.ac.jp/~ntcadm/

[2] Kando, N. Cross-Linguistic Scholarly Information Transfer and Database Services in Japan. Annual Meeting of the ASIS. (Nov. 1997) Washington DC.

[3] Kando, N. et al. NTCIR: NACSIS Test Collection Project [poster], the 20th Annual BCS-IRSG Colloquium, (March 1998), Autrans, France.

[4] Kageura, K. Overview of Term Recognition Task, In the Proceedings of the First NTCIR Workshop, August 30-Septmeber 1, 1999, Tokyo (to appear)

[5] The list of 65 academic societies is available at: URL: http://www.rd.nacsis.ac.jp/~ntcadm/acknowledge/thanks1-en.html

[6] Sakai, T. et al. BMIR-J2: Test Collection for Evaluation of Japanese Information Retrieval Systems. SIGIR Forum (to appear)

[7] Kageura, K. et al. NACSIS Corpus Project for IR and Terminological Research. NLPPR '97, (Dec. 1997), 493-49

[8] Kando, N. et al. NTCIR-1: Its Policy and Practice. IPSJ SIG Notes, Vol.99, No.20, p.33-40 (1999) [in Japanese]

[9] Kuriyama, K. et al. Pooling for a Large Scal Test Collection : Analysis of the Search Resutls for the Pre-test of the NTCIR-1 Workshop. IPSJ SIG Notes, Vol. 99 (May, 1999) [in Japanese]

[10] Kuriyama, K. et al. Construction of a Large Scale Test Collection : Analysis of the Training Topics of the NTCIR-1. IPSJ SIG Notes, Vol. 99 (July, 1999) [in Japanese]

[11] Kando, N. et al. Construction of a Large Scale Test Collection : Analysis of the Test Topics of the NTCIR-1. In the Proceedings of IPSJ Annual Meeting (to appear) [in Japanese]

[12] Kuriyama, K. et al. Pooling for a Large Scale Test Collection : Analysis of the Search Results for the first NTCIR Workshop. In the Proceedings of IPSJ Annual Meeting (to appear) [in Japanese]

[13] Rorvig, M. et al. The N-Gram Hypothesis Applied to Matched Sets of Visualized Japanese-English Technical Documents. In the Proceeding of the ASIS Annual Meeting (to appear)

[14] Voorhees, E.M. Variations in Relevance Judgments and the Measurement of Retrieval Effectiveness. In the Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Melbourne, August, 1998, p.315-323.

Fig 5.1-1. Ad Hoc - All Runs (Relevant) top 12 runs

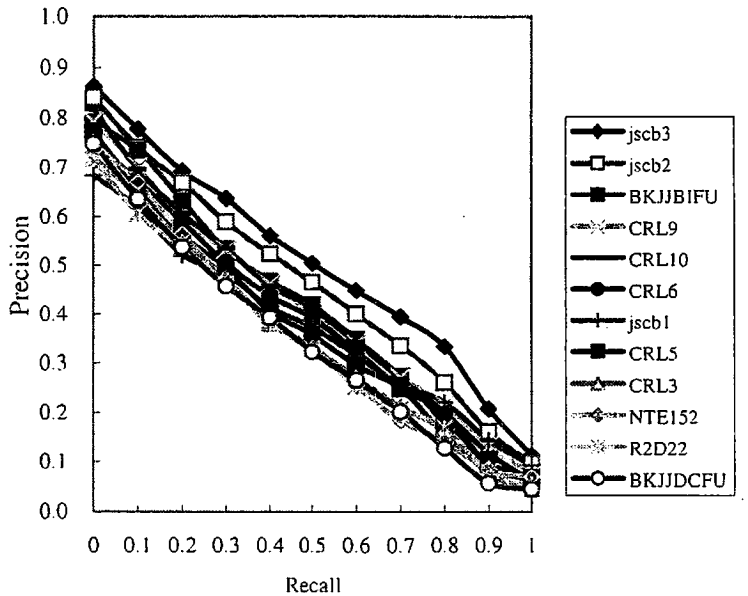

Table5.1-1b Ad Hoc - All Runs (Relevant) top 12 runs

n=48

	top-12 runs											
	jscb3	jscb2	BKJJBIFU	CRL9	CRL10	CRL6	jscb1	CRL5	CRL3	NTE152	R2D22	BKJJDCFU
at 5 docs	0.6679	0.5962	0.5736	0.5660	0.5623	0.5358	0.4528	0.5019	0.4981	0.5132	0.4906	0.5208
at 10 docs	0.5340	0.5170	0.4755	0.4717	0.4774	0.4736	0.3962	0.4283	0.4170	0.4151	0.4094	0.4264
at 15 docs	0.4579	0.4377	0.4214	0.4138	0.4189	0.3962	0.3572	0.3723	0.3597	0.3522	0.3547	0.3686
at 20 docs	0.4066	0.3887	0.3830	0.3679	0.3651	0.3575	0.3245	0.3302	0.3189	0.3038	0.3113	0.3198
at 30 docs	0.3522	0.3327	0.3233	0.3107	0.3038	0.2981	0.2843	0.2792	0.2755	0.2497	0.2629	0.2629
at 100 docs	0.1823	0.1711	0.1477	0.1604	0.1621	0.1509	0.1455	0.1500	0.1372	0.1196	0.1370	0.1277
at 200 docs	0.1133	0.1058	0.0881	0.0980	0.1018	0.0938	0.0919	0.0992	0.0851	0.0743	0.0852	0.0777
at 500 docs	0.0544	0.0512	0.0411	0.0468	0.0508	0.0451	0.0462	0.0491	0.0428	0.0371	0.0422	0.0372
at 1000 docs	0.0296	0.0279	0.0232	0.0253	0.0288	0.0252	0.0259	0.0298	0.0244	0.0214	0.0241	0.0206
r-prec	0.4648	0.4301	0.4160	0.3905	0.3956	0.3815	0.3505	0.3775	0.3310	0.3358	0.3323	0.3656
ave prec	0.4855	0.4436	0.3966	0.3917	0.3889	0.3716	0.3596	0.3479	0.3337	0.3316	0.3289	0.3252

Table5.1-1c Ad Hoc - All Runs (Relevant) top 12 runs

n=48

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb3	interact	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	RF(topXwords,reweighting)
jscb2	auto	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
BKJJBIFU	auto	long	w/concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
CRL9	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL10	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL6	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
jscb1	auto	short	w/o concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
CRL5	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL3	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
NTE152	auto	short+title	w/concept	word	longest match(EDR+internal)	inverted	Q=word	vector space	tf/idf+docLength+cooccurence	QE based on term similarity
R2D22	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	document vector expansion based on document interrelationship
BKJJDCFU	auto	long	w/concept	word+phrase	longest match(chasen+edict)	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	

Fig 5.1-2. Ad Hoc - All Runs (Partial Relevant) top 12 runs

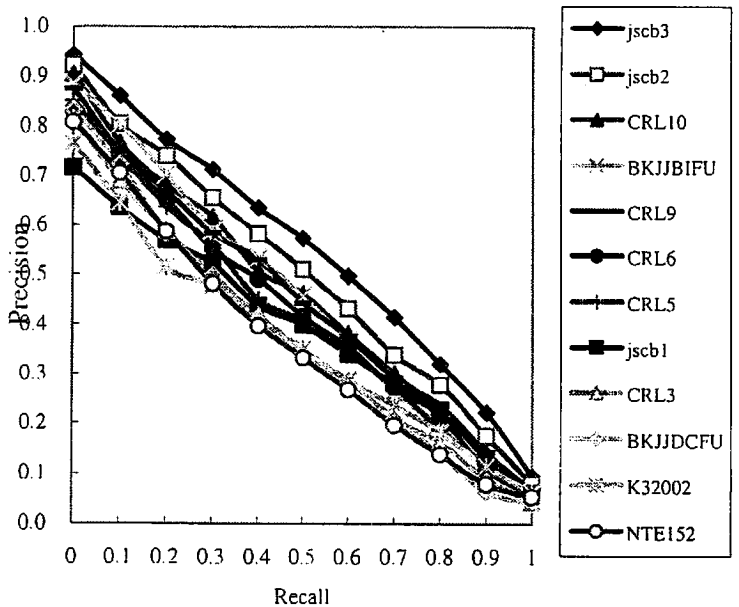

Table5.1-2b Ad Hoc - All Runs (Partial Relevant) top 12 runs n=48

	top-12 runs											
	jscb3	jscb2	CRL10	BKJJBIFU	CRL9	CRL6	CRL5	jscb1	CRL3	BKJJDC	K32002	NTE152
at 5 docs	0.7811	0.6981	0.6717	0.6642	0.6415	0.6264	0.6151	0.5321	0.5623	0.6000	0.5585	0.5925
at 10 docs	0.6434	0.6170	0.5679	0.5830	0.5623	0.5623	0.5170	0.4755	0.4849	0.5075	0.4792	0.4830
at 15 docs	0.5610	0.5358	0.5132	0.5170	0.4956	0.4855	0.4579	0.4239	0.4239	0.4478	0.4189	0.4176
at 20 docs	0.5000	0.4783	0.4538	0.4689	0.4434	0.4358	0.4132	0.3896	0.3802	0.4000	0.3877	0.3698
at 30 docs	0.4358	0.4101	0.3824	0.4000	0.3862	0.3616	0.3547	0.3491	0.3321	0.3346	0.3252	0.3050
at 100 docs	0.2364	0.2221	0.2096	0.1943	0.2089	0.1957	0.1923	0.1887	0.1726	0.1713	0.1711	0.1500
at 200 docs	0.1471	0.1388	0.1322	0.1185	0.1284	0.1240	0.1280	0.1208	0.1095	0.1054	0.1059	0.0947
at 500 docs	0.0695	0.0660	0.0655	0.0554	0.0611	0.0596	0.0633	0.0600	0.0557	0.0501	0.0522	0.0474
at 1000 docs	0.0375	0.0358	0.0365	0.0307	0.0330	0.0327	0.0374	0.0333	0.0316	0.0276	0.0300	0.0272
r-prec	0.5007	0.4673	0.4404	0.4456	0.4354	0.4122	0.4160	0.3804	0.3665	0.3936	0.3657	0.3485
ave prec	0.5358	0.4852	0.4404	0.4350	0.4323	0.4075	0.3922	0.3790	0.3575	0.3536	0.3495	0.3490

Table5.1-2c Ad Hoc - All Runs (Partial Relevant) top 12 runs n=48

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb3	interact	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	RF(topXwords.reweighting)
jscb2	auto	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
CRL10	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
BKJJBIFU	auto	long	w/concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
CRL9	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL6	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL5	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
jscb1	auto	short	w/o concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
CRL3	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
BKJJDCFU	auto	long	w/concept	word+phrase	longest match(chasen+edict)	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
K32002	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
NTE152	auto	short+title	w/concept	word	longest match(EDR+internal)	inverted	Q=word	vector space	tf/idf+docLength+cooccurence	QE based on term similarity

Fig 5.1-3. Ad Hoc - Short Queries (Relevant) top 12 runs

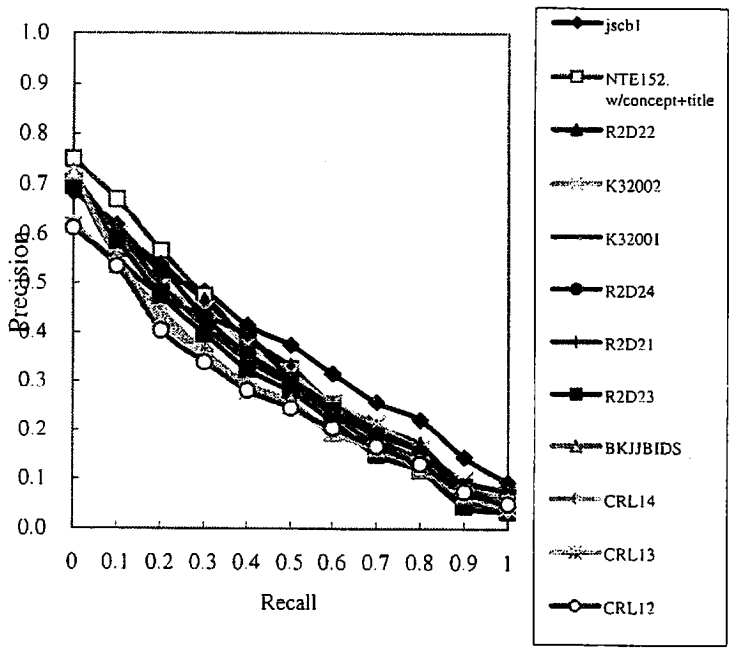

Table5.1-3b Ad Hoc - Short Queries (Relevant) top 12 runs

n=31

	top-12 runs											
	jscb1	NTE152	R2D22	K32002	K32001	R2D24	R2D21	R2D23	BKJJBIDS	CRL14	CRL13	CRL12
at 5 docs	0.4528	0.5132	0.4906	0.4717	0.4566	0.4868	0.4830	0.4755	0.4642	0.4075	0.4113	0.4075
at 10 docs	0.3962	0.4151	0.4094	0.3981	0.3811	0.4094	0.3925	0.3868	0.3623	0.3434	0.3340	0.3245
at 15 docs	0.3572	0.3522	0.3547	0.3484	0.3296	0.3447	0.3346	0.3220	0.3170	0.2943	0.2843	0.2906
at 20 docs	0.3245	0.3038	0.3113	0.3142	0.2981	0.2925	0.2953	0.2840	0.2689	0.2585	0.2557	0.2585
at 30 docs	0.2843	0.2497	0.2629	0.2535	0.2434	0.2434	0.2421	0.2302	0.2239	0.2176	0.2107	0.2057
at 100 docs	0.1455	0.1196	0.1370	0.1298	0.1262	0.1292	0.1215	0.1085	0.1077	0.1119	0.1102	0.1087
at 200 docs	0.0919	0.0743	0.0852	0.0808	0.0798	0.0785	0.0743	0.0701	0.0683	0.0708	0.0708	0.0710
at 500 docs	0.0462	0.0371	0.0422	0.0399	0.0392	0.0389	0.0368	0.0331	0.0341	0.0371	0.0368	0.0364
at 1000 docs	0.0259	0.0214	0.0241	0.0230	0.0226	0.0223	0.0213	0.0196	0.0198	0.0217	0.0216	0.0214
r-prec	0.3505	0.3358	0.3323	0.3273	0.3317	0.3225	0.3017	0.3010	0.2817	0.2872	0.2763	0.2699
ave prec	0.3596	0.3316	0.3289	0.3181	0.3159	0.3103	0.2933	0.2784	0.2715	0.2627	0.2584	0.2575

Table5.1-3 Ad Hoc - Short Queries (Relevant) top 12 runs

n=31

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb1	auto	short	w/o concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	
NTE152	auto	short+title	w/concept	word	longest match(EDR+internal)	inverted	Q=word	vector space	tf/idf+docLength+cooccurence	AutomaticRF
R2D22	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	QE based on term similarity
K32002	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	document vector expansion based on document interrelationship
K32001	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
R2D24	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf+proximity	
R2D21	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	document vector expansion based on document interrelationship
R2D23	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf+proximity	document vector expansion based on document interrelationship
BKJJBIDS	auto	short	w/o concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	document vector expansion based on document interrelationship
CRL14	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL13	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL12	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	

Fig 5.1-4. Ad Hoc - Short Queries (Partial Relevant) top 12 runs

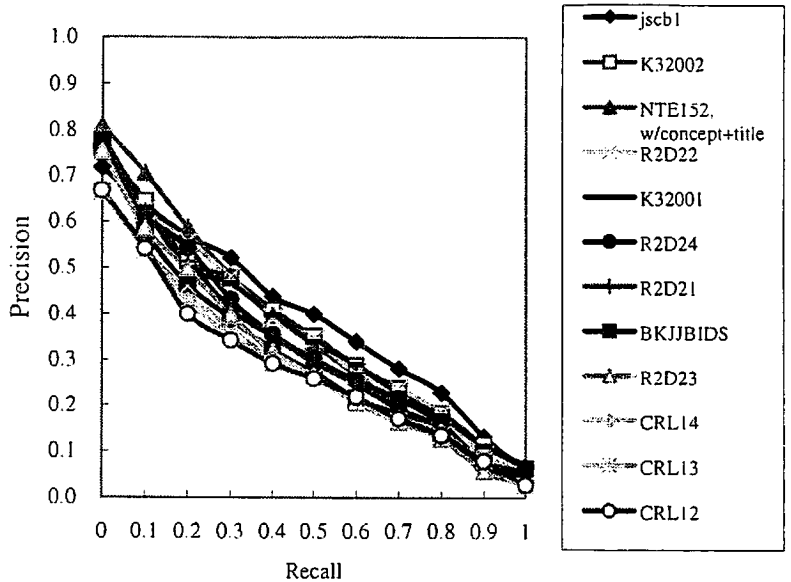

Table5.1-4b Ad Hoc - Short Queries (Partial Relevant) top 12 runs n=31

	top-12 runs											
	jscb1	K32002	NTE152	R2D22	K32001	R2D24	R2D21	BKJJBIDS	R2D23	CRL14	CRL13	CRL12
at 5 docs	0.5321	0.5585	0.5925	0.5509	0.5434	0.5245	0.5358	0.5245	0.5245	0.4491	0.4453	0.4340
at 10 docs	0.4755	0.4792	0.4830	0.4698	0.4623	0.4566	0.4528	0.4264	0.4340	0.3849	0.3774	0.3547
at 15 docs	0.4239	0.4189	0.4176	0.4201	0.4038	0.4038	0.3925	0.3811	0.3736	0.3346	0.3245	0.3270
at 20 docs	0.3896	0.3877	0.3698	0.3802	0.3717	0.3566	0.3528	0.3283	0.3387	0.2943	0.2925	0.2934
at 30 docs	0.3491	0.3252	0.3050	0.3270	0.3126	0.3025	0.2975	0.2799	0.2836	0.2560	0.2484	0.2409
at 100 docs	0.1887	0.1711	0.1500	0.1732	0.1666	0.1626	0.1511	0.1396	0.1358	0.1430	0.1392	0.1377
at 200 docs	0.1208	0.1059	0.0947	0.1072	0.1050	0.0990	0.0928	0.0880	0.0872	0.0921	0.0912	0.0912
at 500 docs	0.0600	0.0522	0.0474	0.0532	0.0512	0.0494	0.0463	0.0438	0.0419	0.0486	0.0479	0.0474
at 1000 docs	0.0333	0.0300	0.0272	0.0302	0.0294	0.0282	0.0269	0.0253	0.0249	0.0284	0.0280	0.0278
r-prec	0.3804	0.3657	0.3485	0.3555	0.3573	0.3328	0.3199	0.3147	0.3139	0.3040	0.2910	0.2753
ave prec	0.3790	0.3495	0.3490	0.3473	0.3454	0.3219	0.3097	0.2927	0.2862	0.2732	0.2675	0.2654

Table5.1-4c Ad Hoc - Short Queries (Partial Relevant) top 12 runs n=31

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb1	auto	short	w/o concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
NTE152	auto	short+title	w/concept	word	longest match(EDR+internal)	inverted	Q=word	vector space	tf/idf+docLength+cooccurence	QE based on term similarity
K32002	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
R2D22	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	document vector expansion based on document interrelationship
K32001	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
R2D24	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf+proximity	document vector expansion based on document interrelationship
R2D21	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	document vector expansion based on document interrelationship
BKJJBIDS	auto	short	w/o concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
R2D23	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf+proximity	document vector expansion based on document interrelationship
CRL14	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL13	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL12	auto	short	w/o concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	

Fig 5.1-5. Ad Hoc - Long Queries (Relevant) top 10 runs

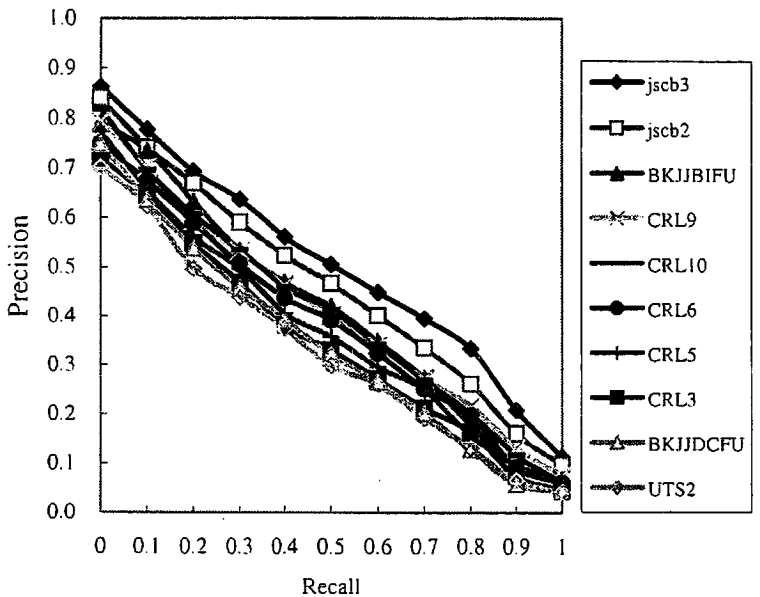

Table5.1-5b Ad Hoc - Long Queries (Relevant) top 10 runs n=17

	top 10 runs									
	jscb3	jscb2	BKJJBIFU	CRL9	CRL10	CRL6	CRL5	CRL3	BKJJDCFU	UTS2
at 5 docs	0.6679	0.5962	0.5736	0.5660	0.5623	0.5358	0.5019	0.4981	0.5208	0.4377
at 10 docs	0.5340	0.5170	0.4755	0.4717	0.4774	0.4736	0.4283	0.4170	0.4264	0.3830
at 15 docs	0.4579	0.4377	0.4214	0.4138	0.4189	0.3962	0.3723	0.3597	0.3686	0.3421
at 20 docs	0.4066	0.3887	0.3830	0.3679	0.3651	0.3575	0.3302	0.3189	0.3198	0.3019
at 30 docs	0.3522	0.3327	0.3233	0.3107	0.3038	0.2981	0.2792	0.2755	0.2629	0.2579
at 100 docs	0.1823	0.1711	0.1477	0.1604	0.1621	0.1509	0.1500	0.1372	0.1277	0.1334
at 200 docs	0.1133	0.1058	0.0881	0.0980	0.1018	0.0938	0.0992	0.0851	0.0777	0.0809
at 500 docs	0.0544	0.0512	0.0411	0.0468	0.0508	0.0451	0.0491	0.0428	0.0372	0.0385
at 1000 docs	0.0296	0.0279	0.0232	0.0253	0.0288	0.0252	0.0298	0.0244	0.0206	0.0214
r-prec	0.4648	0.4301	0.4160	0.3905	0.3956	0.3815	0.3775	0.3310	0.3656	0.3243
ave prec	0.4855	0.4436	0.3966	0.3917	0.3889	0.3716	0.3479	0.3337	0.3252	0.3043

Table5.1-5 Ad Hoc - Long Queries (Relevant) top 10 runs n=17

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb3	interact	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	RF(topXwords.reweighting)
jscb2	auto	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
BKJJBIFU	auto	long	w/concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
CRL9	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL10	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL6	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL5	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL3	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
BKJJDCFU	auto	long	w/concept	word+phrase	longest match(chasen+edict)	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
UTS2	auto	long	w/o concept	n-gram		suffix array		vector space	tf/idf+docLength+stringLength	

Fig 5.1-6. Ad Hoc - Long Queries (Paritial Relevant) top 10 runs

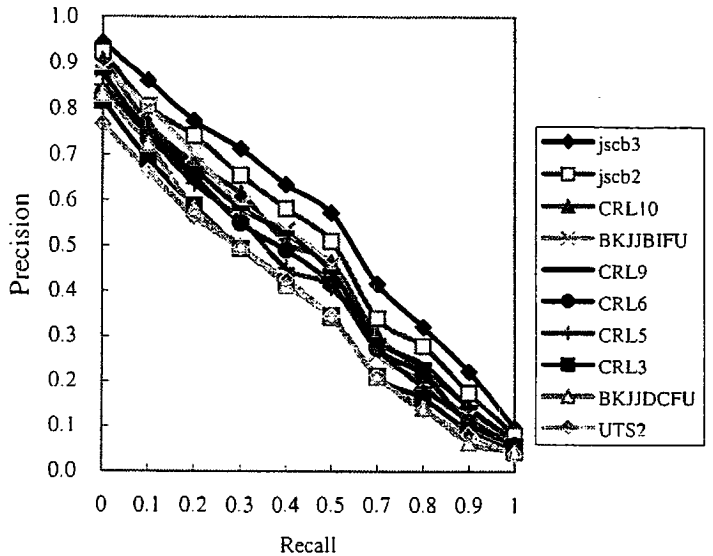

Table5.1-6b Ad Hoc - Long Queries (Partial Relevant) top 10 runs n=17

	top 10 runs									
	jscb3	jscb2	CRL10	BKJJBIFU	CRL9	CRL6	CRL5	CRL3	BKJJDCFU	UTS2
at 5 docs	0.7811	0.6981	0.6717	0.6642	0.6415	0.6264	0.6151	0.5623	0.6000	0.5321
at 10 docs	0.6434	0.6170	0.5679	0.5830	0.5623	0.5623	0.5170	0.4849	0.5075	0.4736
at 15 docs	0.5610	0.5358	0.5132	0.5170	0.4956	0.4855	0.4579	0.4239	0.4478	0.4252
at 20 docs	0.5000	0.4783	0.4538	0.4689	0.4434	0.4358	0.4132	0.3802	0.4000	0.3792
at 30 docs	0.4358	0.4101	0.3824	0.4000	0.3862	0.3616	0.3547	0.3321	0.3346	0.3283
at 100 docs	0.2364	0.2221	0.2096	0.1943	0.2089	0.1957	0.1923	0.1726	0.1713	0.1751
at 200 docs	0.1471	0.1388	0.1322	0.1185	0.1284	0.1240	0.1280	0.1095	0.1054	0.1084
at 500 docs	0.0695	0.0660	0.0655	0.0554	0.0611	0.0596	0.0633	0.0557	0.0501	0.0517
at 1000 docs	0.0375	0.0358	0.0365	0.0307	0.0330	0.0327	0.0374	0.0316	0.0276	0.0287
r-prec	0.5007	0.4673	0.4404	0.4456	0.4354	0.4122	0.4160	0.3665	0.3936	0.3688
ave prec	0.5358	0.4852	0.4404	0.4350	0.4323	0.4075	0.3922	0.3575	0.3536	0.3420

Table5.1-6c Ad Hoc - Long Queries (Partial Relevant) top 10 runs n=17

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb3	interact	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	RF(topXwords.reweighting)
jscb2	auto	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
CRL10	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
BKJJBIFU	auto	long	w/concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
CRL9	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL6	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL5	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL3	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
BKJJDCFU	auto	long	w/concept	word+phrase	longest match(chasen+edict)	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
UTS2	auto	long	w/o concept	n-gram		suffix array		vector space	tf/idf+docLength+stringLength	

Fig 5.1-7. Ad Hoc - Automatic (Relevant) top 12 runs

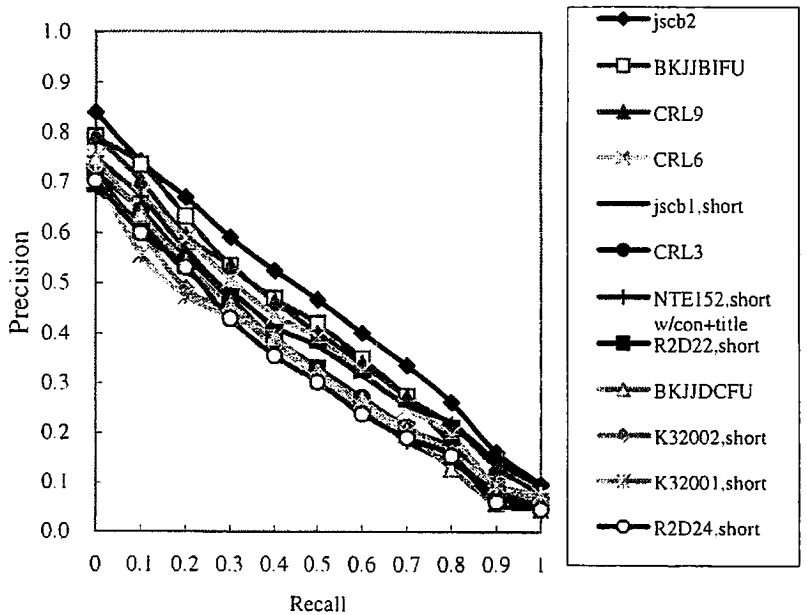

Table5.1-7b Ad Hoc - Automatic (Relevant) top 12 runs n=40

	top-12 runs											
	jscb2	BKJJBIF	CRL9	CRL6	jscb1	CRL3	NTE152	R2D22	BKJJDCFU	K32002	K32001	R2D24
at 5 docs	0.5962	0.5736	0.5660	0.5358	0.4528	0.4981	0.5132	0.4906	0.5208	0.4717	0.4566	0.4868
at 10 docs	0.5170	0.4755	0.4717	0.4736	0.3962	0.4170	0.4151	0.4094	0.4264	0.3981	0.3811	0.4094
at 15 docs	0.4377	0.4214	0.4138	0.3962	0.3572	0.3597	0.3522	0.3547	0.3686	0.3484	0.3296	0.3447
at 20 docs	0.3887	0.3830	0.3679	0.3575	0.3245	0.3189	0.3038	0.3113	0.3198	0.3142	0.2981	0.2925
at 30 docs	0.3327	0.3233	0.3107	0.2981	0.2843	0.2755	0.2497	0.2629	0.2629	0.2535	0.2434	0.2434
at 100 docs	0.1711	0.1477	0.1604	0.1509	0.1455	0.1372	0.1196	0.1370	0.1277	0.1298	0.1262	0.1292
at 200 docs	0.1058	0.0881	0.0980	0.0938	0.0919	0.0851	0.0743	0.0852	0.0777	0.0808	0.0798	0.0785
at 500 docs	0.0512	0.0411	0.0468	0.0451	0.0462	0.0428	0.0371	0.0422	0.0372	0.0399	0.0392	0.0389
at 1000 docs	0.0279	0.0232	0.0253	0.0252	0.0259	0.0244	0.0214	0.0241	0.0206	0.0230	0.0226	0.0223
r-prec	0.4301	0.4160	0.3905	0.3815	0.3505	0.3310	0.3358	0.3323	0.3656	0.3273	0.3317	0.3225
ave prec	0.4436	0.3966	0.3917	0.3716	0.3596	0.3337	0.3316	0.3289	0.3252	0.3181	0.3159	0.3103

Table5.1-7c Ad Hoc - Automatic (Relevant) top 12 runs n=40

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb2	auto	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
BKJJBIFU	auto	long	w/concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
CRL9	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL6	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
jscb1	auto	short	w/o concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
CRL3	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
NTE152	auto	short+title	w/concept	word	longest match(EDR+internal)	inverted	Q=word	vector space	tf/idf+docLength+cooccurence	QE based on term similarity
R2D22	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	document vector expansion based on document interrelationship
BKJJDCFU	auto	long	w/concept	word+phrase	longest match(chasen+edict)	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
K32002	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
K32001	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
R2D24	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf+proximity	document vector expansion based on document interrelationsh

Fig 5.1-8. Ad Hoc - Automatic (Partial Relevant) top 12 runs

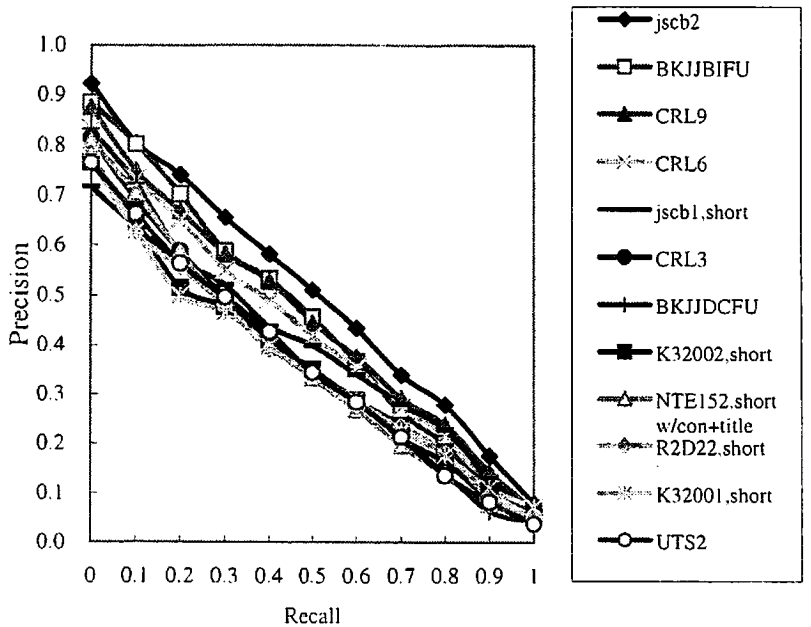

Table5.1-8b Ad Hoc - Automatic (Partial Relevant) top 12 runs n=40

	top-12 runs											
	jscb2	BKJJBIFU	CRL9	CRL6	jscb1	CRL3	BKJJDCFU	K32002	NTE152	R2D22	K32001	UTS2
at 5 docs	0.6981	0.6642	0.6415	0.6264	0.5321	0.5623	0.6000	0.5585	0.5925	0.5509	0.5434	0.5321
at 10 docs	0.6170	0.5830	0.5623	0.5623	0.4755	0.4849	0.5075	0.4792	0.4830	0.4698	0.4623	0.4736
at 15 docs	0.5358	0.5170	0.4956	0.4855	0.4239	0.4239	0.4478	0.4189	0.4176	0.4201	0.4038	0.4252
at 20 docs	0.4783	0.4689	0.4434	0.4358	0.3896	0.3802	0.4000	0.3877	0.3698	0.3802	0.3717	0.3792
at 30 docs	0.4101	0.4000	0.3862	0.3616	0.3491	0.3321	0.3346	0.3252	0.3050	0.3270	0.3126	0.3283
at 100 docs	0.2221	0.1943	0.2089	0.1957	0.1887	0.1726	0.1713	0.1711	0.1500	0.1732	0.1666	0.1751
at 200 docs	0.1388	0.1185	0.1284	0.1240	0.1208	0.1095	0.1054	0.1059	0.0947	0.1072	0.1050	0.1084
at 500 docs	0.0660	0.0554	0.0611	0.0596	0.0600	0.0557	0.0501	0.0522	0.0474	0.0532	0.0512	0.0517
at 1000 docs	0.0358	0.0307	0.0330	0.0327	0.0333	0.0316	0.0276	0.0300	0.0272	0.0302	0.0294	0.0287
r-prec	0.4673	0.4456	0.4354	0.4122	0.3804	0.3665	0.3936	0.3657	0.3485	0.3555	0.3573	0.3688
ave prec	0.4852	0.4350	0.4323	0.4075	0.3790	0.3575	0.3536	0.3495	0.3490	0.3473	0.3454	0.3420

Table5.1-8c Ad Hoc - Automatic (Partial Relevant) top 12 runs n=40

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb2	auto	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
BKJJBIFU	auto	long	w/concept	bi-gram	discard HIRAGANA	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
CRL9	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL6	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
jscb1	auto	short	w/o concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	AutomaticRF
CRL3	auto	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
BKJJDCFU	auto	long	w/concept	word+phrase	longest match(chasen+edict)	inverted	Q=word	probablistic	tf/idf+docLength+Qlength+collectionLength	
K32002	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
NTE152	auto	short+title	w/concept	word	longest match(EDR+internal)	inverted	Q=word	vector space	tf/idf+docLength+cooccurence	QE based on term similarity
R2D22	auto	short	w/o concept	word	morphological(Chasen)+POS+stopword+normalize	inverted	Q=word	vector space	tf/idf	document vector expansion based on document interrelationship
K32001	auto	short	w/o concept	n-gram		B-tree	Q=word	vector space	tf/idf+docLength	
UTS2	auto	long	w/o concept	n-gram		suffix array		vector space	tf/idf+docLength+stringLength	

Fig 5.1-9. Ad Hoc - Interactive (Relevant) top runs

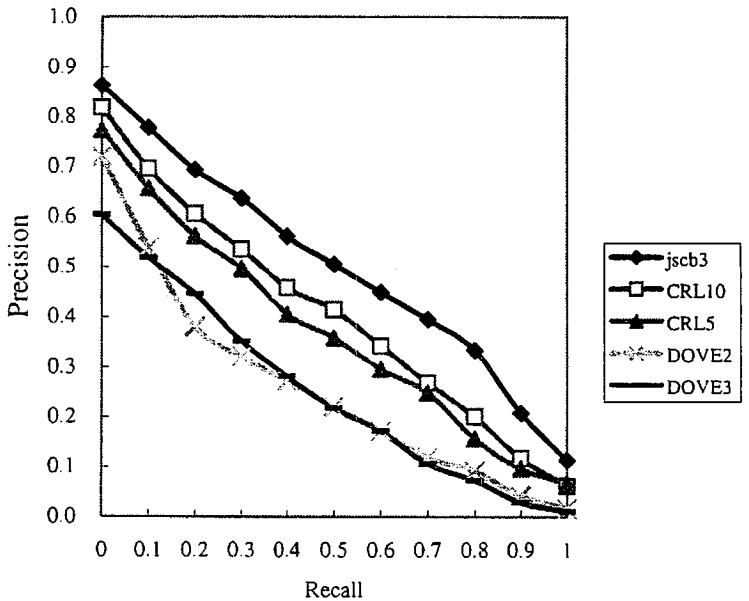

Table5.1-9b Ad Hoc - Interactive (Relevant) top runs

	top 5 runs n=8				
	jscb3	CRL10	CRL5	DOVE2	DOVE3
at 5 docs	0.6679	0.5623	0.5019	0.4075	0.3736
at 10 docs	0.5340	0.4774	0.4283	0.2811	0.3132
at 15 docs	0.4579	0.4189	0.3723	0.2327	0.2591
at 20 docs	0.4066	0.3651	0.3302	0.2075	0.2368
at 30 docs	0.3522	0.3038	0.2792	0.1723	0.1956
at 100 docs	0.1823	0.1621	0.1500	0.1030	0.1017
at 200 docs	0.1133	0.1018	0.0992	0.0670	0.0704
at 500 docs	0.0544	0.0508	0.0491	0.0374	0.0375
at 1000 docs	0.0296	0.0288	0.0298	0.0225	0.0221
r-prec	0.4648	0.3956	0.3775	0.2491	0.2727
ave prec	0.4855	0.3889	0.3479	0.2396	0.2317

Table5.1-9c Ad Hoc - Interactive (Relevant) top runs n=8

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb3	interact	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	RF(topXwords,reweighting)
CRL10	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL5	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
DOVE2	interact	long	w/o concept	word	morphological+POS+stopword+charClass	inverted	Q=word	vector space	tf/idf+docLength	RF(top200words)
DOVE3	interact	long	w/o concept	word	morphological+POS+stopword+charClass	inverted	Q=word	vector space	tf/idf+docLength	RF(userChosen)

Fig 5.1-10. Ad Hoc - Interactive (Partial Relevant) top runs

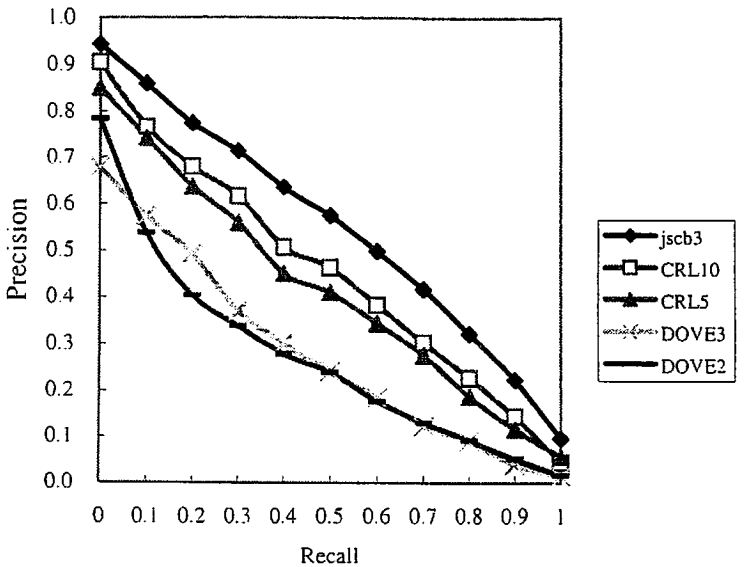

Table5.1-10b Ad Hoc - Interactive (PartialRelevant) top runs

	top 5 runs n=8				
	jscb3	CRL10	CRL5	DOVE3	DOVE2
at 5 docs	0.7811	0.6717	0.6151	0.4528	0.4642
at 10 docs	0.6434	0.5679	0.5170	0.3830	0.3453
at 15 docs	0.5610	0.5132	0.4579	0.3195	0.2830
at 20 docs	0.5000	0.4538	0.4132	0.2962	0.2538
at 30 docs	0.4358	0.3824	0.3547	0.2447	0.2119
at 100 docs	0.2364	0.2096	0.1923	0.1298	0.1315
at 200 docs	0.1471	0.1322	0.1280	0.0924	0.0880
at 500 docs	0.0695	0.0655	0.0633	0.0489	0.0483
at 1000 docs	0.0375	0.0365	0.0374	0.0282	0.0290
r-prec	0.5007	0.4404	0.4160	0.2951	0.2611
ave prec	0.5358	0.4404	0.3922	0.2556	0.2484

Table5.1-10c Ad Hoc - Interactive (Partial Relevant) top runs n=8

RunID	method	topic length		IndexUnit	indexing	indexStruc	queryUnit	Model	ranking	queryExpansion
jscb3	interact	long	w/concept	word+phrase	morphological+POS+stopword	inverted	Q=word+phrase	vector space	tf/idf	RF(topXwords.reweighting)
CRL10	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
CRL5	interact	long	w/concept	phrase?	stemming+dictionary(EDR)	--	Q=word+phrase	probablistic	tf/idf+Qidf	
DOVE3	interact	long	w/o concept	word	morphological+POS+stopword+charClass	inverted	Q=word	vector space	tf/idf+docLength	RF(userChosen)
DOVE2	interact	long	w/o concept	word	morphological+POS+stopword+charClass	inverted	Q=word	vector space	tf/idf+docLength	RF(top200words)

Fig 5.2-1. Cross-Lingual - All Runs (Relevant) top 12 runs

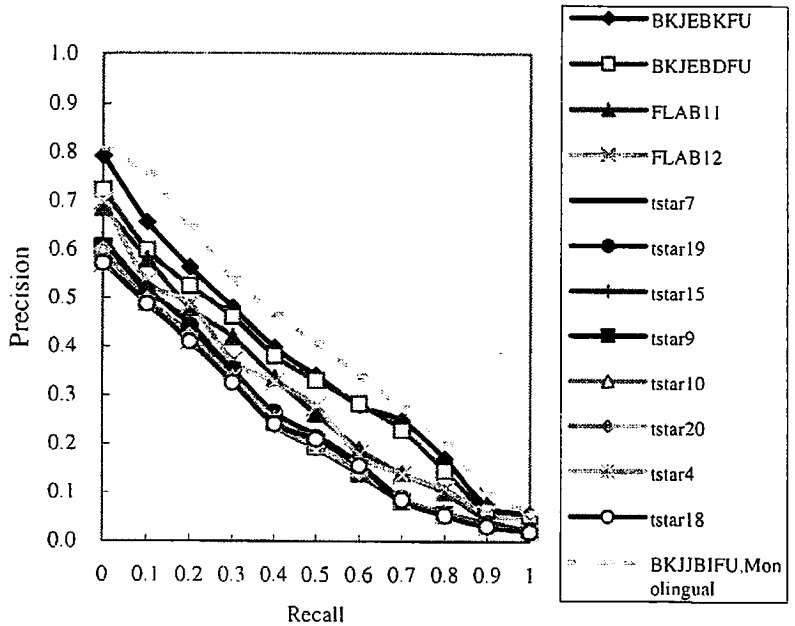

Table5.2-1b Cross-Lingual - All Runs (Relevant) top 12 runs n=69

	top-12 runs												Monolingual
	BKJEBKFU	BKJEBDFU	FLAB11	FLAB12	tstar7	tstar19	tstar15	tstar9	tstar10	tstar20	tstar4	tstar18	BKJJBIFU
at 5 docs	0.4462	0.4359	0.4051	0.3846	0.3897	0.3897	0.4000	0.3897	0.3795	0.3795	0.3487	0.3538	0.6000
at 10 docs	0.3641	0.3641	0.3256	0.2897	0.2667	0.2667	0.2795	0.2615	0.2718	0.2718	0.2718	0.2718	0.5077
at 15 docs	0.3145	0.3060	0.2684	0.2479	0.2274	0.2256	0.2239	0.2205	0.2205	0.2188	0.2171	0.2154	0.4718
at 20 docs	0.2756	0.2615	0.2436	0.2141	0.1910	0.1897	0.1885	0.1859	0.1795	0.1808	0.1782	0.1821	0.4397
at 30 docs	0.2162	0.2094	0.1923	0.1684	0.1564	0.1564	0.1538	0.1513	0.1538	0.1547	0.1487	0.1496	0.3795
at 100 docs	0.0972	0.0962	0.0921	0.0772	0.0797	0.0795	0.0769	0.0779	0.0813	0.0818	0.0805	0.0810	0.1782
at 200 docs	0.0569	0.0555	0.0558	0.0487	0.0503	0.0503	0.0495	0.0497	0.0509	0.0513	0.0512	0.0518	0.1069
at 500 docs	0.0273	0.0266	0.0277	0.0238	0.0253	0.0253	0.0252	0.0248	0.0252	0.0257	0.0252	0.0261	0.0501
at 1000 docs	0.0155	0.0152	0.0155	0.0139	0.0142	0.0142	0.0143	0.0141	0.0143	0.0145	0.0144	0.0147	0.0286
r-prec	0.3371	0.3191	0.2802	0.2694	0.2271	0.2255	0.2305	0.2225	0.2400	0.2390	0.2263	0.2267	0.4121
ave prec	0.3431	0.3216	0.2779	0.2676	0.2299	0.2286	0.2271	0.2202	0.2197	0.2193	0.2179	0.2163	0.3991

Table5.2-1c Cross-Lingual - All Runs (Relevant) top 12 runs n=69

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBKFU	auto	long	w/concept	w/Econ	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
BKJEBDFU	auto	long	w/concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
FLAB11	interact	long	w/concept	w/oEcon	-	-	-	
FLAB12	interact	long	w/concept	w/oEcon	-	-	-	
tstar7	auto	very short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar19	auto	very short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar15	auto	very short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar9	auto	very short	w/concept	w/Econ	dictionary	afterTransl(LOB)	vector space	tf/idf+docLength
tstar10	auto	short+title	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar20	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar4	auto	short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar18	auto	short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength

Fig 5.2-2. Cross-Lingual - All Runs (Partial Relevant) top 12 runs

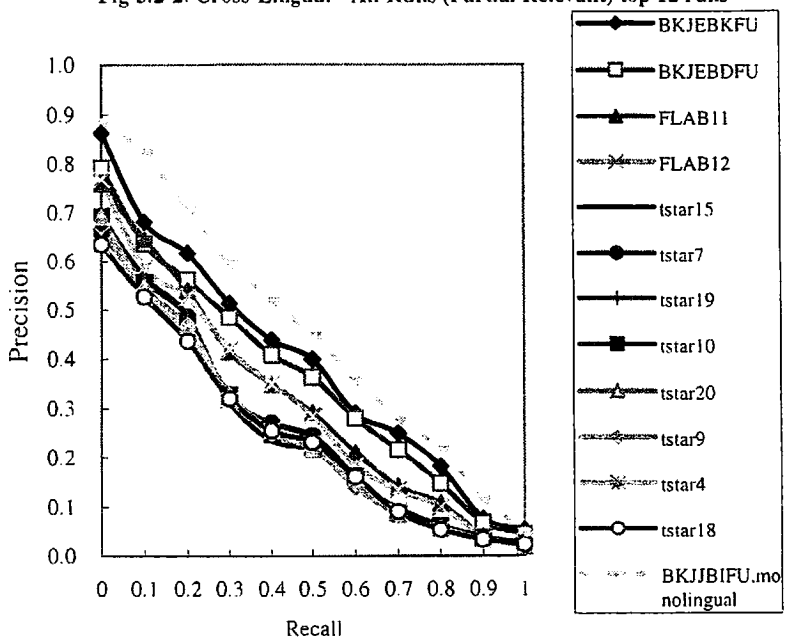

Table5.2-2b Cross-Lingual - All Runs (Partial Relevant) top 12 runs

n=69

	top-12 runs												Monolingual
	BKJEBKFU	BKJEBDFU	FLAB11	FLAB12	tstar15	tstar7	tstar19	tstar10	tstar20	tstar9	tstar4	tstar18	BKJJBII-U
at 5 docs	0.5385	0.5128	0.5077	0.4513	0.4513	0.4462	0.4462	0.4462	0.4462	0.4410	0.4051	0.4103	0.6923
at 10 docs	0.4410	0.4231	0.4077	0.3538	0.3205	0.3077	0.3077	0.3154	0.3154	0.3026	0.3128	0.3128	0.6128
at 15 docs	0.3846	0.3709	0.3402	0.3128	0.2581	0.2581	0.2564	0.2513	0.2496	0.2496	0.2496	0.2479	0.5658
at 20 docs	0.3449	0.3295	0.3038	0.2718	0.2205	0.2231	0.2218	0.2090	0.2103	0.2192	0.2103	0.2141	0.5269
at 30 docs	0.2761	0.2692	0.2462	0.2222	0.1846	0.1872	0.1872	0.1855	0.1863	0.1821	0.1786	0.1795	0.4615
at 100 docs	0.1328	0.1318	0.1244	0.1054	0.0979	0.0987	0.0985	0.1005	0.1010	0.0974	0.1013	0.1018	0.2321
at 200 docs	0.0781	0.0760	0.0749	0.0659	0.0660	0.0653	0.0653	0.0642	0.0646	0.0650	0.0660	0.0667	0.1427
at 500 docs	0.0371	0.0363	0.0374	0.0322	0.0341	0.0338	0.0338	0.0334	0.0339	0.0333	0.0336	0.0344	0.0670
at 1000 docs	0.0207	0.0204	0.0206	0.0186	0.0192	0.0189	0.0189	0.0189	0.0191	0.0188	0.0191	0.0194	0.0374
r-prec	0.3762	0.3572	0.3237	0.3035	0.2537	0.2586	0.2570	0.2581	0.2571	0.2537	0.2437	0.2441	0.4468
ave prec	0.3755	0.3438	0.3033	0.2828	0.2443	0.2437	0.2424	0.2402	0.2398	0.2339	0.2322	0.2315	0.4378

Table5.2-2c Cross-Lingual - All Runs (Partial Relevant) top 12 runs

n=69

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBKFU	auto	long	w/concept	w/Econ	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
BKJEBDFU	auto	long	w/concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
FLAB11	interact	long	w/concept	w/oEcon	-	-	-	-
FLAB12	interact	long	w/concept	w/oEcon	-	-	-	-
tstar15	auto	very short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar7	auto	very short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar19	auto	very short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar10	auto	short+title	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar20	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar9	auto	very short	w/concept	w/Econ	dictionary	afterTransl(LOB)	vector space	tf/idf+docLength
tstar4	auto	short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar18	auto	short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength

Fig 5.2-3. Cross-Lingual - Short w/o Concepts (Relevant) top 12 runs

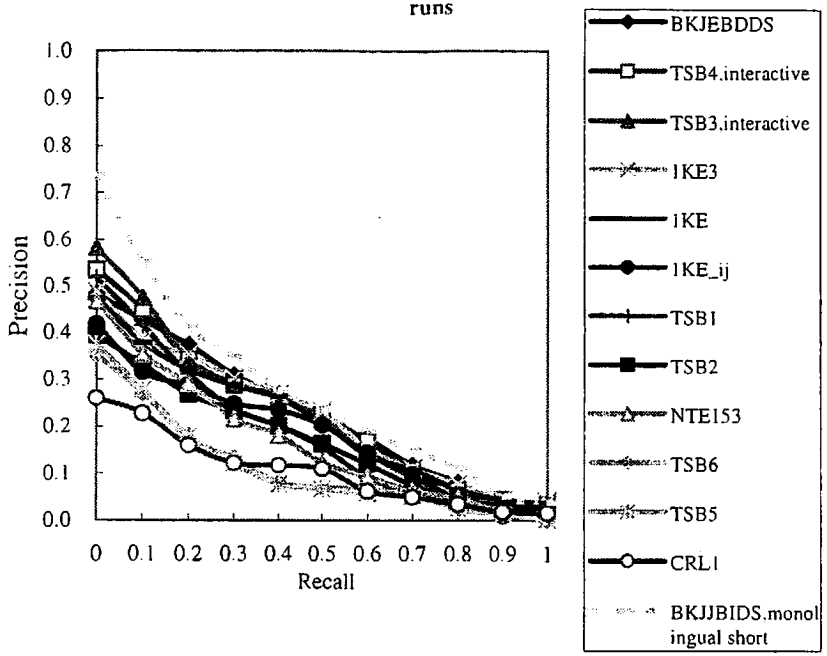

Table5.2-3b Cross-Lingual - Short Query without Concepts (Relevant) top 12 runs n=42

	top-12 runs												Monolingual
	BKJEBDDS	TSB4	TSB3	IKE3	IKE	IKE_ij	TSB1	TSB2	NTE153	TSB6	TSB5	CRL1	BKJJBIDS
at 5 docs	0.3077	0.3077	0.3026	0.2821	0.2872	0.2718	0.2359	0.2051	0.2410	0.2000	0.1949	0.1189	0.4769
at 10 docs	0.2410	0.2513	0.2359	0.2436	0.2333	0.2103	0.2051	0.2026	0.2051	0.1667	0.1564	0.1162	0.4000
at 15 docs	0.2205	0.2120	0.2103	0.1983	0.1880	0.1641	0.1795	0.1880	0.1675	0.1368	0.1265	0.1009	0.3590
at 20 docs	0.1885	0.1987	0.1795	0.1718	0.1603	0.1436	0.1641	0.1654	0.1487	0.1167	0.1064	0.0946	0.3051
at 30 docs	0.1376	0.1547	0.1479	0.1325	0.1282	0.1239	0.1333	0.1291	0.1265	0.0932	0.0846	0.0757	0.2624
at 100 docs	0.0692	0.0769	0.0721	0.0651	0.0605	0.0562	0.0615	0.0595	0.0674	0.0428	0.0403	0.0416	0.1297
at 200 docs	0.0413	0.0469	0.0433	0.0401	0.0372	0.0355	0.0388	0.0383	0.0413	0.0263	0.0249	0.0309	0.0831
at 500 docs	0.0209	0.0229	0.0222	0.0198	0.0199	0.0183	0.0199	0.0191	0.0210	0.0138	0.0137	0.0179	0.0419
at 1000 docs	0.0121	0.0133	0.0130	0.0120	0.0114	0.0113	0.0119	0.0117	0.0115	0.0074	0.0074	0.0112	0.0245
r-prec	0.2225	0.2453	0.2296	0.2223	0.1950	0.1976	0.1816	0.1872	0.1663	0.1235	0.1168	0.0908	0.2826
ave prec	0.2109	0.2090	0.2084	0.2062	0.1940	0.1713	0.1617	0.1524	0.1517	0.0997	0.0965	0.0940	0.2618

Table5.2-3c Cross-Lingual - Short Query without Concepts (Relevant) top 12 runs n=42

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBDDS	auto	short	w/o concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
TSB4	interact	short	w/o concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength
TSB3	interact	short	w/o concept	w/oEcon	MT	local feedback	probablistic	tf/idf+docLength
IKE3	auto	short	w/o concept	w/oEcon	dictionary	QE wcluster	-	tf/idf+docLength
IKE	auto	short	w/o concept	w/oEcon	dictionary	QE wcluster	-	tf/idf+docLength
IKE_ij	auto	short	w/o concept	w/oEcon	dictionary	QE wcluster	-	tf/idf+docLength
TSB1	auto	short	w/o concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength
TSB2	auto	short	w/o concept	w/oEcon	MT	local feedback	probablistic	tf/idf+docLength
NTE153	auto	short	w/o concept	w/oEcon	corpus	noQE	vector space	tf/idf+docLength+proximity
TSB6	auto	short	w/o concept	w/oEcon	MT	noQE	vector space	tf/idf+proximity
TSB5	auto	short	w/o concept	w/oEcon	MT	noQE	vector space	tf/idf+proximity
CRL1	auto	very short	w/concept	w/oEcon	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength

Fig 5.2-4. Cross-Lingual - Short w/o Concepts (Partial Relevant)
top 12 runs

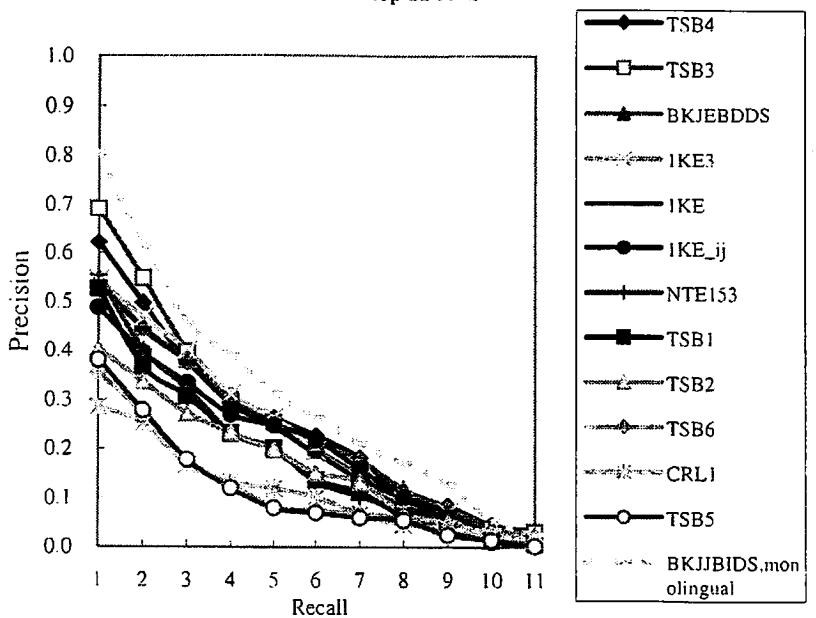

Table5.2-4b Cross-Lingual - Short Query without Concepts (Partial Relevant) top 1 n=42

	top-12 runs												Monolingual
	TSB4	TSB3	BKJEBDDS	IKE3	IKE	IKE_ij	NTE153	TSB1	TSB2	TSB6	CRL1	TSB5	BKJJBIDS
at 5 docs	0.3846	0.3744	0.3744	0.3436	0.3590	0.3282	0.3128	0.2667	0.2256	0.2051	0.1459	0.2051	0.5487
at 10 docs	0.3205	0.3000	0.2795	0.3000	0.2949	0.2590	0.2615	0.2231	0.2231	0.1769	0.1378	0.1692	0.4692
at 15 docs	0.2667	0.2667	0.2581	0.2513	0.2444	0.2085	0.2188	0.1949	0.2051	0.1470	0.1225	0.1350	0.4325
at 20 docs	0.2449	0.2269	0.2244	0.2205	0.2128	0.1846	0.1987	0.1769	0.1808	0.1244	0.1149	0.1128	0.3744
at 30 docs	0.1991	0.1872	0.1701	0.1718	0.1684	0.1615	0.1667	0.1462	0.1410	0.1017	0.0964	0.0915	0.3299
at 100 docs	0.1028	0.0936	0.0867	0.0874	0.0849	0.0741	0.0903	0.0715	0.0677	0.0500	0.0562	0.0469	0.1692
at 200 docs	0.0633	0.0577	0.0535	0.0555	0.0517	0.0477	0.0569	0.0464	0.0436	0.0314	0.0405	0.0291	0.1078
at 500 docs	0.0313	0.0300	0.0269	0.0283	0.0281	0.0243	0.0289	0.0247	0.0226	0.0175	0.0236	0.0172	0.0541
at 1000 docs	0.0180	0.0172	0.0158	0.0166	0.0157	0.0145	0.0163	0.0149	0.0141	0.0095	0.0149	0.0095	0.0314
r-prec	0.2731	0.2486	0.2384	0.2347	0.2183	0.2131	0.1946	0.1932	0.2001	0.1325	0.1043	0.1215	0.3182
ave prec	0.2260	0.2241	0.2205	0.2113	0.2046	0.1936	0.1691	0.1597	0.1558	0.0997	0.0965	0.0965	0.2888

Table5.2-4c Cross-Lingual - Short Query without Concepts (Partial Relevant) top 12 runs n=42

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
TSB4	interact	short	w/o concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength
TSB3	interact	short	w/o concept	w/oEcon	MT	local feedback	probablistic	tf/idf+docLength
BKJEBDDS	auto	short	w/o concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
IKE3	auto	short	w/o concept	w/oEcon	dictionary	noQE		tf/idf+docLength
IKE	auto	short	w/o concept	w/oEcon	dictionary	noQE		tf/idf+docLength
IKE_ij	auto	short	w/o concept	w/oEcon	dictionary	noQE		tf/idf+docLength
NTE153	auto	short	w/o concept	w/oEcon	corpus	noQE	vector space	tf/idf+docLength+proximity
TSB1	auto	short	w/o concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength
TSB2	auto	short	w/o concept	w/oEcon	MT	local feedback	probablistic	tf/idf+docLength
TSB6	auto	short	w/o concept	w/oEcon	MT	noQE	vector space	tf/idf+proximity
CRL1	auto	very short	w/concept	w/oEcon	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
TSB5	auto	short	w/o concept	w/oEcon	MT	noQE	vector space	tf/idf+proximity

Fig 5.2-5. Cross-Lingual - Short or VeryShort w/Concepts (Relevant) top 12 runs

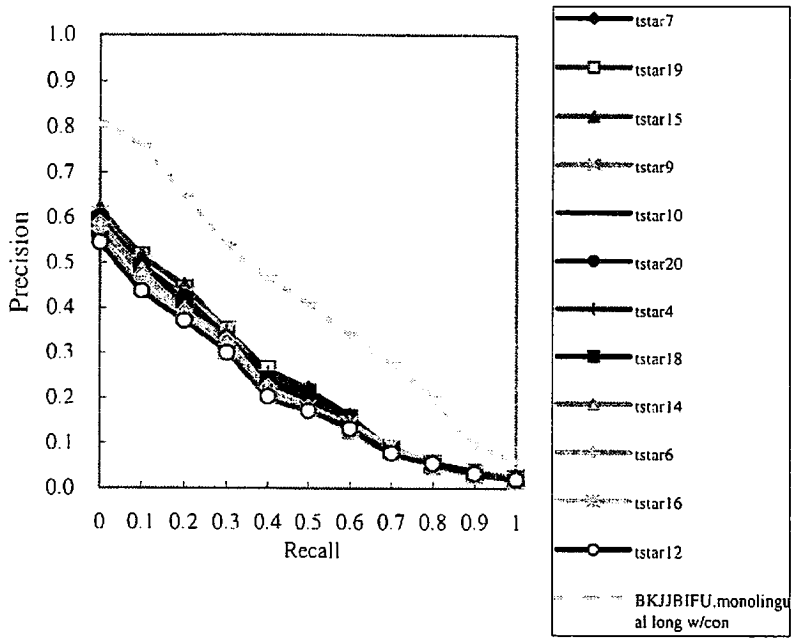

Table5.2-5b Cross-Lingual - Short or VeryShort Query with Concepts (Relevant) top 12 runs n=16

	top-12 runs												Monolingual
	tstar7	tstar19	tstar15	tstar9	tstar10	tstar20	tstar4	tstar18	tstar14	tstar6	tstar16	tstar12	BKJJBIFU
at 5 docs	0.3897	0.3897	0.4000	0.3897	0.3795	0.3795	0.3487	0.3538	0.3590	0.3538	0.3385	0.3487	0.6000
at 10 docs	0.2667	0.2667	0.2795	0.2615	0.2718	0.2718	0.2718	0.2718	0.2769	0.2590	0.2615	0.2564	0.5077
at 15 docs	0.2274	0.2256	0.2239	0.2205	0.2205	0.2188	0.2171	0.2154	0.2171	0.2085	0.2068	0.2051	0.4718
at 20 docs	0.1910	0.1897	0.1885	0.1859	0.1795	0.1808	0.1782	0.1821	0.1846	0.1679	0.1718	0.1731	0.4397
at 30 docs	0.1564	0.1564	0.1538	0.1513	0.1538	0.1547	0.1487	0.1496	0.1487	0.1359	0.1453	0.1444	0.3795
at 100 docs	0.0797	0.0795	0.0769	0.0779	0.0813	0.0818	0.0805	0.0810	0.0787	0.0759	0.0774	0.0769	0.1782
at 200 docs	0.0503	0.0503	0.0495	0.0497	0.0509	0.0513	0.0512	0.0518	0.0517	0.0486	0.0500	0.0485	0.1069
at 500 docs	0.0253	0.0253	0.0252	0.0248	0.0252	0.0257	0.0252	0.0261	0.0263	0.0245	0.0254	0.0243	0.0501
at 1000 docs	0.0142	0.0142	0.0143	0.0141	0.0143	0.0145	0.0144	0.0147	0.0147	0.0143	0.0142	0.0137	0.0286
r-prec	0.2271	0.2255	0.2305	0.2225	0.2400	0.2390	0.2263	0.2267	0.2177	0.2100	0.2166	0.2191	0.4121
ave prec	0.2299	0.2286	0.2271	0.2202	0.2197	0.2193	0.2179	0.2163	0.2059	0.2050	0.2019	0.1920	0.3991

Table5.2-5c Cross-Lingual - Short or VeryShort Query with Concepts (Relevant) top 12 runs n=16

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
tstar7	auto	very short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar19	auto	very short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar15	auto	very short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar9	auto	very short	w/concept	w/Econ	dictionary	afterTransl(LOB)	vector space	tf/idf+docLength
tstar10	auto	short+title	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar20	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar4	auto	short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar18	auto	short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar14	auto	short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar6	auto	short	w/concept	w/Econ	dictionary	afterTransl(LOB)	vector space	tf/idf+docLength
tstar16	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar12	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(LOB)	vector space	tf/idf+docLength

Fig 5.2-6. Cross-Lingual - Short or VeryShort w/Concepts (Paritial Relevant) top 12 runs

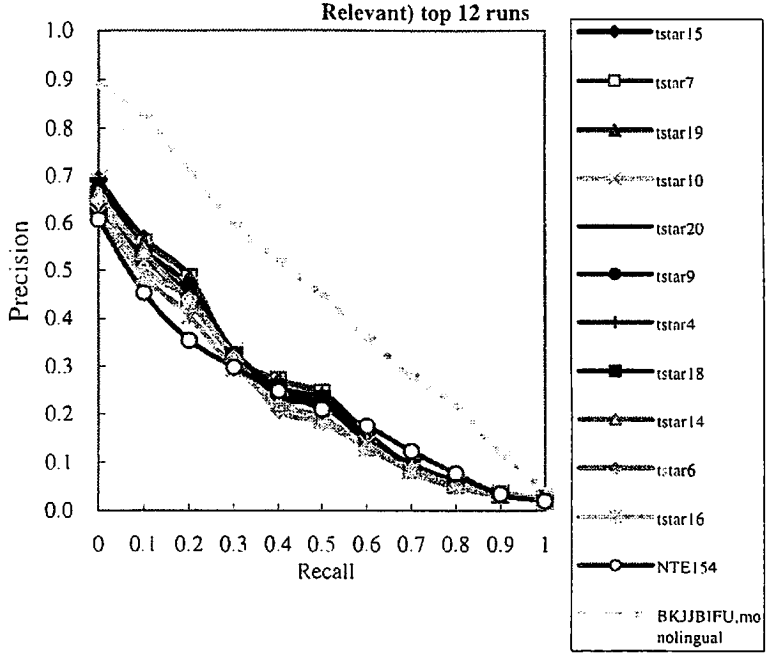

Table5.2-6b Cross-Lingual - Short or VeryShort Query with Concepts (Paritial Relevant) top 12 runs n=16

	top-12 runs												Monolingual
	tstar15	tstar7	tstar19	tstar10	tstar20	tstar9	tstar4	tstar18	tstar14	tstar6	tstar16	NTE154	BKJJBIFU
at 5 docs	0.4513	0.4462	0.4462	0.4462	0.4462	0.4410	0.4051	0.4103	0.4103	0.4000	0.3846	0.3436	0.6923
at 10 docs	0.3205	0.3077	0.3077	0.3154	0.3154	0.3026	0.3128	0.3128	0.3205	0.2974	0.3000	0.2923	0.6128
at 15 docs	0.2581	0.2581	0.2564	0.2513	0.2496	0.2496	0.2496	0.2479	0.2530	0.2376	0.2427	0.2684	0.5658
at 20 docs	0.2205	0.2231	0.2218	0.2090	0.2103	0.2192	0.2103	0.2141	0.2205	0.1987	0.2026	0.2385	0.5269
at 30 docs	0.1846	0.1872	0.1872	0.1855	0.1863	0.1821	0.1786	0.1795	0.1795	0.1667	0.1735	0.1991	0.4615
at 100 docs	0.0979	0.0987	0.0985	0.1005	0.1010	0.0974	0.1013	0.1018	0.1005	0.0974	0.0985	0.1015	0.2321
at 200 docs	0.0660	0.0653	0.0653	0.0642	0.0646	0.0650	0.0660	0.0667	0.0671	0.0638	0.0646	0.0615	0.1427
at 500 docs	0.0341	0.0338	0.0338	0.0334	0.0339	0.0333	0.0336	0.0344	0.0348	0.0329	0.0337	0.0317	0.0670
at 1000 docs	0.0192	0.0189	0.0189	0.0189	0.0191	0.0188	0.0191	0.0194	0.0194	0.0191	0.0189	0.0184	0.0374
r-prec	0.2537	0.2586	0.2570	0.2581	0.2571	0.2537	0.2437	0.2441	0.2431	0.2278	0.2353	0.2371	0.4468
ave prec	0.2443	0.2437	0.2424	0.2402	0.2398	0.2339	0.2322	0.2315	0.2264	0.2174	0.2143	0.2132	0.4378

Table5.2-6c Cross-Lingual - Short or VeryShort Query with Concepts (Paritial Relevant) top 12 runs n=16

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
tstar15	auto	very short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar7	auto	very short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar19	auto	very short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar10	auto	short+title	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar20	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar9	auto	very short	w/concept	w/Econ	dictionary	afterTransl(LOB)	vector space	tf/idf+docLength
tstar4	auto	short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar18	auto	short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar14	auto	short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar6	auto	short	w/concept	w/Econ	dictionary	afterTransl(LOB)	vector space	tf/idf+docLength
tstar16	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
NTE154	auto	very short	w/concept	w/oEcon	corpus	noQE	vector space	tf/idf+docLength+proximity

Fig 5.2-7. Cross-Lingual - Long Query (Relevant) top 10 runs

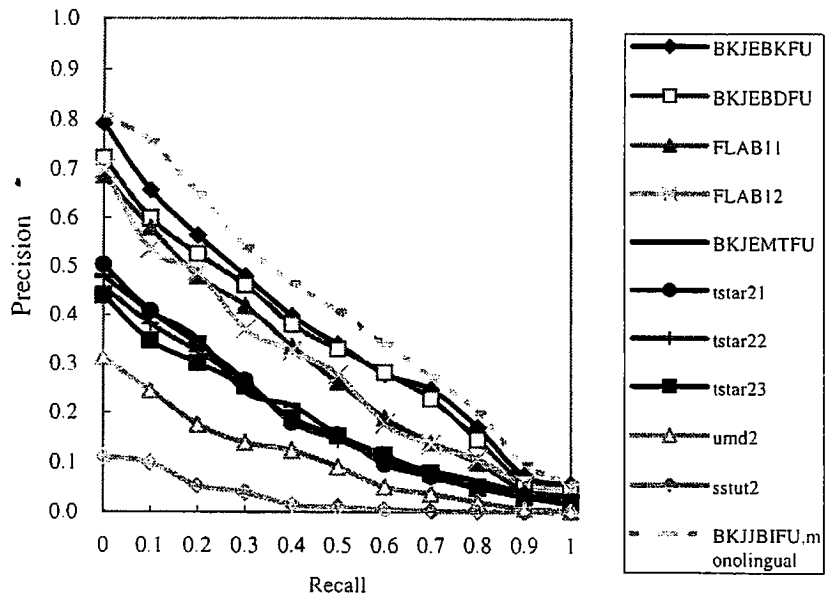

Table5.2-7b Cross-Lingual - Long Query (Relevant) top 10 runs n=10

	top runs										Monolingual
	BKJEBKFU	BKJEBDFU	FLAB11	FLAB12	BKJEMTFU	tstar21	tstar22	tstar23	umd2	sstut2	BKJJBIFU
at 5 docs	0.4462	0.4359	0.4051	0.3846	0.2923	0.2923	0.2769	0.2615	0.1862	0.0211	0.6000
at 10 docs	0.3641	0.3641	0.3256	0.2897	0.2436	0.2026	0.2000	0.1949	0.1552	0.0421	0.5077
at 15 docs	0.3145	0.3060	0.2684	0.2479	0.2034	0.1778	0.1761	0.1573	0.1402	0.0509	0.4718
at 20 docs	0.2756	0.2615	0.2436	0.2141	0.1782	0.1590	0.1628	0.1346	0.1448	0.0513	0.4397
at 30 docs	0.2162	0.2094	0.1923	0.1684	0.1462	0.1325	0.1393	0.1248	0.1172	0.0439	0.3795
at 100 docs	0.0972	0.0962	0.0921	0.0772	0.0733	0.0736	0.0726	0.0703	0.0524	0.0234	0.1782
at 200 docs	0.0569	0.0555	0.0558	0.0487	0.0485	0.0476	0.0469	0.0471	0.0328	0.0162	0.1069
at 500 docs	0.0273	0.0266	0.0277	0.0238	0.0254	0.0241	0.0236	0.0243	0.0178	0.0097	0.0501
at 1000 docs	0.0155	0.0152	0.0155	0.0139	0.0147	0.0138	0.0138	0.0140	0.0107	0.0067	0.0286
r-prec	0.3371	0.3191	0.2802	0.2694	0.2205	0.1943	0.1725	0.1741	0.1197	0.0453	0.4121
ave prec	0.3431	0.3216	0.2779	0.2676	0.1763	0.1750	0.1650	0.1607	0.0968	0.0251	0.3991

Table5.2-7c Cross-Lingual - Long Query (Relevant) top 10 runs n=10

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBKFU	auto	long	w/concept	w/Econ	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
BKJEBDFU	auto	long	w/concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
FLAB11	interact	long	w/concept	w/oEcon	-	-	-	
FLAB12	interact	long	w/concept	w/oEcon	-	-	-	
BKJEMTFU	auto	long	w/concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
tstar21	auto	long	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar22	auto	long	w/concept	w/Econ	dictionary	afterTransl(LOB,etc1-e0,TREC4&5)	vector space	tf/idf+docLength
tstar23	auto	long	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
umd2	auto	long	w/o concept	w/oEcon	MT	noQE	vector space	tf/idf+docLength
sstut2	auto	long	w/o concept	w/oEcon	dictionary	transl to multiple words	vector space	tf/idf

Fig 5.2-8. Cross-Lingual - Long Query (Partial Relevant) top 10 runs

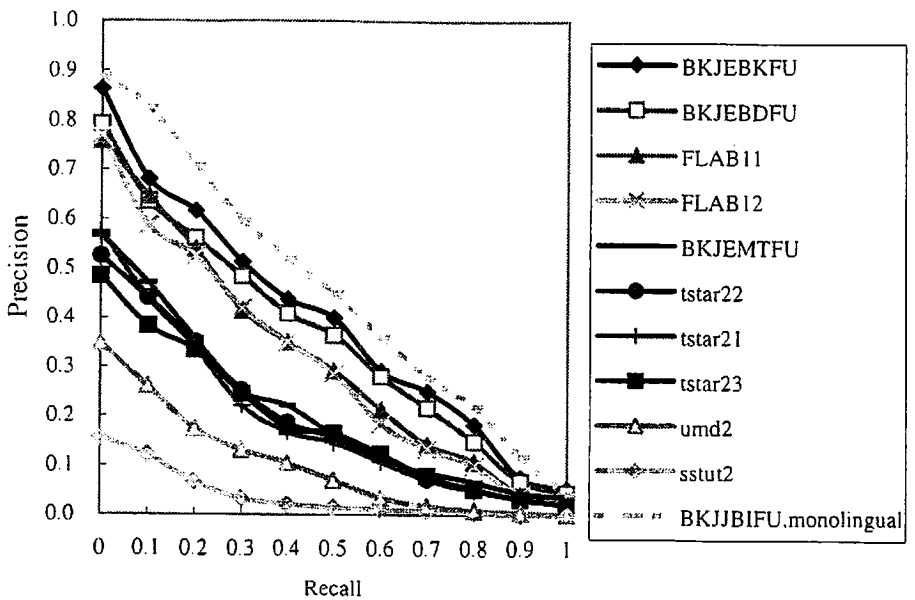

Table5.2-8b Cross-Lingual - Long Query (Partial Relevant) top 10 runs n=10

	top runs										Monolingual
	BKJEBKFU	BKJEBDFU	FLAB11	FLAB12	BKJEMTFU	tstar22	tstar21	tstar23	umd2	sstut2	BKJJBIFU
at 5 docs	0.5385	0.5128	0.5077	0.4513	0.3487	0.3282	0.3231	0.2821	0.2138	0.0474	0.6923
at 10 docs	0.4410	0.4231	0.4077	0.3538	0.2821	0.2410	0.2308	0.2205	0.1690	0.0632	0.6128
at 15 docs	0.3846	0.3709	0.3402	0.3128	0.2393	0.2188	0.2017	0.1829	0.1517	0.0702	0.5658
at 20 docs	0.3449	0.3295	0.3038	0.2718	0.2090	0.2026	0.1846	0.1564	0.1534	0.0737	0.5269
at 30 docs	0.2761	0.2692	0.2462	0.2222	0.1718	0.1735	0.1547	0.1479	0.1264	0.0623	0.4615
at 100 docs	0.1328	0.1318	0.1244	0.1054	0.0918	0.0956	0.0897	0.0918	0.0593	0.0353	0.2321
at 200 docs	0.0781	0.0760	0.0749	0.0659	0.0603	0.0628	0.0588	0.0617	0.0400	0.0246	0.1427
at 500 docs	0.0371	0.0363	0.0374	0.0322	0.0318	0.0322	0.0316	0.0327	0.0225	0.0154	0.0670
at 1000 docs	0.0207	0.0204	0.0206	0.0186	0.0185	0.0185	0.0181	0.0187	0.0139	0.0104	0.0374
r-prec	0.3762	0.3572	0.3237	0.3035	0.2358	0.2006	0.2069	0.1876	0.1376	0.0476	0.4468
ave prec	0.3755	0.3438	0.3033	0.2828	0.1925	#####	#####	0.1723	#####	0.0315	0.4378

Table5.2-8c Cross-Lingual - Long Query (Partial Relevant) top 10 runs n=10

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBKFU	auto	long	w/concept	w/Econ	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
BKJEBDFU	auto	long	w/concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
FLAB11	interact	long	w/concept	w/oEcon	-	-	-	
FLAB12	interact	long	w/concept	w/oEcon	-	-	-	
BKJEMTFU	auto	long	w/concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
tstar22	long	w/concept	w/Econ	dictionary	afterTransl(LOB,etc1-e0,TREC4&5)	vector space		tf/idf+docLength
tstar21	auto	long	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar23	auto	long	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
umd2	auto	long	w/o concept	w/oEcon	MT	noQE	vector space	tf/idf+docLength
sstut2	auto	long	w/o concept	w/oEcon	dictionary	transl to multiple words	vector space	tf/idf

Fig 5.2-9. Cross-Lingual - Automatic (Relevant) top 12 runs

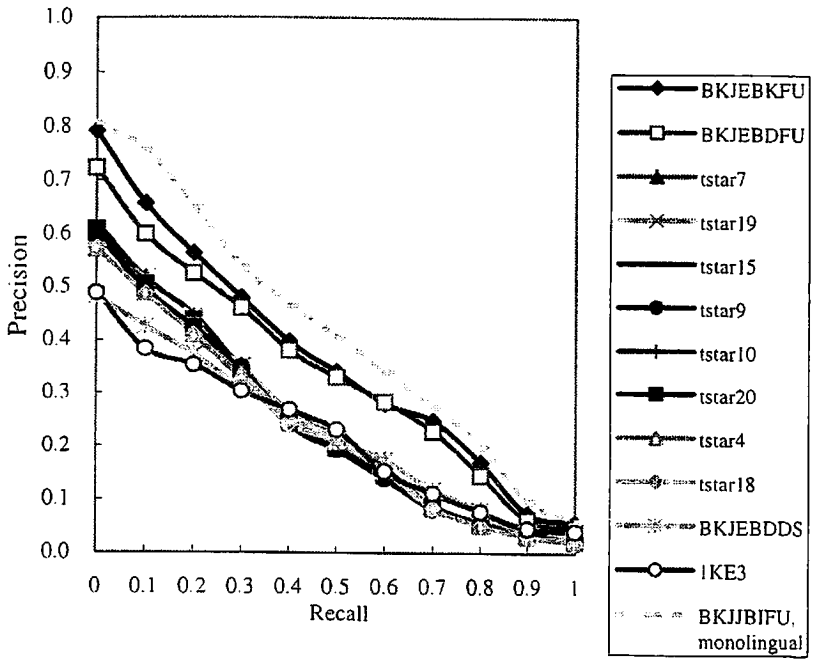

Table5.2-9b Cross-Lingual - Automatic (Relevant) top 12 runs n=65

	top-12 runs												Monolingual
	BKJEBKFU	BKJEBDFU	tstar7	tstar19	tstar15	tstar9	tstar10	tstar20	tstar4	tstar18	BKJEBDDS	IKE3	BKJJBIFU
at 5 docs	0.4462	0.4359	0.3897	0.3897	0.4000	0.3897	0.3795	0.3795	0.3487	0.3538	0.3077	0.2821	0.6000
at 10 docs	0.3641	0.3641	0.2667	0.2667	0.2795	0.2615	0.2718	0.2718	0.2718	0.2718	0.2410	0.2436	0.5077
at 15 docs	0.3145	0.3060	0.2274	0.2256	0.2239	0.2205	0.2205	0.2188	0.2171	0.2154	0.2205	0.1983	0.4718
at 20 docs	0.2756	0.2615	0.1910	0.1897	0.1885	0.1859	0.1795	0.1808	0.1782	0.1821	0.1885	0.1718	0.4397
at 30 docs	0.2162	0.2094	0.1564	0.1564	0.1538	0.1513	0.1538	0.1547	0.1487	0.1496	0.1376	0.1325	0.3795
at 100 docs	0.0972	0.0962	0.0797	0.0795	0.0769	0.0779	0.0813	0.0818	0.0805	0.0810	0.0692	0.0651	0.1782
at 200 docs	0.0569	0.0555	0.0503	0.0503	0.0495	0.0497	0.0509	0.0513	0.0512	0.0518	0.0413	0.0401	0.1069
at 500 docs	0.0273	0.0266	0.0253	0.0253	0.0252	0.0248	0.0252	0.0257	0.0252	0.0261	0.0209	0.0198	0.0501
at 1000 docs	0.0155	0.0152	0.0142	0.0142	0.0143	0.0141	0.0143	0.0145	0.0144	0.0147	0.0121	0.0120	0.0286
r-prec	0.3371	0.3191	0.2271	0.2255	0.2305	0.2225	0.2400	0.2390	0.2263	0.2267	0.2225	0.2223	0.4121
ave prec	0.3431	0.3216	0.2299	0.2286	0.2271	0.2202	0.2197	0.2193	0.2179	0.2163	0.2109	0.2062	0.3991

Table5.2-9c Cross-Lingual - Automatic (Relevant) top 12 runs n=65

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBKFU	auto	long	w/concept	w/Econ	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
BKJEBDFU	auto	long	w/concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
tstar7	auto	very short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar19	auto	very short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar15	auto	very short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar9	auto	short+title	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar10	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar20	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar4	auto	short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar18	auto	short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
BKJEBDDS	auto	short	w/o concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
IKE3	auto	short	w/o concept	w/oEcon	dictionary	QE wcluster	-	tf/idf+docLength

Fig 5.2-10. Cross-Lingual - Automatic (Partial Relevant) top 12 runs

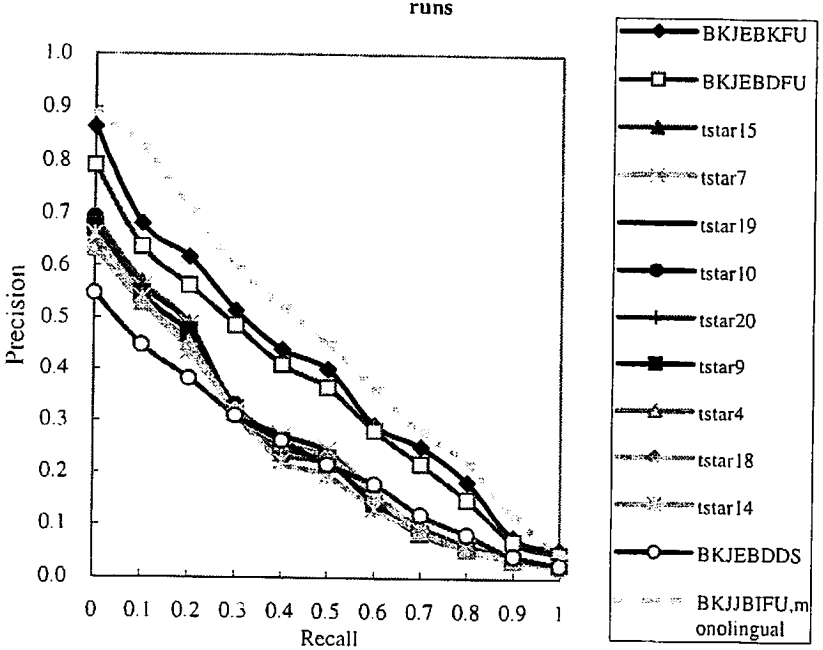

Table5.2-10b Cross-Lingual - Automatic (partial Relevant) top 12 rt n=65

	top-12 runs												Monolingual
	BKJEBKFU	BKJEBDFU	tstar15	tstar7	tstar19	tstar10	tstar20	tstar9	tstar4	tstar18	tstar14	BKJEBDDS	BKJJBIFU
at 5 docs	0.5385	0.5128	0.4513	0.4462	0.4462	0.4462	0.4462	0.4410	0.4051	0.4103	0.4103	0.3744	0.6923
at 10 docs	0.4410	0.4231	0.3205	0.3077	0.3077	0.3154	0.3154	0.3026	0.3128	0.3128	0.3205	0.2795	0.6128
at 15 docs	0.3846	0.3709	0.2581	0.2581	0.2564	0.2513	0.2496	0.2496	0.2496	0.2479	0.2530	0.2581	0.5658
at 20 docs	0.3449	0.3295	0.2205	0.2231	0.2218	0.2090	0.2103	0.2192	0.2103	0.2141	0.2205	0.2244	0.5269
at 30 docs	0.2761	0.2692	0.1846	0.1872	0.1872	0.1855	0.1863	0.1821	0.1786	0.1795	0.1795	0.1701	0.4615
at 100 docs	0.1328	0.1318	0.0979	0.0987	0.0985	0.1005	0.1010	0.0974	0.1013	0.1018	0.1005	0.0867	0.2321
at 200 docs	0.0781	0.0760	0.0660	0.0653	0.0653	0.0642	0.0646	0.0650	0.0660	0.0667	0.0671	0.0535	0.1427
at 500 docs	0.0371	0.0363	0.0341	0.0338	0.0338	0.0334	0.0339	0.0333	0.0336	0.0344	0.0348	0.0269	0.0670
at 1000 docs	0.0207	0.0204	0.0192	0.0189	0.0189	0.0189	0.0191	0.0188	0.0191	0.0194	0.0194	0.0158	0.0374
r-prec	0.3762	0.3572	0.2537	0.2586	0.2570	0.2581	0.2571	0.2537	0.2437	0.2441	0.2431	0.2384	0.4468
ave prec	0.3755	0.3438	0.2443	0.2437	0.2424	0.2402	0.2398	0.2339	0.2322	0.2315	0.2264	0.2205	0.4378

Table5.2-10c Cross-Lingual - Automatic (partial Relevant) top 12 runs n=65

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
BKJEBKFU	auto	long	w/concept	w/Econ	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
BKJEBDFU	auto	long	w/concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLength
tstar15	auto	very short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
tstar7	auto	very short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar19	auto	very short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar10	auto	short+title	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar20	auto	short+title	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar9	auto	short+title	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar4	auto	short	w/concept	w/Econ	dictionary	noQE	vector space	tf/idf+docLength
tstar18	auto	short	w/concept	w/Econ	dictionary	afterTransl(etc1-e0)	vector space	tf/idf+docLength
tstar14	auto	short	w/concept	w/Econ	dictionary	afterTransl(TREC4&5)	vector space	tf/idf+docLength
BKJEBDDS	auto	short	w/o concept	w/oEcon	dictionary	noQE	probablistic	tf/idf+docLength+qLength+collectionLengt

Fig 5.2-11. Cross-Lingual - Interactive (Relevant) top runs

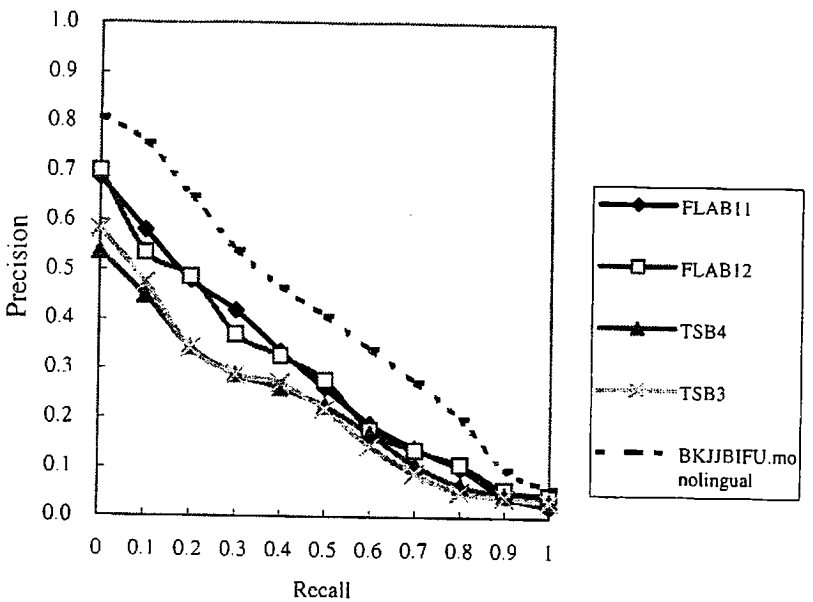

Table5.2-11b Cross-Lingual - Interactive (Relevant) top runs

n=4

	top runs				Monolingual
	FLAB11	FLAB12	TSB4	TSB3	BKJJBIFU
at 5 docs	0.4051	0.3846	0.3077	0.3026	0.6000
at 10 docs	0.3256	0.2897	0.2513	0.2359	0.5077
at 15 docs	0.2684	0.2479	0.2120	0.2103	0.4718
at 20 docs	0.2436	0.2141	0.1987	0.1795	0.4397
at 30 docs	0.1923	0.1684	0.1547	0.1479	0.3795
at 100 docs	0.0921	0.0772	0.0769	0.0721	0.1782
at 200 docs	0.0558	0.0487	0.0469	0.0433	0.1069
at 500 docs	0.0277	0.0238	0.0229	0.0222	0.0501
at 1000 docs	0.0155	0.0139	0.0133	0.0130	0.0286
r-prec	0.2802	0.2694	0.2453	0.2296	0.4121
ave prec	0.2779	0.2676	0.2090	0.2084	0.3991

Table5.2-11c Cross-Lingual - Automatic (Relevant) top 12 runs

n=4

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
FLAB11	interact	long	w/concept	w/oEcon	-	-	-	
FLAB12	interact	long	w/concept	w/oEcon	-	-	-	
TSB4	interact	short	w/o concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength
TSB3	interact	short	w/o concept	w/oEcon	MT	local feedback	probablistic	tf/idf+docLength

Fig 5.2-12. Cross-Lingual - Interactive (Partial Relevant) top runs

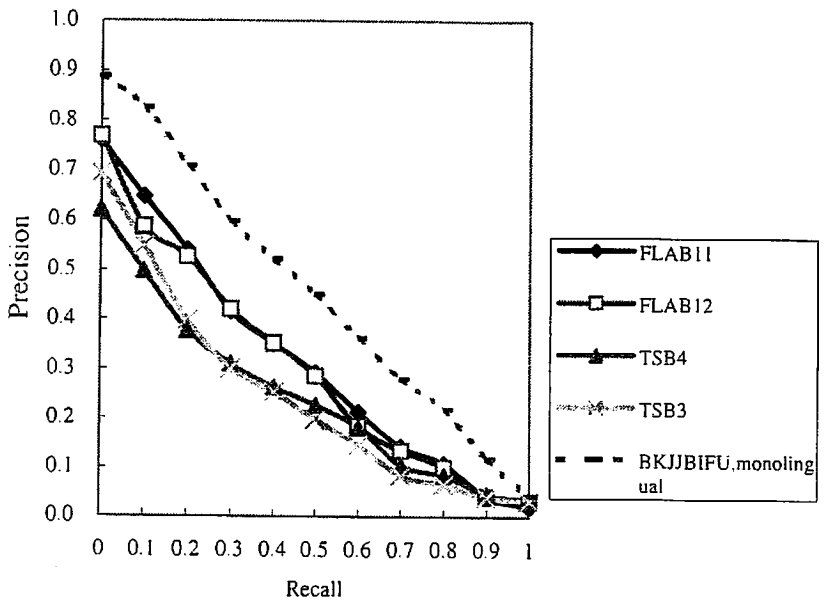

Table5.2-12b Cross-Lingual - Interactive (Paritial Relevant) top runs

n=4

	top runs				Monolingual
	FLAB11	FLAB12	TSB4	TSB3	BKJJBIFU
at 5 docs	0.5077	0.4513	0.3846	0.3744	0.6923
at 10 docs	0.4077	0.3538	0.3205	0.3000	0.6128
at 15 docs	0.3402	0.3128	0.2667	0.2667	0.5658
at 20 docs	0.3038	0.2718	0.2449	0.2269	0.5269
at 30 docs	0.2462	0.2222	0.1991	0.1872	0.4615
at 100 docs	0.1244	0.1054	0.1028	0.0936	0.2321
at 200 docs	0.0749	0.0659	0.0633	0.0577	0.1427
at 500 docs	0.0374	0.0322	0.0313	0.0300	0.0670
at 1000 docs	0.0206	0.0186	0.0180	0.0172	0.0374
r-prec	0.3237	0.3035	0.2731	0.2486	0.4468
ave prec	0.3033	0.2828	0.2260	0.2241	0.4378

Table5.2-12c Cross-Lingual - Automatic (Partial Relevant) top 12 runs

n=4

RunID	method	topic length			queryTranslation	queryExpansion	Model	ranking
FLAB11	interact	long	w/concept	w/oEcon	-	-	-	
FLAB12	interact	long	w/concept	w/oEcon	-	-	-	
TSB4	interact	short	w/o concept	w/oEcon	MT	noQE	probablistic	tf/idf+docLength
TSB3	interact	short	w/o concept	w/oEcon	MT	local feedback	probablistic	tf/idf+docLength