

Overview of Patent Retrieval Task at NTCIR-3

Makoto Iwayama⁺, Atsushi Fujii^{*}, Noriko Kando⁺⁺, Akihiko Takano^{**}

⁺Tokyo Institute of Technology / Hitachi Ltd.

^{*}University of Library and Information Science / Japan Science and Technology Corporation

^{++**}National Institute of Informatics

+ iwayama@pi.titech.ac.jp, * fujii@ulis.ac.jp, ++kando@nii.ac.jp, **takano@acm.org

Abstract

We describe the overview of patent retrieval task at NTCIR-3. The main task was the technical survey task, where participants tried to retrieve relevant patents to news articles. In this paper, we introduce the task design, the patent collections, the characteristics of the submitted systems, and the results overview. We also arranged the free-styled task, where participants could try anything they want as far as the patent collections were used. We describe the brief summaries of the proposals submitted to the free-styled task.

1. Introduction

The goal of Patent Retrieval Task is to provide test collections for enhancing research on patent information processing, from patent retrieval to patent mining. Although there exist many commercial patent retrieval systems and services, patent retrieval has not been paid much attention in the research field of information retrieval. One of the reasons is the lack of test collection on patent. TREC [5] used patent documents as a part of the document collections, but there was no treatment specially applied to the patent collection.

In SIGIR2000, the first workshop on patent retrieval was held [3] and there were many fruitful discussions on the current status and future directions of patent retrieval. The workshop convinced us that there was the need of

test collections specifically for patents.

We then asked for PATOLIS Co. [4] to provide patent collections for the patent retrieval task. Consequently, we could release three kinds of patent collections; those were two years' full texts, five years' Japanese abstracts, and five years' English abstracts. At the same time, we could fortunately have cooperation with JIPA (Japan Intellectual Property Association) [2] in creating search topics and assessing the relevance. Since each member of JIPA belongs to the intellectual property division in her/his company, they are all experts in patent searching. All the above contributions enabled us to kick off the first evaluation workshop designed for patent information processing.

There are various phases and aspects in patent information processing. For example, various kinds of users (researchers, patent searchers, business managers, and so on) search patents for various purposes (technical survey, finding conflicting applications, buying/selling patents, and so on). Corresponding to each situation, the appropriate search model should be developed. The standard of the relevance judgments may also depend on each situation. In some cases, retrieving relevant patents is not enough but further analysis on the retrieved patents might be necessary. For example, creating a patent map of a product would clarify the patent relations

between the techniques used to make the product. Cross-lingual patent retrieval is also important when applying patents to foreign countries. All of these are within scope of our project and this task is the first step toward our goal.

2. Task Design

In this workshop, we focused on a simple task of technical survey. End-users we assumed in the task were novice users, for example, business managers. The major reason of adopting such general task was that we could only use the two years' full texts that were not enough for trying more patent-oriented task like finding conflicting applications from patents.

To fit the task to a real situation, we used Japanese news articles as the original sources of search topics, so the task was conducting cross-database retrieval, searching patents by news articles. The task assumed the following situation. When a business manager looks through news articles and is interested in one of them, she/he clips it out and asks a searcher to find related patents to the clipping. The manager passes the clipping to the searcher along with her/his memorandum, and this clipping with memorandum became the search topic in this task. The memorandum helps the searcher to have the exact information need the manager has, when the clipping contains non-relevant topics or the clipping has little description on the information need. Task participants played the role of the searcher and tried to retrieve relevant patents to the clipping. Since the purpose of the searching was technical survey, the claim part in patent was not treated specifically in assessing the relevance. Patent documents were treated as if those were technical papers.

Cross-database retrieval itself is so general that techniques investigated in the task can be applied to various combinations of databases. This is another purpose of the task.

We prepared search topics in four languages, Japanese, English, Korean, and Chinese (both traditional and simplified). Participants could try

cross-lingual patent retrieval by using one of the non-Japanese topics. Unfortunately, only two groups submitted cross-lingual results and both of them used English topics.

In addition to the technical survey task explained so far, we arranged the optional task, where participants could try anything they want as far as they used the patent collections provided. One of the purposes of this free-styled task is to explore next official tasks.

3. Characteristics of Patent Applications

In this section, we briefly review the characteristics of patent applications (patent documents).

- There are structures, for example, claims, purposes, effects, and embodiments of the invention.
- Although the claim part is the most important in patent, it is written in an unusual style especially for Japanese patent; all the sub-topics are written in single sentence.
- To enlarge the scope of invention, vague or general terms are often used in claims.
- Patents include much technical terminology. Applicants may define and use their original terms not used in other patents.
- There are large variations in length. The longest patent in our collections contains about 30,000 Japanese words!
- The search models would be significantly different between industries, for example, between chemical / pharmaceutical industries and computers / machinery / electric industries.
- Classification exists. IPC (International Patent Classification) is the most popular one.
- The criterion of evaluation depends on the purpose of searching. For example, high recall is required for finding conflicting applications.
- In some industries, images are important to judge the relevance.

Our task focused on few of the above characteristics. We treated patent documents as technical documents rather than legal statements, so we did not distinguish between the claim part and the others in assessing the relevance. High recall was not necessary, so we used the standard averaged precision to evaluate the results. Few groups used structures and classifications. Images were not included in the patent collections provided.

4. Patent Collections

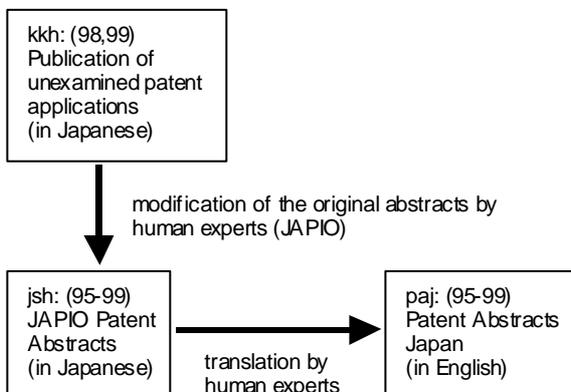


Figure 1: Relationships between the patent collections

PATOLIS Co. provided and we released the following patent collections.

- **kkh**: Publication of unexamined patent applications (1998, 1999) (in Japanese)
- **jsh**: JAPIO Patent Abstracts (1995–1999) (in Japanese)
- **paj**: Patent Abstracts Japan (1995– 1999) (in English)

“Kkh” contains full texts of unexamined patent applications in Japanese. Images were eliminated. “Jsh” contains human edited abstracts in Japanese. Although all the texts in “kkh” have the abstracts written by the applicants, experts in JAPIO (Japan Patent Information Organization) [1] shortened/lengthened about half of them to fit the length within about 400 Japanese characters. They also normalized technical terms if necessary. “Paj” is English translation of “jsh”.

Figure 1 shows the relationships between these three collections. Here, we see parallel relations, for example, full texts vs. abstracts, original abstracts vs. edited abstracts, and Japanese abstracts vs. English abstracts. Researchers can use these parallel collections for various purposes, for example, finding rules of abstracting, creating a term normalization dictionary, acquiring translation knowledge, and so on.

Table 1 summarizes the characteristics of the three collections.

	kkh	jsh	paj
Type	Full text	Abstract	Abstract
Language	Japanese	Japanese	English
Years	98,99	95-99	95-99
Number of documents	697,262	1,706,154	1,701,339
Bytes	18139M	1883M	2711M

Table 1: Characteristics of the patent collections

Field	Explanation
<LANG>	Language code
<PURPOSE>	Purpose of search
<TITLE>	Concise representation of search topic
<ARTICLE>	MAINICHI news article in NTCIR format
<SUPPLEMENT>	Supplemental information of news article
<DESCRIPTION>	Short description of search topic
<NARRATIVE>	Long description of search topic
<CONCEPT>	List of keywords
<PI>	Original patents of news article

Table 2: Explanations of the fields in topics

5. Topics

JIPA members created topics, six for the dry run and 25 for the formal run. Since the topics for the dry run were substantially revised after the dry run, we decided to re-use those in the formal run. In consequence, we had the total 31 topics for the formal run.

Figure 2 is an example of the topics and Table 2 shows the explanations of the fields in the topics. In our task, <ARTICLE> and <SUPPLEMENT> correspond to the news clipping and the memorandum respectively.

```

<TOPIC>
<NUM>
P004
</NUM>
<LANG>
JA
</LANG>
<PURPOSE>
技術動向調査
</PURPOSE>
<TITLE>
バーコードなどの符号を比較し優劣を判定する装置
</TITLE>
<ARTICLE>
<A-DOC>
<A-DOCNO>JA-981031179</A-DOCNO>
<A-LANG>JA</A-LANG>
<A-SECTION>社会</A-SECTION>
<A-AE>無</A-AE>
<A-WORDS>189</A-WORDS>
<A-HEADLINE>エポック社の特許侵害訴訟、バンダイが敗訴 - 東京地裁</A-HEADLINE>
<A-DATE>1998-10-31</A-DATE>
<A-TEXT>
カードゲームの特許を侵害されたとして、玩具(がんく)製造会社のエポック社がバンダイに2億6400万円の損害賠償を求めた訴訟で、東京地裁は30日、約1億1400万円の支払いを命じた。森義之裁判長は、バンダイが1992年7月～93年3月に製造・販売した小型ゲーム機「スーパーバーコードウォーズ」のキー操作などの機能について「エポック社が持つ特許の技術的範囲に属する」と指摘した。
</A-TEXT>
<A-DOC>
</ARTICLE>
<SUPPLEMENT>
バーコードなどを読み込み、これに基づく数値を比較して勝敗を決定していればよい。
</SUPPLEMENT>
<DESCRIPTION>
バーコードなどの符号を複数読み込ませ、これら符号に対応する数値を比較することにより、これらの優劣/勝敗の判定を行うことで対戦を行う装置にはどのようなものがあるか。
</DESCRIPTION>
<NARRATIVE>
「スーパーバーコードウォーズ」とは、小型ゲーム機の一つであり、キャラクターなどが描かれたカードに記録されたバーコードを読み込ませ、プレーヤーが攻撃や防御などのキー操作を行うことで、半リアルタイムに対戦を行うものである。符号の例としては、バーコードや磁気コードなどがあるが、これらに限定するものではない。
</NARRATIVE>
<CONCEPT>
符号 バーコード コード 優劣 勝敗 比較 判定
</CONCEPT>
<PI>
PATENT-KKH-G-H01-333373
</PI>
</TOPIC>
    
```

Figure 2: Example of a topic (Japanese)

The topics also contain <DESCRIPTION> and <NARRATIVE> fields we are familiar with. Since many NTCIR tasks already have the results for using <DESCRIPTION> and <NARRATIVE> fields, we can compare our results of using these fields with the results of other tasks.

Searching by <PI> field can emulate finding conflicting applications by patents. However, note

that the relevance in this task was assessed for the technical survey purpose but not for the conflict finding purpose.

The topics were originally written in Japanese, and translated to English, Korean and Chinese (traditional and simplified).

6. Runs

6.1 Mandatory runs

Participants were obliged to submit at least a run called mandatory run which used only <ARTICLE> and <SUPPLEMENT> fields. Manual searching was allowed and any language could be used for topics. In the mandatory run, the two years' full texts (i.e., "kkh98" and "kkh99") had to be searched and participants were inhibited from using the same years' abstracts (i.e., "jsh98", "jsh99", "paj98", and "paj99") for any purpose.

6.2 Optional runs

In addition to the mandatory run, participants could submit optional runs where they could use any resource. For example, it was allowed to retrieve patents from "jsh" or "paj" abstracts. We recommended for participants to submit runs using <DESCRIPTION> and <NARRATIVE>.

7. Relevance Judgment

Here is the process of collecting the relevance judgments.

1. After/during creating topics, JIPA members conducted manual search to collect as many relevant patents as possible. We call the patent set judged during this manual search as "PJ". Any system/resource was allowed to use because we wanted to have the set retrieved under the circumstance of their daily work of patent searching.
2. Participants submitted runs and we made a pool for each topic by collecting the top 30 patents from every submitted run. We call this pool as "PS".
3. JIPA members assessed the relevance for "PS - PJ" patents, which were the patents

they had not seen in their preliminary search.

Grades of relevance were “A (relevant)”, “B (partially relevant)”, “C (non-relevant judged after reading texts)”, and “D (non-relevant judged by just looking at titles)”.

As opposed to previous NTCIR tasks having done manual search after the pooling to increase the exhaustiveness of the relevant documents, we did manual search before the pooling to enable the comparison of the search ability between the human experts and the submitted systems.

Along with the grade of relevance (i.e., “A”, “B”, “C”, or “D”), each judged patent has a mark (“S”, “J”, or “U”) representing the origin from which the patent was retrieved. Table 3 explains about the marks. For example, a document with “BJ” means that the document was judged as “partially relevant” (i.e. “B-”) and only found by experts in their preliminary search (i.e., “-J”).

Mark	Explanation
S	Found only in “PS”, that is “PS – PJ”.
J	Found only in “PJ”, that is “PJ – PS”.
U	Found both in “PS” and “PJ”, that is “PS PJ”.

Table 3: Explanations of the origin marks

Here, note that all the submitted runs contributed to collecting the “S” patents, but only the top 30 patents for each run were used. Note also that we can restore the patent set retrieved by the manual search (i.e., “PJ” set) by collecting “J” and “U” patents.

8. Results Overview

8.1 Participants

Eight groups submitted the 36 runs. One group submitted runs only for pooling. We briefly describe the characteristics of each group.

LAPIN: This group focused on the “term distillation” in cross-database retrieval, where the difference between the term frequency in source database and that in target database was integrated into the overall term weighting.

SRGDU: This group tried several pseudo relevance feedback methods in the context of patent retrieval. The proposed method using Taylor formula was compared with the traditional Rocchio method.

daikyo: This group made long gram-based index from the patent collections. Compared with the traditional gram-based indexing, proposed method produce more compact index.

DTEC: This group searched various kinds of abstracts rather than full texts, and compared the effectiveness of those. The abstracts were JAPIO patent abstracts and the combinations of “title”, “applicant’s abstract”, and “claims”. Manual and automatic runs were compared.

DOVE: This group also submitted manual and automatic runs. In the manual runs, non-relevant passages in <ARTICLE> were eliminated manually.

IFLAB: This group evaluated their cross-lingual IR system PRIME through several monolingual runs. They also evaluated their translation extraction method by using Japanese-US patent families, which were not provided in this task.

brkly: This group submitted both monolingual and cross-lingual runs. In the cross-lingual runs, words in English topics were translated into Japanese words by using English-Japanese dictionary automatically created by the aligned bilingual corpus (i.e., “paj” and “jsh”). Their method of creating the dictionary is based on word co-occurrence with the association measure.

sics: This group also submitted cross-lingual runs, where they automatically created a cross-lingual thesaurus from the aligned bilingual corpus, “paj” and “jsh”, and used the thesaurus for word-based query translation. The Random Indexing vector-space technique was used to extract the cross-lingual thesaurus. Note that, in both the “sics” and the “brkly” groups, there was no member who understands Japanese.

8.2 Recall/Precision

The recall/precision graphs of the mandatory runs are shown in Figure 3, and those of the optional runs in Figure 4. In each figure, there are both results for the strict relevance (“A”) and the relaxed relevance (“A” + “B”). For each run in the figures, brief system description is specified; the description includes the searching mode (automatic or manual), the topic fields used in query construction, and the topic language.

8.3 Topic-by-topic Results

Figure 5 shows the median of the average precisions for each topic. Figure 6 shows the breakdown of the relevance judgments. Detailed analysis on each topic will be given by JIPA, where it will be discussed about the reasons why systems could not find some patents human experts found and vice versa.

8.4 Recall of the relevant patents retrieved in the preliminary human search

Figure 7 shows the recall of the relevant patents retrieved in the preliminary human search. In the process of making pool, we used only the top 30 documents for each run. Here, we extracted more documents from each run and investigated how many human retrieving relevant patents could be covered by the systems.

9. Optional (Free-styled) Task

The following two groups applied to the optional task.

CRL: This group investigated the method of extracting various rules from the existing alignments in patents. The “diff” command of UNIX was used to find the alignments between JAPIO patent abstracts and the original abstracts by applicants, between claims and embodiments, and between different claims in an application.

TTT: This group focused on the unusual style of Japanese claims, and tried to automatically structure the claims to raise the readability of claims. Rhetorical structure analysis was applied for this purpose.

10. Summary and Future Directions

In this paper, we described the overview of patent retrieval task at NTCIR-3. We are planning to continue our effort for the next patent retrieval task along with the following directions.

- Longer range of years will be covered, for example, nearly ten years’ patents.
- Purpose of search would shift to more real one, for example, searching conflicting applications. This would imply that the structure of patents should be considered in the relevance judgment.
- We may have to explore an evaluation measure tailored to the task.
- We will continue to have free-styled task in parallel with the main task.

Acknowledgements

We are grateful to PATOLIS Co. for providing the patent collections of this task. We also thank all the members of JIPA who created the topics and assessed the relevance. Without their expertise in patent, this task would not be realized. Lastly, we thank all the participants for their contributions to this task.

References

- [1] JAPIO (Japan Patent Information Organization) (<http://www.japio.or.jp/>)
- [2] JIPA (Japan Intellectual Property Association) (<http://www.jipa.or.jp/>)
- [3] ACM-SIGIR Workshop on Patent Retrieval, organized by Mun-Kew Leong and Noriko Kando, 2000. (<http://research.nii.ac.jp/ntcir/sigir2000ws/>)
- [4] PATOLIS Co. (<http://www.patolis.co.jp/e-index.html>)
- [5] TREC (Text Retrieval Conference) (<http://trec.nist.gov/>)

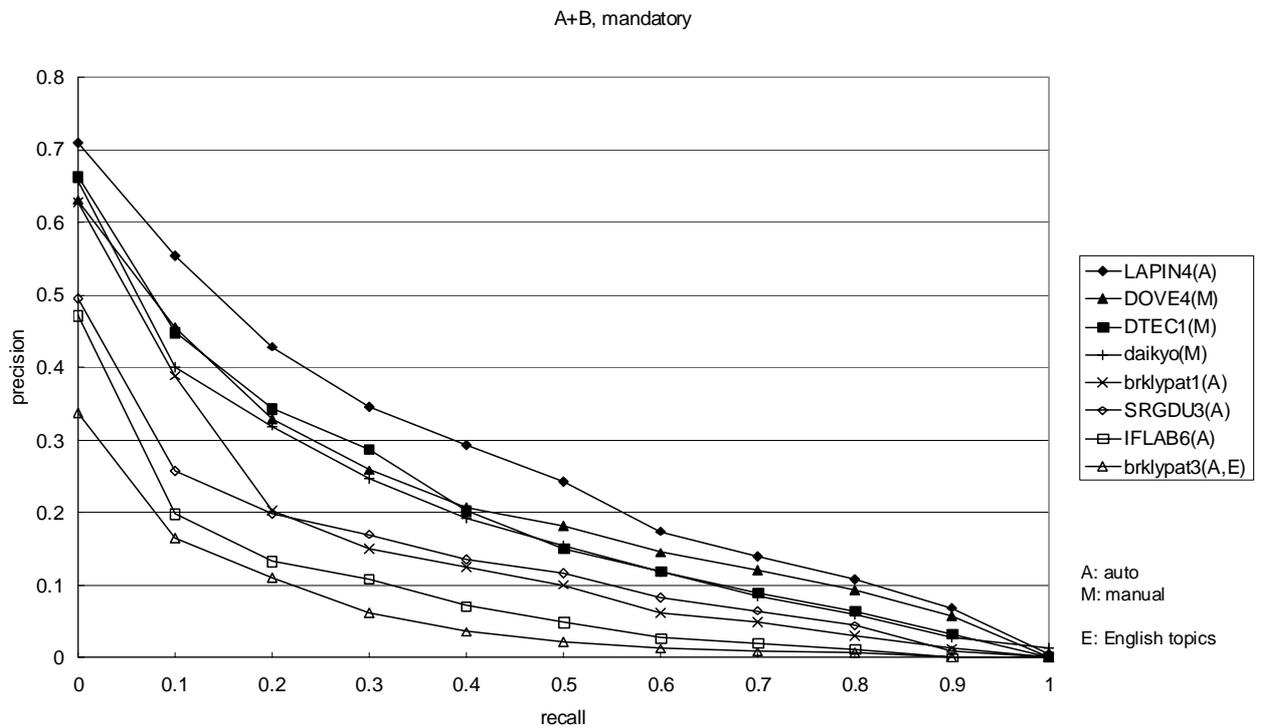
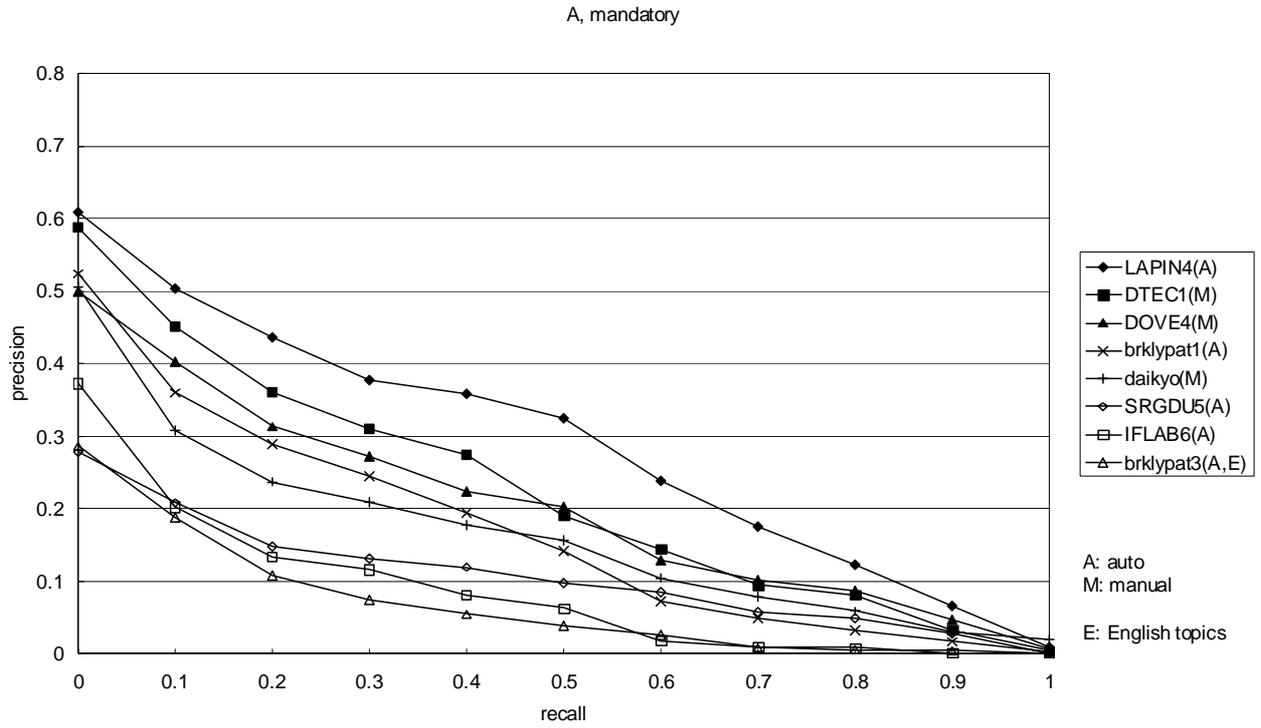
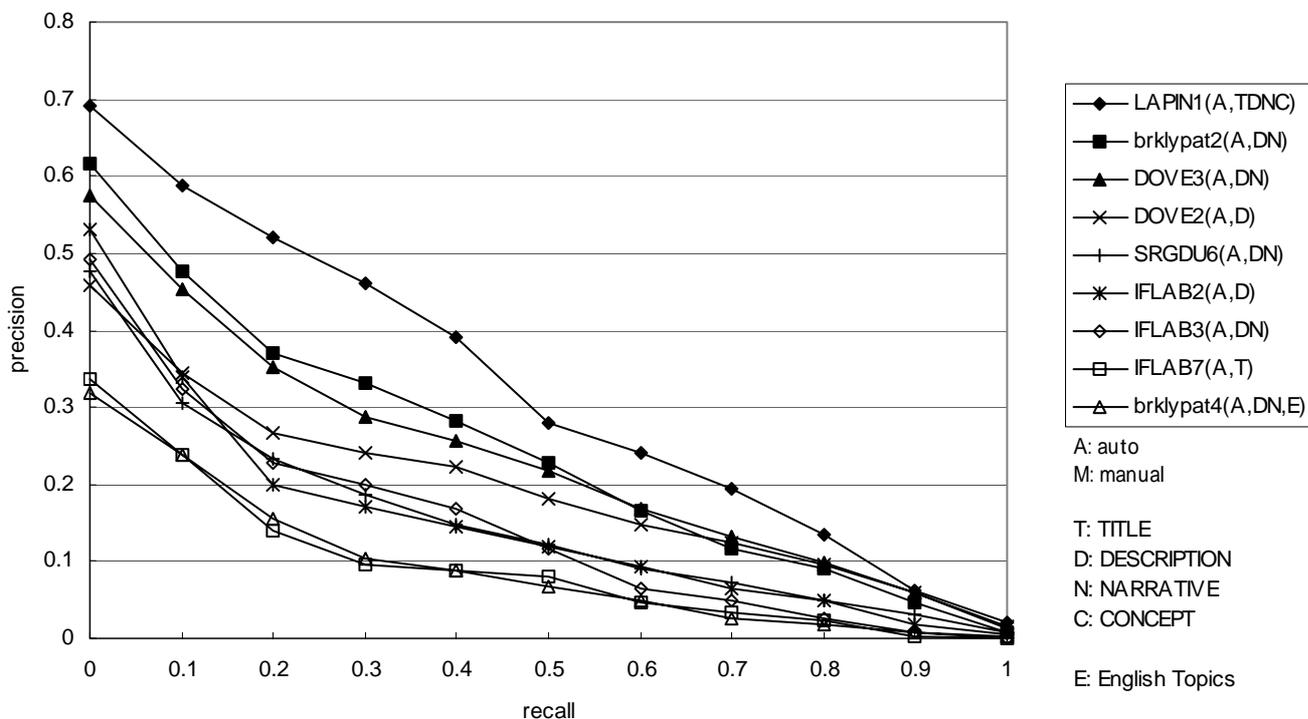


Figure 3: Recall/Precision of mandatory runs

A, optional



A+B, optional

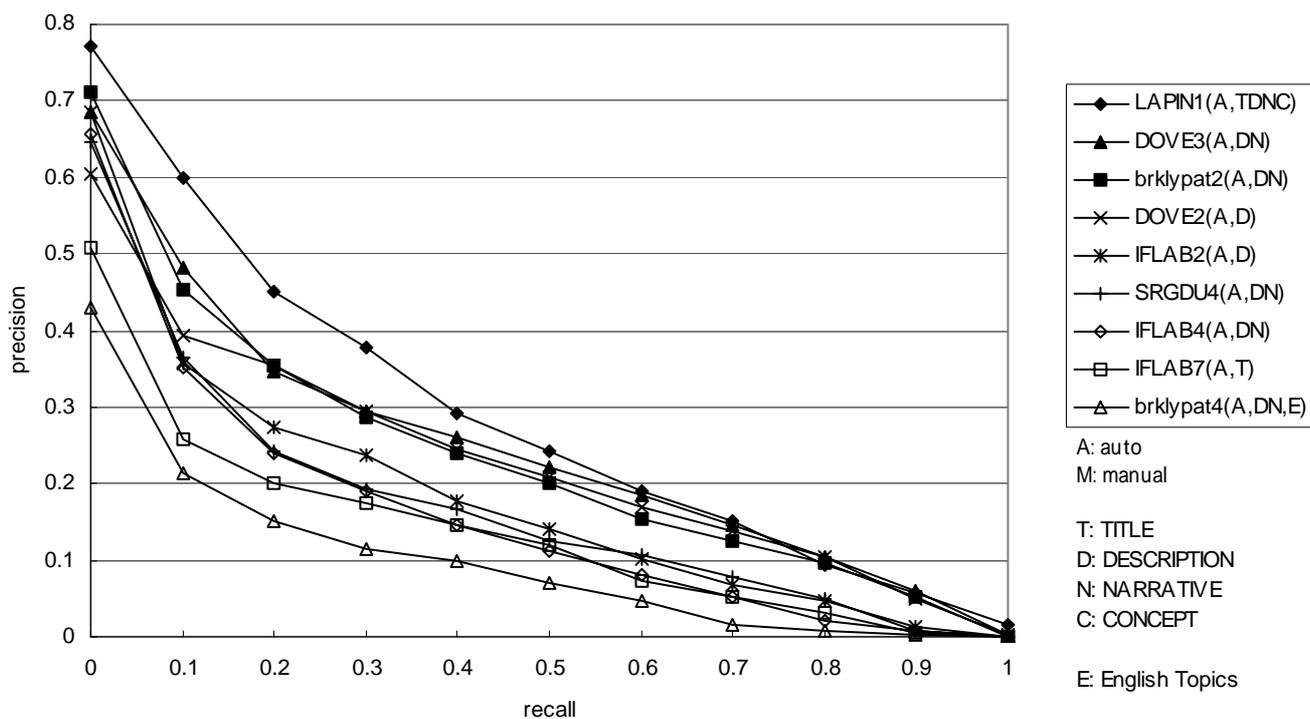


Figure 4: Recall/Precision of optional runs

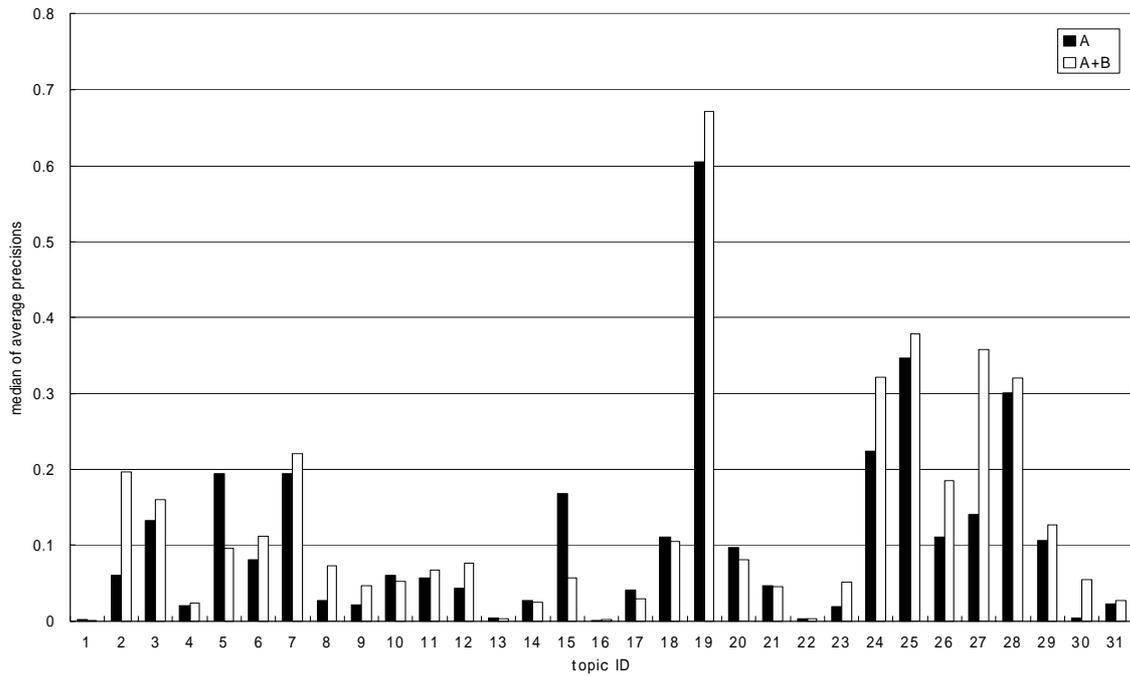


Figure 5: Median of average precisions (all runs)

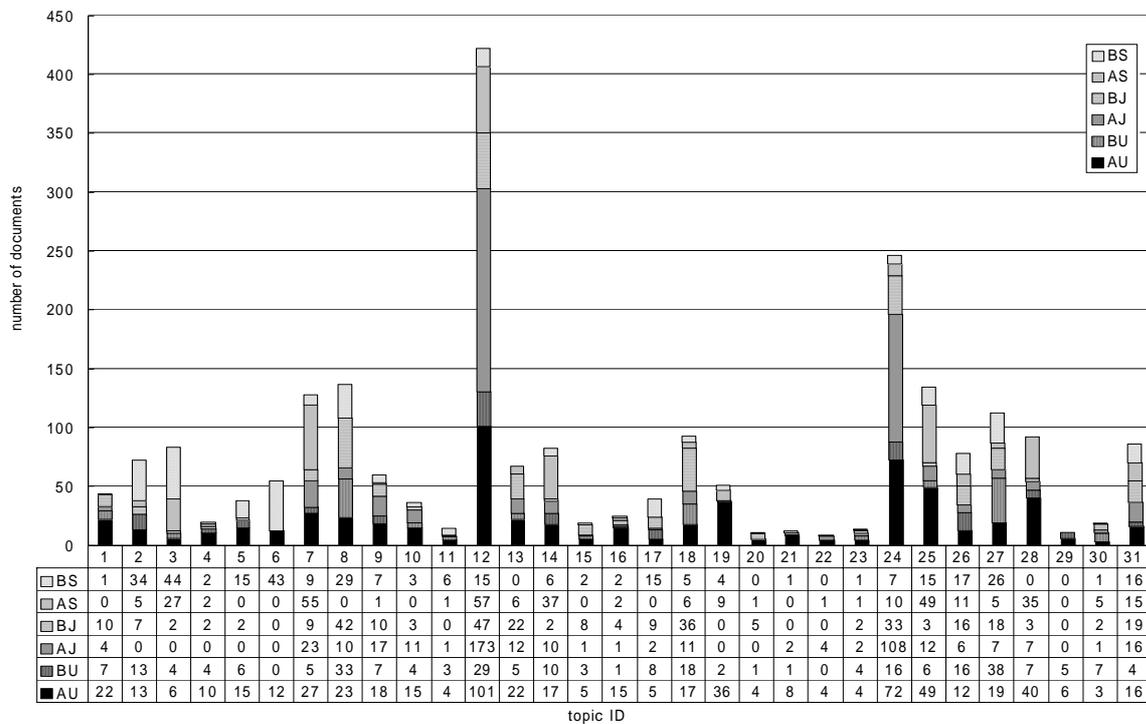


Figure 6: Breakdown of relevance judgments

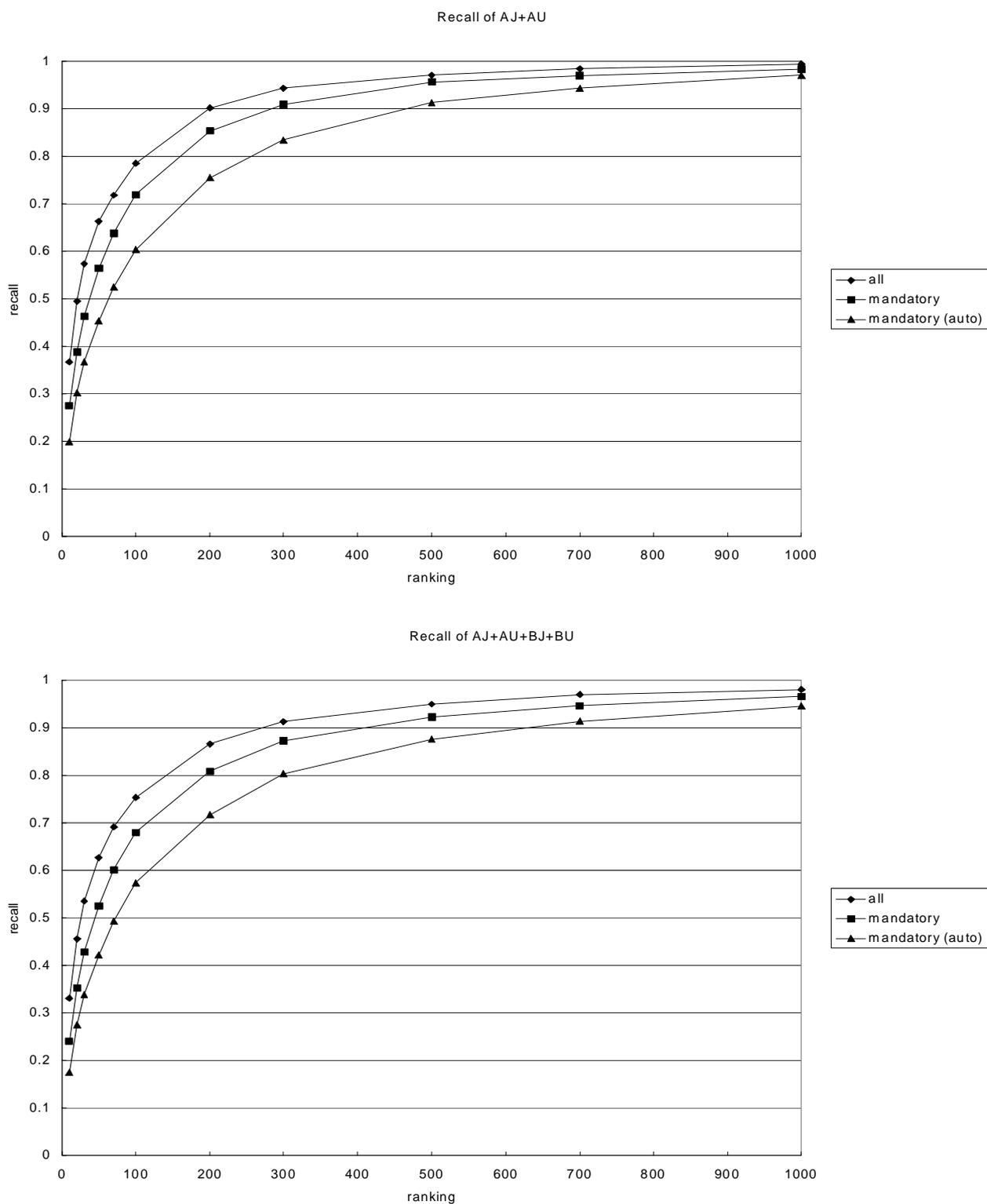


Figure 7: Recall of the relevant patents retrieved in the preliminary human search