

## Patent Mining Task Round Table

2008.3.31

Hidetsugu Nanba (Hiroshima City University)

Atsushi Fujii (University of Tsukuba)

Makoto Iwayama (Hitachi, Ltd.)

Taiichi Hashimoto (Tokyo Institute of Technology)

## Background (1/2)

“Intellectual Property Strategic Program 2006”  
(by Intellectual Property Strategy  
Headquarters Japan) (in page 30)

### (4) Developing the Integrated Search System for Patent and Literature Information

- In research activities carried out at universities and public research institutes, **patent information is as important as literature information.**
- Also, as early as possible in FY2007, the GOJ will **develop an Integrated Search System for Patent and Literature Information** (tentative name), whereby both patent information and literature information can be searched.

## Background (2/2)

### Invalidity Search

To **retrieve research papers and patents** that could invalidate existing claims.

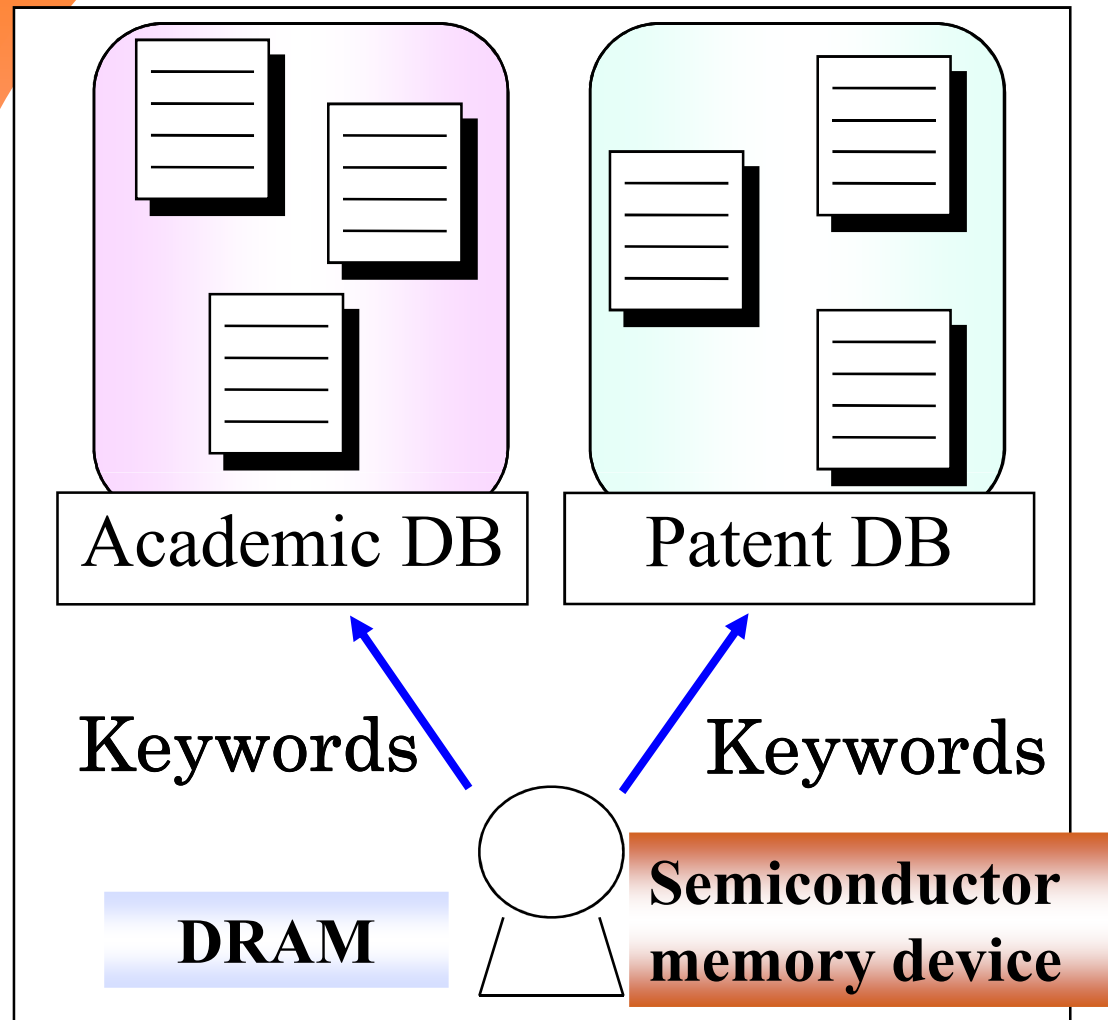
conducted by

- examiners in government patent offices
- searchers in the intellectual property divisions of private companies

## Purpose

To develop several techniques for retrieving and classifying research papers and patents

# The Aim of the Task

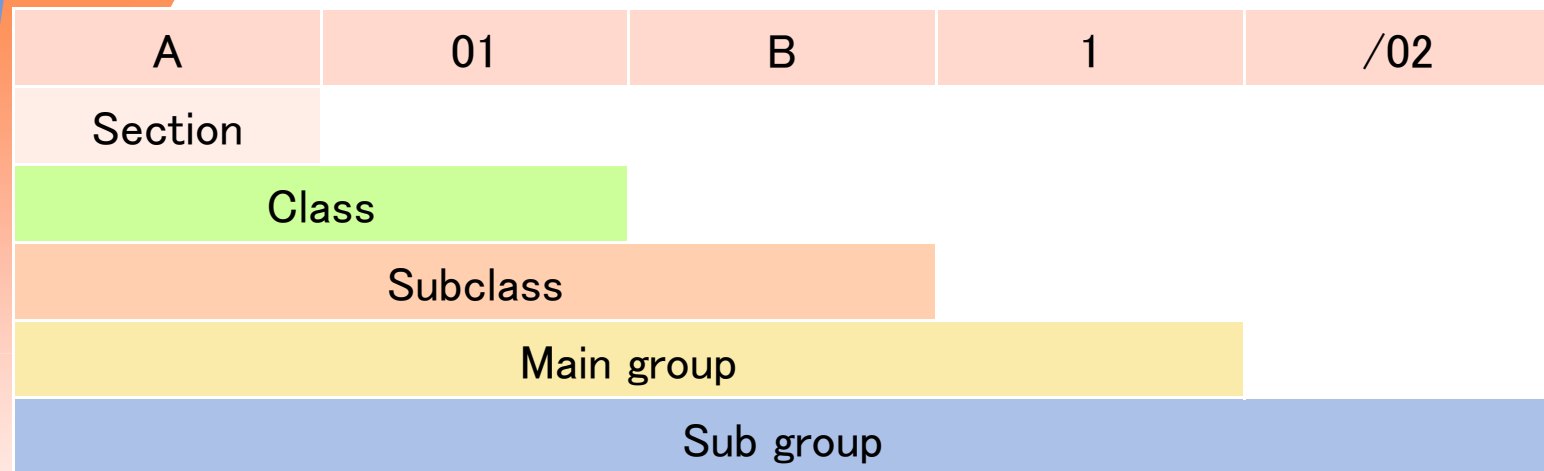


To develop several techniques for retrieving and classifying research papers and patents

# The Task at NTCIR-7

- Categorization of research papers (abstracts) written in Japanese or in English into the International Patent Classification (IPC), which is one of patent classification systems. (Two tasks (English and Japanese) will be conducted.)
- Each group submits a ranked list of the IPC codes for each target research paper.

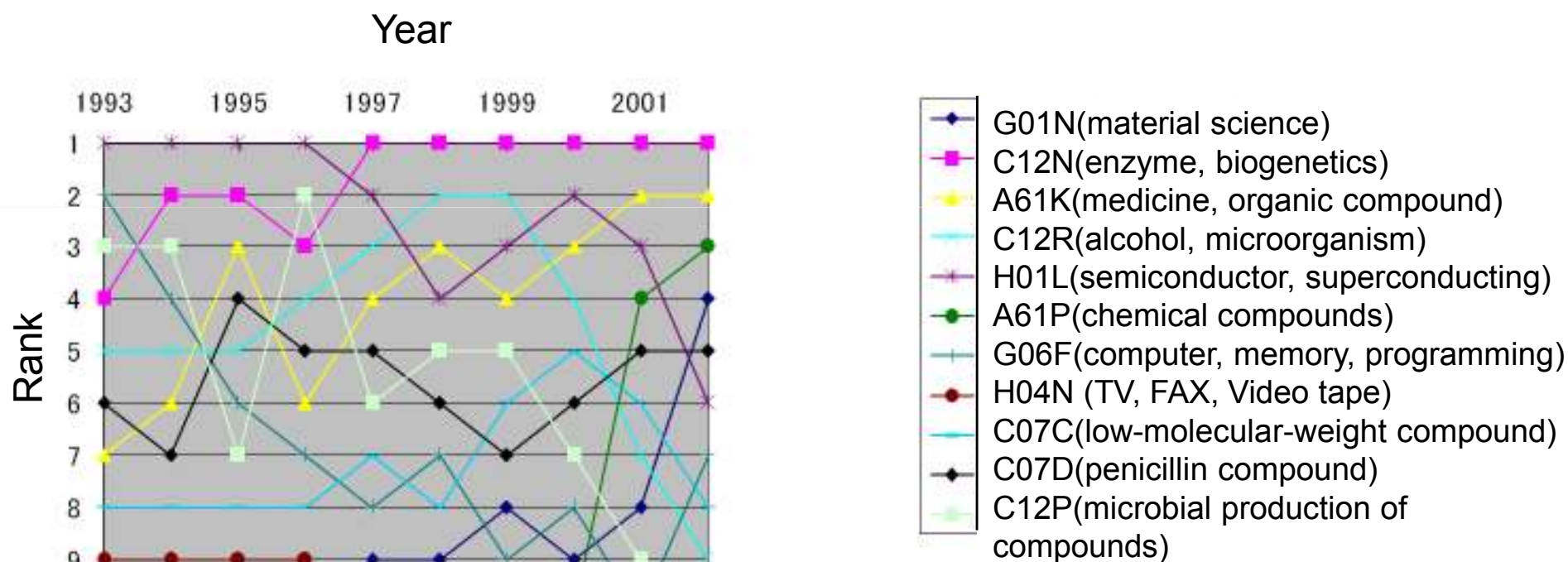
# International Patent Classification (IPC)



A	Section	Human necessities
A01	Class	Agriculture; Forestry; Hunting; Fishing; ...
A01B	Subclass	Soil working in agriculture or forestry; parts, details or accessories of agricultural machines or implements, in general
A01B 1/00	Main group	Hand tools
A01B 1/02	Sub group	Spades; Shovels

## Relevance of Basic Research to Technology for each Domain

The number of cited papers for each IPC code





# Example of an Abstract

<TOPIC>

<TOPIC-ID> 100 </TOPIC-ID>

<TITLE> DTMF (Dual Tone Multi-Frequency) transmission method for a mobile communication system </TITLE>

<ABSTRACT> High efficient speech encoding scheme called VSELP, is adopted for Japanese digital mobile communication systems.

However, DTMP (Dual Tone Multi-Frequency) signals are distorted by using this encoding scheme. This paper presents a DTMF signal transmission scheme. DTMF signals are transmitted in the form of call control messages from mobile station (MS) to mobile control center (MCC). In addition, necessary control capabilities in MS and MCC is described. </ABSTRACT>

</TOPIC>

# 論文抄録データ

<TITL>問題解決能力育成のための教育ツールアルゴブロックのシステム・イメージ</TITL>

<AUPK>鈴木 栄幸 / 加藤 浩</AUPK>

<CONF>春季全国大会</CONF>

<CNFD>1992. 03. 24 - 1992. 03. 27</CNFD>

<ABST>問題解決能力育成のためのツールとしてアルゴブロックのコンセプトを提案する。アルゴブロックとは、物理的ブロックにプログラム言語のコマンド、もしくは機能モジュールを割り当てておき、これらのブロックを組み立てることでプログラムを作成することのできる物理言語である。アルゴブロックの最大の特徴は、プログラミングという内的過程を、物理的なブロックの組み立てという、「見て、触れる」世界に投影していることである。本論では、現在開発中のアルゴブロックのシステムイメージを示すとともに、アルゴブロックの各特徴が、問題解決能力の育成にどのように結びついているか解説する。</ABST>

<SOCN>電子情報通信学会</SOCN>

## The Article 30 in Japanese Patent Law

- The article 30 in Japanese patent law provides a six-month grace period for disclosures made through a publication or a presentation at a conference or an exhibition.
- The applicants need to mention the proceeding titles (or conference names) and the date it was published in “Indication of exceptions to lack of novelty” field (exception field) in the patent.

**[Indication of exceptions to lack of novelty]** The provisions set forth in Article 30, Paragraph 1 in Japanese patent law. Proceedings (Volume 4) of the 60<sup>th</sup> Annual Meeting of the Information Processing Society of Japan, published in March 14, 2000.

# Assignment of IPC Codes into Research Papers

Japanese Patent

=

Japanese Research Paper

IPC

||

English Research Paper

# Patent

【発明の名称】プログラム生成装置及びプログラム生成システム

【出願番号】特願平4-153166

【出願日】平成4年(1992)6月12日

【新規性喪失の例外の表示】特許法第30条第1項適用申請有り  
1992年3月24日～3月27日 社団法人電子情報通信学会開  
催の「電子情報通信学会春季大会」において文書をもって発表

【発明者】【氏名】鈴木 栄幸

【目的】 コンピュータ等をプログラムする場面において、固有の機能が割り当てられている物理的ブロックのいくつかを利用者が組み立てたものを、接続状態をもとにプログラムに変換することで、プログラミング作業または、学習者、初心者のプログラミングの習得を容易にする。

## Corresponding Research Paper

<TITL>問題解決能力育成のための教育ツールアルゴブロックのシステム・イメージ</TITL>

<AUPK>鈴木 栄幸 / 加藤 浩</AUPK>

<CONF>春季全国大会</CONF>

<CNFD>1992. 03. 24 - 1992. 03. 27</CNFD>

<ABST>問題解決能力育成のためのツールとしてアルゴブロックのコンセプトを提案する。アルゴブロックとは、物理的ブロックにプログラム言語のコマンド、もしくは機能モジュールを割り当てておき、これらのブロックを組み立てることでプログラムを作成することのできる物理言語である。アルゴブロックの最大の特徴は、プログラミングという内的過程を、物理的なブロックの組み立てという、「見て、触れる」世界に投影していることである。本論では、現在開発中のアルゴブロックのシステムイメージを示すとともに、アルゴブロックの各特徴が、問題解決能力の育成にどのように結びついているか解説する。</ABST>

<SOCN>電子情報通信学会</SOCN>

# Topics

**Patent–Research Paper Pairs: 976**

- **Dry run : 97 topics**
- **Formal run : 879 topics**

**Average number of correct IPC codes for a topic: 2.3**

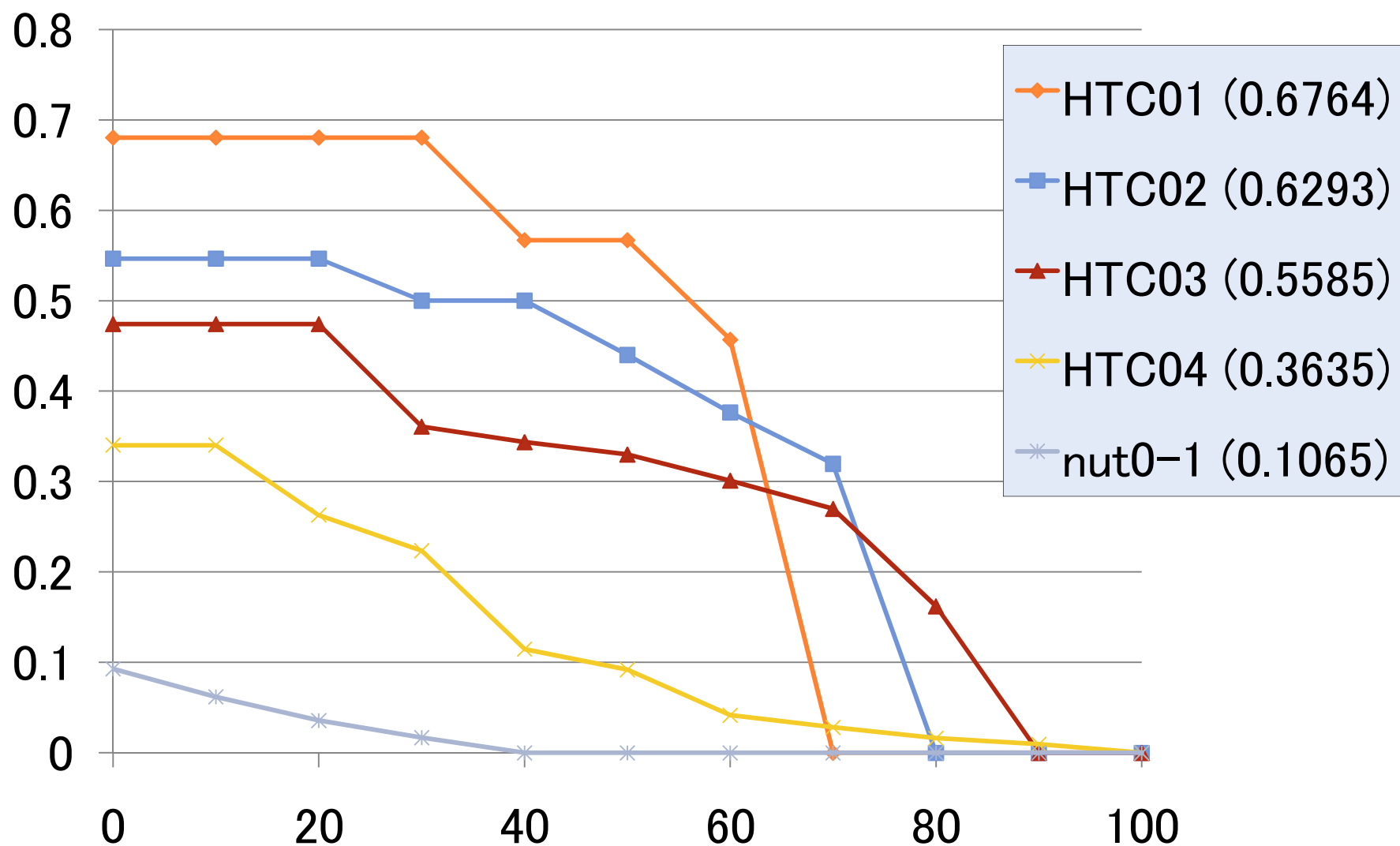
# Submission

- Japanese Subtask  
6 systems from 3 groups
- English Subtask  
14 systems from 8 groups

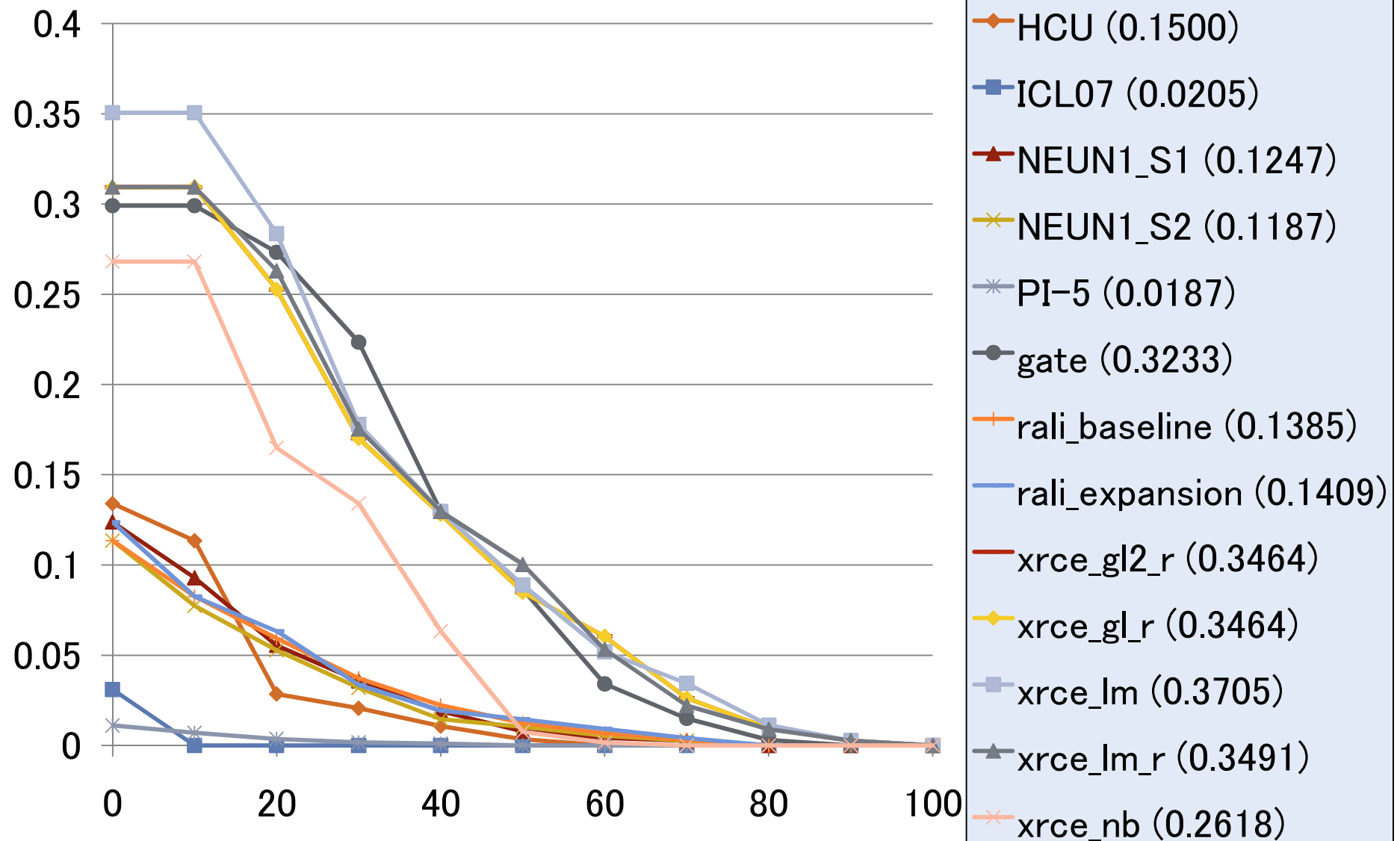
**Total: 20 systems from 10 groups**



# Results of Japanese Subtask



# Results of English Subtask



# Discussion

Task of assigning IPC codes into research papers

→ Task of identifying patents correspond to given research papers

**Different Task!**

## Solution

Remove the patents, corresponding to research papers used for topics, from pseudo-training data

# Schedule

Preparation	Call For Participation	–2007.11.15
	Data release	As needed
Dry run	Topic release	2008.01.31
	Submission deadline	2008.03.01
	Evaluation release	2008.03.10
	Analysis of the evaluation	2008.03.31
Formal run	Topic release	2008.05.01
	Submission deadline	2008.05.31
	Evaluation release	2008.06.10
	Analysis of the evaluation	2008.07
Preparation for meeting	Paper for the proceedings due	2008.10.01
	Camera-ready paper for the proceedings due	2008.11.01
	Workshop meeting at Tokyo	2008.12.16–19