Overview of the NTCIR-10 Cross-Lingual Link Discovery Task

Ling-Xiang Tang†, In-Su Kang‡, Fuminori Kimura*, Yi-Hsun Lee♦, Andrew Trotman♠,
Shlomo Geva†, Yue Xu†
†Queensland University of Technology ‡Kyungsung University *Ritsumeikan University
♦Academia Sinica ♠University of Otago
{l4.tang, s.geva, yue.xu}@qut.edu.au dbaisk@ks.ac.kr fkimura@is.ritsumei.ac.jp
rog@iis.sinica.edu.tw andrew@cs.otago.ac.nz

ABSTRACT
This paper presents an overview of NTCIR-10 Cross-lingual Link Discovery (CrossLink-2) task. For the task, we continued using the evaluation framework developed for the NTCIR-9 CrossLink-1 task. Overall, recommended links were evaluated at two levels (file-to-file and anchor-to-file); and system performance was evaluated with metrics: LMAP, R-Prec and P@N.

Categories and Subject Descriptors
I.2.7 [**Artificial Intelligence**]: Natural Language Processing – *text analysis*.

I.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing – *linguistic processing*.

General Terms
Experimentation.

Keywords
Wikipedia, Cross-lingual Link Discovery, CrossLink-2, Evaluation Tool, Evaluation Metrics.

1. INTRODUCTION
Knowledge discovery is not always an easy task especially in a multi-lingual environment such as in a knowledge base (e.g. Wikipedia) with tens of millions of articles. To provide easy access to the cross-lingual information and break the language barrier, cross-lingual link discovery (CLLD) is concerned with automatically finding potential links between documents in different languages. In contrast to traditional information retrieval tasks where queries are not attached to explicit context, or only loosely attached to context, cross language link discovery algorithms actively recommend a set of meaningful anchors in the context of a source document and establish links to documents in an alternative language. CLLD is helpful for complimentary knowledge discovery in different language domains.

To explore the best approaches to realising CLLD automatically and justify the system performance with proper evaluation methods, CrossLink as a pilot task of NTCIR-9 has been successfully held in 2011. At the end of experimentation season, in total 57 runs from 11 teams were received. Among the submissions, many good approaches were seen. And the developed evaluation framework was proven effective [1, 2].

For the CrossLink-1 task, the main focus was on solving the problems in cross linking documents from English Wikipedia to CJK (Chinese, Japanese, Korean) Wikipedia. For CrossLink-2, we emphasised our study on realising the cross-lingual linking from CJK languages documents to English ones. This time participants will have to deal with an extra problem when trying to cross link documents as there are no word boundaries in Chinese / Japanese text, and in Korean *eojeol*. The need for natural language processing in CJK language to English document linking could make the cross-lingual document linking task very challenging. For example, Figure 1 shows some possible problems of knowledge discovery (from Chinese to other languages) in Wikipedia:

- Most anchors are not really relevant
- Good anchors are not always there, even so, relevant articles may not be available in a language that user prefer.
- Can something more interesting be recommended? For example, what is 聖誕布丁[1] described in the pudding (Chinese: 布丁) article of Figure 1? We know what is bread (麵包) but what is 麵包布丁 in English.
- How can an article be linked to other relevant articles in a different language automatically, dynamically and efficiently.

The goal of the CrossLink-2 task is set to continuing to create a reusable resource for evaluating cross-lingual link discovery approaches which should address the problems discussed before.

NTCIR-10 Workshop Meeting, 2013, Tokyo, Japan

[1]聖誕布丁 is the Chinese word for Christmas pudding

Figure 1. Cross-lingual Linking in Wikipedia

The remainder of this paper is organized as follows: First, the new cross-lingual link discovery tasks are defined in Section 2. The submission specification is outlined in Section 3. The assessment and evaluation frame work is discussed in Section 4. Evaluation results are given in Section 5. We then conclude in Section 6.

2. CROSSLINK-2 TASK

2.1 Task Definition

Comparing with the tasks at NTCIR-9, we setup three similar but different subtasks that require cross-lingual linking with opposite link direction. The new subtasks are listed as follows:

- Chinese to English CLLD (C2E)
- Japanese to English CLLD (J2E)
- Korean to English CLLD (K2E)

These subtasks are not simple replicas of previous CrossLink subtasks. These tasks allow the CLLD approaches evaluated at NTCIR-9 for suggesting good links from English documents to relevant Chinese, Japanese and Korean (CJK) articles to be re-examined in a different linking environment.

Plus, huge amount of effort was committed by the participating teams of the previous task in creating various CLLD systems from scratch. So in order to further evaluate those systems with different topics, runs are also allowed for the subtasks defined in previous evaluation round:

- English to Chinese CLLD (E2C)
- English to Japanese CLLD (E2J)
- English to Korean CLLD (E2K)

Having the same subtasks in the new evaluation round also allows seeing continuous improvement of the existing CLLD systems.

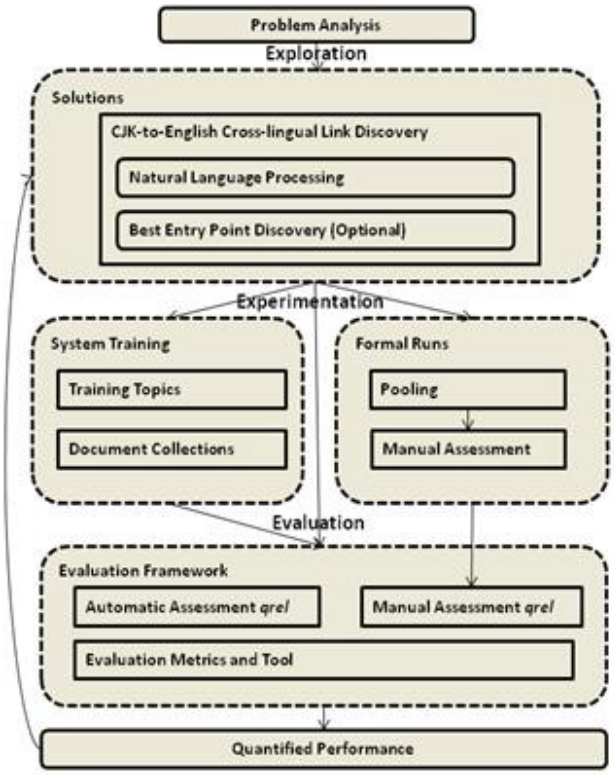


Figure 2. The CLLD Evaluation Procedure

2.2 Methodology

A same evaluation framework was utilised for benchmarking system performance. The whole evaluation procedure can be illustrated in Figure 2. However, to achieve the new goals and have proper system performance benchmarking, the evaluation tool set was improved to address the language and link differences in the new settings.

2.3 Document Collections

For experiment, we created a complete new set of CEJK Wikipedia document collections which were built from recent Wikipedia database dumps. The details of the collections are given in Table 1. There are 5,539,719 articles in total in the entire collections.

Table 1. CJK Wikipedia document collections

Language	# doc	Size	Dump Date
Chinese	404,620	3.6GB	11/01/2012
English	3,581,772	33.0GB	04/01/2012
Japanese	858,610	9.8GB	04/01/2012
Korean	297,913	2.2GB	22/01/2012

2.4 Topics

Four sets of 25 Wikipedia articles in CEJK languages were selected separately as test topics[2]. The details of the topics are given in Table 3. All test topics were provided in their original forms (in XML format) except that previously existing links marked with the *link* tag were removed.

3. SUBMISSION

3.1 General Rules

The basic rules and specifications for final run submission remain unchanged as previously defined in NTCIR-9 task CrossLink-1 [2]. In general, up to 250 anchors are allowed for each topic; and up to 5 targets are allowed for each anchor. Also chronological items shouldn't be included. Anchors should be identified within the text inside the *bdy* element and before any reference or external links section, e.g. <st>References</st>.

3.2 Run Specification

Table 2 gives the full content of the DTD file for the run specification. The source or the target language of link should be given in its language code which could be either *zh*, *en*, *ja*, or *ko*.

The position (zero-based offset) of an anchor is calculated by counting the number of bytes (*not* the number of characters) from the beginning of the topic file. Similarly, the length of an anchor is the number of bytes of anchor text.

Also, note that:

- If an anchor has a name that doesn't match the text in the topic with the given offset and length, it will be discarded in pooling,
- XML tag(s) might be included in anchor text when anchor is identified, but these extra tag(s) must be removed from the specified anchor name. For example, if an anchor "A Sample <it>Anchor" is found, the specification of this anchor should be:

  <anchor offset="768" length="19" name="A Sample Anchor">

  However, if an anchor contains an incomplete XML tag, it will be discarded in pooling too. For example,

  The specification of an anchor - "A Sample <i" is incorrect, even the anchor text may be given correctly as follows:

  <anchor offset="768" length="11" name="A Sample">

To sum up, anchor length should be calculated by counting every character in bytes including all XML tags in the original anchor text identified for a topic file, but the name attribute of an anchor should be only specified with the text after removing all XML tags (including the incomplete ones) due to the XML well-formedness requirement for run file processing.

Table 2. Submission XML File DTD

```
<!ATTLIST crosslink-submission
  participant-id CDATA #REQUIRED
  run-id CDATA #REQUIRED
  task (A2F) #REQUIRED
  source_lang (zh|en|ja|ko) ) #REQUIRED
  default_lang (zh|en|ja|ko) ) #REQUIRED
>
<!ELEMENT details (machine, time)>
<!ELEMENT machine (cpu, speed, cores, hyperthreads, memory)>
<!ELEMENT cpu (#PCDATA)>
<!ELEMENT speed (#PCDATA)>
<!ELEMENT cores (#PCDATA)>
<!ELEMENT hyperthreads (#PCDATA)>
<!ELEMENT memory (#PCDATA)>
<!ELEMENT time (#PCDATA)>
<!ELEMENT description (#PCDATA)>
<!ELEMENT collections (collection+)>
<!ELEMENT collection (#PCDATA)>
<!ELEMENT topic (outgoing)>
<!ATTLIST topic
  file CDATA #REQUIRED
  name CDATA #REQUIRED
>

4.  <!ELEMENT outgoing (anchor+)>

<!ELEMENT anchor (tofile+)>
<!ATTLIST anchor
  name CDATA #REQUIRED
  offset CDATA #REQUIRED
  length CDATA #REQUIRED
>
<!ELEMENT tofile (#PCDATA)>
<!ATTLIST tofile
  bep_offset CDATA #REQUIRED
  lang (zh|en|ja|ko)#REQUIRED
  title CDATA #REQUIRED
>
```

[2] http://crosslink.googlecode.com/files/topics-no-links.zip

Table 3. Test topics of the NTCIR-10 CrossLink-2 task

#	Title	Chinese	Japanese	Korean
1	Toi invasion	女真入侵日本	刀伊の入寇	도이의 입구
2	Tutsi	图西族	ツチ	투치족
3	Papaver rhoeas	虞美人 (花卉)	ヒナゲシ	개양귀비
4	Cellophane noodles	粉丝	春雨 (食品)	당면
5	Traveler's cheque	旅行支票	トラベラーズチェック	여행자 수표
6	Ravioli	意大利餃	ラビオリ	라비올리
7	X264	X264 編碼器	X264	X264
8	Nanako Matsushima	松嶋菜菜子	松嶋菜々子	마쓰시마 나나코
9	Pentaceratops	五角龙	ペンタケラトプス	펜타케라톱스
10	Xylitol	木糖醇	キシリトール	자일리톨
11	Tolui	拖雷	トルイ	툴루이
12	Japanese giant salamander	日本大鲵	オオサンショウウオ	일본장수도롱뇽
13	Chef	廚師	調理師	조리사
14	Tieguanyin	铁观音	鉄観音	철관음
15	Crown Prince Sado	朝鮮莊祖	荘献世子	조선 장조
16	Zhu Bajie	猪八戒	猪八戒	저팔계
17	Mawangdui	马王堆汉墓	馬王堆漢墓	마왕퇴
18	Saṃsāra	輪迴	輪廻	윤회
19	Lee Young Ae	李英愛	イ・ヨンエ	이영애
20	Song Hye-kyo	宋惠教	ソン・ヘギョ	송혜교
21	Tsundere	傲嬌	ツンデレ	츤데레
22	Youtiao	油条	油条	유조
23	Westwood Studios	西木工作室	Westwood	웨스트우드 스튜디오
24	Tom yum	冬蔭	トムヤムクン	똠얌
25	Wasabi	山葵	ワサビ	고추냉이

4.1 Participant Submissions

In total, 67 submissions from 10 groups were received. The names and affiliations of the groups who submitted run(s) are given in Table 4. The statistics of runs for each sub-task is given in Table 5.

System descriptions of runs for tasks with two different link directions are outlined in Table 30 and Table 31 separately.

5. ASSESSMENT AND EVALUATION

5.1 Assessment

The assessment work for both Wikipedia ground-truth and pooled links from participants was organised by the members of organising team from university (academy) of Taiwan[3], Japan[4], and Korea[5] separately according to each different language specific task.

5.2 Evaluation Methods and Metrics

With the availability of *qrels* from different assessment sources, the performance of different systems can be evaluated in both file-to-file (F2F) and anchor-to-file (A2F) levels. In the F2F evaluation, performance is measured based on the ideology of what other relevant articles can be recommended without needing specifying relevant anchors. In A2F evaluation, the relevance of specified anchors must however be considered. For the details of these two kinds of evaluations, it is recommended to refer to our previous work on NTCIR-9 [1, 2].

[3] Academia Sinica
[4] Kyungsung University
[5] Ritsumeikan University

Table 4. CrossLink-2 Participants

GROUP	AFILIATION
DCU	Dublin City University
III	Institute for Information Industry
KECIR	Shenyang Aerospace University
KMI	The Open University
KSLP	Kyungsung University
NTHU	National Tsing Hua University
OKSAT	Osaka Kyoiku University
QUT	Queensland University of Technology
RDLL	Ritsumeikan University
UKP	TU Darmstadt

Table 5. Submission statistics of participants

GROUP	CJK2E			E2CJK		
	C2E	J2E	K2E	E2C	E2J	E2K
DCU	2	0	0	2	0	0
III	3	0	0	1	0	0
KECIR	4	0	0	0	0	0
KMI	3	3	3	2	2	2
KSLP	0	0	1	0	0	0
NTHU	3	1	0	0	0	0
OKSAT	2	2	2	2	2	2
QUT	2	2	2	1	1	1
RDLL	0	5	0	0	0	0
UKP	3	3	3	0	0	0
Sub-total	22	16	11	8	5	5
Total	49			18		

5.3 System Performance Benchmarking

Three metrics, *LMAP*, *P@N*, and *R-Prec* [1, 2] which were previously used in the CrossLink-1 task and the INEX Link-the-Wiki task [3], were employed for system evaluation. Run performance was mainly measured using *LMAP* metric (Link Mean Average Precision, which was originally named MAP but was renamed here to avoid confusion because of the common use of MAP in the information retrieval evaluation [1, 4]).

While *LMAP* is our primary evaluation metric with *Precision-at-5* as second, *R-Prec* scores of submitted runs are also given.

5.3.1 Precision and Recall

Precision and recall are the two key underlying fractions of these three system evaluation metrics (*LMAP*, *P@N*, and *R-Prec*). They are computed for each topic separately, and have to be treated differently for the different evaluation level (F2F or A2F).

File-to-File Evaluation

$$Precision_{f2f} = \frac{number\ of\ relevant\ links}{min\ (1250,\ number\ of\ recommended\ links)} \quad (1)$$

and,

$$Recall = \frac{number\ of\ relevant\ links}{min\ (1250,\ number\ of\ links\ in\ qrels)} \quad (2)$$

Anchor-to-File Evaluation

In anchor to file evaluation, relevance of both anchor and target is considered.

For an identified anchor *i*, if a target document, *j*, is relevant to the anchor then it receives a score of 1, otherwise 0. Thus, the link score can be defined as:

$$f_{link}^{i}(j) = \begin{cases} 1, & \text{if relevant} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

The score of anchor, *i*, is defined as:

$$f_{anchor}(i) = \begin{cases} 1, & \text{if relevant and } \sum_{j}^{L_i} f_{link}^{i}(j) \geq 1 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where L_i is the number of recommended links for anchor *i*. If anchor *i* is considered relevant and it has at least one relevant target, then $f_{anchor}(i) = 1$. Otherwise, the score is 0.

The precision and recall of a topic evaluated in anchor to file level are:

$$Precision_{a2f} = \left(\sum_{i=1}^{n} \left(f_{anchor}(i) \times \frac{\sum_{j=1}^{k_i} f_{link}^{i}(j)}{k_i} \right) \right) / n \quad (5)$$

$$Recall_{a2f} = \left(\sum_{i=1}^{n} \left(f_{anchor}(i) \times \frac{\sum_{j=1}^{k_i} f_{link}^{i}(j)}{k_i} \right) \right) / N \quad (6)$$

where,

$$n = min\ (250,\ the\ number\ of\ identified\ anchors) \quad (7)$$

$$N = min\ (250,\ \#\ \text{of anchors in}\ qrel) \quad (8)$$

$$k_i = min(5,\ \#\ of\ relevant\ links\ for\ anchor_i) \quad (9)$$

5.3.2 System Evaluation Metrics

LMAP is defined as:

$$LMAP = \left(\sum_{t=1}^{n} \frac{\sum_{k=1}^{m} p_{kt}}{m} \right) / n \quad (10)$$

where *n* is the number of topics (source articles used in evaluation); *m* is the number of identified items (articles for F2F or anchors in A2F); and P_{kt} is the precision of the top *K* items for topic *t*.

R-Prec is defined as:

$$R - Prec = \sum_{t=1}^{n} P_t\ @\ R\ / n \quad (11)$$

where *n* is the number of topics; $P_t@R$ is the precision at *R*; and in A2F,
$R = min(250,\ \#\ of\ relevant\ anchors\ in\ qrels\ of\ topic\ t)$,
or, in F2F,
$R = min(1250,\ \#\ of\ relevant\ links\ in\ qrels\ of\ topic\ t)$.

Similarly, *Precision-at-N* is computed using the average precision for all topics at a pre-defined position *N* in the results list. In this evaluation, values of *N* were chosen as: 5, 10, 20, 30, 50, and 250.

5.4 Evaluation Tool

The evaluation tool used for facilitating the system performance benchmarking was updated according to the new settings of the CrossLink-2 tasks. A snapshot of the enhanced tool is given in Figure 3.

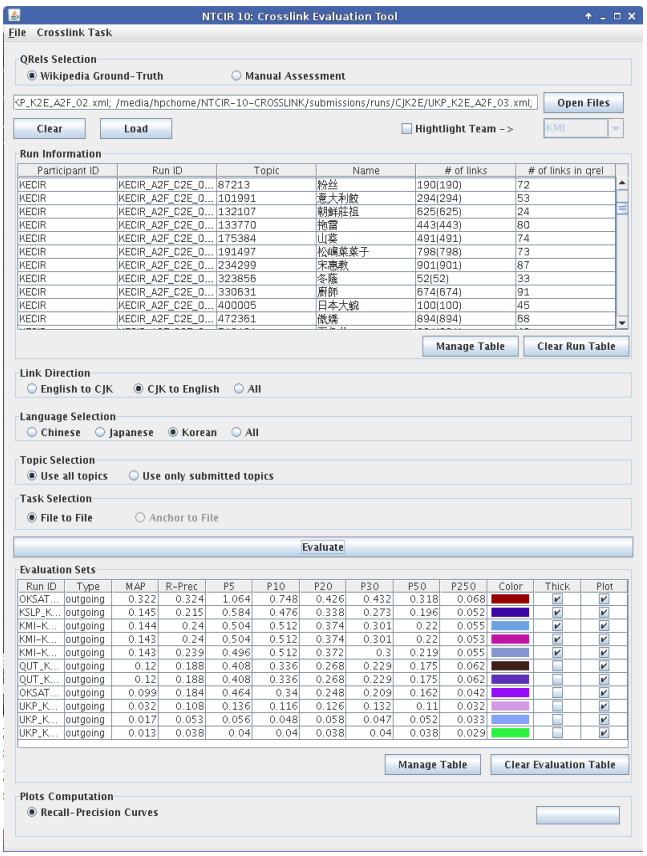


Figure 3. Evaluation Tool

6. EVALUATION RESULTS

This section presents all the evaluation results calculated against the *qrels* created from Wikipedia Ground-truth and manual assessment results for both the English to CJK and the CJK to English tasks. The evaluation includes all 25 test topics. All *LMAP* and *R-Prec* tables showed in this section are sorted on the *LMAP* scores of submitted runs; and P@N tables are sorted on P@5 values.

Generally, there are three different evaluation scenarios:

- F2F evaluation with Wikipedia ground-truth (GT F2F)
- F2F evaluation with manual assessment results (MA F2F)
- A2F evaluation with manual assessment results (MA A2F)

6.1 English to CJK Language Tasks

Table 6 shows the scores of *LMAP* and *R-Prec* of all English to CJK runs evaluated in file to file level with the Wikipedia ground truth for the three different language subtasks: English-to-Chinese, English-to-Japanese, and English-to-Korean respectively.

Similarly, Table 7 and Table 8 show the results (*LMAP* and *R-Prec* scores) of runs evaluated with manual assessment results in F2F and A2F level respectively.

Table 6. E2CJK F2F evaluation results with Wikipedia ground truth: LMAP, R-Prec

English-to-Chinese		
Run-ID	**LMAP**	**R-Prec**
KMI-E2C-A2F-02-ORC	0.404	0.404
OKSAT-E2C-A2F-01-SMP	0.280	0.358
KMI-E2C-A2F-01-ESA	0.249	0.335
OKSAT-E2C-A2F-01-REF	0.071	0.148
QUT_E2C_A2F_01_LinkProbPnCase Sensitive	0.048	0.108
DCU-E2C-A2F-01-NW	0.045	0.129
DCU-E2C-A2F-002-NWE	0.045	0.129
III_E2C_A2F_01_PNM	0.000	0.000
English-to-Japanese		
Run-ID	**LMAP**	**R-Prec**
OKSAT-E2J-A2F-01-SMP	0.423	0.425
KMI-E2J-A2F-02-ORC	0.341	0.341
KMI-E2J-A2F-01-ESA	0.206	0.285
OKSAT-E2J-A2F-01-REF	0.101	0.176
QUT_E2J_A2F_01_LinkProbPnCaseSensitive	0.043	0.098
English-to-Korean		
Run-ID	**LMAP**	**R-Prec**
OKSAT-E2K-A2F-01-SMP	0.623	0.626
KMI-E2K-A2F-02-ORC	0.492	0.492
KMI-E2K-A2F-01-ESA	0.302	0.384
OKSAT-E2K-A2F-01-REF	0.176	0.274
QUT_E2K_A2F_01_LinkProbPnCase Sensitive	0.062	0.116

Table 9 to Table 17 list the scores of runs evaluated using the *Precision-at-N* metric with different *qrels* (Wikipedia ground truth and manual assessment results) in different evaluation levels (file to file and anchor to file) for English-to-Chinese, English-to-Japanese, and English-to-Korean tasks separately.

Figure 4 to Figure 9 show the interpolated precision-recall curves of English to CJK runs in different evaluation scenarios separately. So with these curves, the performance of each system can be easily compared.

Overall, systems of team KMI have the highest scores in multiple evaluation scenarios (measured with different metrics: *LMAP, R-Prec, Precision-at-N* in different evaluation levels against different *qrels*).

Particularly, in the file to file evaluation where system is measured with *LMAP* metric, for the English to Chinese task a KMI system has the best performance. For both the English to Japanese and the English to Korean tasks, team OKSAT however takes the lead.

Table 7. E2CJK F2F evaluation results with manual assessment results: LMAP, R-PREC

English-to-Chinese		
Run-ID	**LMAP**	**R-Prec**
KMI-E2C-A2F-02-ORC	0.133	0.273
KMI-E2C-A2F-01-ESA	0.112	0.275
QUT_E2C_A2F_01_LinkProbPnCase Sensitive	0.099	0.102
OKSAT-E2C-A2F-01-SMP	0.020	0.054
OKSAT-E2C-A2F-01-REF	0.013	0.048
DCU-E2C-A2F-01-NW	0.006	0.026
DCU-E2C-A2F-002-NWE	0.006	0.026
III_E2C_A2F_01_PNM	0.000	0.000
English-to-Japanese		
Run-ID	**LMAP**	**R-Prec**
KMI-E2J-A2F-02-ORC	0.450	0.513
KMI-E2J-A2F-01-ESA	0.383	0.424
OKSAT-E2J-A2F-01-SMP	0.281	0.410
OKSAT-E2J-A2F-01-REF	0.121	0.240
QUT_E2J_A2F_01_LinkProbPnCaseSensitive	0.086	0.114
English-to-Korean		
Run-ID	**LMAP**	**R-Prec**
KMI-E2K-A2F-02-ORC	0.433	0.493
KMI-E2K-A2F-01-ESA	0.424	0.457
OKSAT-E2K-A2F-01-SMP	0.334	0.422
OKSAT-E2K-A2F-01-REF	0.198	0.325
QUT_E2K_A2F_01_LinkProbPnCase Sensitive	0.102	0.144

Table 8. E2CJK A2F evaluation results with manual assessment results: LMAP, R-PREC

English-to-Chinese		
Run-ID	**LMAP**	**R-Prec**
QUT_E2C_A2F_01_LinkProbPnCase Sensitive	0.229	0.245
KMI-E2C-A2F-01-ESA	0.174	0.201
KMI-E2C-A2F-02-ORC	0.168	0.210
OKSAT-E2C-A2F-01-REF	0.020	0.029
OKSAT-E2C-A2F-01-SMP	0.019	0.026
DCU-E2C-A2F-01-NW	0.007	0.008
DCU-E2C-A2F-002-NWE	0.007	0.008
III_E2C_A2F_01_PNM	0.000	0.000
English-to-Japanese		
Run-ID	**LMAP**	**R-Prec**
KMI-E2J-A2F-02-ORC	0.452	0.337
KMI-E2J-A2F-01-ESA	0.440	0.279
OKSAT-E2J-A2F-01-SMP	0.196	0.190
QUT_E2J_A2F_01_LinkProbPnCaseSensitive	0.187	0.125
OKSAT-E2J-A2F-01-REF	0.123	0.115
English-to-Korean		
Run-ID	**LMAP**	**R-Prec**
KMI-E2K-A2F-01-ESA	0.537	0.311
KMI-E2K-A2F-02-ORC	0.533	0.293
OKSAT-E2K-A2F-01-SMP	0.284	0.194
QUT_E2K_A2F_01_LinkProbPnCase Sensitive	0.220	0.127
OKSAT-E2K-A2F-01-REF	0.211	0.147

In the anchor to file evaluation, where the relevance of anchor is considered and assessed by human assessor, a QUT system seemly identifies the most relevant anchors, which leads to increasing the link average precision. Also, with the same *LMAP* metric KMI systems outperform any other systems in both the English to Japanese and the English to Korean tasks.

The top three teams benchmarked with two measures (*LMAP* and *P@5)* in different evaluation levels for different language tasks are listed below:

F2F evaluation with Wikipedia ground truth

- English-to-Chinese

LMAP: KMI, OKSAT, QUT

Precision-at-5: KMI, OKSAT, QUT

- English-to-Japanese

LMAP: OKSAT, KMI, QUT

Precision-at-5: KMI, OKSAT, QUT

- English-to-Korean

LMAP: OKSAT, KMI, QUT

Precision-at-5: OKSAT, KMI, QUT

F2F evaluation with manual assessment results

- English-to-Chinese

LMAP: KMI, QUT, OKSAT

Precision-at-5: KMI, QUT, OKSAT

- English-to-Japanese

LMAP: KMI, OKSAT, QUT

Precision-at-5: KMI, OKSAT, QUT

- English-to-Korean

LMAP: KMI, OKSAT, QUT

Precision-at-5: KMI, OKSAT, QUT

A2F evaluation with manual assessment results

- English-to-Chinese

LMAP: QUT, KMI, OKSAT

Precision-at-5: KMI, QUT, OKSAT

- English-to-Japanese

LMAP: KMI, OKSAT, QUT

Precision-at-5: KMI, OKSAT, QUT

- English-to-Korean

LMAP: KMI, OKSAT, QUT

Precision-at-5: KMI, OKSAT, QUT

Table 9. F2F evaluation results with Wikipedia ground truth: *Precision-at-N* (English-to-Chinese)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-E2C-A2F-02-ORC	1.000	0.988	0.822	0.637	0.403	0.081
OKSAT-E2C-A2F-01-SMP	0.720	0.698	0.620	0.511	0.358	0.074
KMI-E2C-A2F-01-ESA	0.664	0.668	0.554	0.467	0.327	0.081
OKSAT-E2C-A2F-01-REF	0.368	0.384	0.306	0.231	0.144	0.029
QUT_E2C_A2F_01_LinkProbPnCaseSensitive	0.216	0.188	0.150	0.124	0.105	0.042
DCU-E2C-A2F-01-NW	0.112	0.156	0.168	0.145	0.118	0.043
DCU-E2C-A2F-002-NWE	0.112	0.156	0.168	0.145	0.118	0.043
III_E2C_A2F_01_PNM	0.000	0.000	0.000	0.000	0.000	0.000

Table 10. F2F evaluation results with Wikipedia ground truth: *Precision-at-N* (English-to-Japanese)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-E2J-A2F-02-ORC	1.000	1.000	0.856	0.675	0.423	0.085
OKSAT-E2J-A2F-01-SMP	0.992	0.966	0.878	0.728	0.506	0.104
KMI-E2J-A2F-01-ESA	0.696	0.680	0.558	0.473	0.334	0.085
OKSAT-E2J-A2F-01-REF	0.520	0.490	0.420	0.336	0.210	0.042
QUT_E2J_A2F_01_LinkProbPnCaseSensitive	0.160	0.172	0.130	0.125	0.114	0.050

Table 11. F2F evaluation results with Wikipedia ground truth: *Precision-at-N* (English-to-Korean)

Run-ID	P5	P10	P20	P30	P50	P250
OKSAT-E2K-A2F-01-SMP	1.000	0.888	0.826	0.661	0.441	0.088
KMI-E2K-A2F-02-ORC	1.000	0.980	0.746	0.563	0.340	0.068
KMI-E2K-A2F-01-ESA	0.672	0.636	0.528	0.420	0.288	0.068
OKSAT-E2K-A2F-01-REF	0.584	0.520	0.392	0.296	0.186	0.037
QUT_E2K_A2F_01_LinkProbPnCaseSensitive	0.192	0.180	0.136	0.119	0.093	0.036

Table 12. F2F evaluation results with manual assessment results: *Precision-at-N* (English-to-Chinese)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-E2C-A2F-02-ORC	0.352	0.360	0.300	0.273	0.264	0.119
KMI-E2C-A2F-01-ESA	0.104	0.188	0.218	0.248	0.258	0.119
QUT_E2C_A2F_01_LinkProbPnCaseSensitive	0.152	0.112	0.092	0.093	0.103	0.112
OKSAT-E2C-A2F-01-SMP	0.224	0.196	0.142	0.115	0.086	0.017
OKSAT-E2C-A2F-01-REF	0.128	0.120	0.134	0.109	0.073	0.015
DCU-E2C-A2F-01-NW	0.016	0.024	0.028	0.029	0.023	0.020
DCU-E2C-A2F-002-NWE	0.016	0.024	0.028	0.029	0.023	0.020
III_E2C_A2F_01_PNM	0.000	0.000	0.000	0.000	0.000	0.000

Table 13. F2F evaluation results with manual assessment results: *Precision-at-N* (English-to-Japanese)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-E2J-A2F-02-ORC	0.616	0.628	0.542	0.447	0.314	0.077
OKSAT-E2J-A2F-01-SMP	0.576	0.514	0.442	0.339	0.222	0.044
KMI-E2J-A2F-01-ESA	0.568	0.556	0.444	0.384	0.281	0.077
OKSAT-E2J-A2F-01-REF	0.384	0.344	0.270	0.201	0.126	0.025
QUT_E2J_A2F_01_LinkProbPnCaseSensitive	0.136	0.124	0.104	0.107	0.097	0.042

Table 14. F2F evaluation results with manual assessment results: *Precision-at-N* (English-to-Korean)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-E2K-A2F-01-ESA	0.584	0.532	0.432	0.349	0.240	0.061
OKSAT-E2K-A2F-01-SMP	0.544	0.460	0.400	0.308	0.201	0.040
KMI-E2K-A2F-02-ORC	0.488	0.512	0.476	0.377	0.258	0.061
OKSAT-E2K-A2F-01-REF	0.424	0.380	0.282	0.216	0.133	0.027
QUT_E2K_A2F_01_LinkProbPnCaseSensitive	0.192	0.164	0.120	0.103	0.082	0.032

Table 15. A2F evaluation results with manual assessment results: *Precision-at-N* (English-to-Chinese)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-E2C-A2F-02-ORC	0.168	0.176	0.182	0.171	0.160	0.072
QUT_E2C_A2F_01_LinkProbPnCaseSensitive	0.104	0.112	0.176	0.201	0.201	0.097
KMI-E2C-A2F-01-ESA	0.056	0.116	0.120	0.144	0.149	0.070
OKSAT-E2C-A2F-01-SMP	0.048	0.080	0.050	0.039	0.034	0.007
OKSAT-E2C-A2F-01-REF	0.008	0.012	0.046	0.040	0.037	0.008
DCU-E2C-A2F-002-NWE	0.000	0.000	0.000	0.001	0.004	0.007
DCU-E2C-A2F-01-NW	0.000	0.000	0.000	0.001	0.004	0.007
III_E2C_A2F_01_PNM	0.000	0.000	0.000	0.000	0.000	0.000

Table 16. A2F evaluation results with manual assessment results: *Precision-at-N* (English-to-Japanese)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-E2J-A2F-01-ESA	0.336	0.352	0.288	0.264	0.206	0.058
KMI-E2J-A2F-02-ORC	0.296	0.348	0.332	0.305	0.243	0.060
OKSAT-E2J-A2F-01-SMP	0.208	0.208	0.198	0.160	0.106	0.023
QUT_E2J_A2F_01_LinkProbPnCaseSensitive	0.192	0.156	0.120	0.105	0.089	0.028
OKSAT-E2J-A2F-01-REF	0.152	0.136	0.122	0.100	0.070	0.014

Table 17. A2F evaluation results with manual assessment results: *Precision-at-N* (English-to-Korean)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-E2K-A2F-01-ESA	0.344	0.328	0.296	0.243	0.178	0.046
KMI-E2K-A2F-02-ORC	0.224	0.264	0.290	0.264	0.198	0.046
OKSAT-E2K-A2F-01-SMP	0.208	0.224	0.196	0.157	0.110	0.023
QUT_E2K_A2F_01_LinkProbPnCaseSensitive	0.192	0.156	0.124	0.104	0.078	0.022
OKSAT-E2K-A2F-01-REF	0.128	0.144	0.122	0.109	0.078	0.016

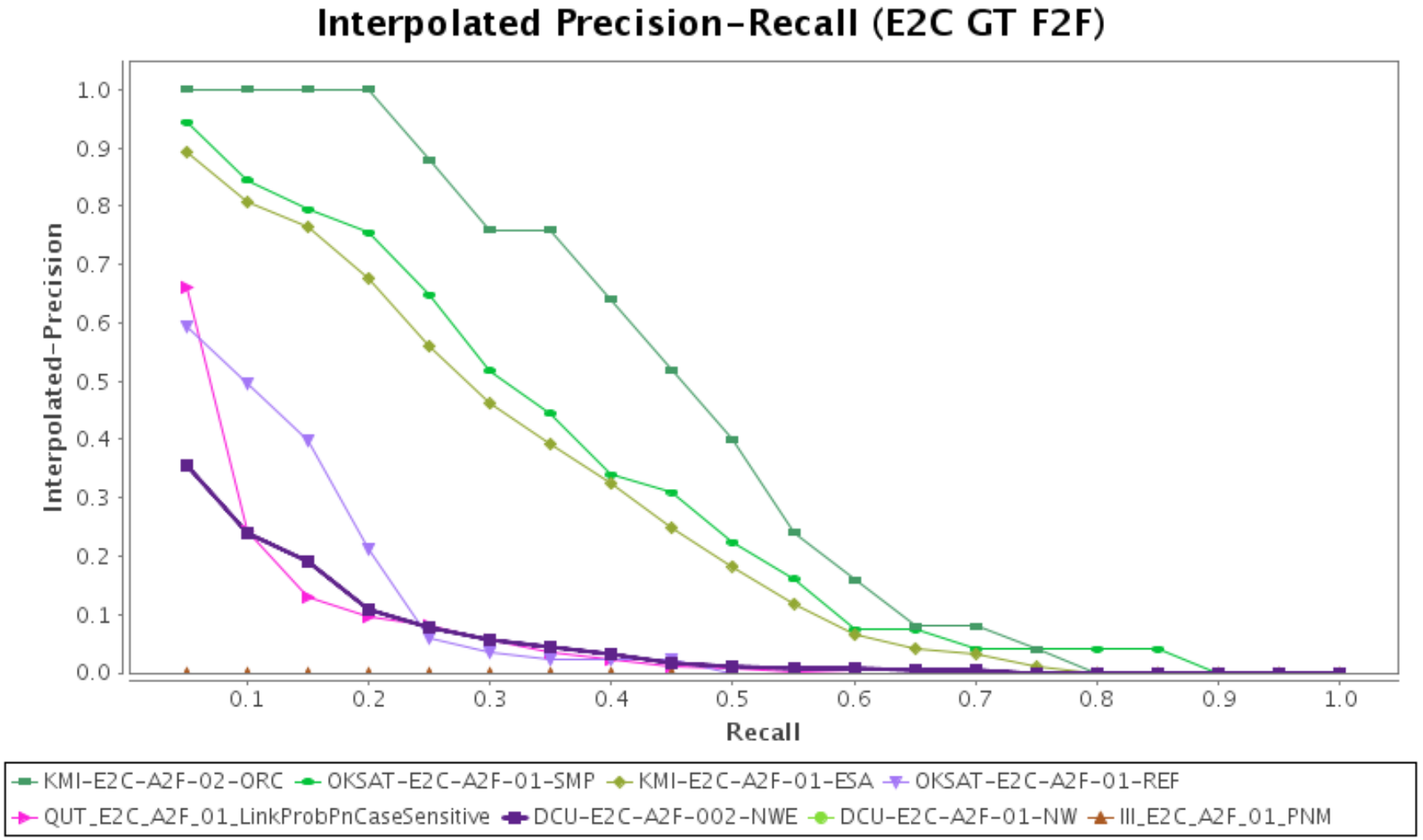


Figure 4. Interpolated Precision-Recall curves of English-to-Chinese runs in F2F evaluation with Wikipedia ground-truth

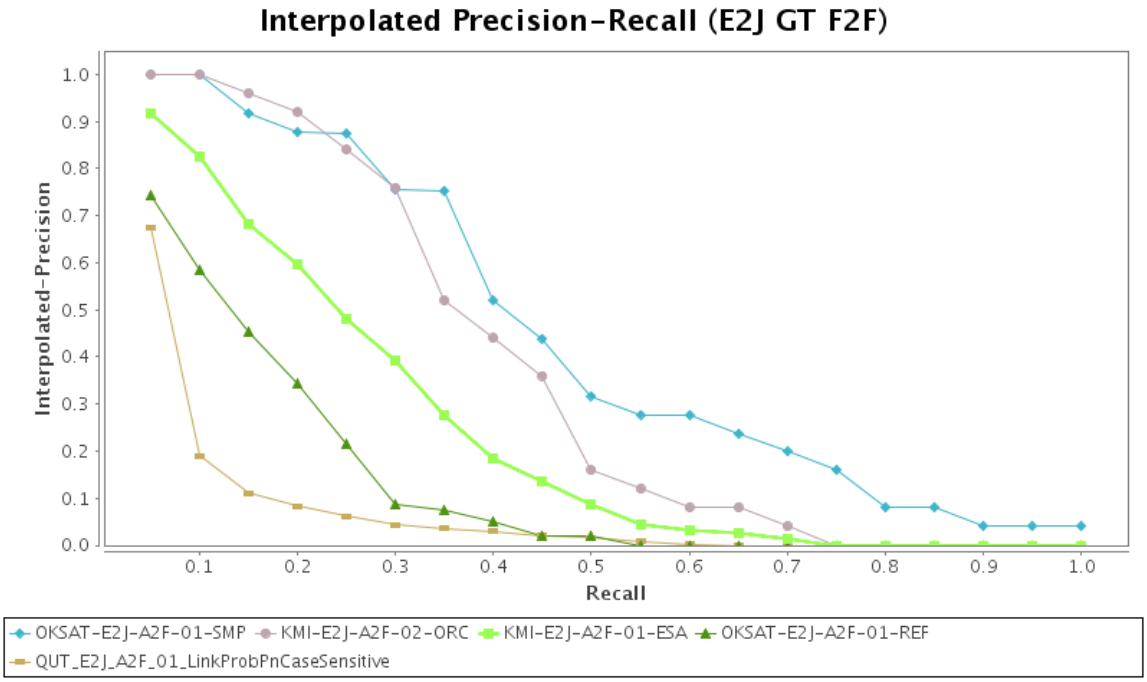


Figure 5. Interpolated Precision-Recall of curves English-to-Japanese runs in F2F evaluation with Wikipedia ground-truth

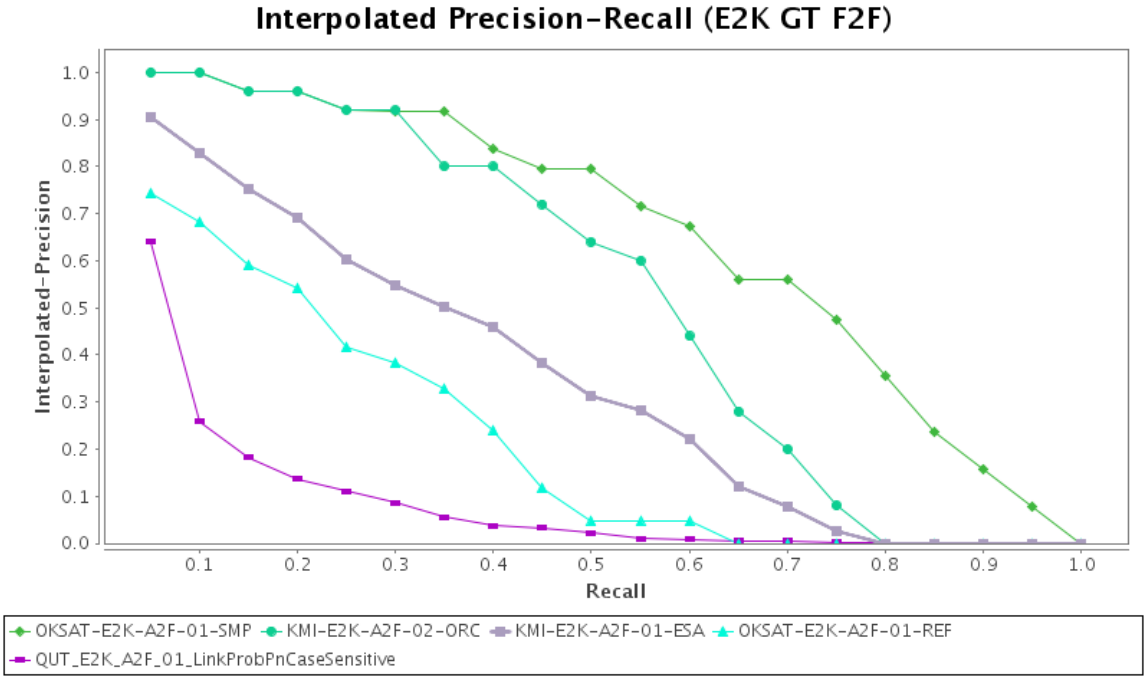
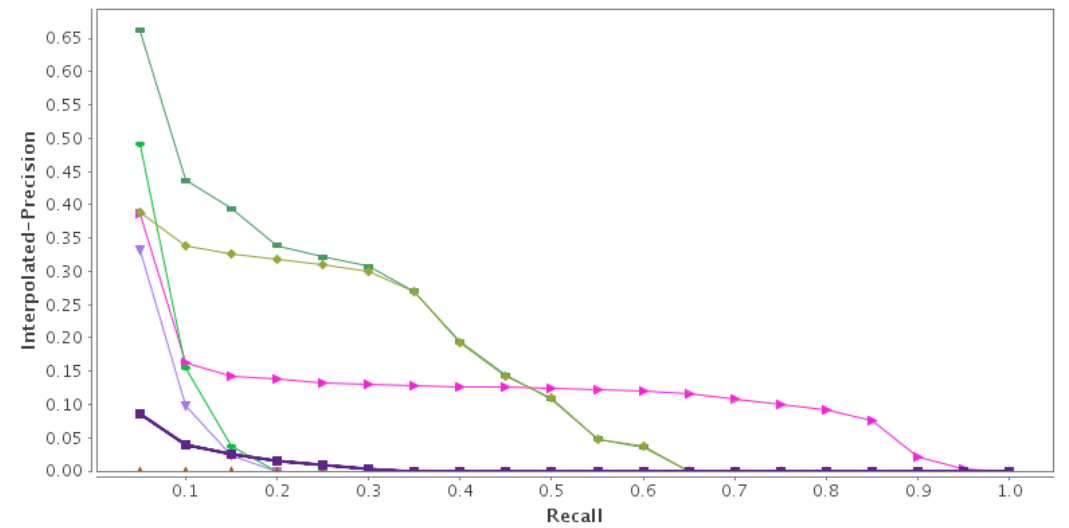


Figure 6. Interpolated Precision-Recall curves of English-to-Korean runs in F2F evaluation with Wikipedia ground-truth

a.



b.

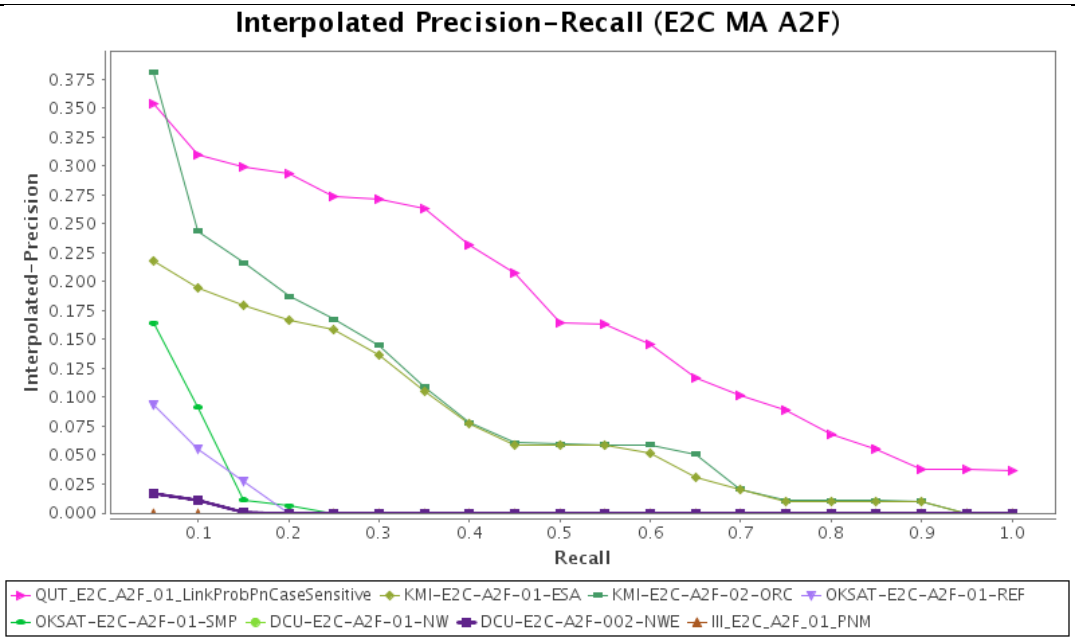
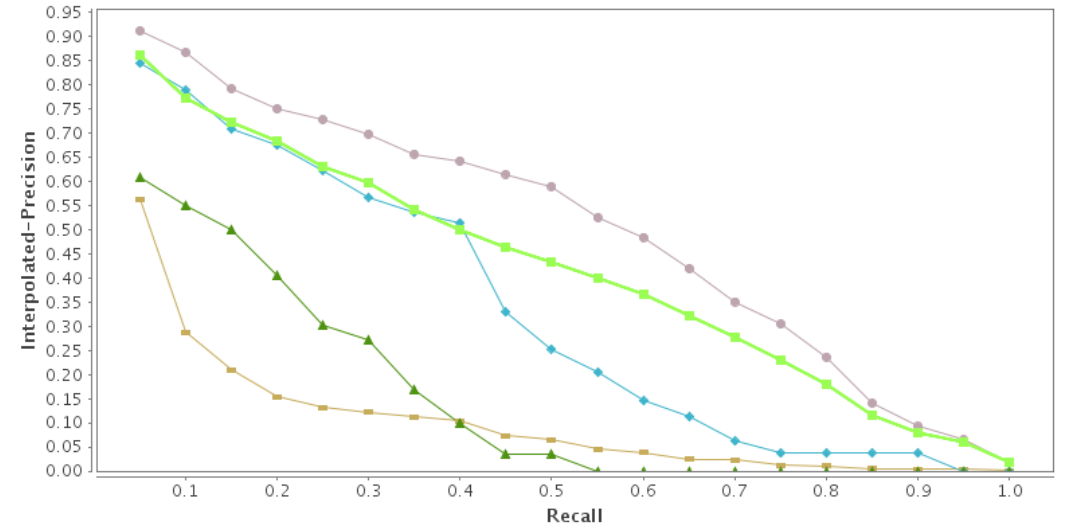


Figure 7. Interpolated Precision-Recall of English-to-Chinese runs evaluated with manual assessment results (plot *a* is P-R curves of runs evaluated in F2F level; plot *b* is P-R curves of runs evaluated in A2F level)

a.



b.

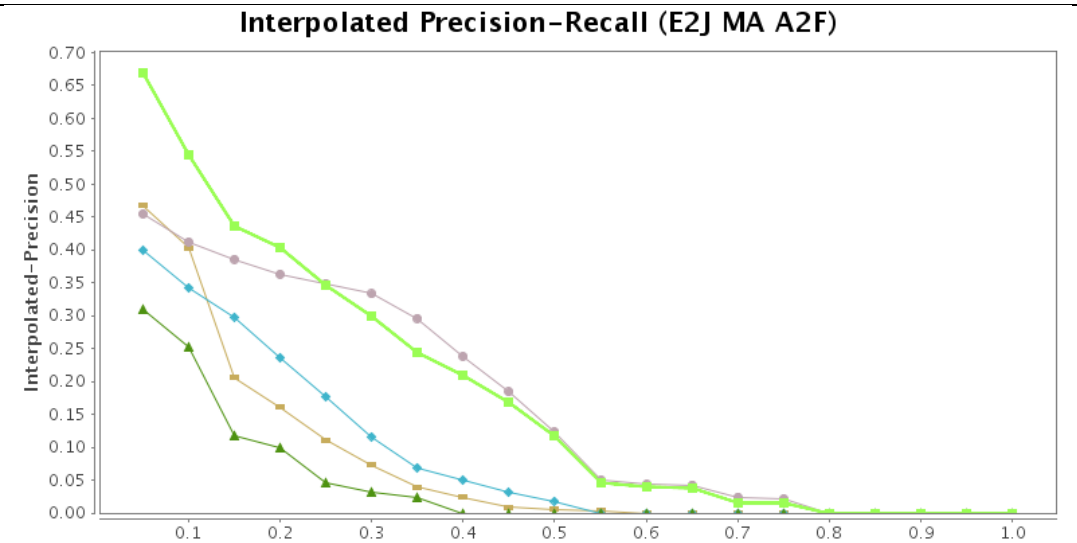
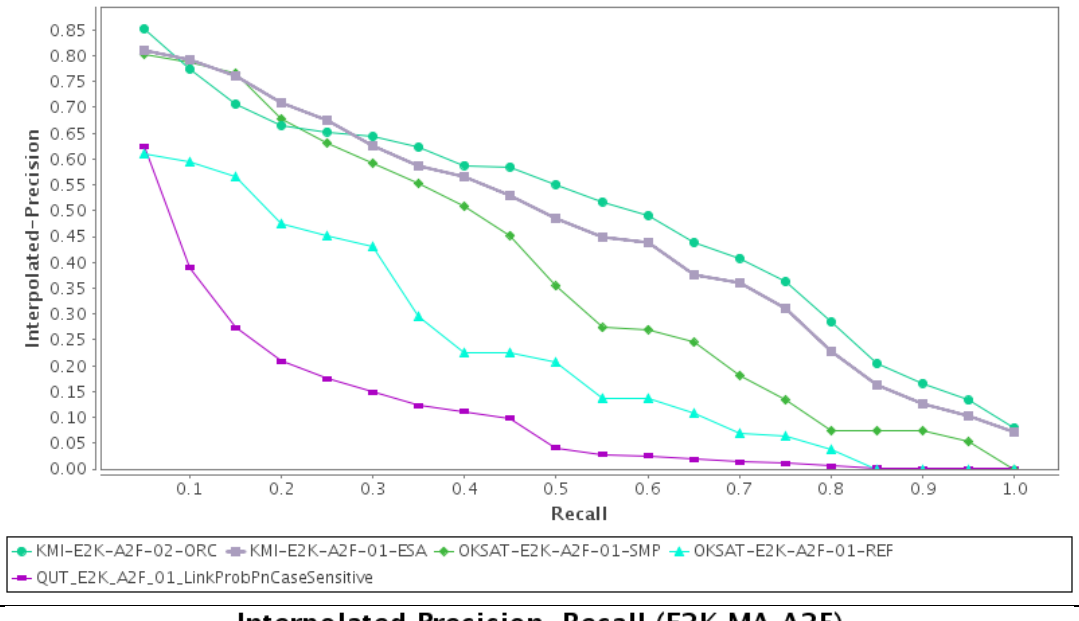


Figure 8. Interpolated Precision-Recall of English-to-Japanese runs evaluated with manual assessment results (plot *a* is P-R curves of runs evaluated in F2F level; plot *b* is P-R curves of runs evaluated in A2F level)

a.



b.

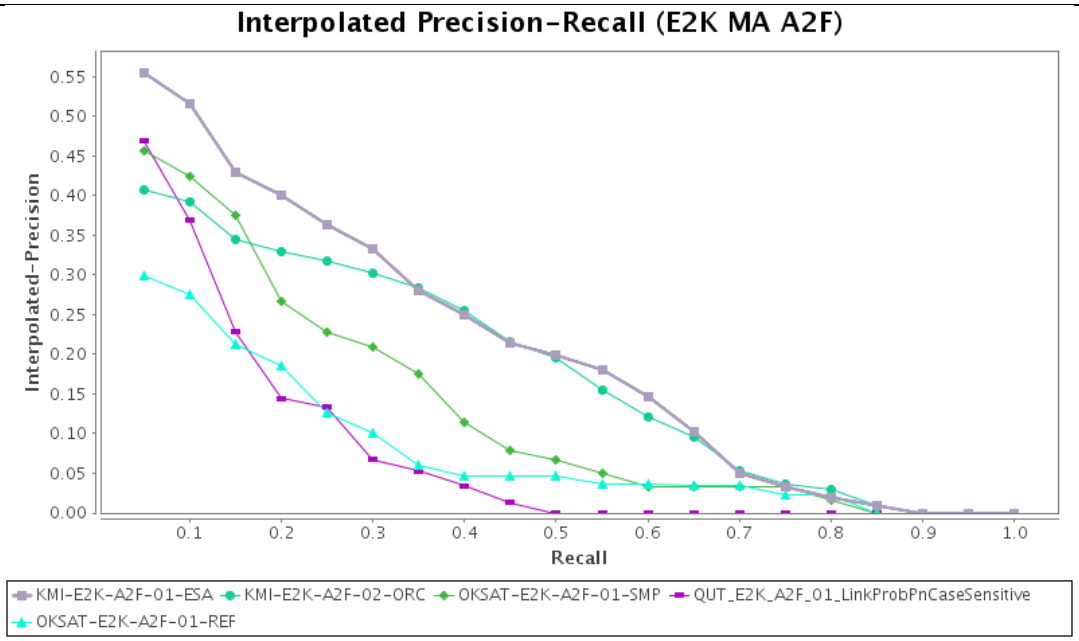


Figure 9. Interpolated Precision-Recall of English-to-Korean links evaluated with manual assessment results (plot *a* is P-R curves of runs evaluated in F2F level; plot *b* is P-R curves of runs evaluated in A2F level)

6.2 CJK Language to English Tasks

Unlike linking English to CJK articles, linking CJK articles to English could require natural language processing which may have to be treated differently. Even the same anchor identification and link recommendation methods are used, results may vary.

Table 18, Table 19 and Table 20 show the scores of *LMAP* and *R-Prec* of all CJK to English runs evaluated in different scenarios (F2F evaluation with Wikipedia ground truth, and F2F & A2F evaluations with manual assessment results) for the three different language subtasks (Chinese-to-English, Japanese-to-English, and Korean-to-English) separately.

Table 21 to Table 29 list the scores of CJK to English runs evaluated using the *Precision-at-N* metric with Wikipedia ground truth and manual assessment results in both file to file and anchor to file levels for the Chinese-to-English, Japanese-to-English, and Korean-to-English tasks separately.

Figure 10 to Figure 18 show the interpolated precision-recall curves of runs in different evaluation scenarios of the CJK to English tasks.

When runs are measured against the *qrel* created from Wikipedia ground truth in file to file level, team OKSAT takes up the first place in all CJK to English tasks, and remains as the top performer in the Japanese to English task when evaluated with the *qrel* built from the manual assessment results. In other F2F evaluations with manual assessment results, team QUT has the highest LMAP score in the Chinese to English task, while team KSLP is ranked number one in the Korea to English task.

When the relevance of recommended anchors is concerned and addressed, the rankings of systems could be different to that in the F2F evaluation. So in the A2F evaluation with manual assessment results, a system of team KECIR in Chinese to English task achieves the highest LMAP score while team QUT in the Japanese to English task along with team KSLP in the Korean to English task have the highest rank.

With the LMAP metric Team OKSAT does very well in the F2F evaluation with Wikipedia ground truth, but has low rankings in the A2F evaluation with manual assessment results. It may be worthy of note that runs of team OKSAT was not pooled along with others for manual assessment due to the late submission.

Although team OKSAT has less satisfactory performance in the evaluation with manual assessment results, they still achieve promising results (ranked number one) when measured with P@N metric in most evaluation scenarios. When measured with P@N metric, team KMI is the top performer of the F2F evaluation with Wikipedia ground truth in the Japanese to English task; a system of KSLP outperform any others in all evaluations with manual assessment results in the Korean to English task.

The top three teams of different evaluation scenarios with two measures (*LMAP* and *P@5)* are given below:

F2F evaluation with Wikipedia ground truth

- Chinese-to-English

LMAP: OKSAT, KMI, UKP

Precision-at-5: OKSAT, UKP, KMI

- Japanese-to-English

LMAP: OKSAT, KMI, UKP

Precision-at-5: KMI, OKSAT, UKP

- Korean-to-English

LMAP: OKSAT, KSLP, KMI

Precision-at-5: OKSAT, KSLP, KMI

F2F evaluation with manual assessment results

- Chinese-to-English

LMAP: QUT, KMI, OKSAT

Precision-at-5: OKSAT, NTHU, QUT

- Japanese-to-English

LMAP: OKSAT, UKP, KMI

Precision-at-5: OKSAT, KMI, UKP

- Korean-to-English

LMAP: KSLP, OKSAT, KMI

Precision-at-5: KSLP, OKSAT, KMI

A2F evaluation with manual assessment results

- Chinese-to-English

LMAP: KECIR, QUT, KMI

Precision-at-5: OKSAT, NTHU, QUT

- Japanese-to-English

LMAP: QUT, UKP, OKSAT

Precision-at-5: OKSAT, RDLL, UKP

- Korean-to-English

LMAP: KSLP, KMI, OKSAT

Precision-at-5: KSLP, KMI, OKSAT

7. CONCLUSION AND FUTURE WORK

In this paper, we present the evaluation of the second cross-lingual link discovery (CrossLink-2) task at NTCIR-10. For the task, new CEJK Wikipedia collections and new test topics were created, and system performance was benchmarked with the previously developed evaluation framework which was subsequently enhanced according to the new settings. This year's task focuses on cross-lingual linking of documents from Chinese, Korean, and Japanese to English. However, evaluation for E2CJK language tasks was also supported.

The evaluations show promising results of participating teams, and some teams achieve very good results. The evaluation framework is continuously proven effective in identifying high performance algorithms for the CJK to English document linking.

We believe there are still much to do in the future CLLD research. At the moment, only general cross-lingual document linking approaches were proposed, and the needs for personalised knowledge discovery are not addressed. Furthermore, these CLLD methods were only experimented on the Wikipedia data. Will they succeed in other knowledge domain such as patent data? How do we do cross-lingual

document linking between different knowledge bases (e.g. patents and Wikipedia)? These still are unanswered questions.

REFERENCES

[1] L.-X. Tang, K. Y. Itakura, S. Geva, A. Trotman, and Y. Xu, "The Effectiveness of Cross-lingual Link Discovery," in *Proceedings of The Fourth International Workshop on Evaluating Information Access (EVIA)*, Tokyo, Japan, 2011, pp. 1-8.

[2] L.-X. Tang, S. Geva, A. Trotman, Y. Xu, and K. Y. Itakura, "Overview of the NTCIR-9 Crosslink Task: Cross-lingual Link Discovery," in *Proceedings of NTCIR-9*, Tokyo, Japan, 2011, pp. 437-463.

[3] W. Huang, S. Geva, and A. Trotman, "Overview of the INEX 2009 Link the Wiki Track," in *Proceedings of INEX 2009*, 2010, pp. 312-323.

[4] R. A. Baeza-Yates and B. A. Ribeiro-Neto, *Modern Information Retrieval*: ACM Press / Addison-Wesley, 1999.

Table 18. CJK2E F2F evaluation with Wikipedia ground-truth: LMAP, R-PREC

Chinese-to-English			Japanese-to-English			Korean-to-English		
Run-ID	**LMAP**	**R-Prec**	**Run-ID**	**LMAP**	**R-Prec**	**Run-ID**	**LMAP**	**R-Prec**
OKSAT-C2E-A2F-01-SMP	0.517	0.520	OKSAT-J2E-A2F-01-SMP	0.548	0.561	OKSAT-K2E-A2F-01-SMP	0.322	0.324
KMI-C2E-A2F-02-ORC	0.221	0.337	KMI-J2E-A2F-02-ORC	0.224	0.224	KSLP_K2E_A2F_01_MLUA	0.145	0.215
KMI-C2E-A2F-01-LIS	0.221	0.336	UKP_J2E_A2F_01_CSC	0.220	0.318	KMI-K2E-A2F-01-ORC	0.144	0.240
KMI-C2E-A2F-03-LIS	0.219	0.336	QUT_J2E_A2F_01_LinkProbPN	0.171	0.281	KMI-K2E-A2F-03-LIS	0.143	0.240
UKP_C2E_A2F_01_CSC	0.181	0.260	QUT_J2E_A2F_02_LinkProbPN2	0.171	0.281	KMI-K2E-A2F-01-LIS	0.143	0.239
QUT_C2E_A2F_01_LinkProbPN	0.158	0.282	OKSAT-J2E-A2F-01-REF	0.150	0.294	QUT_K2E_A2F_02_LinkProbPN2	0.120	0.188
OKSAT-C2E-A2F-01-REF	0.134	0.248	UKP_J2E_A2F_02_CSC	0.117	0.224	QUT_K2E_A2F_01_LinkProbPN	0.120	0.188
UKP_C2E_A2F_02_CSC	0.102	0.218	KMI-J2E-A2F-01-LIS	0.114	0.176	OKSAT-K2E-A2F-01-REF	0.099	0.184
NTHU-C2E-A2F-03-lda-ext	0.082	0.194	KMI-J2E-A2F-03-LIS	0.113	0.176	UKP_K2E_A2F_01_CSC	0.032	0.108
NTHU-C2E-A2F-02-lda-ext	0.082	0.194	UKP_J2E_A2F_03_CSC	0.086	0.173	UKP_K2E_A2F_02_CSC	0.017	0.053
NTHU-C2E-A2F-01-keywordSim	0.080	0.192	NTHU-J2E-A2F-01-keywordSim	0.083	0.189	UKP_K2E_A2F_03_CSC	0.013	0.038
UKP_C2E_A2F_03_CSC	0.075	0.157	RDLL_A2F_J2E_05_tfdiceLL	0.063	0.143			
III_C2E_A2F_01_PNM	0.072	0.172	RDLL_A2F_J2E_02_okapiBM25	0.028	0.082			
III_C2E_A2F_02_PNM	0.071	0.133	RDLL_A2F_J2E_01_tfidf	0.021	0.073			
QUT_C2E_A2F_02_LinkProbPN2	0.059	0.111	RDLL_A2F_J2E_04_tfdice	0.021	0.072			
KECIR_A2F_C2E_02_FSCLIR	0.054	0.119	RDLL_A2F_J2E_03_dice	0.021	0.067			
KECIR_A2F_C2E_01_FSCLIR	0.046	0.105						
KECIR_A2F_C2E_03_FSCLIR	0.036	0.097						
KECIR_A2F_C2E_04_FSCLIR	0.036	0.105						
III_C2E_A2F_03_PNM	0.032	0.091						
DCU-C2E-A2F-03-NW	0.011	0.049						
DCU-C2E-A2F-04-NWE	0.011	0.049						

Table 19. CJK2E F2F evaluation with manual assessment results: LMAP, R-PREC

Chinese-to-English			Japanese-to-English			Korean-to-English		
Run-ID	**LMAP**	**R-Prec**	**Run-ID**	**LMAP**	**R-Prec**	**Run-ID**	**LMAP**	**R-Prec**
QUT_C2E_A2F_01_LinkProbPN	0.069	0.132	OKSAT-J2E-A2F-01-SMP	0.312	0.418	KSLP_K2E_A2F_01_MLUA	0.302	0.301
KMI-C2E-A2F-02-ORC	0.067	0.180	UKP_J2E_A2F_01_CSC	0.220	0.305	OKSAT-K2E-A2F-01-SMP	0.294	0.352
KMI-C2E-A2F-01-LIS	0.067	0.180	KMI-J2E-A2F-02-ORC	0.171	0.271	KMI-K2E-A2F-01-ORC	0.264	0.284
OKSAT-C2E-A2F-01-REF	0.065	0.149	OKSAT-J2E-A2F-01-REF	0.158	0.242	KMI-K2E-A2F-01-LIS	0.262	0.284
KMI-C2E-A2F-03-LIS	0.064	0.180	QUT_J2E_A2F_01_LinkProbPN	0.145	0.161	KMI-K2E-A2F-03-LIS	0.260	0.284
KECIR_A2F_C2E_01_FSCLIR	0.044	0.081	QUT_J2E_A2F_02_LinkProbPN2	0.145	0.161	OKSAT-K2E-A2F-01-REF	0.198	0.190
UKP_C2E_A2F_01_CSC	0.038	0.140	KMI-J2E-A2F-01-LIS	0.138	0.202	QUT_K2E_A2F_01_LinkProbPN	0.196	0.204
KECIR_A2F_C2E_02_FSCLIR	0.037	0.077	RDLL_A2F_J2E_05_tfdiceLL	0.137	0.160	QUT_K2E_A2F_02_LinkProbPN2	0.196	0.204
QUT_C2E_A2F_02_LinkProbPN2	0.037	0.049	KMI-J2E-A2F-03-LIS	0.137	0.202	UKP_K2E_A2F_01_CSC	0.094	0.085
NTHU-C2E-A2F-02-lda-ext	0.034	0.113	UKP_J2E_A2F_02_CSC	0.127	0.168	UKP_K2E_A2F_02_CSC	0.047	0.035
NTHU-C2E-A2F-03-lda-ext	0.034	0.114	NTHU-J2E-A2F-01-keywordSim	0.102	0.138	UKP_K2E_A2F_03_CSC	0.036	0.042
KECIR_A2F_C2E_03_FSCLIR	0.031	0.076	UKP_J2E_A2F_03_CSC	0.100	0.121			
UKP_C2E_A2F_02_CSC	0.029	0.101	RDLL_A2F_J2E_03_dice	0.060	0.083			
KECIR_A2F_C2E_04_FSCLIR	0.028	0.076	RDLL_A2F_J2E_04_tfdice	0.060	0.091			
III_C2E_A2F_02_PNM	0.027	0.090	RDLL_A2F_J2E_02_okapiBM25	0.052	0.080			
NTHU-C2E-A2F-01-keywordSim	0.025	0.096	RDLL_A2F_J2E_01_tfidf	0.046	0.089			
UKP_C2E_A2F_03_CSC	0.024	0.063						
OKSAT-C2E-A2F-01-SMP	0.023	0.059						
III_C2E_A2F_01_PNM	0.011	0.061						
III_C2E_A2F_03_PNM	0.009	0.037						
DCU-C2E-A2F-04-NWE	0.005	0.019						
DCU-C2E-A2F-03-NW	0.005	0.019						

Table 20. CJK2E A2F evaluation with manual assessment results: LMAP, R-PREC

Chinese-to-English			Japanese-to-English			Korean-to-English		
Run-ID	**LMAP**	**R-Prec**	**Run-ID**	**LMAP**	**R-Prec**	**Run-ID**	**LMAP**	**R-Prec**
KECIR_A2F_C2E_01_FSCLIR	0.113	0.147	QUT_J2E_A2F_01_LinkProbPN	0.270	0.068	KSLP_K2E_A2F_01_MLUA	0.208	0.099
QUT_C2E_A2F_02_LinkProbPN2	0.089	0.087	QUT_J2E_A2F_02_LinkProbPN2	0.270	0.068	KMI-K2E-A2F-01-LIS	0.184	0.073
QUT_C2E_A2F_01_LinkProbPN	0.089	0.087	UKP_J2E_A2F_03_CSC	0.187	0.109	KMI-K2E-A2F-01-ORC	0.184	0.073
KMI-C2E-A2F-01-LIS	0.077	0.060	UKP_J2E_A2F_01_CSC	0.187	0.109	KMI-K2E-A2F-03-LIS	0.180	0.073
KMI-C2E-A2F-02-ORC	0.077	0.060	UKP_J2E_A2F_02_CSC	0.187	0.109	OKSAT-K2E-A2F-01-REF	0.139	0.023
KMI-C2E-A2F-03-LIS	0.076	0.060	OKSAT-J2E-A2F-01-SMP	0.159	0.120	QUT_K2E_A2F_01_LinkProbPN	0.137	0.040
KECIR_A2F_C2E_02_FSCLIR	0.065	0.107	RDLL_A2F_J2E_05_tfdiceLL	0.151	0.105	QUT_K2E_A2F_02_LinkProbPN2	0.137	0.040
KECIR_A2F_C2E_03_FSCLIR	0.056	0.090	OKSAT-J2E-A2F-01-REF	0.151	0.046	OKSAT-K2E-A2F-01-SMP	0.131	0.059
OKSAT-C2E-A2F-01-REF	0.053	0.072	NTHU-J2E-A2F-01-keywordSim	0.127	0.074	UKP_K2E_A2F_02_CSC	0.067	0.011
KECIR_A2F_C2E_04_FSCLIR	0.038	0.070	KMI-J2E-A2F-02-ORC	0.072	0.058	UKP_K2E_A2F_01_CSC	0.067	0.011
UKP_C2E_A2F_01_CSC	0.027	0.055	KMI-J2E-A2F-03-LIS	0.062	0.042	UKP_K2E_A2F_03_CSC	0.067	0.011
UKP_C2E_A2F_03_CSC	0.027	0.055	KMI-J2E-A2F-01-LIS	0.062	0.042			
UKP_C2E_A2F_02_CSC	0.027	0.055	RDLL_A2F_J2E_03_dice	0.029	0.031			
NTHU-C2E-A2F-03-lda-ext	0.021	0.036	RDLL_A2F_J2E_04_tfdice	0.029	0.028			
NTHU-C2E-A2F-02-lda-ext	0.020	0.033	RDLL_A2F_J2E_02_okapiBM25	0.027	0.030			
OKSAT-C2E-A2F-01-SMP	0.016	0.022	RDLL_A2F_J2E_01_tfidf	0.023	0.024			
NTHU-C2E-A2F-01-keywordSim	0.012	0.024						
DCU-C2E-A2F-04-NWE	0.007	0.013						
DCU-C2E-A2F-03-NW	0.007	0.013						
III_C2E_A2F_01_PNM	0.000	0.000						
III_C2E_A2F_02_PNM	0.000	0.000						
III_C2E_A2F_03_PNM	0.000	0.000						

Table 21. F2F evaluation with Wikipedia ground-truth: *Precision-at-N* (Chinese-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
OKSAT-C2E-A2F-01-SMP	1.000	0.972	0.916	0.779	0.582	0.123
UKP_C2E_A2F_01_CSC	0.688	0.628	0.528	0.439	0.300	0.063
OKSAT-C2E-A2F-01-REF	0.536	0.468	0.376	0.324	0.257	0.063
KMI-C2E-A2F-02-ORC	0.376	0.412	0.422	0.393	0.334	0.106
KMI-C2E-A2F-01-LIS	0.376	0.412	0.422	0.393	0.334	0.106
KMI-C2E-A2F-03-LIS	0.376	0.412	0.422	0.393	0.334	0.104
UKP_C2E_A2F_02_CSC	0.368	0.312	0.278	0.247	0.222	0.072
QUT_C2E_A2F_01_LinkProbPN	0.352	0.376	0.368	0.332	0.285	0.104
III_C2E_A2F_01_PNM	0.272	0.272	0.254	0.227	0.184	0.043
NTHU-C2E-A2F-03-lda-ext	0.256	0.240	0.230	0.224	0.195	0.070
NTHU-C2E-A2F-02-lda-ext	0.256	0.236	0.228	0.224	0.195	0.070
NTHU-C2E-A2F-01-keywordSim	0.256	0.236	0.226	0.221	0.194	0.068
UKP_C2E_A2F_03_CSC	0.216	0.200	0.180	0.180	0.156	0.072
KECIR_A2F_C2E_02_FSCLIR	0.200	0.164	0.150	0.145	0.126	0.060
III_C2E_A2F_03_PNM	0.168	0.188	0.158	0.131	0.103	0.022
KECIR_A2F_C2E_01_FSCLIR	0.136	0.128	0.138	0.132	0.110	0.055
III_C2E_A2F_02_PNM	0.128	0.156	0.154	0.144	0.126	0.077
DCU-C2E-A2F-04-NWE	0.096	0.104	0.082	0.064	0.047	0.012
DCU-C2E-A2F-03-NW	0.096	0.104	0.082	0.064	0.047	0.012
KECIR_A2F_C2E_03_FSCLIR	0.080	0.108	0.106	0.107	0.102	0.052
QUT_C2E_A2F_02_LinkProbPN2	0.080	0.084	0.104	0.104	0.108	0.080
KECIR_A2F_C2E_04_FSCLIR	0.080	0.112	0.110	0.116	0.110	0.054

Table 22. F2F evaluation with Wikipedia ground-truth: *Precision-at-N* (Japanese-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
KMI-J2E-A2F-02-ORC	0.946	0.877	0.667	0.499	0.315	0.068
OKSAT-J2E-A2F-01-SMP	0.938	0.938	0.915	0.829	0.657	0.178
UKP_J2E_A2F_01_CSC	0.646	0.623	0.550	0.478	0.392	0.107
KMI-J2E-A2F-01-LIS	0.592	0.504	0.388	0.306	0.212	0.063
KMI-J2E-A2F-03-LIS	0.592	0.504	0.388	0.306	0.212	0.061
OKSAT-J2E-A2F-01-REF	0.408	0.396	0.385	0.349	0.291	0.109
UKP_J2E_A2F_02_CSC	0.362	0.315	0.292	0.263	0.231	0.106
QUT_J2E_A2F_02_LinkProbPN2	0.269	0.300	0.315	0.308	0.285	0.136
QUT_J2E_A2F_01_LinkProbPN	0.269	0.300	0.315	0.308	0.285	0.136
RDLL_A2F_J2E_05_tfdiceLL	0.269	0.212	0.190	0.165	0.152	0.062
NTHU-J2E-A2F-01-keywordSim	0.254	0.246	0.233	0.224	0.199	0.084
UKP_J2E_A2F_03_CSC	0.162	0.204	0.196	0.197	0.172	0.094
RDLL_A2F_J2E_02_okapiBM25	0.162	0.123	0.102	0.100	0.091	0.038
RDLL_A2F_J2E_04_tfdice	0.108	0.088	0.094	0.083	0.073	0.033
RDLL_A2F_J2E_03_dice	0.100	0.085	0.087	0.085	0.072	0.033
RDLL_A2F_J2E_01_tfidf	0.069	0.085	0.090	0.082	0.075	0.036

Table 23. F2F evaluation with Wikipedia ground-truth: *Precision-at-N* (Korean-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
OKSAT-K2E-A2F-01-SMP	0.912	0.630	0.544	0.432	0.318	0.068
KSLP_K2E_A2F_01_MLUA	0.584	0.476	0.338	0.273	0.196	0.052
KMI-K2E-A2F-01-ORC	0.504	0.512	0.374	0.301	0.220	0.055
KMI-K2E-A2F-03-LIS	0.504	0.512	0.374	0.301	0.220	0.053
KMI-K2E-A2F-01-LIS	0.496	0.512	0.372	0.300	0.219	0.055
OKSAT-K2E-A2F-01-REF	0.464	0.340	0.248	0.209	0.162	0.042
QUT_K2E_A2F_02_LinkProbPN2	0.408	0.336	0.268	0.229	0.175	0.062
QUT_K2E_A2F_01_LinkProbPN	0.408	0.336	0.268	0.229	0.175	0.062
UKP_K2E_A2F_01_CSC	0.136	0.112	0.130	0.132	0.110	0.032
UKP_K2E_A2F_02_CSC	0.056	0.048	0.058	0.047	0.052	0.033
UKP_K2E_A2F_03_CSC	0.040	0.040	0.038	0.040	0.038	0.029

Table 24. F2F evaluation with manual assessment results: *Precision-at-N* (Chinese-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
OKSAT-C2E-A2F-01-REF	0.384	0.368	0.356	0.320	0.266	0.065
OKSAT-C2E-A2F-01-SMP	0.288	0.212	0.194	0.163	0.112	0.023
NTHU-C2E-A2F-02-lda-ext	0.192	0.136	0.118	0.121	0.126	0.078
NTHU-C2E-A2F-03-lda-ext	0.192	0.136	0.118	0.121	0.126	0.078
NTHU-C2E-A2F-01-keywordSim	0.192	0.136	0.118	0.120	0.123	0.051
QUT_C2E_A2F_01_LinkProbPN	0.176	0.152	0.144	0.135	0.147	0.108
KECIR_A2F_C2E_03_FSCLIR	0.096	0.100	0.088	0.081	0.083	0.071
UKP_C2E_A2F_01_CSC	0.088	0.124	0.150	0.169	0.158	0.065
KECIR_A2F_C2E_01_FSCLIR	0.080	0.088	0.094	0.100	0.090	0.084
QUT_C2E_A2F_02_LinkProbPN2	0.080	0.068	0.060	0.059	0.055	0.060
KECIR_A2F_C2E_02_FSCLIR	0.072	0.092	0.096	0.095	0.091	0.072
III_C2E_A2F_01_PNM	0.072	0.090	0.108	0.104	0.089	0.022
UKP_C2E_A2F_02_CSC	0.064	0.052	0.054	0.068	0.072	0.078
KECIR_A2F_C2E_04_FSCLIR	0.056	0.096	0.088	0.084	0.086	0.066
III_C2E_A2F_02_PNM	0.056	0.056	0.096	0.105	0.104	0.077
KMI-C2E-A2F-02-ORC	0.048	0.072	0.110	0.127	0.142	0.123
KMI-C2E-A2F-01-LIS	0.048	0.072	0.110	0.127	0.142	0.123
KMI-C2E-A2F-03-LIS	0.048	0.072	0.110	0.127	0.142	0.120
UKP_C2E_A2F_03_CSC	0.048	0.040	0.048	0.041	0.047	0.055
III_C2E_A2F_03_PNM	0.040	0.084	0.076	0.072	0.058	0.014
DCU-C2E-A2F-03-NW	0.024	0.016	0.024	0.023	0.021	0.019
DCU-C2E-A2F-04-NWE	0.024	0.016	0.024	0.023	0.021	0.019

Table 25. F2F evaluation with manual assessment results: *Precision-at-N* (Japanese-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
OKSAT-J2E-A2F-01-SMP	0.520	0.460	0.398	0.357	0.267	0.058
KMI-J2E-A2F-01-LIS	0.360	0.312	0.234	0.176	0.119	0.033
KMI-J2E-A2F-03-LIS	0.360	0.312	0.234	0.176	0.119	0.031
KMI-J2E-A2F-02-ORC	0.352	0.356	0.292	0.223	0.140	0.033
UKP_J2E_A2F_01_CSC	0.344	0.360	0.302	0.267	0.212	0.058
OKSAT-J2E-A2F-01-REF	0.272	0.252	0.242	0.223	0.188	0.057
RDLL_A2F_J2E_05_tfdiceLL	0.272	0.180	0.150	0.131	0.114	0.042
UKP_J2E_A2F_02_CSC	0.200	0.176	0.168	0.152	0.133	0.060
NTHU-J2E-A2F-01-keywordSim	0.184	0.164	0.138	0.133	0.123	0.049
QUT_J2E_A2F_01_LinkProbPN	0.168	0.176	0.188	0.175	0.153	0.066
QUT_J2E_A2F_02_LinkProbPN2	0.168	0.176	0.188	0.175	0.153	0.066
RDLL_A2F_J2E_03_dice	0.128	0.088	0.070	0.061	0.052	0.023
RDLL_A2F_J2E_04_tfdice	0.128	0.096	0.080	0.068	0.053	0.023
UKP_J2E_A2F_03_CSC	0.096	0.128	0.120	0.117	0.102	0.056
RDLL_A2F_J2E_02_okapiBM25	0.096	0.080	0.072	0.076	0.062	0.024
RDLL_A2F_J2E_01_tfidf	0.080	0.084	0.076	0.063	0.056	0.023

Table 26. F2F evaluation with manual assessment results: *Precision-at-N* (Korean-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
KSLP_K2E_A2F_01_MLUA	0.416	0.340	0.246	0.207	0.156	0.045
OKSAT-K2E-A2F-01-SMP	0.392	0.260	0.244	0.184	0.135	0.027
KMI-K2E-A2F-01-ORC	0.336	0.328	0.240	0.205	0.156	0.043
KMI-K2E-A2F-03-LIS	0.336	0.328	0.240	0.205	0.155	0.040
KMI-K2E-A2F-01-LIS	0.328	0.328	0.238	0.204	0.155	0.043
OKSAT-K2E-A2F-01-REF	0.320	0.224	0.174	0.149	0.111	0.032
QUT_K2E_A2F_02_LinkProbPN2	0.264	0.184	0.146	0.129	0.102	0.036
QUT_K2E_A2F_01_LinkProbPN	0.264	0.184	0.146	0.129	0.102	0.036
UKP_K2E_A2F_01_CSC	0.136	0.088	0.086	0.088	0.081	0.027
UKP_K2E_A2F_02_CSC	0.056	0.052	0.050	0.037	0.035	0.025
UKP_K2E_A2F_03_CSC	0.032	0.036	0.034	0.035	0.029	0.021

Table 27. A2F evaluation with manual assessment results: *Precision-at-N* (Chinese-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
OKSAT-C2E-A2F-01-REF	0.096	0.072	0.084	0.083	0.082	0.029
OKSAT-C2E-A2F-01-SMP	0.056	0.036	0.040	0.040	0.029	0.007
NTHU-C2E-A2F-03-lda-ext	0.040	0.024	0.026	0.025	0.029	0.013
NTHU-C2E-A2F-02-lda-ext	0.040	0.024	0.026	0.025	0.029	0.012
NTHU-C2E-A2F-01-keywordSim	0.040	0.024	0.024	0.024	0.028	0.008
QUT_C2E_A2F_01_LinkProbPN	0.032	0.048	0.034	0.041	0.048	0.052
QUT_C2E_A2F_02_LinkProbPN2	0.032	0.048	0.034	0.041	0.048	0.052
KECIR_A2F_C2E_01_FSCLIR	0.024	0.036	0.046	0.061	0.074	0.064
KECIR_A2F_C2E_02_FSCLIR	0.024	0.036	0.042	0.047	0.055	0.047
KECIR_A2F_C2E_03_FSCLIR	0.016	0.020	0.024	0.036	0.046	0.040
KECIR_A2F_C2E_04_FSCLIR	0.016	0.020	0.028	0.037	0.038	0.032
UKP_C2E_A2F_02_CSC	0.008	0.004	0.022	0.029	0.031	0.024
UKP_C2E_A2F_03_CSC	0.008	0.004	0.022	0.029	0.031	0.024
UKP_C2E_A2F_01_CSC	0.008	0.004	0.022	0.029	0.031	0.024
KMI-C2E-A2F-02-ORC	0.000	0.008	0.008	0.019	0.029	0.045
KMI-C2E-A2F-01-LIS	0.000	0.008	0.008	0.019	0.029	0.045
KMI-C2E-A2F-03-LIS	0.000	0.008	0.008	0.019	0.029	0.044
DCU-C2E-A2F-04-NWE	0.000	0.008	0.008	0.008	0.005	0.008
DCU-C2E-A2F-03-NW	0.000	0.008	0.008	0.008	0.005	0.008
III_C2E_A2F_01_PNM	0.000	0.000	0.000	0.000	0.000	0.000
III_C2E_A2F_02_PNM	0.000	0.000	0.000	0.000	0.000	0.000
III_C2E_A2F_03_PNM	0.000	0.000	0.000	0.000	0.000	0.000

Table 28. A2F evaluation with manual assessment results: *Precision-at-N* (Japanese-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
OKSAT-J2E-A2F-01-SMP	0.144	0.120	0.110	0.107	0.083	0.019
RDLL_A2F_J2E_05_tfdiceLL	0.112	0.120	0.110	0.100	0.076	0.023
UKP_J2E_A2F_01_CSC	0.080	0.112	0.114	0.097	0.082	0.026
UKP_J2E_A2F_02_CSC	0.080	0.112	0.114	0.097	0.082	0.026
UKP_J2E_A2F_03_CSC	0.080	0.112	0.114	0.097	0.082	0.026
KMI-J2E-A2F-02-ORC	0.080	0.092	0.066	0.056	0.039	0.010
NTHU-J2E-A2F-01-keywordSim	0.064	0.068	0.066	0.064	0.062	0.017
KMI-J2E-A2F-01-LIS	0.064	0.064	0.046	0.040	0.029	0.009
KMI-J2E-A2F-03-LIS	0.064	0.064	0.046	0.040	0.029	0.009
QUT_J2E_A2F_01_LinkProbPN	0.048	0.040	0.072	0.065	0.072	0.037
QUT_J2E_A2F_02_LinkProbPN2	0.048	0.040	0.072	0.065	0.072	0.037
OKSAT-J2E-A2F-01-REF	0.048	0.044	0.052	0.049	0.045	0.022
RDLL_A2F_J2E_02_okapiBM25	0.048	0.036	0.028	0.020	0.019	0.006
RDLL_A2F_J2E_01_tfidf	0.048	0.028	0.026	0.021	0.015	0.006
RDLL_A2F_J2E_03_dice	0.040	0.040	0.042	0.032	0.025	0.006
RDLL_A2F_J2E_04_tfdice	0.040	0.032	0.036	0.032	0.022	0.006

Table 29. A2F evaluation with manual assessment results: *Precision-at-N* (Korean-to-English)

Run-ID	P5	P10	P20	P30	P50	P250
KSLP_K2E_A2F_01_MLUA	0.136	0.108	0.086	0.076	0.058	0.016
KMI-K2E-A2F-01-ORC	0.096	0.084	0.060	0.055	0.051	0.019
KMI-K2E-A2F-01-LIS	0.096	0.084	0.058	0.055	0.050	0.019
KMI-K2E-A2F-03-LIS	0.096	0.084	0.060	0.055	0.051	0.018
OKSAT-K2E-A2F-01-SMP	0.056	0.044	0.044	0.043	0.031	0.009
OKSAT-K2E-A2F-01-REF	0.048	0.028	0.020	0.023	0.022	0.010
QUT_K2E_A2F_02_LinkProbPN2	0.048	0.048	0.042	0.039	0.034	0.014
QUT_K2E_A2F_01_LinkProbPN	0.048	0.048	0.042	0.039	0.034	0.014
UKP_K2E_A2F_03_CSC	0.032	0.016	0.018	0.016	0.018	0.008
UKP_K2E_A2F_01_CSC	0.032	0.016	0.018	0.016	0.018	0.008
UKP_K2E_A2F_02_CSC	0.032	0.016	0.018	0.016	0.018	0.008

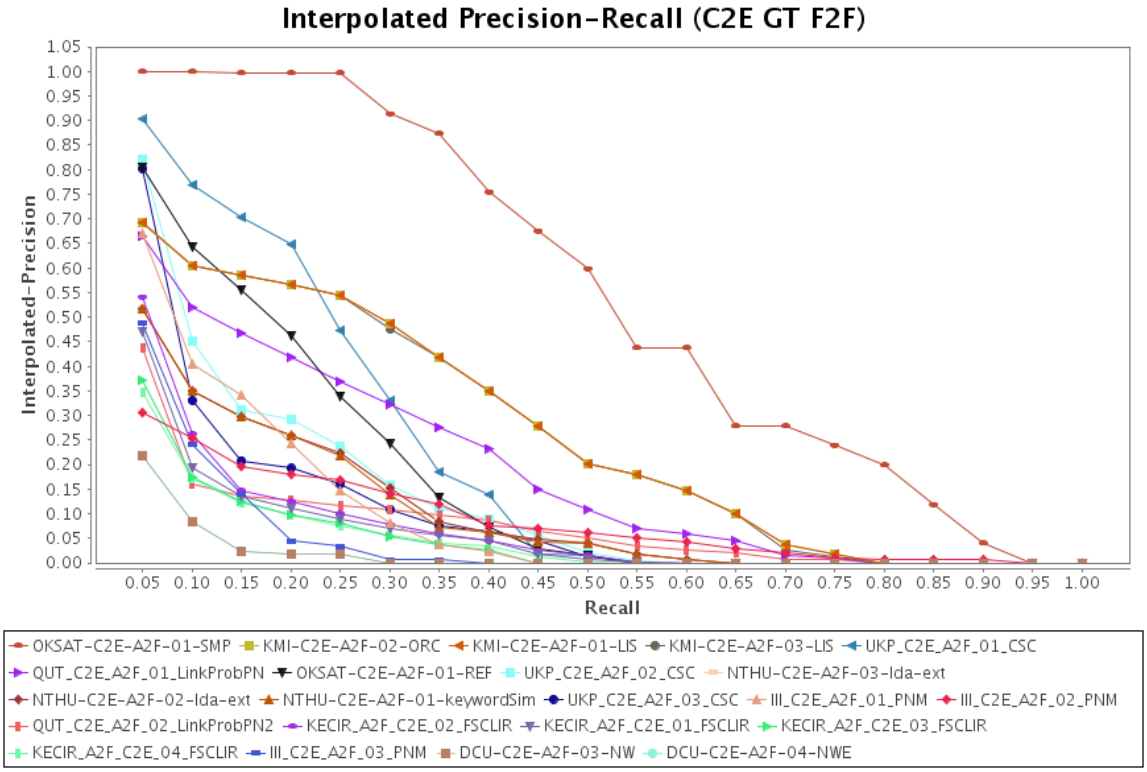


Figure 10. Interpolated Precision-Recall curves of Chinese-to-English runs in F2F evaluation with Wikipedia ground-truth

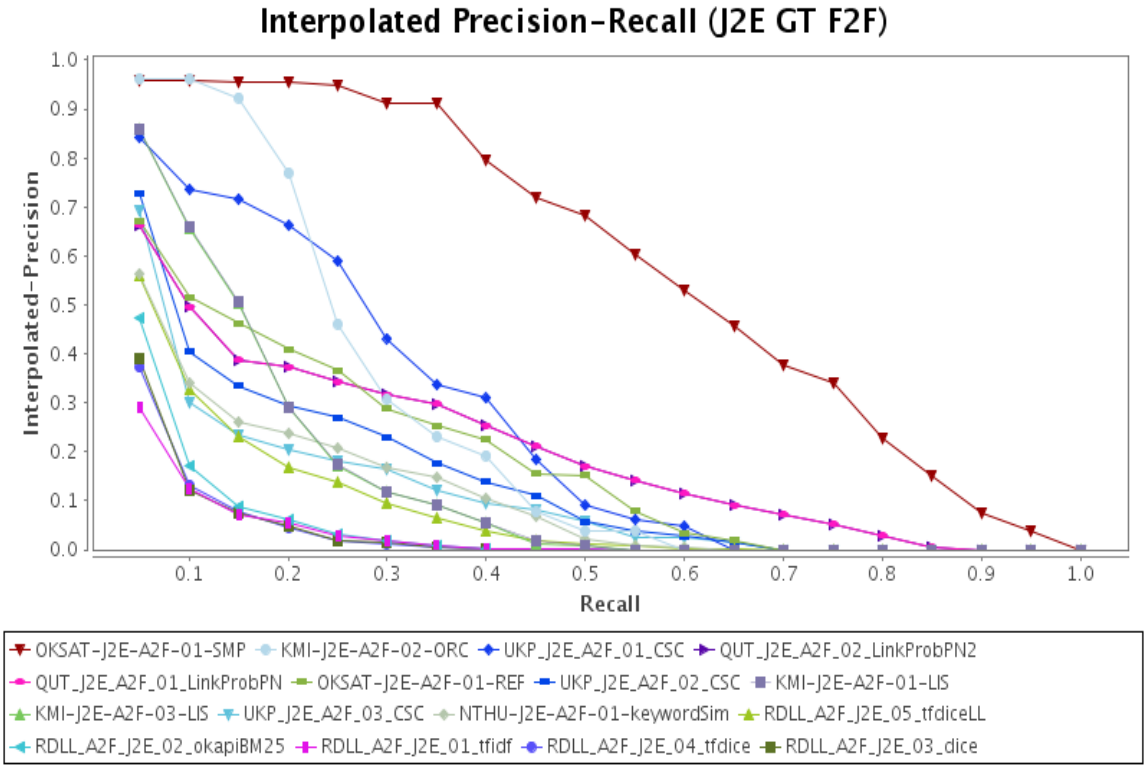


Figure 11. Interpolated Precision-Recall curves of Japanese-to-English runs in F2F evaluation with Wikipedia ground-truth

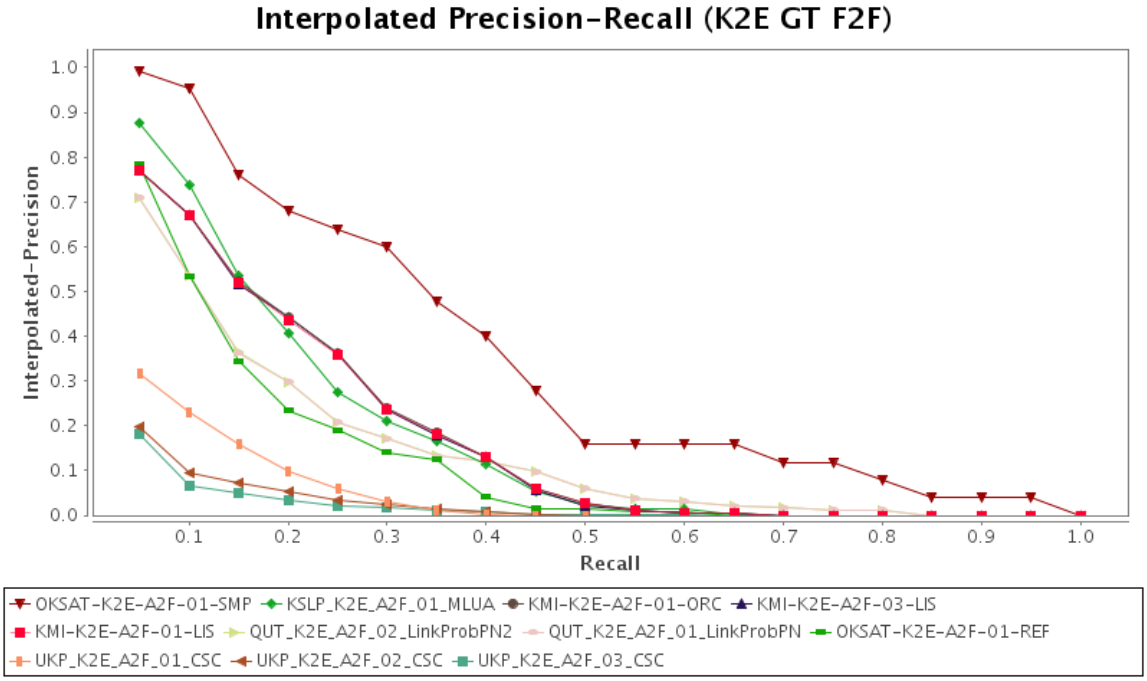
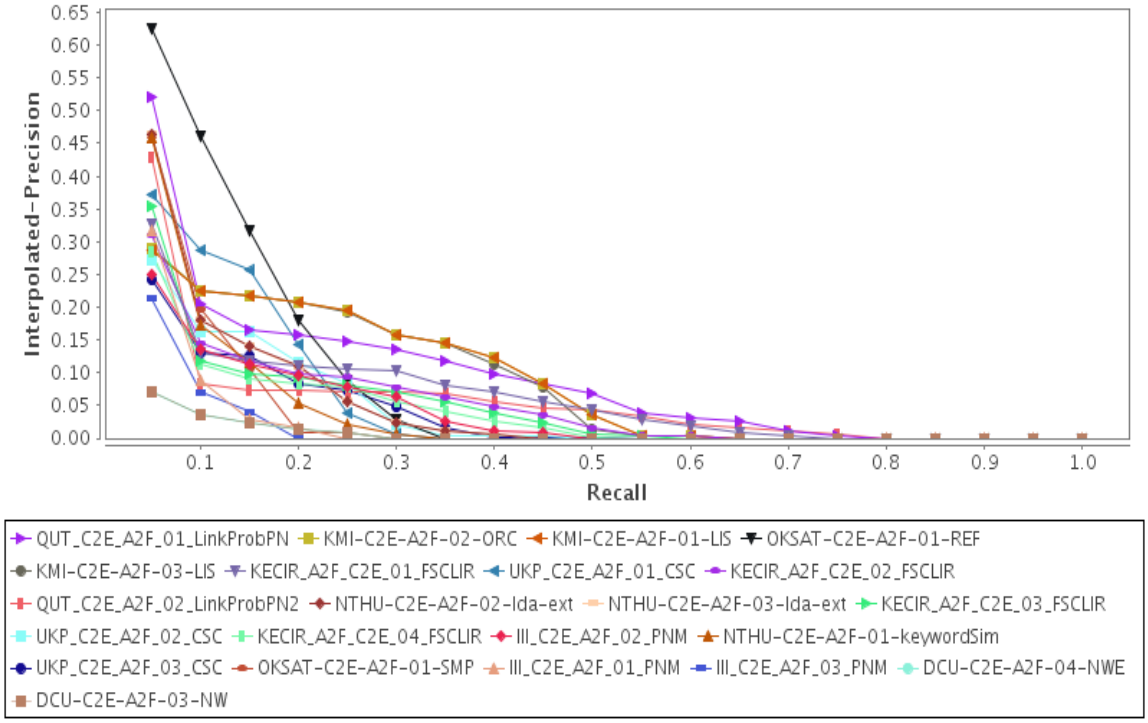


Figure 12. Interpolated Precision-Recall curves of Korean-to-English runs in F2F evaluation with Wikipedia ground-truth

a.



b.

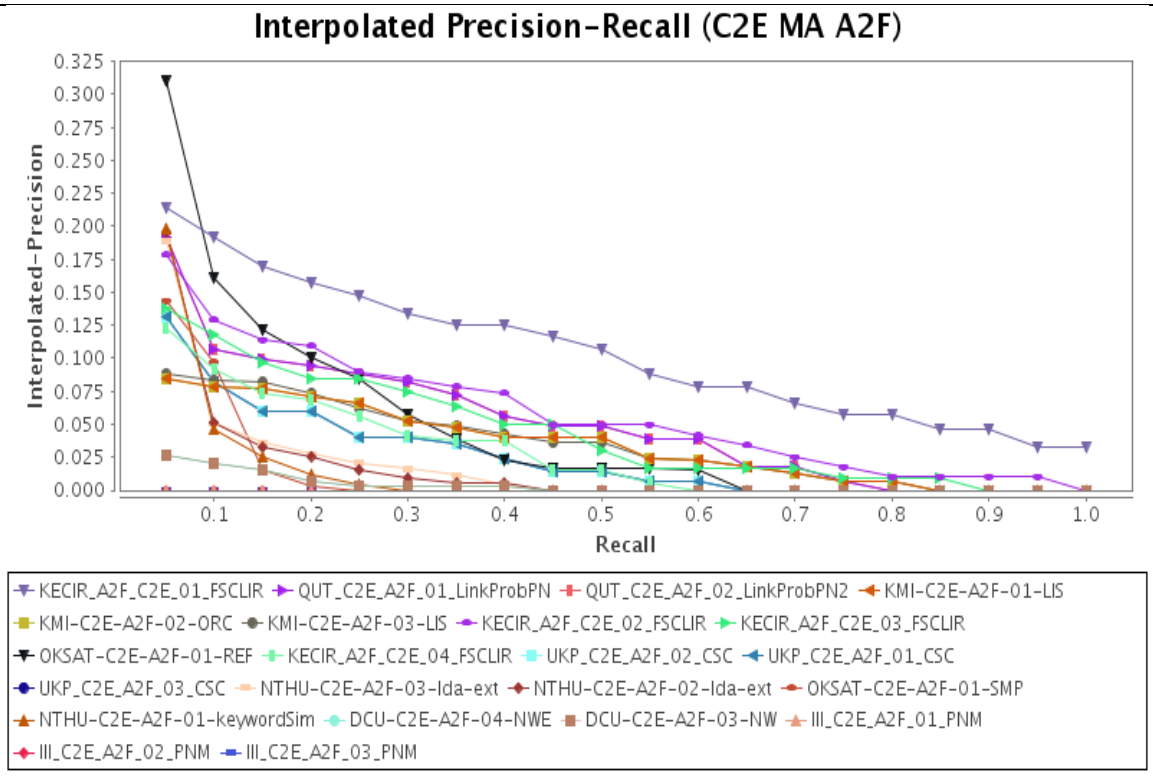
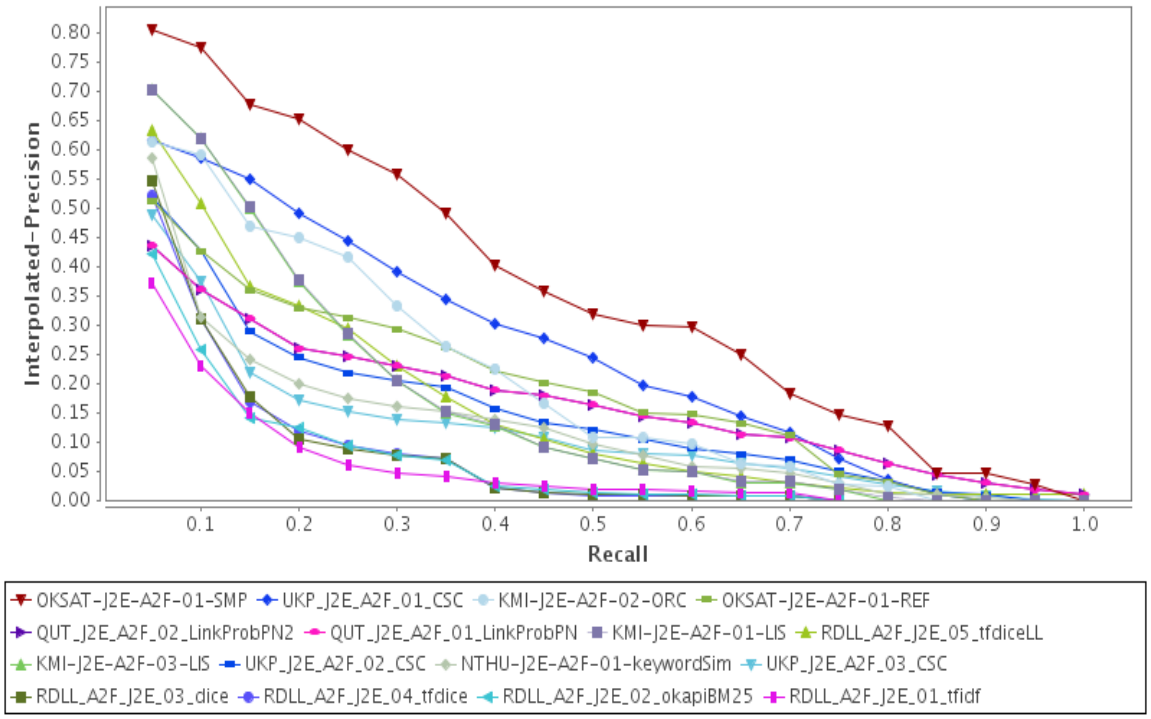


Figure 13. Interpolated Precision-Recall of Chinese-to-English runs evaluated with manual assessment results (plot *a* is P-R curves of runs evaluated in F2F level; plot *b* is P-R curves of runs evaluated in A2F level)

a.



b.

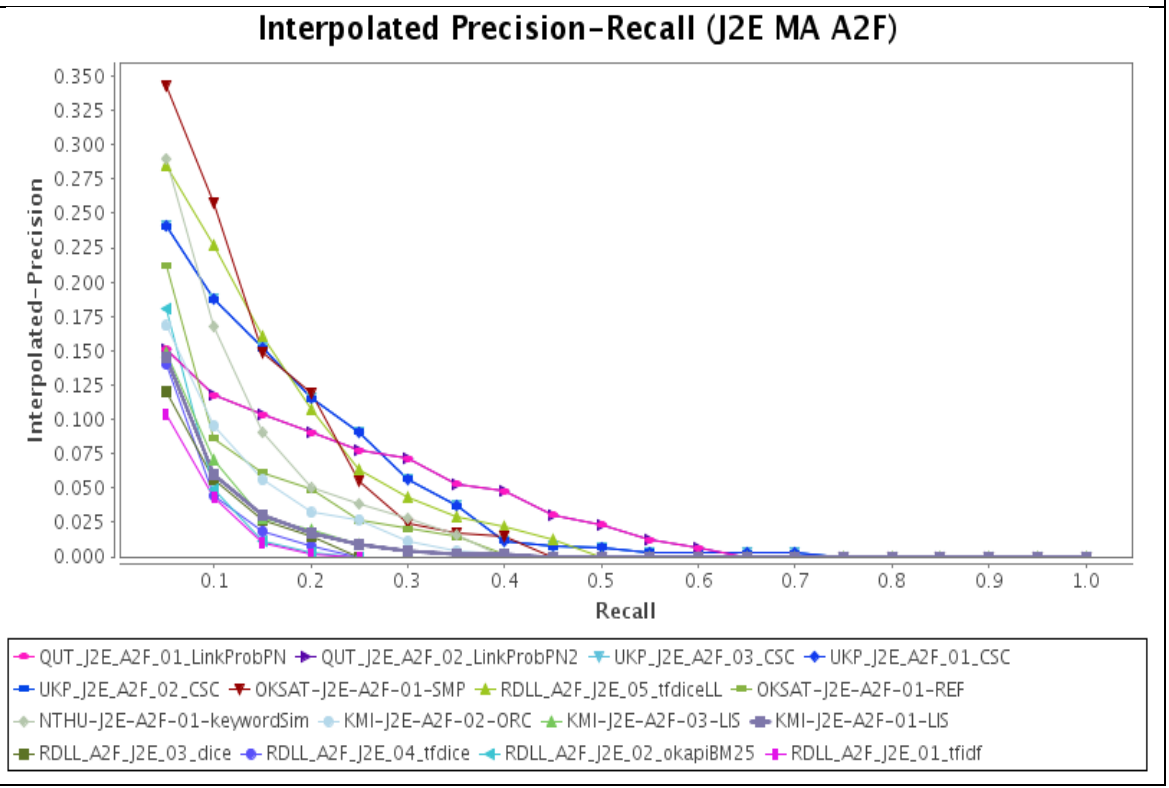
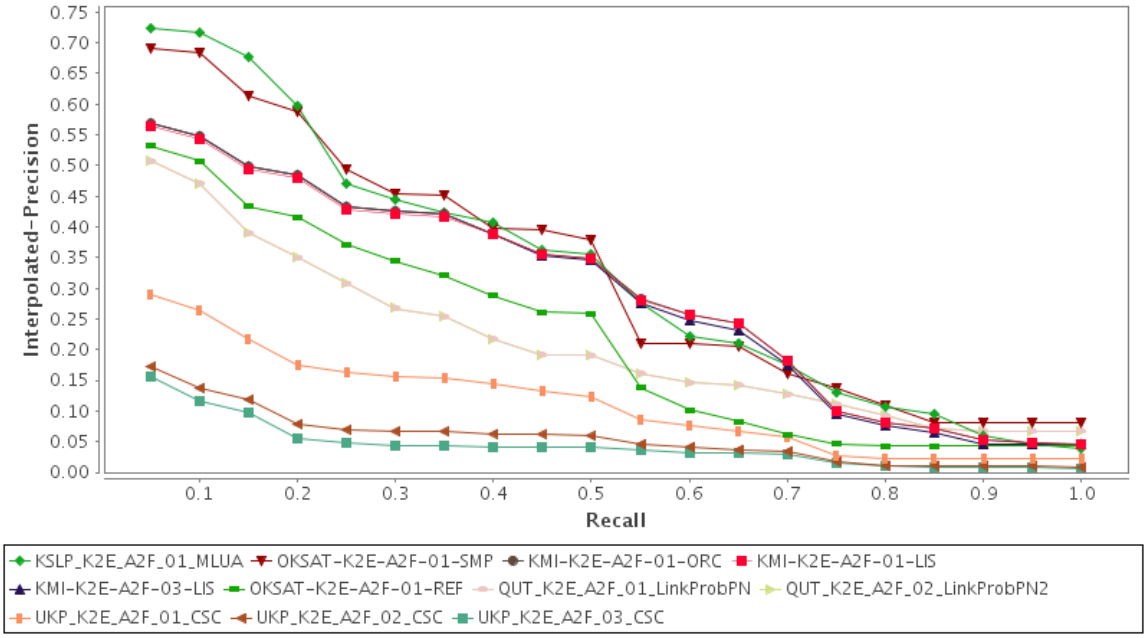


Figure 14. Interpolated Precision-Recall of Japanese-to-English runs evaluated with manual assessment results (plot *a* is P-R curves of runs evaluated in F2F level; plot *b* is P-R curves of runs evaluated in A2F level)

a.



b.

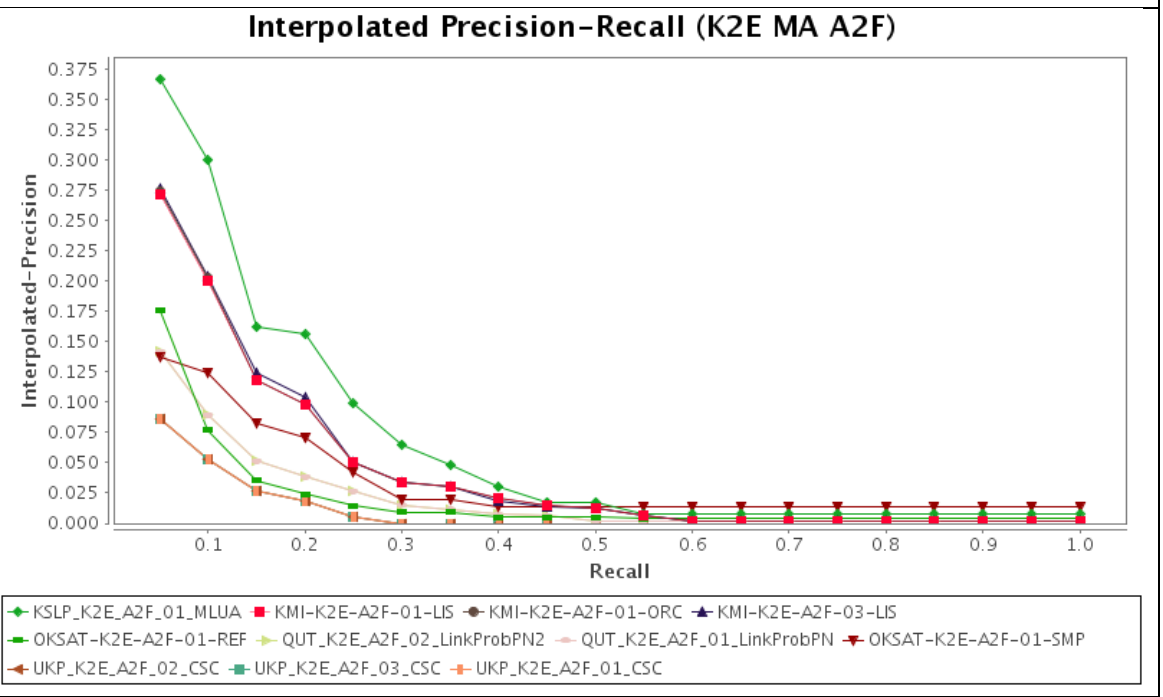


Figure 15. **Interpolated Precision-Recall of Korean-to-English links evaluated with manual assessment results** (plot *a* is P-R curves of runs evaluated in F2F level; plot *b* is P-R curves of runs evaluated in A2F level

Table 30. System descriptions of submitted runs in English-to-CJK subtasks

English-to-Chinese	
KMI-E2C-A2F-01-ESA	1. Wiki Anchor detection and anchor probability filterring 2. ESA and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to anchor probability.
KMI-E2C-A2F-02-ORC	1. Wiki Anchor detection and anchor probability filterring 2. ESA and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to a result set Oracle.
QUT_E2C_A2F_01_LinkProbPnCaseSensitive	This is a run using a link probability algorithm adopted from Dr. Kelly
III_E2C_A2F_01_PNM	Long-term match algorithm with WSD and CLLD
OKSAT-E2C-A2F-01-SMP	Simple anchor tag extraction from topics. And follow cross lingual link.
OKSAT-E2C-A2F-01-REF	Words which matches title tags are extracted. And follow cross lingual link.
English-to-Japanese	
KMI-E2J-A2F-01-ESA	1. Wiki Anchor detection and anchor probability filterring 2. ESA and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to anchor probability.
KMI-E2J-A2F-02-ORC	1. Wiki Anchor detection and anchor probability filterring 2. ESA and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to a result set Oracle.
QUT_E2J_A2F_01_LinkProbPnCaseSensitive	This is a run using a link probability algorithm adopted from Dr. Kelly
OKSAT-E2J-A2F-01-SMP	Simple anchor tag extraction from topics. And follow cross lingual link.
OKSAT-E2J-A2F-01-REF	Words which matches title tags are extracted. And follow cross lingual link.
English-to-Korean	
KMI-E2K-A2F-01-ESA	1. Wiki Anchor detection and anchor probability filterring 2. ESA and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to anchor probability.
KMI-E2K-A2F-02-ORC	1. Wiki Anchor detection and anchor probability filterring 2. ESA and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to a result set Oracle.
QUT_E2K_A2F_01_LinkProbPnCaseSensitive	This is a run using a link probability algorithm adopted from Dr. Kelly. if the number of suggested by the link mining method is less than 250, it will be filled up to 250 through name matching with the algoirthm by Shlomo.
OKSAT-E2K-A2F-01-SMP	Simple anchor tag extraction from topics. And follow cross lingual link.
OKSAT-E2K-A2F-01-REF	Words which matches title tags are extracted. And follow cross lingual link.

Table 31. System descriptions of submitted runs in CJK-to-English subtasks

Chinese-to-English	
III_C2E_A2F_01_PNM	CRF++ -> WSD -> CLLD
III_C2E_A2F_02_PNM	Greedy Algorithm
DCU-C2E-A2F-03-NW	Using N-GRAM to do anchor selection. Wikipedia Miner is used for anchor disambiguration and re-rank. Apache Lucence Software indexes corpora.
DCU-C2E-A2F-04-NWE	Using N-GRAM to do anchor selection. Wikipedia Miner is used for anchor disambiguration and re-rank. Apache Lucence Software indexes corpora.
KECIR_A2F_C2E_01_FSCLIR	It is our first run based on the algorithms of feature selection and cross-lingual information retrieval.
KECIR_A2F_C2E_02_FSCLIR	It is our second run based on the algorithms of feature selection and cross-lingual information retrieval.

KECIR_A2F_C2E_03_FSCLIR	It is our third run based on the algorithms of feature selection and cross-lingual information retrieval.
KECIR_A2F_C2E_04_FSCLIR	It is our fourth run based on the algorithms of feature selection and cross-lingual information retrieval.
KMI-C2E-A2F-01-LIS	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to anchor probability.
KMI-C2E-A2F-02-ORC	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to a result set Oracle.
KMI-C2E-A2F-03-LIS	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Sorting acording to anchor probability.
QUT_C2E_A2F_01_LinkProbPN	This is a run using a link probability algorithm adopted from Dr. Kelly. if the number of suggested by the link mining method is less than 250, it will be filled up to 250 through name matching with the algoirthm by Shlomo.
QUT_C2E_A2F_02_LinkProbPN2	This is a run using a link probability algorithm adopted from Dr. Kelly. if the number of suggested by the link mining method is less than 250, it will be filled up to 250 through name matching with the algoirthm by Shlomo. It is the same with LinkProbPN run, plus fill in all positions with Atire search
UKP_C2E_A2F_01_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
UKP_C2E_A2F_02_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
UKP_C2E_A2F_03_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
III_C2E_A2F_03_PNM	CRF++ WSD CLLD
NTHU-C2E-A2F-01-keywordSim	Use Dice's Coefficient and all context mention matching to compute keyword similarity (gk=0.61, sim=0.025)
NTHU-C2E-A2F-02-lda-ext	Extend the result with machine translation and LDA similarity
NTHU-C2E-A2F-03-lda-ext	Extend the result with machine translation and LDA similarity (Sim. and Tra. Chinese bug fixed)
OKSAT-C2E-A2F-01-REF	Words which matches title tags are extracted. And follow cross lingual link.
OKSAT-C2E-A2F-01-SMP	Simple anchor tag extraction from topics. And follow cross lingual link.
Japanese-to-English	
KMI-J2E-A2F-01-LIS	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to anchor probability.
KMI-J2E-A2F-02-ORC	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to a result set Oracle.
KMI-J2E-A2F-03-LIS	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Sorting acording to anchor probability.
QUT_J2E_A2F_01_LinkProbPN	This is a run using a link probability algorithm adopted from Dr. Kelly. if the number of suggested by the link mining method is less than 250, it will be filled up to 250 through name matching with the algoirthm by Shlomo.
QUT_J2E_A2F_02_LinkProbPN2	This is a run using a link probability algorithm adopted from Dr. Kelly. if the number of suggested by the link mining method is less than 250, it will be filled up to 250 through name matching with the algoirthm by Shlomo. It is the same with LinkProbPN run, plus fill in all positions with Atire search

UKP_J2E_A2F_01_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
UKP_J2E_A2F_02_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
UKP_J2E_A2F_03_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
OKSAT-J2E-A2F-01-REF	Words which matches title tags are extracted. And follow cross lingual link.
NTHU-J2E-A2F-01-keywordSim	Use Dice's Coefficient and all context mention matching to compute keyword similarity (gk=0.61, sim=0.025)
RDLL_A2F_J2E_02_okapiBM25	This RUN uses the Okapi BM25 in order to extract the anchor Translation of the anchor is using Edic and Bing Translator
RDLL_A2F_J2E_03_dice	This RUN uses the dice coefficient in order to extract the anchor Translation of the anchor is using Edic and Bing Translator
OKSAT-J2E-A2F-01-SMP	Simple anchor tag extraction from topics. And follow cross lingual link.
RDLL_A2F_J2E_01_tfidf	This RUN uses the TFIDF in order to extract the anchor Translation of the anchor is using Edic and Bing Translator
RDLL_A2F_J2E_04_tfdice	This RUN uses the dice coefficient and the TF in order to extract the anchor Translation of the anchor is using Edic and Bing Translator
RDLL_A2F_J2E_05_tfdiceLL	This RUN uses the dice coefficient and the TF in order to extract the anchor Translation of the anchor is using Edic,Bing Translator and second language links
Korean-to-English	
KMI-K2E-A2F-01-LIS	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to anchor probability.
KMI-K2E-A2F-01-ORC	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Inclusion of additional low scoring concept disambiguations 6. Sorting acorrding to a result set Oracle.
KMI-K2E-A2F-03-LIS	1. Wiki Anchor detection and anchor probability filterring 2. Link similarity and concept probabilty disambiguation 4. Max concept translation 5. Sorting acording to anchor probability.
KSLP_K2E_A2F_01_MLUA	ML-based Unified Approach
QUT_K2E_A2F_01_LinkProbPN	This is a run using a link probability algorithm adopted from Dr. Kelly. if the number of suggested by the link mining method is less than 250, it will be filled up to 250 through name matching with the algoirthm by Shlomo.
QUT_K2E_A2F_02_LinkProbPN2	This is a run using a link probability algorithm adopted from Dr. Kelly. if the number of suggested by the link mining method is less than 250, it will be filled up to 250 through name matching with the algoirthm by Shlomo. It is the same with LinkProbPN run, plus fill in all positions with Atire search
UKP_K2E_A2F_01_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
UKP_K2E_A2F_02_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
UKP_K2E_A2F_03_CSC	Anchor Probability + Anchor Translation (Cascaded) + Target Discovery (Cascaded) (this official run is from an incomplete system with bugs)
OKSAT-K2E-A2F-01-REF	Words which matches title tags are extracted. And follow cross lingual link.
OKSAT-K2E-A2F-01-SMP	Simple anchor tag extraction from topics. And follow cross lingual link.