AINLP at NTCIR-6: Evaluations for Multilingual and Cross-Lingual Information Retrieval

Chen-Hsin Cheng Reuy-Jye Shue Hung-Lin Lee
Shu-Yu Hsieh Guann-Cyun Yeh Guo-Wei Bian

Department of Information Management
Huafan University, Taiwan, R.O.C.
{m9445121, m9445123, m9445125, m9445130, m9445127}@cat.hfu.edu.tw,
gwbian@cc.hfu.edu.tw

Abstract

In this paper, a multilingual cross-lingual information retrieval (CLIR) system is presented and evaluated in NTCIR-6 project. We use the language-independent indexing technology to process the text collections of Chinese, Japanese, Korean, and English languages. Different machine translation systems are used to translate the queries for bilingual and multilingual CLIR. The experimental results are discussed to analyze the performances of our system. The effectiveness of query translations for bilingual and multilingual CLIR is discussed. In the evaluations, the English version of topics performed better CLIR results to retrieve the Korean text collections than the Chinese version did. However, the Chinese version of topics performed better cross-language information retrieval results to retrieve the Japanese text collections than the English version did.

Keywords: *Cross-Language Information Retrieval, Multilingual Information Retrieval, Bigram.*

1 Introduction

Cross language information retrieval (CLIR) deals with the use of queries in one language to access documents in another. Due to the differences between source and target languages, query translation is usually employed to unify the language in queries and documents. Some different approaches have been proposed for query translation. Dictionary-based approach [1] exploits machine-readable dictionaries and selection strategies like select all, randomly select N and select best N. Corpus-based approaches exploit sentence-aligned corpora and document-aligned corpora. These two approaches are complementary. Dictionary provides translation candidates, and corpus provides context to fit user intention. Coverage of dictionaries, alignment performance and domain shift of corpus are major problems of these two approaches. Hybrid approaches [2, 3, 4, 5] integrate both lexical and corpus knowledge. A synset-based approach [6] is proposed to use an automatically constructed English-Chinese WordNet for Chinese-English information retrieval.

This paper discusses our participation in the Cross-Lingual Information Retrieval (CLIR) task at NTCIR-6 [14]. We participated in monolingual information retrieval (SLIR), bilingual information retrieval (BLIR) and multilingual information retrieval (MLIR) subtasks of the NTCIR-6 CLIR task. Our main goal is to develop a CLIR system which can handle as many languages as possible even with limited resources for query translations. Our system can handle the documents in four languages included Chinese(C), Japanese(J), Korean(K), and English(E) and the multilingual (CJKE) text collections. Since the Asian languages have the different morpheme schemes, different word segmentation systems are used for Chinese, Japanese, and Korean language processing [7, 8, 9, 10, 11, 12, 13, 16].

For CLIR, our system can process queries in Chinese, Japanese, Korean, and English. We submitted the search results for the following combinations in NTCIR-6 CLIR task.

- SLIR: C -> C
- BLIR: C -> J, C -> K
- BLIR: E -> C, E -> J, E -> K
- MLIR: C -> CJK
- MLIR: E -> CJK

As a first-time participant at NTCIR, we focused on the effectiveness of query translations with different machine translation systems for bilingual and multilingual cross-language information retrieval. Our main aims for participating in the BLIR and MLIR tasks are as follows:

• Study the effectiveness of bigram indexing method for Chinese, Japanese, and Korean.

• Study the effectiveness of CLIR using different machine translation (MT) systems.
• Study the effectiveness of Multilingual CLIR (E-CJK and C-CJK).

This paper is organized as follows. Section 2 describes the process of our CLIR system. Section 3 presents the experiments and the evaluation results. Finally Section 4 concludes the remarks.

2 System Description

The system uses bigram-based indexing for Chinese, Japanese, and Korean text collections. Several machine translation systems are used to translate the source languages to target languages. Language model is used for retrieval document scoring, and the pseudo-relevance feedback is used for query expansion. In multilingual IR, the results of SLIR and BLIR for the same query are merged to obtain the retrieval results. For example, Figure 1 shows the processing of Chinese-Japanese cross-language information retrieval.

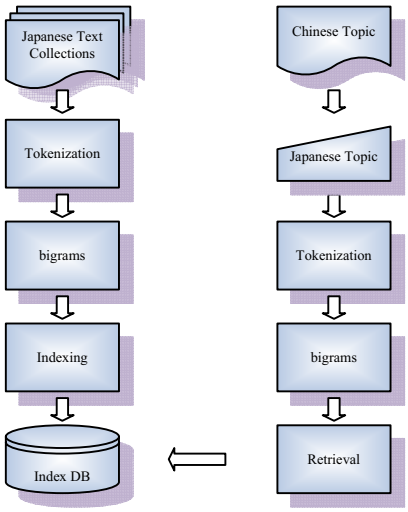

Figure 1. The processing of Chinese-Japanese CLIR

As a newly established research group, we adapted one of the available open source information retrieval systems for our researches. Lemur [15] and Lucene become the candidates for IR search engines. We used the Lemur toolkit developed by the Computer Science Department at the University of Massachusetts and the School of Computer Science at Carnegie Mellon University. There are several reasons to adopt the Lemur toolkit, including:

(1) It supports large-scale text collections to index and retrieve.

(2) This system is designed as a research system, and it accepts the TREC document format. It is very convenient for the TREC-type information retrieval experiments.

(3) It supports the UNICODE coding and UTF-8 document format, which used for the multilingual text collections.

(4) The source codes of the toolkit are developed in C and C++, and supported for different operating systems included UNIX, Linux, and Windows.

2.1 Tokenization

The first task for Chinese, Japanese, and Korean information retrieval, is text segmentation since there are no word boundary in Chinese, Japanese, and Korean texts. The bi-gram text segmentation and word segmentation have been widely used to parse the tokens and words of text collections.

Because the Asian languages have the different morpheme schemes, different word segmentation systems are needed for Chinese, Japanese, and Korean language processing. We adopt the language independent technique of character bigram. The indexing unit is a pair of adjacent characters. For example, the string “漢代文物大展” is indexed as the five tokens “漢代”, “代文”, “文物”, “物大”, and “大展”.

In information retrieval, the punctuation marks and special characters are generally meaningless. Therefore, the system filters out these symbols before indexing and retrieval tasks. Because Chinese, Japanese, and Korean used double-byte language coding, these symbols could be represented in ASCII or in different double-byte codes of these languages. After tokenization, the Lemur toolkit is used to index the document collections.

2.2 Query Processing and Translation

In the monolingual information retrieval, the query is generated from the selected field(s) of the original topic and then parsed as the stream of bigrams. In bilingual and multilingual information retrieval, the topics in source languages are first translated to target languages using different machine translation systems. The Internet Passport MT system is used for Chinese-Japanese, Chinese-Korean, and English-Chinese query translations. The online WorldLingo MT system is used for the English-Japanese, and English–Korean query translations. Because of the coverage of bilingual lexicons, some worlds (e.g. E-Commerce and Nanotechnology) cannot be translated to target languages in these machine translation systems.

3 Experiments

We participated in the STAGE1 and STAGE2 of the NTCIR-6 CLIR task. Our CLIR retrieval experiments consist of the SLIR, BLIR, and MLIR tasks.

3.1 Test Collection

The document sets for STAGE1 and STAGE2 of the NTCIR-6 CLIR task consisted of news articles from 2000 to 2001 in Traditional Chinese, Japanese, and Korean. Table 1 shows the sizes and the numbers of documents for the collections. Figure 2, 3, and 4 are the sample documents for Chinese, Japanese, and Korean. The language of each document is indicated in the <LANG> field.

Table 1. Document sets for STAGE1 and STAGE2 of the NTCIR-6 CLIR Task

Language	Size (in MB)	No. of Documents
Chinese	1113.5	901,446
Japanese	1078.2	858,400
Korean	333.3	220,374

<DOC>
<DOCNO>udn_xxx_20011231_1256511</DOCNO>
<LANG>CH</LANG>
<HEADLINE>傳南部民代兩千萬買回元碟</HEADLINE>
<DATE>2001-12-31</DATE>
<TEXT>
<P>記者林益民／台北報導</P>
<P>遠美鳳被偷拍事件中，外傳被偷拍的版本不只獨家報導所曝光的版本，目前檢警積極追查這些版本之中，有一版本傳被南部某重量級民代以兩千萬元贖回。遠美鳳被偷拍案，隨著檢警偵辦腳步，越來越受矚目。近來道上傳出，流傳在外的版本之一，已被一名南部重量級民代買走，這名民代因是片中男主角，為顧及聲譽而出價買回該版本的母帶。據透露，向其兜售偷拍母帶並索取高價的人，對這名民代的財經實力及地方努力非常了解，整個交易過程非常謹慎，不讓這名民代有報復的機會，這名民代所花的代價並非外傳的六百萬元，而是兩千萬元。</P>
</TEXT>
</DOC>

Figure 2. A sample Chinese document of NTCIR-6

<DOC>
<DOCNO>JA-001231147</DOCNO>
<LANG>JA</LANG>
<SECTION>特集</SECTION>
<AE>無</AE>
<WORDS>538</WORDS>
<HEADLINE>〔女の気持ち〕「女の気持ち」と３０年　滋賀県栗東町　伊勢田恵子・５９歳　主婦　【大阪】</HEADLINE>
<DATE>2000-12-31</DATE>
<TEXT>
　朝食の片付けを終えると熱いコーヒーを前に、まず「女の気持ち」を読む。３０年間、この習慣は変わっていない。
　このコラムがスタートしたのは１９５４年秋だから新世紀には“４７歳”になる。小さなコラムだけれど、一冊の本より重く感じられた日もあった。女性たちのかぼそい声でも数十年の声が集まると大合唱になる。
　８７年には二十数年の間に掲載された作品の中から「戦争・平和編」や「子ども・教育編」など全５巻「『女の気持ち』３０年」（新評論刊）として出版された。全巻を通読したとき、これらは２０世紀がどのような１００年であったかを証する歴史ともなっているのだと思った。
　会ったこともない人たちだけれど、草の根にも似る女性たちは私の仲間なのである。私も「女の気持ち」に投稿を始めて１０年余りになる。大半は没となるが、書いている間に気持ちの整理がついて清書するころにはすっかり落ち着いていることが多い。書くことは癒（いや）しになるようだ。
　これからも、読んでいると美しい風景が浮かんできたり、心がほかほかしてくるような作品に出合いたい。ユーモアにあふれる作品ならもっとうれしい。毎日新聞がある限り、このコラムも定位置にあることだろう。
　新年と「女の気持ち」に乾杯しよう！
</TEXT>
</DOC>

Figure 3. A sample Japanese document of NTCIR-6

<DOC>
<DOCNO>HKIB2000_40303</DOCNO>
<LANG>KR</LANG>
<HEADLINE> 한국일보 : 신경영기법 '스피드 카드제' 도입 </HEADLINE>
<DATE>20000131</DATE>
<TEXT>
LG 텔레콤이 신경영 혁신기법 중 하나인 '스피드카드제'를 도입해 업계의 화제가 되고 있다. 스피드카드제는 조직이나 다른 팀원 또는 본인 때문에 업무의 진도가 지연될 경우 스피드카드를 제시하면 지위고하를 막론하고 업무에 적극 협조하는 제도.
신속한 의사결정으로 업무효율을 높인다는 평가를 받고 있다. LG 텔레콤은 스피드카드제를 전 사업장으로 확대할 계획이다.
</TEXT>
</DOC>

Figure 4. A sample Korean document of NTCIR-6

3.2 Tokenization and Indexing

Table 2 shows the sources, the number of documents, the number of bigram tokens, and the size of bigram for the STAGE1. The document collection consisted of the news articles from various news agencies. Table 3 shows the sources, the number of documents, the number of bigram tokens, and the size of bigram for different topic sets of STAGE2.

Table 2. The Statistics of Document Collection for STAGE1 and STAGE2

Sources	No. of Docs	No. of Bigram	Size of Bigram
Chinese 2000-01	901,446	385,901,067	2,080.6 MB
Japanese 2000-01	858,400	388,357,968	2,231.5 MB
Korean 2000-01	220,374	78,993,015	495.6 MB

Table 3. The Statistics of Document Collection for STAGE2

(a) For NTCIR-5 Topics Sets
See Table 2.

(b) For NTCIR-4 Topics Sets

Sources	No. of Docs	No. of Bigram	Size of Bigram
Chinese 1998-99	381,681	190,424,639	1028.1 MB
Japanese 1998-99	596,058	299,222,825	1704.4 MB
Korean 1998-99	254,438	95,273,991	616.2 MB

(c) For NTCIR-3 Topics Sets

Sources	No. of Docs	No. of Bigram	Size of Bigram
Chinese 1998-99	381,681	190,424,639	1028.1 MB
Japanese 1998-99	220,078	110,103,227	623.7 MB
Korean 1994	66,146	19,335,891	77.5 MB

3.3 Queries

We participated the SLIR, BLIR, and MLIR tasks for the multilingual cross-language information retrieval. The Chinese and English versions of the topics are used for BLIR tasks (E-C, E-J, E-K, C-J, C-K) and MLIR tasks (E-CJK, and C-CJK). Figure 5 lists the Chinese and English versions of the topic 004.

```
<TOPIC>
<NUM>004</NUM>
<ONUM>NTCIR4-004</ONUM>
<SLANG>CH</SLANG>
<TLANG>CH</TLANG>
<TITLE>花蝴蝶，葛瑞菲絲</TITLE>
<DESC>查詢花蝴蝶葛瑞菲絲的相關介紹</DESC>
<NARR>
<BACK>有「花蝴蝶」和「世界上跑得最快的女人」之稱的美國田徑女將葛瑞菲絲．喬納，於 1998 年九月二十一日因心臟病突發死亡。她曾經在 1988 年的漢城奧運會以十秒四九跑完百米打破世界紀錄。請查詢花蝴蝶葛瑞菲絲體壇的重要記事，包括得獎記錄，以及禁藥疑雲之相關報導。</BACK>
<REL>相關資料為葛瑞菲絲體育生涯的重要大事，如得獎記錄，1988 年漢城奧運的禁藥風波等報導。僅報導各界對於她的死亡所發表的看法及追悼活動則為無關。
</REL>
</NARR>
<CONC>花蝴蝶，田徑女將，葛瑞菲絲，禁藥，漢城奧運，得獎記錄</CONC>
</TOPIC>
```

(a) Chinese version

```
<TOPIC>
<NUM>004</NUM>
<ONUM>NTCIR4-004</ONUM>
<SLANG>CH</SLANG>
<TLANG>EN</TLANG>
<TITLE>Florence Griffith Joyner, FloJo</TITLE>
<DESC>Find articles introducing Florence Griffith Joyner</DESC>
<NARR>
<BACK>Florence Griffith Joyner, known as "FloJo," the world's fastest woman, died of a heart attack on Sep. 21st, 1998. She broke the 100 meter dash world record with a time of 10.49 seconds in the Trials for the 1988 Seoul Olympic. Please query important athletic events of "FloJo", Florence Griffith Joyner, including reports about her awards and suspicion of illegal drug use.</BACK>
<REL>Documents about important events of Florence Griffith Joyner's athletic career such as awards and reports about suspicion of illegal drug use at the 1988 Olympic in Seoul, South Korea are relevant. Documents containing only public opinions on her death and memorial activities are not relevant.</REL>
</NARR>
<CONC>FloJo, Track and Field Female Athlete, Florence Griffith Joyner, Illegal Drugs, Olympic in Seoul, Awards</CONC>
</TOPIC>
```

(b) English version

Figure 5. Chinese and English versions of the topic 004

Two different queries are derived from the same topic to compare the retrieval performance.

T-run: the short query from the topic's title, i.e., the content of the <title> field;

D-run: the long query from the topic's description, i.e., the content of the <desc> field

Table 4. Examples of Chinese Queries

Original Query	Bigram
漢代文物大展	漢代 代文 文物 物大 大展
台灣加入 WTO	台灣 灣加 加入 WTO
電子商務，網路，虛擬，交易，購物，電腦	電子 子商 商務 網路 虛擬 交易 購物 電腦
秋鬥，訴求，勞工，抗議，台灣	秋鬥 訴求 勞工 抗議 台灣

In monolingual information retrieval (SLIR), the queries are parsed to generate the bigram patterns for retrieving the relevant documents. Table 4 shows the results of some examples of Chinese queries.

3.4 Results and Discussion

Experimental results are retrieved using the Okapi model with pseudo relevance feedback. Because of the first participation and the coding issues of the text collections in our experiment included three different languages (Chinese, Japanese, and Korean), we spent lots of time to solve the problem of language coding and translate the queries for BLIR tasks (E-C, E-J, E-K, C-J, C-K) and MLIR tasks (E-CJK, and C-CJK). For STAGE1, only two runs are obtained for Chinese-Japanese CLIR. The results of our experiments are shown in Table 5. The relevance judgments provided by NTCIR are at two levels: rigid relevance and relax relevance, the former is strictly relevant but the last is likely relevant.

Table 5. Official evaluation results of STAGE1

AINLP runs	C-J-T-01	C-J-D-02
Relax Judgment (MAP)	0.1270	0.0851
Rigid Judgment (MAP)	0.1051	0.0619

For STAGE2, some tools have been developed to perform more runs of the BLIR and CLIR tasks. The official evaluation results of STAGE2 are shown in Table 6. In our experiments, 8 runs are submitted for NTCIR-6 N3 topics, 14 runs are submitted for NTCIR-6 N4 topics, and 8 runs are submitted for NTCIR-6 N5 topics. In order to evaluate the MLIR, our experiments obtained the SLIR results first and then the results of BLIR tasks. For example, 2 runs are performed for Chinese SLIR of the N4 topics. For C-J and C-K BLIR, 4 runs are performed. The results of C-C, C-J, and C-K runs are merged to obtain the retrieval results of MLIR (C-CJK) task. The raw-score merging strategy is used to sort the multilingual results by their original similarity scores.

The Internet Passport MT system is used for the bilingual Chinese-Japanese, Chinese-Korean, and English-Chinese query translations. The online WorldLingo MT system is used for the English-Japanese and English–Korean query translations. From the viewpoint of cross-language information retrieval, WorldLingo system performed better English–Korean translation than the Chinese–Korean translation using the Internet Passport MT system. Especially in N5 topics, the performances of English-Korean BLIR using WorldLingo MT system is twice of the ones of Chinese-Korean BLIR using Internet Passport MT system. But the Internet Passport MT system performed better Chinese–Japanese translation than the English-Japanese translation using WorldLingo MT system. Comparing the results of SLIR, the differences of the performances of the

short queries (T-runs) and the long queries (D-runs) are not significant for bigram indexing. Our experiments have the better performances in the C-C, E-C, C-K, C-CJK, E-K, E-CJK tasks. Because of the coverage of bilingual lexicons in the MT systems, the translations of unknown words introduced the problems in BLIR and MLIR.

4 Conclusion

In this paper, we discuss the effectiveness of query translations with different machine translation systems for bilingual and multilingual cross-language information retrieval. The language-independent technology - bigram indexing method, is used to process the text collections of various languages. In the experimental results, we can find that the English version of topics performed better cross-language information retrieval results to retrieve the Korean text collections than the Chinese version did. However, the Chinese version of topics performed better cross-language information retrieval results to retrieve the Japanese text collections than the English version did. In the future, we will involve combining the word-based indexing methods, the dictionary-based query translations, and the translation disambiguation using co-occurrence relationships to improve our multilingual (E-CJK and C-CJK) cross-language information retrieval system.

Reference

[1] Ballesteros, L. and Croft, W.B. "Dictionary-based Methods for Cross-Lingual Information Retrieval." Proceedings of the 7th International DEXA Conference on Database and Expert Systems Applications, pp.791-801, 1996.

[2] Ballesteros, L. and Croft, W.B. "Resolving Ambiguity for Cross-Language Retrieval." Proceedings of 21st ACM SIGIR , pp.64-71, 1998.

[3] Bian, G.W. and Chen, H.H. "Integrating Query Translation and Document Translation in a Cross-Language Information Retrieval System." Machine Translation and the Information Soap (AMTA '98), D. Farwell, L Gerber, and E. Hovy (Eds.), Lecture Notes in Computer Science, Vol. 1529, Springer-Verlag, pp. 250-265, 1998.

[4] Bian, G.W. and Chen, H.H. "Cross language information access to multilingual collections on the Internet." Journal of American Society for Information Science, 51(3), pp. 281-296, 2000.

[5] Chen, H.H.; Bian, G.W.; and Lin, W.C. "Resolving translation ambiguity and target polysemy in cross-language information retrieval." Proceedings of 37th Annual Meeting of Association for Computational Linguistics, pp.215-222, 1999.

[6] Chen, H.H.; Lin, C.C.; and Lin, W.C. "Construction of a Chinese-English WordNet and Its Application to CLIR" Proceedings of 5th International Workshop on Information Retrieval with Asian Languages, Hong Kong, pp. 189-196, 2000.

[7] Chen, J.; Li, R.; and Li, F. "Chinese Information Retrieval Using Lemur: NTCIR-5 CIR Experiments at UNT", Proceedings of NTCIR-5, 2005

[8] Gey, F.C. "How similar are Chinese and Japanese for Cross-Language Information Retrieval?", Proceedings of NTCIR-5, 2005.

[9] Kamps, J.; Bruggen, M.; and Rijke, M. "The University of Amsterdam at NTCIR-5", Proceedings of NTCIR-5, 2005.

[10] Kwok, Kui-Lam; Choi, Sora; Dinstl, Norbert; and Deng, Peter. "NTCIR-5 Chinese, English, Korean Cross Language Retrieval Experiments using PIRCS", Proceedings of NTCIR-5, 2005.

[11] Lin, W.C. and Chen, H.H. "Description of NTU Approach to NTCIR3 Multilingual Information Retrieval", Proceedings of NTCIR-3, 2003.

[12] Min, J.; Sun, L.; and Zhang, J. "ISCAS in English-Chinese CLIR at NTCIR-5", Proceedings of NTCIR-5, 2005.

[13] Nakagawa, T. "NTCIR-5 CLIR Experiments at Oki", Proceedings of NTCIR-5, 2005.

[14] NTCIR Project, http://research.nii.ac.jp/ntcir/

[15] The Lemur Toolkit, http://www.lemurproject.org

[16] Tomlinson, S. "CJK Experiments with Hummingbird SearchServerTM at NTCIR-5", Proceedings of NTCIR-5, 2005.

Table 6. Official evaluation results of STAGE2

	runs	AINLP map		ALL							
		Relax	Rigid	Relax				Rigid			
				min	max	med	ave	min	max	med	ave
N3	C-C-T-01	0.2924	0.2318	0.1123	0.3645	0.2997	0.2847	0.0952	0.2825	0.2349	0.2232
	C-C-D-02	0.2753	0.2204	0.1102	0.3953	0.2637	0.2822	0.0942	0.3247	0.2098	0.2291
	C-J-T-01	0.1593	0.1225	0.0000	0.3855	0.1166	0.1807	0.0000	0.3276	0.1192	0.1580
	C-J-D-02	0.1684	0.1309	0.0000	0.4285	0.2667	0.2362	0.0000	0.3616	0.2107	0.1977
	E-C-T-01	0.1509	0.1213	0.0741	0.2027	0.1202	0.1237	0.0588	0.1609	0.1086	0.1035
	E-C-D-02	0.1242	0.0938	0.0663	0.2145	0.1305	0.1293	0.0514	0.1753	0.1048	0.1047
	E-J-T-01	0.1294	0.1022	0.0000	0.3811	0.3060	0.2383	0.0000	0.3348	0.2547	0.2040
	E-J-D-02	0.2098	0.1753	0.0000	0.4386	0.3189	0.2802	0.0000	0.3701	0.2672	0.2357
N4	C-C-T-01	0.2021	0.1687	0.1093	0.3085	0.2319	0.2269	0.0912	0.2524	0.1861	0.1841
	C-C-D-02	0.1905	0.1424	0.1041	0.3174	0.2145	0.2212	0.0817	0.2538	0.1713	0.1746
	C-J-T-01	0.1608	0.1164	0.0000	0.4450	0.1073	0.1756	0.0000	0.3629	0.0780	0.1383
	C-J-D-02	0.1537	0.1072	0.0000	0.4193	0.2340	0.2086	0.0000	0.3295	0.1710	0.1582
	C-K-T-01	0.0325	0.0314	0.0325	0.0549	0.0437	0.0437	0.0314	0.0504	0.0409	0.0409
	C-K-D-02	0.0313	0.0304	0.0313	0.0422	0.0368	0.0368	0.0304	0.0377	0.0341	0.0341
	C-CJK-D-02	0.1074	0.0683	0.1074	0.1074	0.1074	0.1074	0.0683	0.0683	0.0683	0.0683
	E-C-T-01	0.0935	0.0834	0.0776	0.1847	0.1048	0.1138	0.0606	0.1506	0.0926	0.0985
	E-C-D-02	0.0729	0.0619	0.0729	0.1924	0.0956	0.1089	0.0535	0.1427	0.0822	0.0894
	E-J-T-01	0.1381	0.0989	0.0000	0.4610	0.3627	0.2834	0.0000	0.3599	0.2681	0.2171
	E-J-D-02	0.1453	0.1132	0.0000	0.4512	0.3848	0.2852	0.0000	0.3500	0.2993	0.2218
	E-K-T-01	0.0417	0.0387	0.0417	0.0417	0.0417	0.0417	0.0387	0.0387	0.0387	0.0387
	E-K-D-02	0.0647	0.0583	0.0647	0.0647	0.0647	0.0647	0.0583	0.0583	0.0583	0.0583
	E-CJK-D-02	0.0860	0.0609	0.0860	0.0860	0.0860	0.0860	0.0609	0.0609	0.0609	0.0609
N5	C-C-T-01	0.3642	0.3050	0.3233	0.4929	0.3795	0.4013	0.2767	0.4361	0.3341	0.3483
	C-C-D-02	0.3652	0.2970	0.2953	0.4889	0.3756	0.3850	0.2384	0.4240	0.3112	0.3233
	C-K-T-01	0.1339	0.1107	0.0233	0.1339	0.0786	0.0786	0.0203	0.1107	0.0655	0.0655
	C-K-D-02	0.0852	0.0732	0.0234	0.0852	0.0543	0.0543	0.0162	0.0732	0.0447	0.0447
	E-C-T-01	0.1799	0.1609	0.1262	0.3060	0.1759	0.1848	0.1150	0.2379	0.1469	0.1544
	E-C-D-02	0.1864	0.1549	0.1401	0.2933	0.1833	0.1966	0.1279	0.2423	0.1532	0.1659
	E-K-T-01	0.2708	0.2225	0.2708	0.5441	0.4075	0.4075	0.2225	0.4912	0.3569	0.3569
	E-K-D-02	0.3148	0.2780	0.3148	0.5571	0.4360	0.4360	0.2780	0.4936	0.3858	0.3858