

Multi-Engine Based Chinese-to-English Translation System

Yuncun Zuo Yu Zhou Chengqing Zong

National Laboratory of Pattern Recognition, Institute of Automation
Chinese Academy of Sciences, Beijing 100080, China
{yczuo, yzhou, cqzong}@nlpr.ia.ac.cn

Abstract

This paper describes a Multi-Engine based Chinese-to-English spoken language translation system. The design and implementation of the system is given in detail. Three different translation engines are employed in the system and a very simple way is proposed to select the best translation from all the outputs generated by them. The evaluation results from IWSLT2004 are reported and analyzed in detail. The results prove that the Multi-Engine based system is practical.

1. Introduction

Spoken language translation (SLT) technologies attempt to cross the language barriers between people with different native languages who each want to engage in conversation by using their mother-tongue. The importance of these technologies is increasing because there are many more opportunities for cross-language communication in face-to-face and telephone conversation, especially in the domain of tourism. Our work described in this paper is focused on the translation from Chinese spoken language to English, an important part of the multi-lingual information service system oriented to the 28th Beijing Olympic Games.

The remainder of this paper is organized as follows: Section 2 gives the related work of Multi-Engine based spoken language translations; Section 3 describes the overview of our system in detail, and then, Section 4 presents the evaluation results and finally Section 5 gives the conclusion.

2. Related work

Many approaches have been proposed and practiced in the SLT research, such as template based method, statistical method and so on. Different translation methods are integrated together to get better results for that each machine translation method has its own strengths and weaknesses. At present, there are two main approaches to use different translation methods in one system. The great difference

between the two approaches lies in that the output of the first one is combined by the best translation parts generated by various engines but the output of the second approach is a whole result selected from the results generated by each engine [1, 2, 3, 4].

The center idea of the first one is that: different translation engines run together and each engine generates part rather than whole results of an input sentence. The part results created by one engine may be used by the other engines when they generate new part results. Finally a selector or searcher selects part results from all the created part results to compose a new result for the whole input sentence as the system output. This method is very complex and the builder must be familiar with each translation method. The centre idea of the second one is that: different translation engines run independently and get their own result for a complete input sentence, and then a selector selects the best result as the system output. This method does not create new result. It is comparatively easy to be implemented and the builder need not be familiar with every translation method. So it is easy to integrate new translation engines.

Our system integrates three different translation engines which are developed independently and each can run separately. The three engines are: template-based translator (TBMT), inter-lingua based translator (IBMT) and statistic translator (SMT). The three engines all generate a result for a complete input sentence and they can not communicate with each other inside the system, so we select the second approach to accomplish our system.

3. Overview of system

Based on the motivation mentioned above, our system is constructed by three different components: 1) the preprocessor; 2) a center controller; 3) translation engines (TBMT, IBMT and SMT). Our architecture is some different from that mentioned above because of the particularity of the

approach to select results. Fig. 1 shows the architecture of our system. Every component is described in detail next.

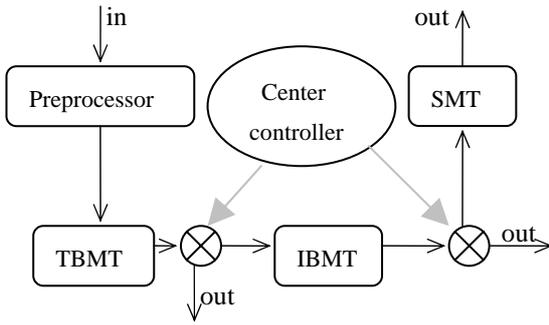


Figure 1: System architecture

3.1. Preprocessor

The preprocessor is designed to process the input sentences before translation. It completes the following functions:

- 1) To delete all repeated words except some special adverbs like “非常(very)”, “十分(very)” etc, this is very important for translation, especially for the template based translator because the repeated words in spoken language may influence the match between the input sentence and the templates.
- 2) To recognize and analyze the numerals and numeral phrases (QP) in the input sentences and translate the Chinese numerals into Arabic numerals.
- 3) To recognize and understand the time words and time phrases (TP) and translate them into English expression.

3.2. Center controller

In our system, the basic function of center controller is to select the system output from the results translated by different translator engines. There is some difference between the selector’s function proposed by [3]. The selector in [3] works after all the translation engines separately have a result, but the center controller in our system works when any translator finished its work and it controls the whole program.

According to our investigation, the results from the TBMT are generally of higher accuracy than the results from the SMT but its coverage is much smaller than that of the statistic method because no proper template can be found for a lot of sentences. The IBMT also has high accuracy but

now it is only oriented to the hotel reservation domain which is only a small subset of tourism domain. So its contribution to the whole system is very limited.

Our center controller is designed very simple. If a sentence is inputted to be translated, the work processes of our system are given as follows: First, the sentence is translated by the template based engine (TBMT). If almost the whole sentence (more than 80% in length computed by words) matches a template, the controller outputs the results as the system output and ended the whole translation process. Otherwise the sentence is sent to the inter-lingua based translator (IBMT). If a result is created here, the center controller outputs the result and finishes the whole process. Otherwise the sentence is passed to the statistic translator and the statistic translator translates the sentence and the result is output by the center controller.

3.3. Translation Engines

3.3.1. Template based translators

The template-based machine translator (TBMT) is the first translation engine in our system. It uses flexible expression format to describe the template condition. The template is designed as:

$$C_1 C_2 \dots C_n \Rightarrow T \quad (1)$$

Where, n is an integer ($n \geq i \geq 1$), C_i is a component which expresses a condition that the input utterance of source language has to meet. The utterance may be a Chinese word, a variable signed as N (noun), V (verb), A (adj) or other symbols, and also can be TP or QP recognized by the processor. T is the output result corresponding to the input and it also contains the variables, TP and QP appeared on the left side. The formula (1) means that if an input sentence of source language meets the conditions expressed by the left side, it will be translated to the target language expressed by the right side T . On the right side, the symbols TP, QP, N, V and other variants are replaced with their corresponding target language expression. For details of the template-based translation translator, please refer to [5].

3.3.2. *Interlingua-based translation method*

The inter-lingua based machine translator (IBMT) works after the template based translator in our system. IF (Interchange Format) developed by C-STAR (Consortium for Speech Translation Advanced Research international) is used as our inter-lingua. An IF consists of four parts: speaker, speech-act, concept, arguments. For example, the Chinese sentence “我想在北京预定一家旅馆” whose responding English is “I would like to make a hotel reservation in Beijing” is expressed by the following IF: c: give-information+ disposition +reservation +accommodation (disposition=(desire,who=I),reservationspec=(reservation,id entifiability=no),accommodation-spec=hotel,location=name-beijing). For detailed information about IF, refer to [6].

In the inter-lingua based translator, there are two key components: a spoken Chinese analyzer and an IF based English generator. The analyzer translates Chinese sentences to IF and it is based on the combination of statistical and rule method. The analyzer first analyzes the sentence into semantic chunks using rule-based method, and then analyzes the sentences using HMM-based statistical method. The approach is of the merits of rule-based method to analyze the deep semantic structure of sentence and it may keep the good robustness of statistical method. For detailed information about the analyzer, please refer to [7]. The generator is used to generate English sentences from IF. Our generator employs a hybrid approach in combination of template-based and feature-based generation methods. Templates containing variables are used to generate English from those fixed expressions in IF, and for the flexible expressions the feature-based generation method is used. The combination of the two methods makes the generator have a good tradeoff between efficiency and flexibility. For detailed information about the generator, please refer to [8]. Now, the inter-lingua based engine is only developed in hotel reservation domain – a small subset of tourism domain, so, in the IWSLT2004 evaluation, its uses are very limited.

3.3.3. *Statistical translation method*

Statistical machine translator (SMT) is the third translation engine in our system. The biggest advantages of the statistical method are their trainability, coverage and robustness. In our system, the SMT translator is based on the

classic statistical translation model-IBM Model 2 [9]. However, here the chunk-based translation is instead of the word-based translation aiming to overcome the shortcomings of word-based methods to lead to more fluent translations since chunk-based translations capture local reordering phenomena. However, until now, most phrase alignment algorithms have been based on complex syntax information e.g. by incorporating parsing technology with crossing constraints or have been narrowly focused on certain special kinds of phrases. These methods have proven to yield poor performance when dealing with long sentences. Further, the methods heavily depend on the performance of associated tools. In order to address these shortcomings effectively, here a new algorithm called multi-layer filtering (MLF) was proposed for automatically aligning bilingual chunks. Multiple layers are used to extract bilingual chunks according to different features of chunks in the bilingual corpus. And the alignment chunks are one-to-one corresponding with each other. The chunking and alignment algorithm doesn't rely on the information from tagging, parsing, syntax analyzing or segmenting for Chinese corpus as most conventional algorithms do. The detailed account of the method may be seen in the related paper [10].

4. Experiment results

Our training corpus is proposed by IWSLT2004 (<http://www.slt.atr.jp/IWSLT2004>), which contains 20K bilingual sentences selected randomly from the BTEC corpus. The BTEC corpus contains 160k bilingual sentences in tourism domain and each sentence is less than 12 Chinese words in general. We use the 20k bilingual corpus to train the statistic translator and extract 800 templates from the Chinese sentences for the template based translator. The test corpus is 500 Chinese sentences supplied by IWSLT2004, and they are also selected randomly from the BTEC corpus. In the Multi-Engine based translation results, 238(47.6%) sentences are translated by template based translator, 237 (47.4%) translated by SMT, 25 (5%) is translated by IBMT. The experimental results are evaluated automatically and subjectively. For details about the evaluation parameters please refer to <http://www.slt.atr.jp/IWSLT2004>. The automatic evaluation results of Multi-Engine machine translation (MEMT) are shown in Table 1. In order to give a comparison between MEMT and the other individual

machine translators, we also use the same 500 test sentences to do the experiment on TBMT and SMT and their automatic evaluation results are also shown in Table 1. Because the coverage of inter-lingua based translator is too small, here we doesn't test this engine individually. We have tested it before using other 100 sentences of hotel reservation domain, and the accuracy is 89.5%.

Table 1: Automatic evaluation results

Para Engi	BLUE	NIST	WER	PER	GTM
SMT	0.2835	6.1412	0.6767	0.6011	0.5426
TBMT	0.1711	1.2343	0.6860	0.6520	0.4187
MEMT	0.3113	5.9217	0.5788	0.5310	0.5639

From the results we can see 1) that the evaluation results of the template based translator are the worst. The reason is that template based translation has low coverage – there are almost one-half of the sentences can not be translated; 2) that the Multi-Engine system has better results than each individual engine as we expected.

The subjective evaluation results of multi-engine translator (MEMT), which consist of adequacy (ADEQ) and fluency (FLUE) are showed in table-2, In order to analysis the traits of each translator in the multi-engine based translation system, we extract out the evaluation results of each sentence and compute each engine's evaluation results, in table-2, the results of SMT are computed out based on evaluation results of 237 sentences translated by the statistic method in the MEMT and the results of TBMT are computed out based on the evaluation results of 238 sentences translated by template based translator in MEMT. This table also gives some automatic evaluation results computed out using each sentence's evaluation results.

Table 2: Evaluation result

Para Engi	ADEQ	FLUE	WER	PER	GTM
SMT	2.5316	3.6123	0.6123	0.5454	0.5546
TBMT	3.1343	3.2700	0.5338	0.5007	0.5999
MEMT	2.8000	3.4000	0.5788	0.5310	0.5639

From Table 2 we can see 1) that the results of template translator are of higher accuracy than statistic translator; 2) that the results translated by statistic translator are more fluency than template based translator.

From the evaluation results, we can conclude that the Multi-Engine machine transition system gets better results than each individual translation engine. The simple approach to selecting results is effective.

5. Conclusions

This paper presents a Multi-Engine based spoken language translation system, in which three different translators are integrated. The evaluation results from IWSLT2004 are also reported and analyzed in detail. This approach to integrating different translators has following advantages: 1) translators are independent from each other that make new translators are easy to be added; 2) It doesn't need the complex selector using the different traits of different translators in coverage and accuracy. However, this approach also has some worseness: the best results may not be selected because of the selecting approach is too simple. Especially if a new translator is added, this problem may be more severe.

The future work is to improve the performance of each engine and find a more effective way to select the best results.

6. Acknowledgements

This work is sponsored by the Natural Sciences Foundation of China under grant No.60375018, as well as the outstanding overseas Chinese Scholars Fund of Chinese Academy of Sciences under grant No. 2003-1-1, and also the PRA project under grant No. SI02-05.

7. References

- [1] Robert, Frederking and Sergei, Nirenburg, Three heads are better than one, In Proceeding of the Fourth Conference on ANLP-94, Stuttgart, Germany, 1994.
- [2] Christopher, Hogan, and Robert, E., Frederking, An Evaluation of the Multi-engine MT Architecture. AMTA 1998 pages113-123
- [3] Tadashi, Nomoto, Predictive Models of Performance in Multi-Engine Machine Translation. <http://www.amtaweb.org/summit/MTSummit/FinalPap>

- [4] Qun, Liu, A Chinese-English Machine Translation System Based on Micro-Engine Architecture, An International Conference on Translation and Information Technology, Hong Kong , Dec. 2000
- [5] Chengqing, Zong, Taiyi, Huang and Bo, Xu, An Improved Template-Based Approach to Spoken Language Translation. In proc. ICSLP2000, vol. , p440-443, Beijing, 2000.
- [6] L. Levin, Donna Gates, Alon, Lavie, Fabio Piansi, DorcasWallace, Taro, Watanable, Monika, Woszczyna, Evaluation of a Practical Interlingua for Task-Oriented Dialogue, Workshop of the SIG-IL, NAACL 2000.
- [7] Guodong, Xie, Chengqing, Zong and Bo, Xu, Chinese Spoken Language Analyzing Based On Combination of Statistical and Rule Method, In Proc.ICSLP2002. p613-616, Beijing, 2002.
- [8] Wenjie, Cao, Chengqing, Zong and Bo, Xu, Research of Chinese and English Language Generation Method Based on IF (in Chinese).To appear in Journal of Chinese Language and Computing 2004.
- [9] Peter, F., Brown, S.A., Della, Pietra, V.J., Della, Pietra, R.L., Mercer, The Mathematics of Statistical Machine Translation: Parameter Estimation, Computational Linguistics, 1993, 19(2): 263~311.
- [10] Yu, Zhou, Chengqing, Zong, and Bo, Xu, Bilingual Chunk Alignment in Statistical Machine Translation, 2004 IEEE International Conference on Systems, Man and Cybernetics(SMCC'04),Hague,Netherlands,2004.
- [11] Rile, Hu, Chengqing, Zong and Bo, Xu, Approach to Automatic Translation Template Acquisition Based on Unannotated Bilingual Grammar Induction. to appear in International Workshop on Machine Translation and Multilingual Information Retrieval.2004