

AN IMPROVED TEMPLATE-BASED APPROACH TO SPOKEN LANGUAGE TRANSLATION¹

Chengqing ZONG, Taiyi HUANG and Bo XU

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

ABSTRACT

In this paper, we describe an improved template-based approach to Chinese-to-English Spoken Language Translation (SLT) and present experimental results. The improved template-based translation approach uses flexible expression format to describe the template condition. The condition of a template may consist of keywords, parts-of-speech and also semantic features, so the input may be matched with a template from shallow level to deep level. In the condition of a template, the distance between two fixed keywords is stretchable, thus some needless words in the input utterances may be skipped in matching operation. And also the translation results of the same template are alterable. The proper results are finally generated according to the specific context. That is, the relation between a template and translated utterance is one-to- n (where, n is an integer and $n \geq 1$). The experiments were performed with input of both text transcription and results of speech recognition. The preliminary experimental results have proven the approach is practical.

1. INTRODUCTION

In this paper, we describe an improved template-based translator employed in the Chinese-to-English SLT system, which is developed at National Laboratory of Pattern Recognition, CAS. The system is limited in the domain of hotel reservation, and the Chinese vocabulary is 12000 words.

The method of template-based translation is easy to realize and the translation result must be correct if the input is matched properly with a template. So, the template-based translator is often employed as one of translation engines in Machine Translation (MT) systems^[1,2]. However, this method is of its congenital weakness. For example, the templates are generally fixed, and the flexibility to match the inputs and to generate the target language is often limited. So, the function and the effectiveness of a template-based translator are very little in SLT systems. In this paper we present an improved template-based approach which is robust in some extent for the spoken Chinese language translation.

Remainder of the paper is organized in the following way: We first describe the improved approaches used in our system, and then report some recent experimental results. Finally, the

concluding remarks are summarized and our future research topics are discussed.

2. THE IMPROVED TEMPLATE-BASED APPROACH TO TRANSLATION

2.1. Principle

It is our goal to extend the flexibility and the robustness of template-based translation approach and make the approach more effective in SLT system. So, we think the approach should be of the following characteristic features:

- Relation between the template and input utterance is one-to- n . Where, n is an integer and $n \geq 1$. That is the templates are expressed not only by the determinate keywords, but also expressed by variables or candidate components. Hence, a template may translate n utterances, which have similar structures and express similar meanings.
- In the template the distance between the two fixed keywords is stretchable. So some needless words in the input utterances may be skipped in matching operation.
- Although the same template may translate n similar utterances, the translation results of the utterances are not completely the same. That is to say, the translation result will be generated according to the specific utterance. This thus avoids the weakness that translation results from template-based translator are hidebound.

Based on the points listed above, we are to describe the template by use of five different types of components.

2.2. Format of the Template

The template is designed as:

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

Where, n is an integer and $n \geq 1$. C_i ($n \geq i \geq 1$) is a component which expresses a condition that the input utterance of source language has to meet. T_l is the output result corresponding to the input. The formula (1) means if an input utterance of source

¹ The research work described in this paper is supported by the National Key Fundamental Research Program (the 973 Program) of China under grant G1998030504, the National Natural Science Foundation of China under grant 69835003 and also the National 863 Hi-Tech Program under grant 863-306-ZT03-02-2.

language meets the conditions $C_1 C_2 \dots C_n$, the input utterance will be translated into the target language expression T_i . In our system the source language is the Chinese language, and target language is the English. A component may be expressed as the following five types:

1) Keyword

Any component C_i may be expressed as a fixed Chinese word. For example, “房间(room)”, “费用(expense)”, “有没有(whether to have)” etc.

2) Parts-of-Speech and Semantic Features

Any component C_i may be expressed as a part-of-speech, phrase name, a part-of-speech with semantic features or a phrase name with semantic features. In our system, the part-of-speech of the Chinese language is divided into 20 kinds including noun (N), verb (V), adjective (A), adverb (D) and preposition (R) etc. If two or more than two same parts-of-speech are used in the same template, the same parts-of-speech are given serial numbers (please see the *Example 2* below). The semantic features are divided into three levels. The first level is divided into 17 kinds, the total number of semantic features on the second level is 144, and the total number of semantic features on the third level is about 1600.

For instance, a C_i is expressed as NP(nso). It means the component is a noun phrase with semantic feature “nso”.

The phrase names are generated by use of parsing rules. The parsing rule is expressed as the following format^[3]:

$$A_i(s_1, \dots, s_{m_i}) \dots A_k(s_1, \dots, s_{m_k}) \rightarrow XP, !H | P \dots (2)$$

Where, k , m_i , and m_k are all integers and $k \geq 1$, $m_i \geq 1$, $m_k \geq 1$. A_i ($k \geq i \geq 1$) is a part-of-speech, s_i ($\geq i \geq 0$) is a semantic feature. XP is the new phrase name. H is the part-of-speech of the head word in the phrase, and H is one of A_i . P is the probability of the rule.

Some parts-of-speech or phrase names that are used in the components may appear in the translation result T_i . In T_i !N+s means the position will be replaced with the English word that corresponds to the Chinese noun of the input utterance. And the English noun may be changed into plurality according to specific context. N appears in the left of the template. Please see the *Example 1* and *Example 2* below. Similarly, !V+ing means the position will be replaced with the English verb that corresponds to the Chinese verb of the input utterance. And the English verb will be changed into gerund. V appears in the left of the template. !A+er means the position will be replaced with the English adjective that corresponds to the Chinese adjective of the input utterance. The English adjective will be changed into comparative degree.

3) * Variable

A C_i may be expressed as *. ‘*’ is a variable, and it means any word or phrase may appear at the position, or nothing appears.

4) Logical Expression of Candidate Components

A component C_i may be expressed as a logical expression of candidate components except * variable. For example, ‘ $C_1 | C_2$ ’ means the condition is one of C_1 and C_2 . The operator ‘|’ means logical OR.

5) Dispensable Components

The dispensable components are allowed to appear in the template and the components are written in square brackets. Such as, ‘[N]’ means a noun may appear or not appear at the position.

Except the * variable and dispensable components, the other ones are called **fixed components**.

According to description above, we give some examples of template as follows:

◇ *Example 1:*

#我想预订 QP 个 N => I want to reserve !QP !N+s.

QP is a number phrase. By use of the template, the input Chinese utterance ‘我想预定三个单人间’ will be translated into ‘I want to reserve 3 single rooms.’.

◇ *Example 2:*

#QP 个 带有 N1 的 N2 => !QP !N2+s with !N1.

By use of the template, the input Chinese ‘一个带空调的单人间’ will be translated into ‘One single room with air condition.’, and the input Chinese ‘两个有电话的双人间’ will be translated into ‘Two double rooms with telephone’.

◇ *Example 3:*

#* 太 A 了 => It is too !A.

The template is designed to translate the Chinese adjective phrases mainly. Such as, ‘太贵了’ is translated into ‘It is too expensive’. Any words before the Chinese character ‘太’ are ignored. For example, the input Chinese utterances ‘价钱太贵了’, ‘费用太贵了’ and ‘简直太贵了’ are all translated into the same result.

◇ *Example 4:*

#[I] [您|你] [是] W|N(nso) 吗 => !I, is that !W|!N?

Where, W is the part-of-speech of Chinese place words, and N(nso) means the noun is an unit name or a department name. Hence, by use of the template, the input ‘您是中科院吗’ will be translated into ‘Is that the Chinese Academy of Sciences?’. The expression is often used in Chinese when someone calls another.

2.3. Pre-Processor

The pre-processor is designed to process the utterances before the template-based translator works. In our system, the pre-processor completes the following three functions mainly:

- To delete all repeated words except some special adverbs. In the same Chinese sentence, except some special adverbs like ‘非常 (very)’, ‘特别 (very, in particular)’ etc., the same words are generally not repeated continually. Even if some special words are repeated continually, the number of times of repeated words is usual not big than two. So, we designed a special base to store the special words. The pre-processor will deletes all extra repeated words in the input utterances according to record in the base.

- To recognize and analyze the number words and number phrases in the input utterances. The number phrases are reduced to the phrase name QP. And the Chinese number words are translated into Arabic numerals.
- To recognize and understand the time words or time phrases in the input utterances. The time phrases are reduced to the phrase name TP and translated into English^[4].

2.4. The Matching Algorithm

Suppose any input utterance U consists of k ($k \geq 1$) Chinese words, and $U = W_1W_2 \cdots W_k$. The template format is as formula (1). The matching algorithm is described as follows.

Algorithm 1:

```

Set the pointer  $T_{pter}$  point to the first template;
Set the pointer  $I_{pter}$  point to the beginning of the input
utterance.
fd = 0; // a counter for number of matched templates.
While( $T_{pter} \neq \text{NULL}$ ) {
  Get_a_Template(); // To get a template and the
  // condition of the template is signed as C.
  n = Count_Num_FC(); // To count the number
  // of fixed components in C.
  if (n > k) continue; // Next template.
  pf = Match_TU(); // To match U with C from the
  // current position of  $I_{pter}$ .
  if (pf) { // pf=TRUE means it is matched successfully.
    fd++;
    if (U is matched completely) {
      TL_Generator(); // To generate the result.
    }
    else {
      Reset  $T_{pter}$  point to the first template;
      continue;
    }
  }
  else { // pf=FALSE means no proper template is found.
    continue;
  }
}
if (fd) {
  Output the translated results;
}
else {
  Give message;
}
return;

```

2.5. The Target Language Generation

When a template is matched successfully with the input utterance, the function of the target language generation is called by *Algorithm 1* to generate the translated results. The input of the target language generator is the serial number of the template that has been matched successfully. Suppose the serial

number of the template is i , and the right side of the template i is marked as T_{ii} . The algorithm is described as follows.

Algorithm 2:

```

Result = "\0"; // To initialize the variable.
word = GetWord(); // To get a word from  $T_{ii}$ .
if (word[0] != '!') {
  Append the word to Result;
}
else {
  Replace the part-of-speech or phrase name behind
  '!' with the corresponding English word;
  if (word = 'N+s') || (word = 'A+er') || (word = 'V+ing')
    Change the termination of the English word;
}
Output the Result;
return;

```

3. EXPERIMENTAL RESULTS

For analysis of the characteristics of the spoken Chinese language in the domain of hotel reservation, we have collected 94 dialogs^[7], which includes 3013 utterances. The dialogs were between clients and the clerks of a hotel. Except the clerks of hotel, the number of speakers is not less 50. The total number of utterances that clients or askers said is 1514. From the 1514 utterances, 106 templates are summarized and about 450 parsing rules are abstracted.

For evaluation of the template-based approach to the spoken Chinese language translation, we tested our system by use of 70 utterances, which are not included in the 1514 utterances. The 70 utterances are uttered by two native Chinese speakers. We tested our system by two ways. The first way is that the translator was tested by use of the results from the Chinese speech recognizer, which is developed at our laboratory. The second way is that the utterances were written down and stored in a text file, and then the utterances in the text file were sent to the translator one by one. The average length of the test utterances is 6.3 Chinese words. The following tables show the experimental results. *Table 1* shows results of the translator with text input.

Results	Correct	Part of Result is Correct	No Result
Num. of Utter.	46	6	18
Ratio (%)	65.7	8.6	25.7

Table 1. The Results of Text Input

To evaluate the translation performance, we took the top ten results of our Chinese speech recognizer as input of the translator. The Chinese speech recognizer had been trained by use of the 1514 utterances, and *Table 2* shows the performance of our Chinese speech recognizer.

CASES	WORD		SENTENCE	
	Top=1	Top=10	Top=1	Top=10
CORRECT RATE (%)	90	95	70	85

Table 2 Performance of the Chinese Speech Recognizer

Table 3 shows the translation results with input of top ten results of speech recognition.

Results	Correct	Part of Result is Correct	No Result
Num. of Utter.	36	7	27
Ratio (%)	51.4	10.0	38.6

Table 3. Translation Results with Input of Top Ten Results of Speech Recognition

As one of multiple translation engines employed in our SLT system, the template-based translator gets 51.4% correct results and 10.0% partial correct results. Though the correct rate is not very high, the authors still think the results of the improved template-based translator are heart-stirring. It greatly reduces the burden of the following translators, including rule-based translator and statistic-based translator.

4. CONCLUSION

The work described in the paper is part of our current work to develop the Chinese-to-English SLT system for the domain of hotel reservation. The preliminary experimental results have proven that the improved template-based translator may work effectively and robustly in some extent as a direct translator. However, we are also working on a number of advanced extensions to the translation system. These include the analysis of more advanced disambiguation techniques, and combination of the template-based method with our partial parser.

As you already saw that in the template some components are prolix. Especially, some phrase names may cause ambiguity sometimes. So, we use the phrase name as few as possible except the numeral phrases and the time phrases. Anyway, if a template is used to translate some utterances or part of utterances, the translated results are undoubtedly correct. What's more, the template-based translator is easy to develop and is of good transportability. So, we think the current results are worthy optimistic.

5. REFERENCES

- [1] Alon Lavie, Lori Levin et. al. The JANUS-III Translation System: Speech-to-Speech Translation in Multiple Domains. *Report of C-STAR Workshop*. Schwetzingen, Germany, Sept., 1999.
- [2] Zheng Baoshan, Liu Qun and Zhang Xiang, Establishing Principle and Method of the Template for Machine Translation (in Chinese). In *Proceedings of Computational*

Linguistics. Tsinghua University Press. 1999. Pages 299-304.

- [3] Chengqing Zong, Research on the Strategies of Spoken Language Translation (in Chinese). *Report for finishing post-doctoral program*. Institute of Automation, CAS. May 2000.
- [4] Chengqing Zong, Taiyi Huang and Bo Xu, Approach to Recognition and Understanding of the Time Constituents in the Spoken Chinese Language Translation. To appear in *Proceedings of the Third International Conference on Multimodal Interfaces (ICMI'2000)*. Springer-Verlag Press, Oct. 2000.
- [5] Satoshi Sato, MBT2: a Method for Combining Fragments of Examples in Example-based Translation. *Artificial Intelligence* 75(1995) 31-49.
- [6] Eiichiro Sumita, Setsuo Yamada and Kazuhide Yamamoto, Solutions to Problems Inherent in Spoken-language Translation: The ATR-MATRIX Approach. *Proc. MT Summit VII*, Sept. 1999. 229-235.
- [7] Chengqing Zong, Hua Wu, Taiyi Huang and Bo Xu. Analysis on Characteristics of Chinese Spoken Language. In *Proceedings of 5th Natural Language Processing Pacific Rim Symposium (NLPRS'99)*, pp. 358-362.