YanFa: An Online Automatic Scoring and Intelligent Feedback System of Student English-Chinese Translation

Yan Tian
School of Foreign Languages
Shanghai Jiao Tong University
`tianyan@sjtu.edu.cn`

Abstract

Online learning calls for instant assessment and feedback. YanFa is a system developed to score online English-Chinese translation exercises with intelligent feedback for Chinese non-English majors. With the aid of HowNet and Cilin—Chinese Synonym Set (Extended Version), the system adopts the hybrid approach to scoring student translation semantically. It compares student translation with model translation by Synonym Matching, Sentence-pattern Matching and Word Similarity Calculating respectively. The experiment results show that the correlation ratio between the scores given by the system and by human raters is 0.58, which indicates that the algorithm is able to fulfill the task of automated scoring. YanFa is also able to provide feedback on syntactic mistakes made by students through interacting with them. It asks students to analyze the English sentence elements. Then it compares the student analyses with those of the parser and points out the parts which might lead to their wrong understanding as well as their wrong translating.

1 Introduction

Online language learning and instructing are popular in the era of the Internet which calls for instant automated assessment and intelligent feedback. How to provide online translation exercises with immediate scoring and intelligent feedback is a challenging task. Although some researchers (Wang & Chang, 2009; Wen, et al., 2009) have investigated ways to score student translation, they did not aim at fully automated scoring of translation, nor did they try to serve online exercise scoring. Wang & Chang discussed methods of the human-aided automated assessment of translation tests in final exams, and Wen adopted bilingual alignment technology to score translation in language testing. However, online fully automated scoring of translation exercises has its own characteristics. Besides, providing online instant intelligent feedback for students presents another challenge to natural language processing. Up to now very little research, if any, has addressed this topic. In order to meet the demand of online automated scoring of translation exercises and to help students with intelligent feedback, an online automated scoring and intelligent feedback system, called YanFa, has been developed.

This paper aims to outline the framework of YanFa. The paper addresses this by explaining two modules of YanFa, namely, the automatic scoring module and the intelligent feedback

module. In order to test the accuracy of YanFa, a study with 200 college students was carried out at Shanghai Jiao Tong University. The research intends to verify whether YanFa is able to undertake the task of online automated scoring of student English-Chinese translation as well as the task of providing students with feedback on the mistakes in their comprehending of English sentences, which might lead to their wrong Chinese translation. This paper begins with an introduction, followed by the explanation of the two modules. The experiment is also described. The research findings suggest that YanFa is eligible not only to score student online translation, but also to provide feedback on student syntactic mistakes in their understanding.

2 Automatic Scoring Module

"Translating means translating meaning." (Nida, 1986) Thus, ideally, automated translation scoring should be done at semantic level. Namely, the system should be able to judge whether the student translation is correct in conveying the original meaning to the target language. Therefore, the scoring module should be able to analyze the meaning of student translation which includes word meaning, phrase meaning as well as sentence meaning because translation involves two kinds of transfer: lexical transfer and structural transfer (Hutchins, 1992). Another consideration of building the module is to simulate the manual translation scoring practice in which the sentences are scored according to the correct translation of language points (words and phrases) and that of sentence structures. Usually, 3/4 scores are given to language points and 1/4 to sentence structures.

The automatic scoring module is composed of two parts: the databases and the automatic scoring system. The databases are English Passage Pool, English Sentence Pool, Model Translation Pool, Model Sentence Pattern Pool, Student Translation Pool. The automatic scoring system is composed of a Chinese Parser (SharpICTCLAS.net with precision rate of 97.58 % and recall rate of 90%), a Word Analyzer, a Sentence Pattern Analyzer, a Rater. Besides, Chinese resources, HowNet and Cilin—Chinese Synonym Set (Extended Version by the Lab of Information Retrieval at Harbin Institute of Technology), are also adopted.

First, student translations are parsed by SharpICTCLAS. Then the parsed sentences are sent to Word Analyzer to be compared with the pre-parsed model translations by the same parser. Three different approaches are taken to deal with different parts of speech respectively: nouns are compared with the synonyms in Cilin, of which the seed nouns are from the model translations; verbs, adjectives and adverbs are compared by calculating the word similarity with the aid of HowNet. Similarly, the seed verbs, adjectives and adverbs also come from the model translations. The rest parts of speech, including idioms, are dealt with by key word matching method. After word processing, Sentence Pattern Analyzer compares the sentence patterns of student translations with the model sentence patterns. Last, the results of both analyzers are sent to the Rater which calculates the final score of a student translation. The formulas are as follows:

The formula for Word Analyzer:

Processing of nouns with Cilin:

$$sem\_cl(W_k) = \begin{cases} \dfrac{l}{\lambda\alpha}, & W_k \in C(W_s) \\ 0, & W_k \notin (W_s) \end{cases}$$

where $sem\_cl(W_k)$ refers to the score of a noun in student translation, W_k stands for a noun in student translation, l is the number of parsed parts of speech in model translation, C is the

synonym set of Cilin which embraces the noun appeared in student translation, W_s is a noun in model translation, λ is the total score of the sentences, α is a constant.

Processing of Verbs, Adjectives and Adverbs with HowNet:

$$simhn(W_k) = \arg \max_{1<i<m} (sim(W_i, W_k))$$

where $simhn(W_k)$ is the maximum value of a primitive, W_i is the primitive in HowNet, W_k is a word in student translation, $1<i<m$ means i is bigger than 1, but less than m (m is the number of primitives).

$$sem\_hn(W_k) = \frac{simhn(W_k)}{\lambda\alpha}$$

where $sem\_hn(W_k)$ refers to the score of a word.

Processing of other parts of speech:

$$sem\_st(W_k) = \begin{cases} \dfrac{l}{\lambda\alpha}, & W_k \in T(W_s) \\ 0, & W_k \notin T(W_s) \end{cases}$$

where $sem\_st(W_k)$ means the score of other parts of speech, T refers to the set of other parts of speech.

The formula for Sentence Pattern Analyzer:

$$sim\_pat = \begin{cases} (1-\alpha)\lambda, & AnsTran \in reg(Std\operatorname{Re}g) \\ 0, & AnsTran \notin reg(Std\operatorname{Re}g) \end{cases}$$

where sim_pat stands for the score of the sentence pattern of a sentence, $reg(Std\operatorname{Re}g)$ refers to the set of model translation (standard version) annotated by regular expression, $AnsTran$ means student translation.

The formula for each sentence:

$$score = \alpha\lambda sim\_sem(w_k) + (1-\alpha)\lambda sim\_pat$$

The formula for the total score of a passage (with 5 sentences to be translated as in YanFa system):

$$Totalscore = (\alpha\lambda sim\_sem(W_k) + (1-\alpha)\lambda sim\_pat) \times 5$$

3　Intelligent Feedback Module

It is believed that comprehending of a source language plays a crucial role in its translation, especially when the source language is a foreign language to a translator. Accordingly, correct understanding of English sentences is essential to its translating into Chinese. Therefore, the intelligent feedback module focuses on whether students could correctly understand the English sentences. Specifically, feedback on correct understanding of clauses is provided rather than that of phrases because wrong translation occurs frequently on linguistic units larger than phrases when complex sentences are to be translated by Chinese college students.

The Intelligent Feedback Module is composed of three parts: parsing of the original English sentences, comparing student parsing results with those of the parser, providing feedback to students.

3.1　Parsing

The module employs the English parser by Carnegie Mellon University (free online parser) to parse the original English sentences. It takes the advantage of the "SBAR" sign as the marks of clauses. For example, following is the parsed result of a sentence: "Because I believe that love begins at home and if we can create a home for the poor, I think that more and more love will spread."

```
<S <SBAR <SBAR Because
            <S <NP I>
               <UP believe
                  <SBAR that
                        <S <NP love>
                           <UP begins
                              <PP at
                                 <NP home>>>>>>>>
         , and
         <SBAR if
               <S <NP we>
                  <UP can
                     <UP create
                        <NP a home>
                        <PP for
                           <NP the poor>>>>>>>
   <S <NP I>
      <UP think
         <SBAR that
               <S <NP <NP more>
                     and
                     <NP more>>
                  <UP love
                     <NP <NP will>
                        <UP spread>>>>>>>
```

3.2 Comparing

The module asks students to mark clauses of the English sentences. Then it compares the marked clauses with the results parsed by the parser through string matching. If the matching fails, the module comes to the decision that a wrong understanding happens.

3.3 Providing feedback

The module is able to provide students with the comparison of their parsing with the right parsing of the whole sentences. If requested, the module is also able to present students with the comparison of their wrongly parsed clauses with the right ones.

4 Experiment

In order to test the accuracy of the automatic scoring system, 200 non-English majors at Shanghai Jiao Tong University were invited to try the online scoring system at "Shanghai Jiao Tong University English Learning Center" (http://english.sjtu.edu.cn). Then the scores were compared with those of two human raters. The correlation ratio is around 0.58 (the correlation ratio between the human raters is 0.67). Also, an online questionnaire was delivered to those who have tried the system to learn their opinions on the scores and the feedback they got. The statistics show that most of the students gave positive responses.

5 Conclusion

YanFa has been developed to score Chinese college students' online English-Chinese translation exercises as well as to provide feedback on their mistakes of understanding the English sentences. Semantic scoring has been explored on lexical level with such resources as HowNet and Cilin. While the scoring on sentence level has to yield to sentence pattern matching due to the unsatisfactory performance of Chinese syntactic parsers. Although this pilot research has achieved its initial purpose, yet it is far from satisfactory. Further efforts should be made in increasing the scoring accuracy and more detailed feedback.

References

Waard, J.D. & Nida, E.A., 1986. *From One Language to Another.* Thomas Nelson Publishers, Tennessee, U.S.A

Wang, L. & Chang, B.B., 2009. *Research on the Human-aided Auto-assessment for Translation Tests in College English.* CAFLEC, No. 128,17-21

Wen, Q.F. et al., *Application of Bilingual Alignment Technology to Automatic Translation Scoring of English Test.* CAFLEC, No. 125, 3-8

Hutchins,W. J. & Somers H. L. ,1992. *An Introduction to Machine Translation*, ACADEMIC PRESS LIMITED, Printed in Great Britain at the University Press, Cambridge. 110-111

D. Callear, J. Jerrams-Smith, and V. Soh, "*CAA of Short Non-MCQ Answers*", Fifth International Computer Assisted Assessment Conference, Loughborough University, 2001.

Brown, H.D.1987. *Principles of Language Learning and Teaching* [M]. Egnlewood Cliffs, NJ: Prentice Hall.