Application of fuzzy neural network model in the evaluation of translation quality of scientific and technical documents

Dongliang Ding

Department of College English Teaching and Studies, Nanyang Normal University, Nanyang, Henan 473061, China

Abstract

Through summing up the literature on translation standard associated with science and technology area, the paper proposes five factors throwing much influence on translation quality, including faithfulness and precision, expressiveness and intelligibility, logical clarity, proper use of terms and the principle of catering to professional readers. As is discussed by translation researchers, translation evaluation is agreed to be a mathematical system characterized by complexity and non-linear nature since these factors are inner connected in a way devoid of evidence. Neural Network is a better tool in resolving complex non-linear problems. In the study, the author makes the five factors an index system in order to get a quantified data. First, they are classified into five scales, excellent, good, medium, pass, and failure with each representing a digital interval. Then a questionnaire is conducted by marking five translated versions according to the offered indexes by thirty participants consisting of translation theorists, translators and professional readers. Then we use Back Propagation network to train the evaluation model with the data gained from these readers. By self-learning and adjusting priorities in translation evaluation, the model is finally capable of producing synthetically assessment on translated versions. The results obtained by means of two different methods are compared and analysed in the paper too. It is found that the difference between them is quite big. Only 60% of them are concord, which means the traditional evaluation by human is less objective and lower in accuracy and fails to produce answer close to the fact. On the contrary, due to the objectivity and artificial intelligence possessed by NN, the result with this approach is proved to be more reliable and accurate.

Keywords: Fuzzy neural network model, evaluation, translation quality, scientific and technical document

1. INTRODUCTION

Translation evaluation is an important part of translation theory, which contributes a lot to the development of theory. However, it is also found to be a less-developed part in the whole translation research. In the 1990’s, Brown et al. (IBM Watson Centre) proposed a source-channel model for statistical Translation evaluation, which made significant improvement over the traditional rule based translation methods. Ochs et al. (RWTH Aachen University) released an open-source toolkit named GIZA++ for word alignment according to Brown’s work, since when statistical Translation evaluation became a popular issue in the research field of natural language processing.

Translation evaluation is a natural language processing, one of the important branches, the purpose is to use computer to translate words or voice from a natural language into another kind of natural language, such as translate Chinese into English. Early translation evaluation with the aid of transformational grammar based on grammar rules, on the basis of translation dictionary to realize the source language into the target language translation. This method can achieve good results in specific areas, such as the weather forecast. Because the comparatively translation is regular and easy to sums up the relevant translation mode in the field. For more complex areas, however, such as news or tourism, summarizes the translation model is difficult to guarantee coverage and accuracy (Brown et al., 1990).

Since the end of last century, with the improving of the computer computing power and storage capacity, Translation evaluation method in statistical method began to boom. At present in Translation evaluation, not only a limited domain statistical method is a method of better performance, but the basic idea of statistical Translation evaluation through the study of the statistical analysis is a large number of parallel corpora,
constructing statistical translation model, and then use it combined with language model for translation (Translation evaluation, translation process can be as a decoding process). From the development phase, statistical Translation evaluation is given priority to with the method based on word, the earliest is now fully transition to translation method based on the phrase. Now the research of statistical Translation evaluation began to appear a lot of syntactic information fusion method, in order to realize the further improve the accuracy of translation. But compared with the method based on phrases, the method of translation model based on syntactic large will be enormous, the corresponding translation speed will be much slower. At present practical statistical Translation evaluation systems are mainly focused on the method based on phrases, now some of the commercial Translation evaluation system and the method based on phrases, such as Google translation system.

The first priority of statistical Translation evaluation is that construct a reasonable statistical model, which is created for the language and based on the statistical model and defined to estimate the model parameters and design the parameter estimation algorithm. Early statistical Translation evaluation based on word USES is noise channel model, and using the maximum likelihood criterion for unsupervised training, but in recent years the phrase based statistical Translation evaluation method commonly is used to distinguish sexual training, generally need a supervised training reference corpus (Xiong et al., 2010).

Statistical Translation evaluation system for translation effect has relationship with bilingual parallel corpora used for training model of the scale. The general consensus is the size of the parallel corpus; the greater the translation results become more accurate. In fact, Google translation system is based on a large scale of bilingual parallel corpora trained. But corpus increase the resulting problem is to get the scale of the model will lead to greater. Under the condition of the same computing power, the translation model, the greater the speed will be slower. In order to solve this problem, many researchers began to consider how to reduce the scale of the model, for example, from the point of view of the mobile phone mobile devices such as reduction as much as possible.

2. MATERIALS AND METHODS

T as a language through a channel noise and distortion, and at the other end of the channel appears as a word S in another kind of language, translation is, in fact, according to the observed S, restore the most possible T problem. Translation evaluation process just like the decoding process, the starting point is the Bayesian formula:

\[
P(T / S) = \frac{P(S / T) \times P(T)}{P(S)}
\]

(1)

For a given source language sentence S, P(S) is a random quantity, so neglected. So the process of translation turned into a maximum conditional probability problem. According to the Bayes formula can be derived:

\[
T = \arg \max P(T | S) = \arg \max P(T) P(S | T)
\]

(2)

This formula is known as the Fundamental Equation of Statistical Translation evaluation. In the formula, P(T) is the probability of the target language text T appear, known as the language model. P(S | T) is by the target language text translated into the probability of the source language text S T, called a Translation model. For translation Model P(S, T), IBM put forward 5 kinds of increasing complexity, the mathematical Model for short IBM Model 1 to 5 Model only considers the probability of the mutual translation of word and the word 1

\[
P(S / T) = \frac{\varepsilon}{(l + 1)^m} \prod_{j=1}^{m} \sum_{l=0}^{i} P(s_j / t_l)
\]

(3)

Model 2 in 1 model on the basis of considering the words in the translation process the change of position, deformation probability is introduced D (aj | j, m, 1), m and 1, respectively, the length of the target language and source language sentence, j is the location of the target language words, aj is its Corresponding to the location of the source language words "as shown in type (4):
\[ P(S/T) = \varepsilon \prod_{j=1}^{m} \sum_{t=0}^{l} P(s_j/t_a_j) d(a_j/j, m, l) \quad (4) \]

Model 3 in the first two models on the basis of considering a word translated into more than one word, reproduction is introduced Probability of \((n, t_a, s)\), translated into \(n\) to said words, the probability of the target language the word "so the three model considers the probability of literal translation. The probability of reproduction and deformation, such as type (2.12) is shown:

\[ P(S/T) \approx \prod_{j=1}^{m} n(\phi|t_a_j) \prod_{j=1}^{m} t(s_j|t_a_j) \prod_{j=1}^{m} d(a_j/j, m, l) \quad (5) \]

The model 4 and 5 are improvements on the basis of model 3. Here is no longer in detail.

Brown, put forward the earliest statistical Translation evaluation model - noise channel model (Brown et al., 1990), assuming that the source language sentence \(S\), after a noisy channel distortion \(T\) for target language sentence, the goal of translation is as much as possible to \(T\) reduction for \(S\), namely to decrypt \(T\) (Decoding). The process of translation can be regarded as a given \(T\) target language sentence, the process of solving the most likely source language sentence \(S\):

\[ \hat{S} = \arg \max_s P(S|T) = \arg \max_s P(S) \cdot P(T|S) \quad (6) \]

Formula (6) with two models to describe the target language sentence for the probability of source language sentence \(S \cdot T\) translation. Among them, \(P(S)\) refers as the Language Model (Language Model) represents the source Language Sentence \(S\) appearance probability; \(P(T|S)\) called Translation Model, said the source language the probability that the sentences \(T\) \(S\) Translation for the target ones.

In order to make the results more in line with the word order, translation researchers have proposed all sorts of sequence model (Xiong et al., 2010), in a nutshell, used in statistical Translation evaluation model can be divided into language model, translation model and adjustable model.

Computational linguists Zhiwei Feng in the present situation and problems of the Translation evaluation (Xu et al., 2004) a pyramid model of the development of the mentioned Translation evaluation, as shown in figure 1-1, the pyramid describes the developing course of statistical Translation evaluation, at the bottom of the pyramid model is the most simple, easy to implement, the pyramid is upward development, need of the more complicated model. The ultimate goal of statistical Translation evaluation is to realize the automatic translation based on intermediate language. The current statistical Translation evaluation model and a lot of research focused on syntax and phrase between models, and on the study of semantics and intermediate language is still in the exploration phase. Statistical Translation evaluation in the source language word pyramid as shown below:

Statistical Translation evaluation model, roughly experienced models based on word, phrase based model, based on three stages of syntactic model. Translation model based on word in the context of information too little, the translation effect is not good, but it is the foundation of the other two models; Based on model scalability of the phrase is very good, as long as there is sufficient corpus, it can get very good translation model, but this kind of model is not usually a good deal with the order; Based on syntactic model is usually the result of the translation of sequence has good guidance, but because of its reliance on syntax analyser, scalability than a model based on the phrase. Researchers hope translation model can integrated the advantages of the above two kinds of models, therefore proposed the BTG grammar (Wu, 1995).
Intermediate language
source language semantic
structure
source language syntax tree
phrase source language ...
source language phrase
source language words

target language semantic
structure
target language syntax tree
target language level phrases
target language phrases
target language words

Word model
phrase model
Hierarchical phrase model
tree-string mode string-tree mode
Tree-tree model
Semantic model

Figure 1. The pyramid of statistical Translation evaluation

ITG grammar phrases hierarchy model (Chiang, 2007), and tree - string model (Vaswani et al., 2011), string - tree model (Liu et al., 2011), the dependence tree model (Zhao et al., 2011) and the model of the forest. Statistical Translation evaluation model based on whether the translation process has been adopted the hierarchy, which generally can be divided into two categories, phrase model and syntax. The history of the statistical Translation evaluation model based on phrases can be traced back to 1996, Wang put forward an improved model for IBM statistical alignment model, which be called the alignment translation model based on structure. This model is divided into two levels: Rough Alignment model and Detailed Alignment model. First of all, the source language and target language phrases through a coarse alignment model similar to the IBM model 2 phrase alignment of the sentence level, and then the words through a fine alignment within the phrase alignment algorithm based on model (Gremundo et al., 2011).

Similarly, Ochs presents an alignment template model, the model is divided into levels phrase alignment and word alignment, the difference is Ochs will align phrase alignment of generalization based on parts of speech of the template, and using the linear logarithmic model framework as a whole. Koehn in drab phrase model on the basis of considering the adjustable factors put forward a phrase translation model based on word alignment. And according to the model implements the translation system based on phrases Moses Pharaoh and upgrade version.

Moses had become a contrast experiment was carried out in the current statistical Translation evaluation of one of the most popular baseline system. Marco et al., the joint probability instead of conditional probability, the model was proposed based on joint probability of the phrase. This model is similar to the model of IBM; the different is that it USES a combined strategy to combine the phrase. Statistical translation model based on syntactic research have also made great progress in recent years, Wooldridge presented the first bilingual syntactic parsing grammar BTG and ITG, and the thoughts of the two methods from the context-free grammar, context-free grammar used in single language syntax analysis, and BTG and ITG grammar will this idea to promote to the bilingual.

BTG and ITG grammar for the first time will be the process of statistical Translation evaluation in the form of syntactic analysis, is just the end of the two grammatical operators are words. DeyiXiong made the BTG grammar model extended to phrase and put forward a kind of syntactic level limit, reward continuous phrases translation. Wei Jiang inspired by BTG grammar puts forward a model of hierarchical phrase, phrases hierarchy model used is a bilingual context-free grammar probability of synchronization. On the rules of a longer phrase, using a placeholder X instead of shorter translate phrases on both ends of the rules, to generate the translation
templates. By a bonding rules code translation as a result, to ensure that the entire translation process can generate a complete analysis of bilingual syntactic tree.

Yamata tree-string model is proposed, the model is the earliest transcription tree model, using the syntax tree of the source language and target language sentence for training. His translation process can be divided into three steps: the sequence, inserted, translation. First use of the source language syntax tree in the sequence and then on the syntax tree node insert some may be ignored the utterances of target languages, finally use the target language words to replace the source language words syntax tree. Yang Liu in addition kind of tree-string model is put forward, in the process of model training using the source language syntax tree, and through to remove certain syntactic nodes to translation model generalization. By iterating through the source language in the process of decoding, the syntactic nodes using the rules of syntax tree cover in the translation table to get result of translation.

Galley put forward a new string-tree model, based on his model can describe the layers of tree structure transformation rules. Galley method of extraction and the source language phrase, the target language structure tree and the alignment between them to maintain consistent minimum rules, they cannot be broken up into other rules. Galley further put forward the method of complex rules extraction by minimum rules and found that the compound rules to promote the performance of the system has a lot of help.

To some extent, can put the minimum rules for the word analogy, analogy and composite rules for phrases, based on the phrase of SMT, which are better than those based on words, because the phrase bound more context information, such as local word choice and phrase, the word order of compound rules relative to the minimum rules, also contains more context information, which is a major cause of composite rules of translating system performance.

Macro presents a string-tree model, an application of actual transcription and tree model, training to use the source language and target language syntax tree extract rules, when decoding by producing in the target language syntax tree rules cohesion was adjusted order. Graehl gave a detailed description to the tree transcription model framework, and put forward a set of general training series-tree model and the method of tree-a tree model. M waves proposed a more complex forest model. The basic idea is to source language sentence syntactic analysis of n-best results for compression forest, on the basis of bilingual word alignment from the forest of the source language syntax and sentence on the extraction of the target language translation rules. Decoding when first produced in the source language sentence n-best syntactic analysis, and then iterate through all the syntactic nodes, using the rule table to cover other way to the tree, the last in the forest in search of the highest scoring compression results. Forest model on syntactic nodes only when extracting rules extraction, Hui Zhang extends the model that the rule extraction stage allows to extract multiple syntactic tree node sequence (tree-sequence).

The Bayes error and model error is automatic translation problems inherent errors, it is difficult to improve. This article obtains the mistakes from the training; analysis of statistical Translation evaluation model construction process, each link may be introduced by the error, and discusses some effective ways to reduce these errors. In the phrase based statistical Translation evaluation model training process as shown in figure 2.
Model training process, first of all, starting from the bilingual sentence alignment corpora, after word alignment, phrase translation extraction rules, several steps such as short tone sequence rules extraction in bilingual translation model, at the same time in the target language training in monolingual corpora language model, at last, through training methods to adjust model parameters in the model the weight of each feature.

Experiment by alignment of this section is to filter the confusion degree of training corpus, get confused degree of five-sixteen respectively to filter the training corpus; the filtering result is shown in figure 3.

To test corpus filtering method based on degree of confusion on the result of automatic translation, with filtering of the proceeds of the 12-training corpus structure phrase model and sequence model respectively. Experiment on 12 phrase model and the sequence model of NIST 2005, 2006, 2008 Chinese-English translation corpora were tested, the result is shown in figure 4.
Experiment proves that training corpus alignment errors in the words of is it going to affect the results of the automatic translation. In confusion degree of training corpus filtration experiment, dramatically reduce the training corpus is started in the confusion degree threshold value is less than 12, in get BLEU - 4 highest score on the test set is a range of degree of confusion for 7-8. This suggests that the alignment degree of confusion between 8 and 12 aligned about 400000 words of no help to the improvement of the performance of the translation system, not only will have the opposite effect. And although alignment confusion degree under 7 to is a good translation of a sentence to, but the clauses of proportion in the whole training corpus is too small, not enough to higher effect of automatic translation. Experiments prove that degree of confusion between 7 and 8 alignment sentence is a critical point, for just to join in the training corpus alignment confusion degree under 7 other pair, large may improve the performance of the translation system; On the contrary, to join in the training corpus alignment confusion degree in eight or more other pair, most likely will cause adverse effect to the translation of system performance.

Use manual annotation of word alignment as a standard, automatic word alignment accuracy of the results, the recall rate, alignment error rate such as formula (7), respectively (8), (9).

\[
\text{Precision}(P, S; A) = \frac{|A \cap P|}{|A|} \quad (7)
\]

\[
\text{Recall}(P, S; A) = \frac{|A \cap S|}{|S|} \quad (8)
\]

\[
\text{AER}(P, S; A) = 1 - \frac{|A \cap S| + |A \cap P|}{|A| + |S|} \quad (9)
\]

In formula, the P contains the manual alignment of a collection of all possible and determines the word alignment, and S is the only contains the manual alignment of all collection of word alignment determined, A is the result of automatic word alignment.

3. RESULTS AND DISCUSSION

Model corresponding phrase, a phrase model based on factor, built-in phrase model of part of speech tagging model is built respectively and built-in syntactic tagging phrases linear sequence model, the sequence in the model contains three position \{M, S, D\} \{previous and next\} the two directions, a total of 6 kinds of sequence probability. The phrase model based on factor with built-in phrase model of sharing the same part of speech tagging based on linear tuning sequence model of part of speech tagging. After training of minimum error rate parameters adjustment, the four kinds of translation model and the corresponding adjustment sequence model BLEU score on the test set as table 1.
The sequence model of the number of rules to some extent, which can reflect the sequence model of sparse data and can be seen in the table 1, and based on a sequence of part-of-speech tagging of the number of sequence model of adjustable sequence rules about two-thirds of short tone sequence model, and based on syntactic tagging sequence of sequence rules in the model of 1/2 of the shortage of short tone sequence model. Sequence model various tuning sequence calculation, the various characteristics and translation model calculation, which are based on a phrase for occurrences in the whole training corpus based on statistics.

Table 1. Blue-4 scores on test sets by different translation models

<table>
<thead>
<tr>
<th>Translation model</th>
<th>Rule number sequence model</th>
<th>BLEU-4 scores</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NIST05</td>
<td>NIST06</td>
<td>NIST08</td>
<td></td>
</tr>
<tr>
<td>Phrase model</td>
<td>82.27M</td>
<td>0.3094</td>
<td>0.2880</td>
<td>0.2223</td>
<td></td>
</tr>
<tr>
<td>Factor phrase model</td>
<td>58.74M</td>
<td>0.3137</td>
<td>0.2901</td>
<td>0.2193</td>
<td></td>
</tr>
<tr>
<td>POS phrase model</td>
<td>58.74M</td>
<td>0.3119</td>
<td>0.2972</td>
<td>0.2284</td>
<td></td>
</tr>
<tr>
<td>Syntax phrase model</td>
<td>38.61M</td>
<td>0.3214</td>
<td>0.3081</td>
<td>0.2357</td>
<td></td>
</tr>
</tbody>
</table>

In training model of phrase extraction stage, and all kinds of translation model in table 1 the same with the phrase is abstracted from the training corpus of number, just at the time of calculating the sequence characteristics, side with the target language and source language exactly the same sequence rules are merged. Only phrase translation model with a short tone sequence contains the number of rules is same, the tuning sequence model based on part-of-speech tagging and the tuning sequence model based on syntactic tagging contains the number of rules is less than its corresponding translation model, it shows that with a small number of annotation symbols instead of words, which can make the order more fully to rule of statistical model.

IBM researchers based on the idea of the statistical Translation evaluation, by British and French bilingual Canadian parliamentary debates as bilingual corpus, developed a method of Candide English Translation evaluation system.

Table 2. Distribution of Educational Level by Age in each Population Group (%)

<table>
<thead>
<tr>
<th></th>
<th>Fluency</th>
<th>Adequacy</th>
<th>Time Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systran</td>
<td>.466</td>
<td>.540</td>
<td>.686</td>
</tr>
<tr>
<td>Candide</td>
<td>.511</td>
<td>.580</td>
<td>.575</td>
</tr>
<tr>
<td>Transman</td>
<td>.819</td>
<td>.838</td>
<td>.837</td>
</tr>
<tr>
<td>Manual</td>
<td>.833</td>
<td>.840</td>
<td></td>
</tr>
</tbody>
</table>

Chart is the result of ARPA test, of which the first line is the translation of the famous Systran system as a result, the second line is Candide translation as a result, the third line is the result of Candide plus artificial proofreading, and the fourth line is the result of pure human translation. Evaluation index has two: Fluency and Adequacy (Transman is developed by IBM after a translation editing tools). Time the wire shows with Candide Transman artificial check the Time and the proportion of pure manual translation Time.

4. CONCLUSION

In this paper, the translation evaluation is introduced in system research method, based on rules and statistical translation evaluation method and the instances were analysed and evaluated; At the same time introduces the translation evaluation method based on hybrid strategy, the application of statistical machine learning method in translation evaluation are summarized; translation evaluation techniques were analysed. In the decades of the development of translation evaluation, translation evaluation has made great progress, especially the recent years; all kinds of translation evaluation technology appear constantly. The rise of the network, but also provides a new application background for translation evaluation, some practical translation evaluation system appeared on the market, such as online translation, web pages, E-mail, etc. At the same time, translation evaluation still faces many problems, such as knowledge acquisition, ambiguity and how to better understanding
the language characteristics of the law, etc. Until now, ALPAC report is still worth our reflection. We should re-examine the end goal of translation evaluation, namely automatic access to high-quality translation, because language involves the differences between language and culture, the differences of social background, completely in the sense of automatic translation is not a realistic goal, we need to rethink the target location. At this stage of translation evaluation, translation evaluation should play the role of a secondary artificial, in some breakthrough in the restricted area, translation evaluation should be as a tool, not an independent system, to other tasks of natural language processing services, such as data mining, information retrieval, etc., to seek new application scenarios, such as mobile phone text message translation, TV to subtitle translation, all kinds of information of the release, etc. translation evaluation is always a challenging task, worthy of our perseverance to in-depth study.

5. ACKNOWLEDGMENTS

The authors acknowledge the social sciences project of Planning Office in Henan province(Grant: 2015CWX027), the sciences project of Nanyang Normal University(Grant: QN2016028).

REFERENCES
