Research on Machine Translation Sentence Pattern Transformation Method Based on BP Neural Network

Ning Wang, Meng Wang

Agricultural University of Hebei, Baoding, Hebei 071000, China

Abstract

Machine translation methods are mainly based on rule-based machine translation and template-based machine translation. But both of that are based on the language rules which are complex, difficult to generalize in essence. Sentence conversion refers to the conversion of exclamatory sentences and various interrogative sentences into declarative sentences. From the perspective of artificial intelligence, this paper introduces the BP neural network to study the Chinese-English machine translation of the process language in order to improve the speed and efficiency of translation. At the same time, an integrated BP network model is given and the network structure, parameter processing, learning algorithm and learning sample are discussed. This work is useful to improve the machine translation of intelligence.

Keywords: BP neural network, Machine translation, Sentence conversion.

1. INTRODUCTION

Machine translation has made great progress through the development of several decades. Some special translation systems have been widely used and have achieved very good economic and social benefits (Najafi et al., 2011; Wu et al., 2014; Zhi and Zhou, 2008). At present, machine translation mainly adopts two technical routes: based on rules and based on the template, essentially based on rules. But the template is a more sophisticated rules, and the constraint is smaller. Whichever method is used, a detailed statistical summary of the language rules is required. As the human language structure is very complex, there are great differences between different languages in order to complete the regularization of various languages, standardization. And at this stage it is almost impossible, which requires the introduction of machine translation knowledge of artificial intelligence. Such artificial neural networks are used to improve the machine translation intelligence.

At present, the application of neural networks has penetrated into various disciplines and research fields. Such as intelligent control, pattern recognition, computer vision, adaptive filtering and signal processing, nonlinear optimization, continuous speech recognition, knowledge processing and biomedical engineering (Yu and Meng, 2009). We can find examples of successful application of neural networks in all these fields. It is expected that the artificial neural network will have a good prospect in future development.

BP feed-forward neural network means that structure uses feed-forward network, learning methods uses BP (Back Propagation) algorithm of neural network (Xu et al., 2016; Zhang et al., 2016). BP neural network is a neural network with three or more hidden layers, including input layer, middle layer hidden layer and output layer. Full connectivity achieves between the upper and lower layers, and there is no connection between each layer of neurons.

BP neural network learning can be divided into two processes: working signal of the positive spread of the error signal and the reverse spread (Yuan, 2015; Yang and Chen, 2014). (1) The working signal of the positive propagation: the input signal from the input layer is through the hidden layer unit, and then passes to the output layer where the output signal generates. The weight of the network is fixed during the forward transmission of the signal, and the state of each layer affects only the state of the next neuron. If the desired output is not available at the output layer, the transition error signal propagates backwards. (2) Error signal back propagation network: the difference between the actual output and the desired output is the error signal, and the error signal from the output layer began to spread forward, which is the error signal back propagation. In the process of error signal back-propagation, the weights of the network are adjusted by error feedback. The actual output of the network closer is to the desired output through the constant correction of the weight.
Sentence transformation is the key of machine translation. Technology language is simple, standard, less ambiguous word characteristics, more conducive to the realization of machine translation. Sentence conversion is a transitional stage in machine translation that facilitates the deep analysis, transformation and generation of sentences. According to sentence structure and punctuation, sentence types are converted into statement structures according to the grammatical model of tube theory, and relevant information of sentences is recorded for the purpose of deep analysis and target generation (Shi et al., 2005). The correct and effective sentence pattern conversion plays an important role in the analysis of the whole sentence. On the basis of previous work, we try to introduce an artificial neural network technology into the Chinese-English sentence pattern conversion of the process language. In this paper, the network model of process language transformation is established, and the problems of network structure, parameter processing, learning algorithm and learning sample are discussed in detail.

This paper will discusses the application of BP neural network in machine translation.

2. MATERIALS AND METHODS

2.1 BP learning algorithm

Figure 1 shows a BP forward neural network model with a hidden layer (Yu et al., 2015).

![Figure 1. BP forward feedback network with a hidden layer](image)

The input layer is I, where is I input signals. Any of input signal with i said. Hidden layer is J, that is yi, J neurons, any neuron with j said. Output layer is K, and there are K output neurons, one of which is denoted k by any output neuron. The connection weight between the input layer and the hidden layer is Wji, and the connection weight between the hidden layer and the output layer is Wkj.

The superscript M denotes different training samples, and M denotes the number of training samples, respectively. The selected hidden layer and the output layer of the excitation function, where here are used functions. For a given input sample m, the input to the hidden layer neuron j is

\[ I_{m}^{j} = \sum_{i=1}^{I} W_{ji} X_{i}^{m} \quad (1) \]

Output of hidden layer j neurons is:

\[ \text{Out}_{m}^{j} = f_{1}(I_{m}^{j}) = f_{1}(\sum_{i=1}^{I} W_{ji} X_{i}^{m}) \quad (2) \]
Output of hidden layer k neurons is:

\[ \hat{y}_k^m = f_x^m \left( \sum_{i=1}^{l} w_{kj}^m f_x^m \left( \sum_{i=1}^{l} w_{ji}^m x_i \right) \right) \] (3)

Output of output layer k neurons is:

\[ E_k^m(w) = \frac{1}{2} \left( y_k^m - \hat{y}_k^m \right)^2 = \frac{1}{2} \left( y_k^m - f_x^m \left( \sum_{i=1}^{l} w_{kj}^m f_x^m \left( \sum_{i=1}^{l} w_{ji}^m x_i \right) \right) \right)^2 \] (4)

The on-line error function of the k-th neuron in the output layer is

\[ E_k^m(w) = \sum_{i=1}^{l} E_k^m(w) = \frac{1}{2} \sum_{i=1}^{l} (y_k^m - \hat{y}_k^m)^2 \] (5)

Its total error function is:

\[ e_k^m = y_k^m - \hat{y}_k^m \] (6)

The process of the language sentence structure is very complex. It can be divided into adverbial part, predicate part and object part. If the same network is used, the network is large in scale and complex in structure. The learning and application process is very long (Liu et al., 2007). Therefore, the network is designed as an integrated BP network, and the structure is shown in Figure 1. Each sub-BP network corresponds to a sentence component, the design of the various parts of the sample, a mentor to learn to obtain the appropriate weight to the form of data files stored. Network work, the total control module according to different parts of the sentence, run the appropriate sub-network. This is shown in Fig.2.

In the process of language conversion, the object is the most complex, and the translation is the most difficult. The structure of the network is also more complex. So the following emphasis on the object sentence conversion BP network is discussed. In the object, we use the number of nouns to further subdivide the conversion network into a noun BP network, two noun BP networks and three noun BP networks, which is limited to the length of the article. The following emphasis on two nouns object-oriented transformation BP network is analyzed.

According to the network model, \( x_i \) (i=1, 2, ..., 15) is used as the input variable to design the learning sample with the translation rule of the technology language as the expected output. In order to verify the reliability of the network, different object sentences are selected for verification. Examples show that the results of network decisions are consistent with the results of manual decisions.

It is useful to understand some of the structural features of Chinese, and some of the features of Chinese are summarized below (Wu, 2004). Some of these features may directly affect the network topology and operation rules.
(1) Chinese has a huge character set. "Chinese Dictionary" receives more than 56,000 words. This is different from the Chinese phonetic alphabet of the most significant features, such as English has only 26 basic letters. But there are only a few thousand commonly used Chinese characters. According to statistics, the cumulative frequency of 3072 commonly used Chinese characters is 99.7%.

(2) Chinese characters are morphemes. 93.2% of Chinese characters and morphemes have a one-to-one relationship.

(3) Compared with foreign languages, Chinese characters (morphemes) have two characteristics: freedom and productivity. The so-called freedom is the freedom of the location of Chinese characters which can be pre-can also be post-home, and is not positioned. For example, "产生" and "生产" are two different words. The so-called productivity refers to the commonly used Chinese word formation function which is very high. For example, a Chinese character can act with many other characters to form words. Such as the ability to form the strongest word "child" words constitute up to 704 words.

(4) The ability of Chinese characters to form words is not uniform. There are 3785 modern Chinese characters with word formation ability, and 94164 words can be used. Among them, the first 50 words constitute 52%.

(5) In the common Chinese words, the disyllabic words occupy the majority (74%), which is followed by the monosyllabic words. Words with more than five syllables are very few (0.1%) and can often be re-decomposed. Monosyllabic words and di-syllable words accounted for 96.35%, syllable words 2.21% and less than 0.3% of the five-syllable words.

Morphemes, words, phrases, and sentences are of four different grammatical units. Morpheme is the smallest grammatical unit of sound and meaning, which is constrained by word formation. According to the second feature of Chinese, it can be approximated that Chinese characters and morphemes are composed of words and phrases. They are grammatical units capable of expressing a complete meaning and expressing a specific intention of the speaker. In the written language, the end of the sentence is expressed as a pause (.), a question mark (?), or an exclamation point (!). The processing of Chinese word segmentation is a written text. Whether a Chinese character string can be used to express the meaning of a text, it is bounded by the syntax and semantics of the context(Liang, 2006). With artificial neural network for Chinese automatic word segmentation, a processing unit can be used to represent a morpheme, and the weight of the join between the units represents the constraints of lexical, syntactic, semantic and other aspects.

2.2 Translation sentence pattern transformation based on BP neural network

The structure of the whole sentence converter is shown in Fig. 3.

![Figure 3. Sentence flow chart](image-url)

This flow chart comprises an interrogative sentence processing module, an exclamation sentence processing module and a sentence modal information processing module(Cheng and Yu, 2012). First of all, determine the
sentence, and then a variety of sentences are processed, and finally in the modal information mark verb tense, voice and other information. The reason why make this division is to take full account of the degree of coupling between the various sentence patterns, and strive to reduce the complexity of the analysis. For the exclamatory sentences, the system transformed and translated directly to the example of pattern matching. For the anti-interrogative question, just remove the post-clause. For other questions, if the instance fails to match the conversion to the converter, the treatment of general interrogative sentences is the basis for the transformation of special interrogative sentences and selective interrogative sentences. Special interrogative sentence to remove the sentence after the first component is the general interrogative, and select the first part of the interrogative sentence for the general interrogative sentences or special questions.

In the process of object-language translation, the position of each word is changed, while the part of speech remains unchanged. Therefore, we can build a mapping according to the part of speech of English and Chinese words without regarding to each specific word. According to the characteristics of the process language, we have two nouns of the object of the part of speech in the order attributed to Table 1. The translated words are classified into Table 2.

Table 1 Technology language Chinese object words lexical order table

<table>
<thead>
<tr>
<th>Part of Speech</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjective</td>
<td>1</td>
</tr>
<tr>
<td>Conjunction</td>
<td>1</td>
</tr>
<tr>
<td>Numeral</td>
<td>1</td>
</tr>
<tr>
<td>Noun</td>
<td>1</td>
</tr>
<tr>
<td>Conjunction</td>
<td>1</td>
</tr>
<tr>
<td>Dayton</td>
<td>/</td>
</tr>
<tr>
<td>Adjective</td>
<td>2</td>
</tr>
<tr>
<td>Preposition</td>
<td>1</td>
</tr>
<tr>
<td>Numeral</td>
<td>1</td>
</tr>
<tr>
<td>Position noun</td>
<td>1</td>
</tr>
<tr>
<td>Auxiliary verb</td>
<td>1</td>
</tr>
<tr>
<td>Verb</td>
<td>1</td>
</tr>
<tr>
<td>Noun</td>
<td>2</td>
</tr>
<tr>
<td>Numeral</td>
<td>2</td>
</tr>
<tr>
<td>Position noun</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2 Post - translating Object Words

<table>
<thead>
<tr>
<th>Part of Speech</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>The</td>
<td></td>
</tr>
<tr>
<td>Position noun</td>
<td>1</td>
</tr>
<tr>
<td>adjective</td>
<td>1</td>
</tr>
<tr>
<td>adjective</td>
<td>2</td>
</tr>
<tr>
<td>verb</td>
<td>1</td>
</tr>
<tr>
<td>Of the noun</td>
<td>1</td>
</tr>
<tr>
<td>At the noun</td>
<td>1</td>
</tr>
<tr>
<td>Of the noun</td>
<td>2</td>
</tr>
<tr>
<td>adjective</td>
<td>1</td>
</tr>
<tr>
<td>conjunction</td>
<td>1</td>
</tr>
<tr>
<td>The noun</td>
<td>1</td>
</tr>
<tr>
<td>numeral</td>
<td>1</td>
</tr>
<tr>
<td>conjunction</td>
<td>2</td>
</tr>
<tr>
<td>Dayton</td>
<td>1</td>
</tr>
<tr>
<td>The noun</td>
<td>2</td>
</tr>
<tr>
<td>numeral</td>
<td>2</td>
</tr>
</tbody>
</table>
3. RESULTS AND DISCUSSION

3.1 The treatment of exclamatory sentences

As the exclamatory sentence structure is relatively fixed, the system model match the way that the example is of conversion and generation. In the example model library, the exclamatory sentences are modeled as follows in accordance with the following form, directly to the target language generation.

If the system input “How sweet the apple is!” after shallow syntax analysis, the sentence becomes “How * AjP (sweet) * NP”. The sentence is sent to the model library to match, so the sentence is translated as “how sweet this apple ah!”, directly to complete the exclamatory sentence processing.

3.2 The treatment of anti-interrogative questions

The first half of the interrogative sentence is a declarative sentence. When the questioning question is converted, it can be removed directly after the clause, and when the target is generated according to the type of doubt to add a different clause after the question word. If the last 4 Chunk of the interrogative sentence are “comma + auxiliary verb / Be verb + not ten pronoun”, then it is a positive / negative type. If the last three Chunk are “comma + auxiliary verb / Be verb + pronoun”, “No / Ken” type is anti-interrogative question.

3.3 The treatment of general interrogative

General interrogation of the conversion structure is shown in Fig. 4.

Figure 4. The Process of General Interrogative

If it is a negative question, then system removes the negative word first. For the auxiliary verb do / does directive to guide the general question, remove the sentence directly to the first auxiliary verb, or the need to swap the subject part and the location of the first composition. Finally, record the type information of the sentence. The key to the transformation of general interrogative sentences is the determination of the subject. When the position verbs of EB verbs or auxiliary verbs are restored, they are placed in every possible position, and judged by BP. If they meet the grammatical requirements, they will be sent to the next module. Otherwise, they will be released to the next possible location until the analysis is correct. If it does not meet the requirements of the location, the system does not carry out the next step analysis, direct error.

Be verbs or auxiliary verbs may lead to specific structural ambiguity. For example, when transforming a structure “*Be + NP + Ving + Ven”, it can be reduced to “NP + *Be + Ving + NP + Ven” or “NP + Ving + NP + *Be + Ven” to meet the requirements of syntactic structure. Using BP, it can not only be syntactically restricted, but also can be semantically restricted and disambiguated. For example: “is the boy reading a book given a
lesson by his friend?” Be verbs exist in two possible locations, "reading" and "given" are in front. Using BP neural network to analyze it, first give the verb read and give the argument structure:

\[
\text{read}, V, \text{Agent}[\text{Human}], +[\_\text{NP}\{\text{Reader()}\}] \quad \text{[Dative()]} \\
\text{give}, V, \text{Agent}[\text{Human}], +[\_\text{NP}\{\text{Human}\} \quad \text{PP\{place\}}] \quad \text{[Dative]} \quad \text{[Goal]}
\]

First, the verb Be is placed in front of "reading", then give the subject box (Dative) for the book, and give the request of the lattice of the semantic attribute for the Human, and book semantic attribute Reader. It does not meet the requirements. Release the next position, in line with the syntax and semantic requirements. The result of the conversion is "The boy reading a book is given a lesson by his friend."

3.4 The processing of selective interrogative

There are two types of interrogative sentences: the form of a general interrogative sentence and the form of a similar interrogative sentence. For the form of similar general interrogative sentences, respectively, it is before or after the conjunction or part of the analysis. The first part of the general query is in accordance with the treatment, and the latter part of the sentence was not convert, otherwise the same in accordance with the general treatment of interrogative. For the form of similar special questions, only the first part of the special interrogative sentence completes the sentence conversion.

3.5 Modality information processing

Sentence modality information can be exhaustive, therefore, the system lists the English modality information in all possible situations and deals with them. The symbols used in modal information processing are described below.

Predicate verb type: VP, VB
Aspect: 1 _Now, 2 _Past, 3 _Future, 4 _Past Future;
Tense: 1 _general, 2 _completed, 3 _carried out, 4 _to complete;
Person (person): 0 _can not determine the person, 1 _first person, 3 _third person;
Number (number): 0 _can not determine the singular and plural, 1 _odd;
Passive: o _active voice, 1 _passive voice;
Module: 0 _modal verb is not translated at the time of generation, 1 _when the modal verb needs to be translated.

For example, "WILL / HAVE / BEEN / VING" is denoted as "VP, 3,4,0,0,0,0.

4. CONCLUSION

Machine translation is a very complicated process. It is a new field of research to introduce artificial intelligence technology into machine translation. We use artificial neural network technology, and expect to improve the speed and efficiency of sentence translation in the process language translation. This research is still in the exploratory stage, and this paper analyzes the research method, develops the prototype translation system in the study. We will refine the system in future. The word segmentation algorithm is a combination of natural language understanding of neural network segmentation algorithm. In one sentence, the semantic analysis, the internal structure of a sentence, and the relationship between words and words can be seen at a glance, but the segmentation of ambiguous fields requires the use of word segmentation algorithm to deal with. The algorithm uses the different combinations of parts of speech to reflect the different grammatical rules, so the words in the sentence are categorized into the code design, and the input vector is expressed in code form. By training a large number of samples, the BP network learns the grammatical rules contained in each ambiguous field and gives the correct response to untrained samples during testing. The purpose of segmenting ambiguous field by neural network is achieved. In addition, through the improvement of the BP algorithm, the convergence speed has been improved and the word effect has been significantly improved.

REFERENCES