УДК 004.89

THE DEVELOPMENT OF ANSWER VERIFICATION COMPONENT

Qian Longwei, Li Wenzu
Belarusian State University of Informatics and Radioelectronics
e-mail: qianlw1226@gmail.com, lwzzggml@gmail.com

Summary. *With the rapid development of artificial intelligence technology, many intelligent tutoring systems (ITS) based on artificial intelligence technology have been developed. The most basic 2 functions of the ITS are: 1. ability to convert user answers described in natural language into fragment of knowledge base of ITS; 2. knowledge-based representation rules automatically verify the correctness and completeness of user answers.*

Converting user answers in natural language into fragment of knowledge base of ITS. The main difficulties of transition come from two aspects: the structure analysis of natural language texts; the analysis of fragment of knowledge base represented in internal language. In order to solve these two aspects problems, we proposed use OSTIS Technology [1] to develop component of ITS, which is able to extract sc-construction from natural language texts. ITS developed based on OSTIS Technology is a kind of ostis-system. The fragment of its knowledge base is the sc-construction represented in SC-code, which is an internal language used for knowledge representation. SC-code is the form of semantic network language with the basic set-theoretic interpretation. There are several ways of external representation such as SCg, SCn, SCs, which will be shown blow.

In order to convert natural language texts into the sc-constructions, it is necessary to describe the syntactic and semantic knowledge for text analysis, as well as the construction of extraction rules. Since each external language can have its own set of unique features, within the framework of OSTIS Technology, concept **Subject Domain** is proposed to describe knowledge of a certain of specific domain. We attempt to construct ontologies of specific natural language to implement conversion from open natural language text corpora.

Based on the principles of OSTIS Technology, an ontology is interpreted as a specification of the system of concepts of the corresponding subject domain. In each subject domain various distinguished ontologies are described to reflect a certain set of the concept features of the subject domain. For analysis of Chinese language texts, the **Section. Subject domain of lexical analysis**, the **Section. Subject domain of syntactic analysis** and the **Section. Subject domain of semantic analysis** need to be constructed to describe specification of a system of concepts, logical rules (e. g., extraction rules) and even other knowledge from the lexical aspect, syntactic aspect and semantic aspect of the Chinese language, respectively.

Aspect that construction of ontologies, the problem solvers of ostis-system also need to be considered, it can utilizes the knowledge base of linguistics to extract sc-construction from natural language free texts based on a series of text processing technologies, i. e., lexical analysis, syntactic analysis, semantic analysis, and extraction rules. The problem solver is a hierarchical system of agents of knowledge processing which oriented on solving specific problem. The general structure for developed problem solver for translating external texts into the fragment of knowledge base is represented below in SCn-language. The structure of developed problem solver corresponds to conversion steps from natural language sentence to sc-construction represented in SC-code.

Problem solver translating external texts into the fragment of knowledge base
<=decomposition of an abstract sc-agent:*

{
- *Abstract sc-agent of lexical analysis*
- *Abstract sc-agent of syntactic analysis*

- *Abstract sc-agent of semantic analysis*
- *Abstract sc-agent of generating sc-construction*
- *Abstract sc-agent of verifying the knowledge base*

}

Automatic verification of user answers

After converting the user answer described in natural language into sc-construction that is fragment of knowledge base of ITS, the correctness and completeness of the user answer can be verified based on semantics. Therefore, based on the OSTIS Technology and the SPICE approach (Semantic Propositional Image Caption Evaluation) proposed in the literature [2], this article proposes an approach to develop a sc-agent for verifying user answers in the ostis-systems. The basic principle of automatically verifying the correctness and completeness of user answers in the ostis-systems is to calculate the similarity between the standard answer in the form of the semantic networks and the user answer in the form of the semantic networks (built using SC-code). Therefore, the basic function of this sc-agent is to calculate the similarity between the standard answer and the user answer.

Next, we will introduce in detail the working algorithm of the sc-agent for verifying user answers.

Input: Semantic fragment of standard answers and semantic fragment of user answer.

Output: The similarity between fragments, and the sc-node used to record the matching status of substructures.

1. checking whether the fragment of standard answer and the fragment of user answer exist at the same time, if so, go to step 2), otherwise, go to step 10);

2. according to the rules of knowledge representation (various types of sc-constructions), the semantic fragments of standard answer and user answer are decomposed into substructures;

3. iteratively traverse each substructure of the standard answer and user answer, classify them according to the type of substructure (three element sc-construction, five element sc-construction, etc.), and count the number of all substructures;

4. one type of substructure is randomly selected from the set of recorded standard answer substructure types;

5. according to the standard answer substructure type selected in step 4), a corresponding type of substructure is selected from the set of recorded user answer substructure types;

6. iteratively compare each substructure with the same substructure type between the standard answer and the user answer, and record the number of matched substructures and the matched substructures. The criterion for judging the matching of the same type of substructures is that the sc-nodes at the corresponding positions between the two substructures have the same main identifier. If the substructure contains sc-links, the contents of the sc-links at the corresponding positions must be also the same;

7. repeat step 4 --- step 6 until all types of substructures have been traversed;

8. calculating similarity, and generating semantic fragments for recording sc-agent running results;

9. removing all temporary sc-elements created while the sc-agent is running;

10. exit the program.

This article roughly introduces the basic process of automatically verifying the correctness and completeness of user answers in the ostis-systems. The first part mainly introduces the approach to converting user answers represented in natural language into knowledge construction of ostis-systems. In the second part, an approach to automatically verifying the correctness and completeness of user answers based on semantics and internal knowledge representation rules of the system is introduced.

References

1. Tseng, Yuen-Hsien. Chinese Open Relation Extraction for Knowledge Acquisition~/ Yuen-Hsien Tseng, et al. // Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics. − Gothenburg, 2014, − P.~12−16.

2. Li, W., Qian, L. Development of a problem solver for automatic answer verification in the intelligent tutoring systems / W. Li, L. Qian // International Conference on Open Semantic Technologies for Intelligent Systems. – Minsk, 2021. – P. 169–178.

УДК 316.42

AN IMPROVED DECODING METHOD OF HAMMING PRODUCT CODES

Ren Xunhuan
Belarusian State University of Informatics and Radioelectronics
e-mail: renxunhuan@bsuir.by

Summary. *Product codes are preferred in high data rate wireless communication systems to achieve good error performance. However, to conceive a proper decoding method, which can correct all the potential errors pattern that within the half of the code distance for the product code, is not easy. In this paper, an improved decoding method of hamming product code has proposed. The proposed method can correct all the combinations of errors that less than four. As a result, it has a good performance in BER and WER.*

Over half a century, the using of error-correcting coding (ECC) is one of the effective methods which improves the performance of noise immunity in information transmitted of communication systems and networks. The most widely used coding are those methods based on Hamming codes, various cyclic codes, BCH codes, product codes and others, allowing to correct random and dependent errors [1]. In Error control coding, parity check bits are calculated based on the input data. The input data and parity check bits are transmitted across a noisy channel. In the receiver, an ECC decoder is used to detect or correct the errors induced during the transmission. The number of parity bits depends upon the number of information bits. At present, the most successful coding schemes are turbo codes and low-density parity-check codes, since their excellent capability, closely to the Shannon limit. Under some specific requirement (typically, code-rates near to the unity and low error rates required), product codes may turn into competitive. Product codes [2, 3], which can be easily realized by concatenating simple component codes, have a good protection capability against both random and burst errors. Product codes, whose component codes are Hamming or extended Hamming product codes (BCH product codes), BCH product codes can be constructed to improve the error correction capability, but a more complex decoding process is required. The simplest two-dimensional product codes are single parity check (SPC) product codes, assured to correct only one error by inverting the intersection bit in the erroneous row and column [3].

Product codes are formed in the verification codes $C_1(n_1, k_1, d_1)$ and $C_2(n_2, k_2, d_2)$ for the rows and columns of the source code matrix $n_1 \times n_2$, respectively, where n is the length of the code; k is the number of information symbols. The random-error-detecting and random-error-correcting capabilities of code are determined by its minimum distance $d_{\min}$, if the component codes C_1 and C_2 have minimum Hamming distances d_1 and d_2, accordingly, then the minimum Hamming distance of the two-dimensional product codes C_{pc} is the product d_1 and d_2 ($d_{pc} = d_1 \times d_2$), and at the same time $d_{pc} \geq 2t + 1$, t is the multiplicity of corrected errors [1], which greatly increases the error correction capability.

According to the coding theory, it is need to add supervisory symbols to the information symbol sequence for reducing the appearance error. The result after conducting such opera-