AN APPROACH TO CALCULATING THE SIMILARITY BETWEEN SEMANTIC SEGMENTS IN THE INTELLIGENT TUTORING SYSTEMS

Li Wenzu, Qian Longwei.
Department of intellectual information technologies,
Belarussian state university informatics and radioelectronics
Minsk, Belarus
E-mail: lwzzggml@gmail.com, qianlw1226@gmail.com

This article proposes an approach to develop a sc-agent for calculating the similarity between semantic fragments described based on factual knowledge in the intelligent tutoring systems.

Introduction

With the rapid development of artificial intelligence technology, many intelligent tutoring systems (ITS) based on artificial intelligence technology have been developed. One of the most basic functions of ITS is to automatically verify the correctness and completeness of user answers. In the ITS, according to the type of knowledge describing the answer to the question, answer verification can be divided into: 1. answer verification of factual knowledge type (factual knowledge does not contain variables represents facts); 2. answer verification of logical knowledge type (logical knowledge contains variables, used to describe logical expressions). The most critical step in answer verification in ITS developed using OSTIS Technology (ostis-systems) is to calculate the similarity between the semantic fragment of the standard answer and the semantic fragment of the user answer. Due to the limitation of article capacity, in this article we only introduce the approach to calculating the similarity between semantic fragments based on factual knowledge description. In the ostis-systems, the sc-agent is the only entity that performs transformations and information processing in sc-memory. Therefore, this article proposes an approach to develop a sc-agent for calculating the similarity between semantic fragments described based on factual knowledge in the ostis-systems [1,2].

I. Existing approaches and problems

SPICE (Semantic Propositional Image Caption Evaluation) approach is mainly used to evaluate the quality of automatically generated image caption. The basic principle of this approach is to convert the automatically generated image caption and the image caption labeled by the user into the representation form of the semantic networks, and then calculate the semantic similarity between them. When calculating the semantic similarity, this approach first decomposes the semantic fragments representing the image caption into sub-structures according to the representation rules of knowledge, and then the final similarity is obtained according to the matching relationship of the sub-structures between the semantic fragments.

Although the SPICE approach calculates the similarity between semantic segments based on semantics, there is no uniform rule standard for constructing the semantic networks in this approach, so this approach cannot convert complex image captions into semantic network forms [1].

II. Proposed approach

Based on the SPICE approach and the OSTIS Technology used to develop semantic intelligence systems and the corresponding platforms, tools and approaches, an approach to developing a sc-agent for calculating the similarity between semantic fragments described based on factual knowledge is proposed in this article. The main function of this sc-agent is to use the semantic approach to calculate the similarity between any two semantic fragments described based on factual knowledge in the knowledge base of the ostis-systems. Next, we will introduce in detail the development process of the sc-agent for calculating the similarity between semantic fragments described based on factual knowledge [2].

The sc-agent for computing semantic similarity between semantic fragments described based on factual knowledge needs to complete the following tasks:

1. according to the representation rules of factual knowledge, the standard answers and user answers in the form of semantic networks are decomposed into substructures;
2. using formulas (1), (2), and (3) to calculate the precision P_{sc}, recall R_{sc} and similarity F_{sc} between semantic fragments.

$$P_{sc}(u,s) = \frac{|T_{sc}(u) \otimes T_{sc}(s)|}{|T_{sc}(u)|} \qquad (1)$$

$$R_{sc}(u,s) = \frac{|T_{sc}(u) \otimes T_{sc}(s)|}{|T_{sc}(s)|} \qquad (2)$$

$$F_{sc}(u,s) = \frac{2 \cdot P_{sc}(u,s) \cdot R_{sc}(u,s)}{P_{sc}(u,s) + R_{sc}(u,s)} \qquad (3)$$

The main calculation parameters in the formulas include:

- $T_{sc}(u)$ — all substructures after the decomposition of the user answers u;

- $T_{sc}(s)$ — all substructures after the decomposition of the standard answers s;
- $\otimes$ — binary matching operator, which represents the number of matching substructures in the set of two substructures.

Since the sc-agent developed in this article will be used for answer verification in the future, we will regard the two semantic fragments used to calculate similarity as the semantic fragment representing the standard answer and the semantic fragment representing the user answer.

Next, we will introduce the working algorithm of this sc-agent in detail:

Algorithm 1 — The working algorithm of sc-agent for computing semantic similarity between semantic fragments described based on factual knowledge

Input: Semantic fragment of standard answers and semantic fragment of user answer. The condition of the sc-agent response is that two semantic fragments that use factual knowledge to represent the answer appear in the sc-memory, and the similarity between them has not been calculated.

Output: The precision, recall and 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 factual 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. using formulas (1), (2), (3) calculate precision, recall and similarity, and generate 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.

Figure 1 shows an example of calculating the similarity between semantic fragments.

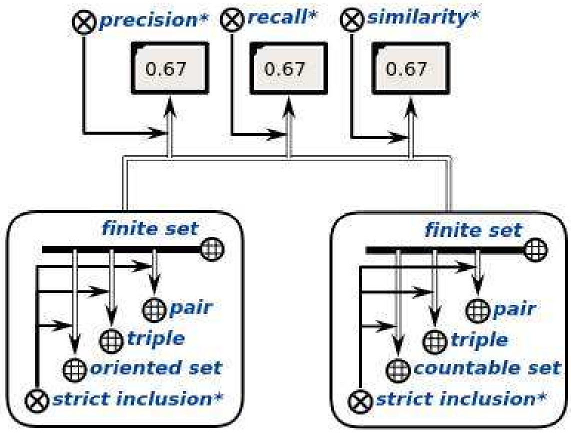

Рис. 1 – An example of calculating the similarity between semantic fragments

III. Conclusion

Based on the SPICE approach and OSTIS Technology, the sc-agent for computing semantic similarity between semantic fragments described based on factual knowledge in the ostis-systems is developed in this article. The working algorithm of this sc-agent is introduced in detail in the article. Since a unified knowledge representation approach and knowledge processing tool are provided by OSTIS Technology, the developed sc-agent can calculate the similarity between semantic fragments with very complex semantic structures.

IV. List of references

1. 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.
2. Golenkov, V. V., Guljakina, N. A. "Proekt otkrytoj semanticheskoj tehnologii komponentnogo proektirovanija intellektual'nyh sistem. chast' 1: Principy sozdanija project of open semantic technology for component design of intelligent systems. part 1: Creation principles]," / V. V. Golenkov, N. A. Guljakina // Ontologija proektirovanija [Ontology of design]. –2014. –№ 1. – P. 42–64.