

# Establishment an automatic grading system based on segmentation technology and the text similarity

Liping Zhao <sup>1, a</sup>, Qiliang Shu <sup>2, b</sup>

<sup>1</sup>school of information and engineering Jingdezhen ceramic institute, Jingdezhen 333403, China

<sup>2</sup>College of technology and art Jingdezhen ceramic institute, Jingdezhen 333401, China

<sup>a</sup>52778885@qq.com, <sup>b</sup>86673488@qq.com

---

## Abstract

Recent years, online exam system has become the hot research in colleges and universities, involved in the establishment and maintenance of question bank, the automatic form test papers, and automatic judge test papers, etc. Objective evaluation part of the automatic judge system is relatively easiness, but automatic judge of the subjective item has hindered the pace of the online examination system. In order to solve the bottleneck problem of network examination system, giving the model to judge subjective item to improving reading efficiency and reduce the artificial interferences in this paper. This paper focuses on the "Communication" problem between the computer and human. Chinese word segmentation technology is utilized to extract the keywords, and computing similarity between words to obtain the final score.

## Keywords

Word segmentation technology, syntax similarity, automatic grading.

---

## 1. The ideas of subjective topic automatic scoring algorithm

In artificial grading of subjective questions, teachers usually make the following steps: (1) set criteria; (2) find out the key part and key words of standard answer, assigned the total score to the keywords, namely scoring points; (3) look for the scoring points, check the number of scoring points in the students' answer; (4) if the students' expression is clear and smooth; (5) composite gives a relatively objective and fair score.

There are two factors have important impact on the subjective item judge: (1) the standard answer in test database and set of keywords; (2) the similar degree of students answer and standard answer in the test database. Therefore, in grading module, can put keywords extracted in the standard answer, according to its importance in the answer, set to different score; When grading, search and the standard answer from the students answer the corresponding keywords and extracted, and finally calculated with each degree of keywords in the standard answer, and accumulative every little score that matched with keyword. This method is similar to teacher's manual process of grading, evaluation method is proper.

Based on this analysis, this paper will do the following preparatory work:

(1) to clauses dealing with the standard answer, determine the key words of each clause and its weight; Establish near-synonyms table of each keyword.

(2) The clauses pretreatment the answers of student when grading, pattern matching the standard clauses in the standard answer with the clauses in the student's answer to retain useful clauses, and get the similarity values of clause grammar; The clauses that preserved used the maximum forward

matching algorithm to compare with the existing thesaurus and make automatic segmentation processing; Get the keywords score by fuzzy matching processing to the students answer after word segmentation processing and key words in the standard answer; The final score is the sum of the key words and syntax similarity score.

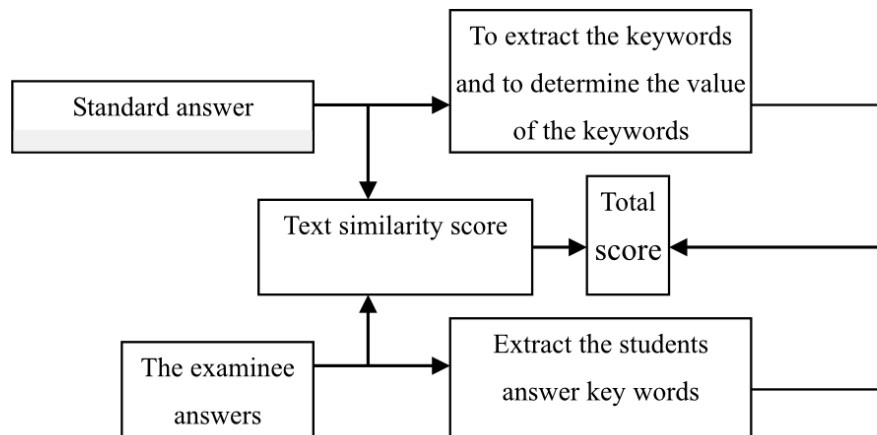


Fig.1 The flow chart of grading

## 2. correlation technique

### 2.1 Word segmentation technology

Useful clauses are preserved in the student answer after the pattern recognition. Word segmentation is an important part of the key words fuzzy matching algorithm and directly affect the score of keyword matching. Therefore, the subjective topic grading algorithm automatically in this paper studied whether success a lot depends on the correctness of the segmentation results.

Chinese word segmentation technology belong to the category of natural language processing technology, the existing word segmentation method can be divided into three types: (1) mechanical word segmentation method that is also called string matching method of word segmentation. There are four common solutions according to the priority that matches the word length and scanning direction: positive maximum matching, the minimum matching, reverse maximum matching, reverse the minimum matching. (2) The understanding of word segmentation method based on. (3) The word segmentation method based on statistical. The basic principle of Chinese word segmentation technology is pretreatment to the input string (remove the punctuation, English characters and Numbers), word segmentation operation and output word. The model of word segmentation as shown in the following figure.

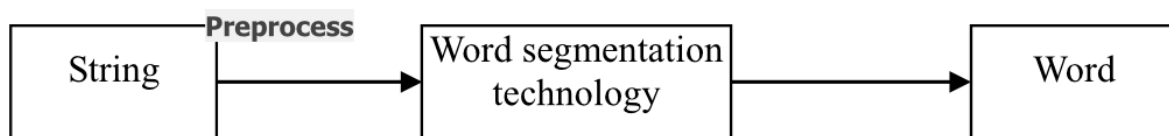


Fig.2 Word segmentation technology

### 2.2 sentence similarity

Sentence similarity is a standard of two statements similar degree. The higher the similarity between two statements means the more similar. In this paper the statement similarity is used to evaluate the examinee answers with degree of closeness to the standard answer. Current method for calculating the semantic similarity of words can be divided into two classes: based on word similarity calculation methods of the semantic dictionary and word semantic similarity calculation method based on corpus. The former is based on hyponymy and synonymous relations between concepts.

Get the similarity by calculation the distance of two concepts in hierarchy tree. This method is simple and effective, but it depends on the relatively complete according to the hierarchical relationship between the concepts of large semantic dictionary; The latter as contextual information of words a reference basis for semantic similarity calculation use of large-scale corpus. But two words similar in if and only if they are out of context based on this hypothesis, is dependent on the merits of the corpus and bad anti-jamming.

Based on the idea of artificial grading, in this paper, we design the statement similarity measurement method based on multilevel features a comprehensive statement on morphology, words length and word order similarity characteristic and reasonable expands the sentence similarity measurement. This algorithm not only consider the local similarity but also consider the similar of macroexamination statements on the whole text. To improve the performance of similarity calculation and the ability of subjective item automatically and accurately.

### **3. Implementation of algorithm**

Usually ,teachers first check the students answer and scoring more scored points more high score when judging the students' subjective item; Then check close degree of the student answer and the standard answer and more close more scores. We can conclude that there are two main factors influencing the score: one is the score point, another is close degree of student answer and the standard answer.

This paper simulation to achieve the teacher manual grading process based on the characteristics of the teacher's manual grade. First analysis the close degree of students' keyword and standard keyword in students' answer. Then analysis the close degree of semantic between student answer and the standard answer. Finally decided to the final grade. Algorithm of this paper mainly contains four parts: preprocessing, segmentation process, an improved string matching method and the similarity calculation process.

#### **3.1 Pretreatment**

As shown in figure 2, the pretreatment is the first step of this algorithm. This paper adopted literature statement of pretreatment methods. We according to guobiao code storage structure to difference between Chinese characters and isn't Chinese characters. We scan the text two times. The first scan to punctuate for the text to be processed according to the punctuation; Follow on the second scan, distinguish between Chinese characters and isn't Chinese characters. And automatic identification numbers and strings in English on isn't Chinese characters. The Chinese character string is the object of Chinese word segmentation.

#### **3.2 Word segmentation process**

The algorithm in the project is mainly used in multimedia short-answer questions (a phrase to describe). Standard answer and students answer is relatively simply, there is no context relation, so we selected the string matching participle technique that relatively simple and easy to implement. Vocabulary plays a very important role in the answer, basically constitutes the subject knowledge, so the word -word segmentation to be very accurate in the process of word segmentation professional. To satisfy algorithm on the time efficiency as high as possible, and the requirement of professional word non-professional word segmentation accuracy high, this article adopted reverse maximum matching method, according to a professional word dictionary extraction.

#### **3.3 Improve string matching algorithm**

On the basis of mechanical string matching was improved. According to the maximum matching keywords, when the match is successful, record the string; When the match is not successful, the key to left to delete a single character, to match; Until the remaining two characters. So, when the student's

answer is not quite right, still can get some points, this kind of treatment is to simulate human thinking, are reasonable.

### 3.4 Sentence similarity calculation process

Sentence similarity decided by words similarity, word order, sentence length similarity whereby words similarity plays a major role and secondary statement length similarity and word order similarity function minimum. Based on this, we gives different weights to different part in order to improve the accuracy of the sentence similarity calculation.

## 4. The design and implementation

- 1) Exam content randomly generated to reduce students' cheating in examinations.
- 2) Test paper submission: The students handed paper in or automatically submit by the system when the time is up.
- 3) Automatic mark examination papers.

### 4.1 System algorithm

We conclude the formula of subjective item grading from above-mentioned idea:

$$S = P * \sum_i^n = 1(\Phi A_i * \varphi_i) + (1 - P) * T_0 (0 < P \leq 1)$$

$$T_0 = \begin{cases} 0(T_0 < T) \\ 1(T_0 \geq T) \end{cases}$$

Denote: S: students' actual score; P: the proportion of the keyword; 1 - P: the proportion of the syntax similarity;  $\phi A_i$ : the score of the *i*th keyword similarity;  $\varphi_i$ : the weight of the *i*th a keyword; T0: syntax similarity value of 1 or 0.

Note: T is a threshold of syntax similarity. T0 = 1 when the result of syntax similarity is greater than or equal to T and show the grammar of the students answer conforms to the standard answer, vice T0=0. We can control the grading requirements by adjusting the value of T.

We can change the score proportion of keyword and syntax similarity by adjust P. The stringency ofgrading is controlled by P.

Program understanding consists of three parts:

- (1) To analyze source code and convert it to a abstract representation have control flow and data flow;
- (2)To reduce the complexity by standardized treatment the intermediate representation;
- (3) Treat with analytical and reasoning to the result of the specification process.

Programming follow the rules too when automatic grading.

### 4.2 System design

Marking system can be divided into: the subjective item grading and grading for programming.

The grading to subjective questions using the above introduced algorithm.

To programming problem: first check the running result whether correct, then check whether there is a syntax and logic errors. Score program begin from the results file, if there are any results file, then check to see if the examinee's output with the results of the standard answer files are same exactly, if the answer is "yes" and the code analysis credible then he will get full marks. Some answer can run but result is error, and we can make the following treatment: (1) The most program is correctly but have small careless mistakes, 2) Algorithm error, some programs didn't realize the topic request, 3) Some students just write a few sentences grammatically correct but beside the mark.

Code analysis is when the examinee result file does not exist or incorrect use of a source program files by processing the examinee the scoring method. Matching and knowledge matching is divided into scale. Size match basically see the examinee the reserved word of source documents and the degree of

---

match the standard answer reserved word. Knowledge points matching is check to see if the examinee file contains the keyword in the standard answer.

## 5. Conclusion

In this paper, we presented a subjective item automatic grading system based on the word segmentation technology and text similarity from the idea of subjective item manual grading and provide the key parameters involved in the process of automatic grading. This paper can automatic evaluation of short-answer questions and programming problem, considered whether to have the results, the results of the produce process is reasonable and the lack of results and so on. Because the individual discovers language is very complex, the study of examination system automatic grading to examination system needs further research and discussion in the future work.

## References:

- [1] Yu-hong li, Chai Linyan. Zhang qi word segmentation technology and subjective topic automatically mark of the sentence similarity algorithm. Computer engineering and design. 2010.11
- [2] Chuan-peng cheng, JiHuiWen. This similarity calculation in the application of the subjective topic grading. Computer engineering. 2012.3
- [3] WeiLiang Zha. Subjective item analysis and implementation of automatic scoring algorithm. Software Tribune, 2011.9
- [4] Zhang hongwei. Based on the semantic understanding of programming problem research and realization of the automatic scoring system (a master's degree thesis) dalian: 2010
- [5] Yang wei wei. One-way degree matching method application in the subjective topic automaticgrading. Mathematical techniques and application. 2010