Design and Implementation of Computer Aided English Composition Marking System

Zhao Hua

hopefull7781@163.com

Yinchuan University of Science and Technology, Ningxia, China

Abstract: With the rapid development of automatic marking technology for English writing, the design of auxiliary marking system based on this technology has gradually attracted people's attention and has been widely used in college English teaching. Based on this, this paper introduces the importance of college English composition marking, and introduces the design and implementation of English composition marking system from two aspects: word error correction module and front-end function realization of the system, hoping to provide corresponding reference and help for college English teaching.

Keywords: computer technology; English composition; marking system

1 Introduction

English writing and correction is an indispensable key component in college English teaching, and it is also the most time-consuming and energy-consuming link for teachers. In composition teaching, students may make some simple mistakes, such as spelling and grammar mistakes, which require teachers to invest a lot of time in correcting. Therefore, colleges and universities can try to use computer-aided English composition marking system to complete the organic combination of computer objective evaluation and teachers' subjective evaluation, and improve the quality and efficiency of writing correction.

2 Functions of college English composition marking system

2.1 It is beneficial for teachers to master the students' current English knowledge learning situation

In the process of college English learning, basic knowledge such as vocabulary and grammar is the key to students' learning. Single choice, reading and cloze can train and test students' English level, and writing is also an important way to reflect students' comprehensive English ability. In the process of arranging writing assignments for students, teachers often need to consider the contents taught at this stage and the students' knowledge mastery. Because teachers have strong pertinence in the process of arranging homework, they can understand students' shortcomings in the process of marking writing achievements, which can not be directly demonstrated by other assessment methods. In the actual writing process, students mainly use the key words they have learned recently. When a word is forgotten, they can look it up in the book, which is the process of vocabulary review. Not only that, writing needs constant application of grammar, and

the process of composition marking can help students find grammatical errors. It can be seen that marking homework can help teachers master students' English knowledge learning at this stage.

2.2 It is conducive to helping teachers to establish reasonable evaluation standards

In the process of marking writing assignments, teachers not only need to correct all the mistakes in the composition, but also need to ensure that the corrected content has certain emphasis, such as focusing on word spelling, grammar or the overall structure of the text. The first task of marking is grading, that is, to determine an evaluation standard and through which standard to confirm the composition score. This includes not only words and grammar, but also other aspects that can be used as criteria for evaluating English composition, so that students can find their own writing problems in combination with the final evaluation results. This requires teachers to always adhere to personal evaluation criteria in the evaluation process to ensure a more fair and objective score. Reasonable and unified standards have a high guiding influence on students' writing. Students can constantly supplement the loopholes in their knowledge system and improve their writing works in combination with teachers' evaluation standards, thus ensuring their personal English writing ability. Moreover, students with high scores can set an example for other students to learn, and communication between students will help improve the overall learning efficiency of the class [1].

2.3 It is beneficial to the communication between teachers and students

By marking homework, teachers can give appropriate comments and suggestions according to students' learning situation at this stage, and at the same time, they can communicate with students. Teachers can observe whether all students have made progress through writing, and some words that motivate students do not apply to all students, but teachers can give appropriate guidance and encouragement through the emotions in students' articles to guide students to learn more actively. Moreover, appropriate comments can make students feel that teachers pay attention to themselves. In the process of teaching, teachers not only need to impart theoretical knowledge to students, but also need to educate people. Evaluation is a traditional but human method. The arrangement of English courses in colleges and universities is doomed that teachers can't take care of all students' living conditions and learning conditions, so computer-aided English composition marking system is needed to realize the marking of all students' English writing works. However, teachers should pay attention to the fact that teachers should not rely too much on the computer-aided English composition marking system, but take the system as an aid, focusing on observing students' learning situation at this stage and giving corresponding help in time.

3 Structural design of computer-aided English composition marking system

3.1 Storage layer design method

The system uses separate storage to save system information data, which is basically divided into front-end cache, system file storage, system information database and other related information data storage, vocabulary and text error correction results storage and text type

division model data storage. The convolutional neural network model of text type division is used to store the related model of English writing text training results. The system information data storage module loads the brand-new training result model into the system after the model training, and at the same time removes the original neural network results from the memory [2].

3.2 Structural design of vocabulary error correction module

Vocabulary error correction system realizes unlabeled analysis through text segmentation, analyzes words in different batches, and searches whether the vocabulary exists in the English dictionary. If there is no such vocabulary in the dictionary, it is searched whether there is such vocabulary in the error correction dictionary, and if there is, the error correction result is fed back to the background server for analysis, and at the same time, it is continuously fed back to the error correction system in combination with the student feedback result to update the dictionary database [3]. The vocabulary error correction module uses BK-tree to construct a tree structure for English vocabulary, and uses the thinking of data structure to store data, and then carries out word retrieval through the retrieval tree.

3.3 Design of text classification module

The text classification module of the system can be subdivided into the training content of convolutional neural network model and the English text type classification module with text classification interface is externally supplied. The former is used to carry out neural network training for the English text database in the system and constantly update the text training model in the system server.

4. The Realization of the English composition marking system

4.1 Word preprocessing and dictionary construction

The word text transmitted from the front end to the back end contains HTML tags at a certain probability. In order to solve the problem of word text, it is necessary to give priority to text segmentation. Regular expression processing can be adopted, that is, a specific string processing logic formula is formed by JAVASCRIPT regular expression engine, and the HTML tags and related attributes are all set to null. The rest is English plain text, and then the English text is divided into independent words by combining the spaces between words for storage, and then it is corrected. After the word division is completed, the system will assign numbers to all the divided words, which is convenient to bind the error correction results and the error correction words into one and feed them back to the client in time.

Systematic error dictionaries can be subdivided into correct, wrong, replaceable and grammatically incorrect dictionaries. The correct vocabulary dictionary comes from the British dictionary and teachers' input, and at the same time, some relatively easy-to-appear English words are preserved. The wrong vocabulary dictionary saves frequently used dictionaries and common mistakes in the form of dictionaries, and edits the weights for words that are prone to errors, and saves the weight information in the error-prone data dictionary [4].

The storage format of the replaceable vocabulary dictionary is as follows:

Abatement abate,5,to decrease,4,become less strong,3

Abduction abduct,5,5to carry someone off by force,4

The symbols used above, such as commas, spaces, etc., provide convenience for later text analysis. The storage of error-prone words is no more than five, and reducing the data volume can effectively save a lot of memory space. The continuous updating and storage of vocabulary dictionaries are realized through basic data structures, regular expressions are analyzed and fed back to students' clients. Teachers' feedback and the analysis results of error correction algorithms are synchronized into dictionaries. With the increasing application of students, the accuracy of error correction results will be improved continuously.

The alternative vocabulary storage data structure can be represented by the following code:

```
typedef struct DictionaryNodc

{

int * index; //Error word serial number

char wrong word; //Error word

char exwordone; //The first replaceable word

int*exwordoneweight; // The weight of the first replaceable word

Char cxwordtwo; //The first replaceable word

int*exwordtwoweight; // The weight of the first replaceable word

char exwordthree://The first replaceable word

int*exwordtwoweight; // The weight of the first replaceable word

char exwordthree;//The first replaceable word

int*exwordthreeweight; // The weight of the first replaceable
word
}DietionaryNode:
```

The goal is to construct the values containing the word number, the wrong word, the values of the first to fifth replaceable words and the weights of each replaceable word. The weight represents the degree of substitution, and the higher the weight value, the higher the degree of substitution [5].

4.2 Word error correction

There are two main ways to correct and check English words: one is to preserve words by absolute method, and the other is to use relative method. Relative method is the likelihood function of vocabulary, but its overall efficiency is low and its structure is very complicated, which is unfavorable for text analysis and processing. Because there are obvious differences between English vocabulary error correction and ordinary text error correction, we need to think about vocabulary types, such as the use of articles or grammatical errors [6].

The word error correction method of computer-aided English composition marking system uses BURKHARD-KELLER algorithm. The usage is to find an approximate string from the string and apply it to spelling check. The key of this algorithm lies in the editing distance between

words, which represents the shortest distance between two words, such as the complexity of the transformation operation necessary to transform word B into word C. Moreover, the editing distance between many words waiting to be processed conforms to the metric space and the following data properties:

l（a,b）=0，a=b

l（a,b）=l（b,a）

l（a,b）+l（b,c）=l（a,c）

If you want to analyze the distance between absence and absent, you need to adopt the following methods:

Construct a table with 8 columns and 9 rows, and fill in the data from 2*2, and start calculation according to the following rules, and get the minimum value in the data: First, the leftmost character is the same as the top character, for example, the data edited by c=c is the upper left data, otherwise the upper left number is added by 1; Second, the number on the upper left needs to be increased by 1; Third, the number above needs to be increased by 1.

Through the above-mentioned loop operation, Table 1 can be obtained, and the set value of editing distance is L (absent,absence) = 7. BK-tree needs to select a word as the root node first, and then insert a new word as a child node. If the distance is the first time, it needs to create a new child node and insert it recursively downward along the edge. The details are as follows: First, choose GAME as the root node, and then insert FAME. Because the distance is set to 1, it is necessary to create a new node, and then add GAIN, AIM, SAME, GATE, etc. The structural design of all BK-tree nodes is shown in Table 1, in which index is the index of the current node, prenode is the index of the previous node, data represents the lexical content of the current node, and blownodeone represents the location of the indexes to the next five nodes of this node [7].

Table 1 The lexical similarity comparison operations

		a	b	s	e	n	t
	0	1	2	3	4	5	6
a	1	0	1	2	3	4	5
b	2	2	0	1	2	3	4
s	3	3	3	0	2	3	4
e	4	4	4	4	0	3	4
n	5	5	5	5	5	0	4
c	6	6	6	6	6	6	0
e	7	7	7	7	6	7	7

Error correction retrieval, combined with the third feature l(a,b)+l(b,c)=l(a,c) of BK-tree, if the wrong word abbss is recorded as a, the root node is recorded as b, the node to be retrieved is recorded as c, the editing distance is recorded as n, and the editing distance between words a that need to be retrieved and waiting for error correction is recorded as m, then the retrieved error correction vocabulary results need to be defined as m+l (b, c) = n. Therefore, the searchable vocabulary is l(b,c)=n-m, so it can be confirmed that the editing distance from the root node b is not greater than n-m points [8]. Under the background of relatively low requirements for the

total amount of feedback data, the retrieval data of BK-tree is controlled at 8%, which is in line with the requirements.

4.3 Module defects and subsequent improvement

The existing defects of error correction programs and algorithms are as follows: first, the construction of vocabulary dictionary depends on manual input and students' application of error feedback to build error correction dictionary database, which is relatively passive; Second, when BK-tree has a relatively large vocabulary, the time is relatively complicated, and it takes more time to correct vocabulary; Third, it can't complete the error correction of vocabulary according to the context, and can only simply check spelling problems, but can't check English grammar and articles.

In order to effectively deal with the problem of time complexity under the background of large vocabulary, the following solutions can be tried. The first is to set a specific threshold of editing distance. Second, the error-prone vocabulary is saved in the error correction vocabulary, and the results can be directly fed back without error correction algorithm. The error correction vocabulary is only updated in the background, and the new error correction results will be used after the update. Third, it is only compared with words whose word length gap is less than n, which can reduce the time complexity and speed up the page loading [9].

5 Conclusions

In the teaching of English writing in colleges and universities, composition writing and marking are both essential activities and effective measures and ways to improve students' English ability. In order to further improve the efficiency of teachers' composition marking, colleges and universities should actively introduce computer-aided English composition marking system to improve work efficiency and provide students with more objective and comprehensive correction suggestions. Teachers also need to give full play to the value of computer-aided English composition marking system to ensure the overall efficiency and quality of marking.

References

[1]Gu Feng. Exploration on the First-class Course Construction of English-Chinese/ Chinese-English Translation for English Majors in Colleges and Universities Based on the OBE Concept[J]. Journal of North China University of Water Resources and Electric Power(Social Sciences Edition),2023,39(03):39-45.

[2]Zhao Qian, Oscar Velasquez Arriaza. Practice and Thinking of Danish Teacher-led Online English Course Construction —— Taking Social Pedagogy as an Example[J]. Policy & Scientific Consult,2023(05):59-61.

[3]Cao Ying. Practical Exploration of Intelligent Writing System Assisting Mixed Teaching of English Writing[J]. The Guide of Science & Education,2022(18):22-25.

[4]Zhang Ying, Wang Dawei. Design and Implementation of Computer-based English Composition Marking System[J]. Techniques of Automation and Application,2021,40(02): 160-162+166.

[5]Xie Ting. Research on English Teaching by Parallel Processing of Online Courses and Classroom Teaching[J]. China Educational Technology & Equipment,2019(04):112-113+123.

[6]Li Qiaoli. Research on the Design of Computer-aided English Composition Marking System Based on B/S[J]. Techniques of Automation and Application,2018,37(11):55-58.

[7]Chen Yidong. Research and Design of Computer Aided English Composition Marking System[D]. University of Science and Technology of China,2016.

[8] Hurskainen A .Intelligent Computer-Assisted Language Learning: Implementation to Swahili[J].

[9]Boticario, Jesus G ,Rodriguez-Sanchez,et al.Toward interactive context-aware affective educational recommendations in computer-assisted language learning[J].New Review of Hypermedia & Multimedia, 2016.