Design and Realization of Nature Language Processing in The Application of Computer-assisted Teaching

Xianming Meng

ShanDong water polytechnic,Rizhao 276826,China
Email: mingmingmymail@163. com

Feng Kong

ShanDong water polytechnic,Rizhao
276826,China

Abstract–with the development of computer technology, computer-assisted teaching has become an important aspect in modern education application. More and more educational software plays a major role on helping people learn the language. In this paper, we combine natural language process technology to multimedia technology, and develop a computer-assisted teaching system based on natural language process technology to further intellectualize. Firstly, we investigate on the reliability and feasibility of the application of automatic Chinese text segmentation technologies. And then we probe into the readability of Chinese text, and analyze the factors influencing readability from the perspective of vocabulary. Finally, we propose the readability determination technique based on statistics of graded vocabulary, and establish the standards for readability determination. The feasibility of this technique has been verified by experiments. The exploration provides theoretic analysis and experimental evidence for the application of computer-assisted teaching.

Index Terms–computer-assisted teaching; natural language processing; word segmentation;

I. INTRODUCTION

With the theory of computer-aided instruction and computer network technology continues to develop, Computer-aided, baking in the foreign language teaching is playing an increasingly important role. At the same time, the extensive application of computers to make it from the previous data processing, information processing and development to the present knowledge processing, natural language information processing[1]. Language is mankind's most important communication tool, the load of information were most important, people language processing requirements of increasing the depth and breadth.

In China, the computer is used in transaction processing, office automation, printing and publishing, information retrieval, machine translation, assistant teaching, man-machine dialogue, etc., are inseparable from the Chinese, because all these information are in Chinese as carrier. Chinese information processing covers the words, phrases, sentences, chapters, and other levels of information processing tasks [2].

Multimedia teaching refers to the process of teaching based on teaching goals and teaching characteristics of the object, through the instructional design, a reasonable choice of symbols, language, text, sound, images, graphics and other elements of media information, human-computer interaction through a variety of ways effect, showing various media content, to complete the process of teaching [3]. It conforms to the needs of language teaching, in line with human cognitive characteristics of a multimedia computer and network communications, and other information tools for the technical means, the process of education and educational resources on the design, development, management, use and evaluation to optimize the educational process and educational resources, improving education quality and efficiency [4].

In this paper, we combine natural language process technology to multimedia technology, and develop a computer-assisted teaching system based on natural language process technology to further intellectualize. Firstly, we investigate on the reliability and feasibility of the application of automatic Chinese text segmentation technologies. And then we probe into the readability of Chinese text, and analyze the factors influencing readability from the perspective of vocabulary. Finally, we propose the readability determination technique based on statistics of graded vocabulary, and establish the standards for readability determination. The feasibility of this technique has been verified by experiments. The exploration provides theoretic analysis and experimental evidence for the application of computer-assisted teaching.

II. RELATED WORK

Automatic word segmentation is the basis for natural language processing, and natural language processing is the basis for intelligent computer technology [5]. Since the proposed automatic information processing in Chinese word segmentation has been published in many automatic segmentation methods. Although the name of a variety of different segmentation methods, segmentation is not the same speed, but essentially they can be grouped into two categories: One is segmentation based on statistical methods of mechanical, the other is a rules-based segmentation method of knowledge [6]. Department of Computer Science University of the West developed ABWS automatic segmentation system, and the system uses the "two scans Lenovo - back" method, using a greater amount of lexical, syntactic knowledge, their

knowledge base with eight categories of rules [7]. Tsinghua University, Chang-Ning Huang, Ma Yan and other in the SUN workstation using C language implementation of segmentation based on evaluation of the whole system of Chinese word segmentation SEG, the correct segmentation rate of 99.3% [8]. MM university, HangZhou university improved segmentation system that uses a modified MM segmentation algorithm to the theory of the mechanical sub-word segmentation accuracy is increased to 99.73% [9]. System accuracy results of 95%, is lower than the theoretical value of 99.73%.

Nanjing University, Chi-Shang Wang, Wang Xijiang completes segmentation system WSBN. Using the modified maximum matching, he combined with sub-lexical signs, supplemented by lexical and semantic knowledge [10]. Fudan University segmentation system developed consists of four modules: pre-processing module; ambiguous identification module; ambiguous field processing module; unknown word recognition module [11]. SLEX algorithm model system integrates a variety of data organization and search techniques to achieve low cost high-speed space-time matching and search, while using modern statistical methods of computational linguistics, the use of Markov process of parts of speech tagging and removing ambiguity of algorithm efficiency and stability are made possible optimization.

Probability and data-driven approach has almost become a standard method of natural language processing. Second, the improvement of computer processing speed and storage capacity increases, making speech and language processing in a number of sub-areas, especially in speech recognition, spell check, grammar check these sub-areas, can be commercially developed. Finally, the network technology enables language-based information retrieval and information extraction of the need to become more prominent. This is the natural language processing research to a new phase.

III. THE DESIGN OF COMPUTER ASSISTED TEACHING SYSTEM

Natural language processing technology is the basis for man-machine communication, you can see that it plays in computer assisted instruction is also very important position. However, how the natural language processing technology into computer-aided teaching, has been one focus of the study.

A The Functional Vision on Computer Assisted Teaching System

Learning in the learner's reading, the words become the biggest obstacles that affect their learning, for learning a second language especially in foreign language learners. How to help learners to read during the same time, the effective query words appear to improve the efficiency of learning to read, is to design the starting point for auxiliary reading system.

Read the article supporting system with the level of difficulty to determine function and dictionary search functions. Learners learn to read. The system first to the ease of reading material is judged to see whether the level of the learners themselves close to read. Second, In the reading

process, the learners encounter new words, the system can be queried and the results timely feedback to learners, making learning to read smoothly, learners in the reading process a smooth and effective access to information and new words to memory, improving reading efficiency.

In the preparation of teaching materials, now mostly rely on traditional manual way of operation. Although the textbook writers spent a great time and effort, still can not avoid mistakes and omissions. Compiled material system with auxiliary material for the article to determine the level of difficulty, the generation of text and the vocabulary associated word mark, including pronunciation, parts of speech and meaning and other information. Preparation of teaching materials in the choice of material needed to achieve the right side, science, rather than the previous rule of thumb selection. In the process of compiling teaching materials, rapid, and the correct word list generated, and the associated label. In the same time reducing the workload of teachers and improve the operating speed, effectively avoid possible errors in manual operations, thereby enhancing the efficiency of learning and teaching materials.

B Test Ease Analysis System

Assisted teaching system in use for teaching activities (ie, learning to read or learning and teaching materials), access to meet the material needs of the learner or teacher (reading materials or materials used to write text data) is to enable the smooth conduct of the premise. Update rate of information in today's society very quickly, explosive trend, how to face a flood of information in a short time select the appropriate information for learners and teachers, this is another problem they face.

This requires a tool, he must have been designed according to the difficulty of criterion to measure the specific text materials, and analyze the degree of difficulty of text material, and can result feedback to the user, allowing users to choose the right their level or the required material. The correct choice of materials is the prerequisite for reading and writing materials.

When the user use this tool to post (or text material) to determine the difficulty level, the tool will first use of automatic segmentation system for text input material automatic segmentation, the text automatically cut each sentence into one word between words and vocabulary will be separated by spaces. This form of the composition of the article to the previous sentences, paragraphs, became a word formed by the relatively independent, this time an article is equivalent to a large vocabulary, and word order is chaos. Then, the tool will follow the existing HSK word list, a good article for the segmentation of words automatically marked HSK difficulty levels, to be marked after the completion of various difficulty levels for statistical analysis of words and calculate the various difficulty levels in this article the word proportion. Finally, according to the criteria of the given, the level of difficulty of the article to determine the results that

is, the difficulty level of the article back to the user in selecting material for their reference.

IV. THE ANALYSIS METHOD ON ASSISTED READING SYSTEM

Learning process in reading is the influence of learners learning to read a number of factors. Can be divided into and reading-related factors and factors associated with the reader. Reading-related factors with the written word can be perceived, language code system integrity, legibility and meaning of reading materials. The factors with the readers of perception and cognitive are ability of readers and readers of non-linguistic knowledge, language knowledge and the reader readers reading motivation, purpose and interest. In the reading process, particularly in the independent reading learning, learners choose their own levels in addition to reading materials there are problems, but during the process of reading the article, there is a lot of words but also increased the difficulty for the learner. The main purpose for learning to read is to obtain information, to increase knowledge and expand your vocabulary. In the process of reading a large number of new words, for learners, not only affects the reading accuracy of the information, and continuous query word, will interrupt the learner's thinking, weakened the confidence of learners to learn to read, of reading efficiency.

A System Structure and Function

The system consists of two parts: auxiliary reading tools (tools and dictionaries to determine the level of article tools), and reading library. With this system, the level of decision tools, foreign language learners, you can get the reading material, especially for information on the network mass selective and targeted learning. With a dictionary tool, learners can put aside the shackles of words on the available materials, especially from the network for effective reading materials, and can continue to get the latest information about Chinese. Contact as much as possible a variety of information, not only exercise for the learner information from the mass of useful information capacity, but a lot of reading but also conducive to the acquisition of vocabulary and reading ability.

1) The Composition of Assisted Reading System

Assisted reading system is a secondary learning system. This system includes secondary level learners to read the article to determine the tools and dictionary tools, and fitness for a particular level of the learner's reading library. Article tools and dictionaries are to determine the level of tools for learners according to their level of reading material to choose their own learning materials and smooth. Useful for learners to find their own materials or interested in the future, the first tool can be used to determine the level of statistical analysis of the material contained in the words of the difficulty level and distribution, and thus determine the whole point of vocabulary difficulty of the material. If the material is more difficult for learners, you can begin reading, and used in reading the article in the dictionary, words were timely and effective queries. Moreover, in order to promote learner vocabulary learning, dictionary, should also have to study history query.

Reading library is a major for the beginning of a collection of intermediate learners of reading material. The material in the text material library has been pre-processed by tools to determine the level of the material to determine the level and degree of difficulty by showing classified storage. In addition, the text of the word is connected through the dictionary, the dictionary, so read the article at the same time is able to query word meaning. After selecting the material learners can start reading to learn.

2) The workflow of system

First, word segmentation system using text input material automatically. Segmentation of the text in the same time, the system can automatically slit a good word mark, indicating the part of speech of each word and other information. Tool will then graded according to the existing HSK vocabulary, as has a good article in the segmentation of words automatically marked HSK difficulty levels, to be marked after the completion of the vocabulary of each difficulty level for statistical analysis and calculated the article (or text material) the level of difficulty of each word in the proportion. Finally, according to the criteria of the given, the article (or text material) to determine levels of difficulty, and the results, that is, the difficulty level of the article back to the user in selecting material for their reference. Determine the level of results in the display, the display interface and function of the text box by box material composition. The former is used to display the selected text materials; the latter term is used to display statistics at all levels to determine the proportion and level of results.

B The Implementation of Assisted Reading System

In the design of foreign language teaching assistant to read the system process, application of Chinese word segmentation of the main idea is to use natural language processing technology and database query techniques, especially the Chinese sentence Computer automatic segmentation, POS automatic tagging, Pronunciation automatic annotation technology, with a bilingual dictionary machine, dictionary and corpus sentences, the process of achieving the learners to read words to the query on semi-automatic, auxiliary learners learn to read. When learners use the tools to determine the level of the level found in line with their own articles later, assisted reading tools can be used to read to learn.

In the face of new words, learners only need to use the mouse to click on the word, the system will first of all are reading the article text automatically. The text of the sentence will automatically cut the words apart, so that between the words and terms to be separated by spaces. Composition in the form of articles for the previous sentences, paragraphs, became the word formed by the relatively independent. Then article is equivalent to a large vocabulary. Then, the system will make the words and the link between various types of dictionaries for the word's part of speech, pronunciation, definition, and listen to mark and will follow the learners choose to display the results. Finally, a word encountered in the learners understand the relevant information, you can continue reading.

C Effective Use of Auxiliary System for Teaching Reading

Assisted reading system designed primarily to help learners learn to read, especially self-study. Its purpose is to enable learners to choose to use it to learn to read the article, and add the words in the reading process, and improve reading ability. In order to achieve better learning outcomes, according to the level of Chinese language learners, computer skills, and learning autonomy can be divided into the following application approaches.

1) Teaching learning material specified

In the teaching process, teachers can select materials in the material library, specify the materials for learners of a chapter to learn. Requires learners to use dictionary tools for preparation, advance preview the pronunciation and meaning of the word. If you use this mode, all learners should use the same dictionary from the dictionary of information tools, the class standard uniform. For other learners, they can reduce the learning time consuming to find dictionaries and attention to concentrate on reading. In addition, you can also put certain stories on the Internet or articles designated as learning materials for learners using assistive reading tools for learning. However, teachers in advance to make each word to find related fully prepared before commencement of the teaching.

2) Specify the scope of teaching learning

In this teaching model, teachers can provide a range of learning materials, learners can choose according to interests and needs of the material for their own learning. And through the use of information on the Internet, you can create a more open learning environment. Learners who want to understand their own learning vocabulary when learning history as long as the transfer out of it. History by generating word list learning allows the learner to the actual situation according to their pronunciation and meaning of the word to learn.

Diverse requirements of learners need a different approach to different levels of learners are individualized. Therefore, the effective use of dictionary, learning the history function of the individual learner guidance makes sense.

V. CONCLUSION

With the rapid development of information technology, multimedia, computer technology has been widely used in foreign language teaching. It provides foreign language teaching has brought new hope and prospects for new foreign language teaching situation, broaden the idea of change.

Compared with traditional teaching methods, computer-assisted teaching methods are in many ways irreplaceable advantages, which make up many of the traditional teaching methods of the deficiencies in foreign language teaching plays an important role. With the application of computer technology in teaching a wide range of in-depth, will bring the content of foreign language teaching, models and methods to influence and great changes, and would also promote foreign language teaching have far-reaching reform of the deep significance.

REFERENCES

[1] Jianhua Tao; Fang Zheng; Aijun Li; Ya Li. Advances in Chinese Natural Language Processing and Language resources. International Conference on Oritental Speech Database and Assessments. 2009. pp:13-18

[2] Li Zhao; Feng Li. Statistical Machine Learning in Natural Language Understanding: Object Constraint Language Translator for Business Process. IEEE International Syposium on Knowledge Acqusition and Modeling Workshop. 2008. pp: 1056-1059

[3] Siasar djahantighi, F.; Norouzifard, M.; Davarpanah, S.H.; Shenassa, M.H. Using natural language processing in order to create SQL queries. International Conference on Computer and Communication Engnieering. 2008. pp: 600-604

[4] Mai Fanjin; Yue Xiaoguang; Zhao Ziqiang. Research on chinese subjective questions scroing algorithm based on natural language processing. 2nd International Conference on Advanced Computer Control. 2010. pp: 106-108

[5] D. Tikk; Z.T. Kardkovacs; Z. Andriska; G. Magyar; A. Natural language question processing for hungarian deep web searcher. 2nd IEEE International Conference on Computational Cybernetics. 2004. pp: 303-308

[6] Fromm, P.; Drews, P. Natural language processing for dynamic environments. 24th Annual Conference of the IEEE Industrial Electronics Society. 1998. pp:2018-2021

[7] Bravo, M.; Montes, A.; Reyes, A.. Natural Language Processing Techniques for the Extraction of Semantic Information in Web Services. Seventh Mexican International Conference on Artificial Intelligence. 2008. pp:53-57

[8] Lei Li; Yanquan Zhou; Lin Zhang; Yixin Zhong. Enhancing Speech Based Information Service with Natural Language Processing. IEEE International Conference on Industrial Informatics. 2006. pp: 1057-1062

[9] Dura, E.; Gawronska, B. Natural language processing in information fusion terminology management. 11th International Conference on Information Fusion. 2008. pp: 1-8

[10] Karanastasi, A.; Christodoulakis, S. Semantic Processing of Natural Language Queries in the OntoNL Framework. International Conference on Semantic Computing, 2007. pp: 686-693

[11] Omar, N.; Razik, N.A. Determining the basic elements of object oriented programming using natural language processing. Internatioanl Symposium on Information Technology. 2008. pp:1-6.