



AN OFFICE AUTOMATIC MARKING SYSTEM RESEARCH AND IMPLEMENTATION

¹ YOUCHAN ZHU, ² FENG SHEN

¹ Prof, Information and Network Management Center North China Electric Power University, HeBei China

² Stu, North China Electric Power University, BaoDing HeBei, China

E-mail: ¹ zyc_hd@sina.com, ² shenfeng-ssff@163.com

ABSTRACT

The Microsoft Office Automation Marking System has two ways for marking, recording the operating steps and analyzing the result documents. The method of recording the operating steps has a disadvantage of difficult programming and low adaptability. And the other method has many implementations, but also has the shortage of examination restriction because of the uncompleted analysis of the office objects and low speed. Taking all of these into consideration, this paper addresses an implementation of using c# language and COM technology to analyze the interface methods of each application on the platform of Microsoft Visual Studio 2008, and aiming at the technical difficulties exist in the development process, such as the organization of knowledge, the transfer mode of test points, the positioning and extraction of test points during marking. It also gives the solution, and explains the logic and implementation steps in detail taking word as an example. Relative to recording the operating steps, it has the advantages of easy programming, easy system maintenance, non-limit of operating scope and good adaptability.

Keywords: *Office Automation marking, COM technology, Data Base Access*

1. INTRODUCTION

Currently, about Office skill 's training become more and more. This leads to an increasing assessment scale and marking workload. In order to meet this demand, the study of establishing an intelligent automatic marking system is a worthy subject[1]. There are mainly two ways to achieve the scoring results. One is to record the candidates' operating steps then make analysis, and the other is to analyze the resulting document[2]. As for the first method, we can construct the simulation system of the software or use the macro recording function provided by Microsoft Office to record the candidates' operating steps to achieve the goal. While the construction of the simulation system is a large project, and the updating of the software is very hard, considering of the environmental constraints, this method has a poor adaptability. The macro recording function has a low accuracy for it misses the operating steps and causes a miscarriage of justice very often. There are many practical implementations to mark the papers by analyzing the candidates' documents, such as analyzing the documents' prescription by using VBA, analyzing office documents by programming by using Active X. But these two methods both

have a disadvantage of partial analysis of office objects, and this restricts on examination[3]. In this paper, we will give a solution about the technology difficulties, and take word as an example to explain the implementation logics. At the following part we introduce the related concepts and the advantages of them solve the problems in the traditional system.

2. RELATED CONCEPTS

2.1 Com Technology

COM technology is a specification on how to create dynamic interconvert ion components, and provides some binary system standard and network standard for clients and components to guarantee interoperation. With this standard, any two components can communicate with each other regardless of their operating environments, their developing languages, and whether they run in the same computer[4].

Microsoft office is a software system which is constructed based on COM technology. It provides IDispatch interface and COM automation technology. As for the restrictions of the operating scope and the document type, we can completely resolve the COM interface methods provided by each Office application and get all documents'



properties through the study of the open object library. And the technical difficulties are the complete analysis of the Office object library[1].

The COM technology's advantage in the system is we can use it resolve the Office's knowledges in a maximum extent.

2.2 Database Access technology: ADO.NET

ADO.NET is a database access interface used in NET environment, which allows the interaction among different kinds of data sources and databases. The typical database accessing process by using ADO.NET is as follows[6]:

- a) Import into corresponding namespace
- b) Connect to database using Sql Connection and OleDb Connection Object
- c) Execute the SQL command to operate the database, such as query, insert, update and delete.
- d) Fill up the DataSet Object with the record dataset returned from the database.
- e) Close the database connection
- f) Operate the DataSet as needed
- g) Write the updated data back to database if necessary and not as an independent document. Please do not revise any of the current designations.

Using the technology we can solve the problem which the marking part can't pick up the knowledge completely.

3. PROGRAM OF SYSTEM TECHNICAL DIFFICULTIES

Traditional system's problem such as false positives is high, impelment complex and so on, always because they can't sort out and induce the Office knowledge and not store the knowledge using a rational mode[5]. At the following we will introduce the new system's method to solve the problem.

3.1 Organizational Forms of Knowledge Points

3.1.1 Coding of the Knowledge Points

The points encoding is by the use of three level encoding. The so-called three level encoding is made up of class name, class attributes and class attribute value, which is presented by a string of five letters and numbers. Take Word for example. Word knowledge analysis for the following categories: fonts, backgrounds, border and shadings, bullets and numbering class paragraphs class, picture class, table class, the header and footer class, page setup class, annotation formula super link class. The first bit of the five bits string are used to represent the above categories, the second bit stands for classes of property, followed

by three bits which indicates the property value. For example, the imitation Song-Dynasty-style typeface in the Cfont is encoding as aa006, of which the first letter "a" represents the Font class, the second letter "a" represents the Chinese font, "006" stands for the imitation Song-Dynasty-style typeface, the attribute value of Chinese font. Other categories are also organized in turn. Considering the constraints of the paper length, the collation of various test points is summarized in the other articles.

3.1.2 Database Storage

The system have used the B/S(Brower/Server) mode, so it interacts in the process of setting topics through the pages and the database. By analyzing the knowledge points of Office, it will store all the knowledge points that can be analyzed in the database dictionary table through the methods of above encoding.

The formation of the papers is also a reasonable choice on the knowledge points of pages, and saved to the database for the application of examination and marking. The table structure of papers is as shown:

TABLE I

ShiJuanBH	Paper No.
KaoTiBH	Topic No.
TiGan	Topic
CaoZuoTH	Operating question
CaoZuoFW	Operating range
CaoZuo1	Operating requirements No.1
CaoZuo2	Operating requirements No.2
CaoZuo3	Operating requirements No.3
CaoZuo4	Operating requirements No.4
CaoZuo5	Operating requirements No.5

ShiJuanBH is used to mark the papers, which is useful for the saving, reusing and looking up of the papers. KaoTiBH is used to mark the base documents for the application of the examination. TiGan stands for the path of the base documents that stored on server. CaoZuoTH stands for the numbers of a certain kind of questions on the papers which is important for the paper marking part to check the papers in order. The CaoZuoFW field is mainly used to locate the test points during the paper marking. According to the collation of knowledge points, it finds out the setting scope of one word document, including one entire document, a paragraph, a title, a sheet and so on. Paper marking part can find the location of the knowledge points in the document through the

field, which is a problem of choosing the specific object interface when it comes to specific program development. CaoZuoi, $i = 1, 2, 3, 4, 5$, stands for the encoding of the aforementioned knowledge points. Paper marking part decides what test points appear in the question though these fields and use the DataTable object in the C # language to bind out all the test points. Then they will be separated by using particular symbols and become character strings for the paper marking part to check under the condition of opening candidates' document.

3.2 The design Thread of Marking

3.2.1 Marking Process

Apply COM automation technology to boot word processor transparently and open the word document. Invoke a series of methods through the provided COM interface, and extract the attributes of the open document[2]. Match the test subjects and extract the test point information that set in the test process from the database. Open the examinees' answer document simultaneously and grade according to the scoring criteria. The flow chat is shown in figure 1:

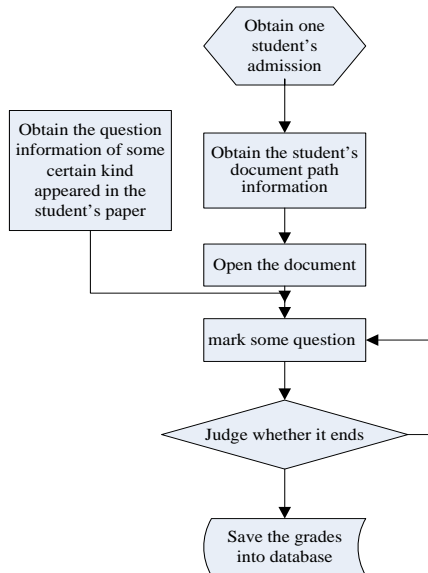


Figure 1. Marking Flow Chart.

3.2.2 Open Document

Prior to add references in visual studio2008, choose Microsoft Office 11..0 Object Library and Microsoft.Office.Interop.Word in the COM options so that you can access to the classes and functions of the components .Open document by using New word.ApplicationClass().applicationClass.Document ts.Open(ref ,filename,ConfirmConversions,ReadOnly,AddToRecentFiles,PassWordDocument,PassWordTemplate,

Revert,WritePassWordDocument,WritePassWordTemplate,Fromat,Encoding,doc_Visible),which filename is the name of the file which contain the filepath, ConfirmConversions decides whether to display the dialog of the file conversion , ReadOnly decides whether to open the files in read-only access, AddToRecentFiles decides whether to add the file to the recently used document list at the bottom of the file menu, PassWordTemplate is the password used while opening the template, Revert decides whether to open the document again when it has been opened, WritePassWordDocument is the password needed while saving the document , WritePassWordTemplate is the password needed while saving the template, Format points to the file converter when opening the document, Encoding refers to the document coding page for use, Doc_Visible decides whether the opened document window is visible [6].

3.2.3 Positioning and Obtaining of The Test Points

The key technical problem in the above process is the positioning and obtaining of the test points. The test point positioning means to find the accurate location of each testing point's information in the document, and it is very critical. In this paper, the test point positioning method is to be associated with a database table fields as well. You should first ensure the scope of the testing points in the process of setting papers, then select the testing points information.For example, the scope of the word testing points are divided into several parts such as the title, the whole document, the first paragraph, the second paragraph, the third paragraph, the table, the first row of the table, the second row of the table. The testing points are nowhere more than these locations, so it must be a person who is familiar with office and clearly know the testing points' classification to set questions. We will take an example, to set an operation question about the word art, to make it clear. When we set the question, set the title "the independent character" for the art word in three lines and four columns, we should choose "title" in the scope of the testing points of the window, then set art word value for "three lines and four columns" in the art word class which belongs to the font class of the window, and then save. Finally, the data appears just as the table 2 in the "tk_shijuanZGTPF" table of the database.



Table II: Field Table Of Subjective Scoring Part Of The Database

ShiJuan BH	KaoT iBH	Ti G an	CaoZ uoTH	CaoZuo FW	Cao Zuo1
2	80000 0069		1	a10000 02#the independ ent charater	ab00 35

According to the CaoZuoFW field values, judge whether it is a1000002. If it is, it can reflect that it's the title and the following behind # is the content of the title. Loop through CaoZuoi,i=1,2,3,4,5 in the table , and then extract out all information and bound to the object in Datatable. Loop through the virtual data table and link all of the testing points to a string. Using the string resolve strategy to resolve the string for the candidates' documents one-by – one marking in the process of scoring. This is to reduce the cycle operation of the database and enhance the efficiency of the system.

Control the cycle index by Word_.document.Shapes.Count, retain word.S-shapeRange sr by using doc.Shapes.Range(ref (object)i), and then decide whether TextEffect.PresetShape in sr matches with the name of the right answer. If matching, add it to the right item list of the marking parameter set in advance, else add to the wrong item list. This method is reading information of the right answer directly from the database. Compared with the method which comparing the standard answer document with the result document after saving the testing points, it's more accurate, for the latter may lead to inaccurate marking because of the standard answer errors.

4. CONCLUSION

The main work is the implementation of the framework of office automation applications based on B / S model, the office automatic marking system based on COM automation technology and the setting and extraction of the standard answer by using the way of interacting with the database. Compared to the method of recording steps to mark, this method is more convenient for developing, since it is not subjected to the operating environment and it reflects the students' proficiency in office applications more realistically. Compared to the method comparing with the standard document, it's more accuracy and the error rate is smaller. What's more, it's easy to maintain and extend. After several large-scale applications, this

system is demonstrated that it has the above advantages and it is stable and efficient which well solves the problem of computer automatic marking of the office operation questions.

REFERENCES:

- [1] YuFu, YuYe, The implementation of office operation automatic marking, 2010.6 Volume10.3
- [2] KemingTang, LingChen, The research and implementation of Word automatic marking, The Computer Engineering and Applications , 2008.44(35) ,P69-72
- [3] DingQiangDu, The office operation automatic marking using VBA technology , Computer Knowledge and Technology Vol.5,No.36, December 2009,p.10267-10268
- [4] ShuXiang Yu,Tingrong Xu, The research and implementation of Office automatic marking based on COM technology,2007.9 Volucom 10 ,3.P1-3
- [5] GuiChengLi, YongKuiZhang, RuiYan, The research and implementation of Word operation, 2002.4. Volume23, 4.P29-30
- [6] KemingTang,LingChen, The interface parsing of Word server tribute extraction,The Computer Engineering and Applications,2008,44(28)p79-82