

# 主观操作题的 XML 形式化表达及自动评阅<sup>①</sup>

陈 洁

(中华女子学院 计算机系, 北京 100101)

**摘 要:** 提出一种基于 XML 技术实现 Access 操作题自动评阅的策略及实现方法. 从数据库中提取对象的操作信息后, 首先参照操作点配置文件, 对操作信息进行过滤, 然后用 XML 形式化表示, 并将 XML 标记与数据库对象属性自动关联, 通过对 XML 文档的解析实现自动评阅. 详细介绍了配置文件的设置、XML 的形式化表达、适应操作题灵活性多样性的评阅策略和错误的定位反馈. 评阅系统在实际教学中得到有效应用, 不仅适合考试题, 也适合平时实验作业的评阅.

**关键词:** XML; Access 数据库; DAO 对象模型; Access 对象库; 自动评阅; 操作题

## Formal Expression of Subjective Operation in XML and its Auto-Checking

CHEN Jie

(Department of Computer, China Women's University, Beijing 100101, China)

**Abstract:** An automatic checking strategy and realization method for Access operations based on XML technique was given. First, operation information was extracted into a text file. Then, the data of txt format was expressed in simple and easy XML as filtered by the configuration file recording operation points, with the tag was related to database Attributes. Finally, the XML documents from the model and the student were analyzed and compared to implement checking. In the article the method of the extraction, the construction of the configuration, the XML representation, the checking strategies accommodated to flexibility and diversity, and the errors locating was discussed in details. The system has been used in teaching effectively. It is applied to the test as well as homework.

**Key words:** XML; access database; DAO object model; access object library; auto-checking

Access 数据库课程是我国很多高校开设的一门计算机公共基础课程, 操作性强且操作灵活, 为实现同一要求, 可以采用不同的步骤, 使用不同的命令式子, Access 操作题自动评阅方法的研究成为很多学者关注的问题<sup>[1-4]</sup>.

文献[1]、[2]介绍了获取数据库对象操作信息的方法, 可以将 mdb 格式中的数据库操作信息存储到 txt 格式的文本文件中, 通过检索文本文件的内容实现自动评阅; 由于获取的对象信息繁杂且不是所有数据都是需要评阅的, 因此, 直接检索文本文件效率较低. 文献[3]提出基于 XML 技术实现机器自动评阅, 根据给定的 DTD 规范建立 XML 文档, 存储表和查询的答案信息, 再与 DAO 对象读取的操作信息进行比对; 该

方法只能用于评阅表和查询对象, 并且代表对象属性的标记名都是自定义的, 不利于应用的扩展和程序的维护.

Access 中的表、查询、窗体、报表和宏 5 种对象都是通过设置属性的方式来设计的, 且每个对象都是由若干元素组成的, 各元素之间形成一定的层次关系, 而 XML 具有自描述性、可扩展性和层次性等特点, 非常适合表达这些数据库对象.

针对上述问题, 本文以 Access 2010 中建立的数据库为研究对象, 通过分析 Access 数据库各对象的结构特征, 将 XML 标记与数据库对象属性相关联, 利用提取的数据库操作信息自动生成 XML 文档, 这种表达方式具有清晰的语义并能保持数据库对象的结构特征,

<sup>①</sup> 基金项目: 中华女子学院科研基金资助重点项目(JS2013-02009)

收稿时间: 2013-11-25; 收到修改稿时间: 2013-12-26

易于检索, 适合机器评阅, 也适合错误信息的定位与反馈。

## 1 设计思想

评阅系统分为 3 个部分: 操作信息的提取、XML 的形式化表达和操作信息的评阅, XML 的形式化表达是整个评阅系统的核心。

XML 技术在计算机自动评阅系统中得到广泛应用<sup>[5-8]</sup>。XML 以一个统一、开放、基于文本格式的模式来描述和交换数据, 具有强大的信息描述能力, 不仅能够表示文档内容, 还能够表示文档的结构。

系统的逻辑框架如图 1 所示。其中, 样文数据库是存储正确操作结果的数据库, 作业数据库是学生提交的数据库。

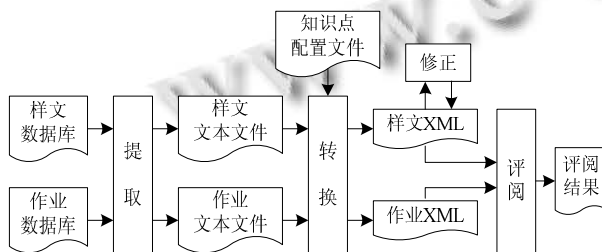


图1 自动评阅系统的逻辑框架

首先, 利用 DAO 对象模型和 Access 对象库提取数据库对象的操作信息<sup>[1,2]</sup>, 然后用 XML 形式化表达, 对数据库文件的评阅就转换为对 XML 的 DOM(Document Object Model, 文档对象模型)解析。

## 2 操作信息的提取

操作信息的有效提取是进行系统评阅的前提, 对于不同的 Access 对象采用不同的提取方式, 并以文本格式存储。

### 2.1 表和查询对象信息的提取

Access 表对象中包含了数据表、字段、索引、关系等数据库的元数据, 并用相关属性描述;此外, 表对象中还存储了用户输入的记录。

利用 DAO 对象模型, 可以获取 Access 表对象和查询对象的信息。一个 Database 对象对应一个数据库, TableDef、Field、Index、Relation、QueryDef、Recordset 等子对象分别对应表、字段、索引、关系、查询、记录集等数据库元素。表对象的操作代码片段如下:

```
Set db = OpenDatabase("e:\accTest\jxgl.mdb")
```

```
FileNum = FreeFile
```

```
Open "e:\accTest\tbl.txt" For Append As #fileNum
```

```
For Each tbl in db.TableDefs '遍历数据库中的表
If tbl.Attributes = 0 Then
```

```
For Each p in tbl.Properties '遍历表的属性
```

```
If p.Name = "Name" Then
```

```
Print #fileNum, "#Table =" & p.Value '表名
```

```
Else
```

```
If p.Value <> "" Then '其他属性
```

```
Print #fileNum, "#" & p.Name & "=" & _
```

```
p.Value
```

```
End If
```

```
End If
```

```
Next
```

```
'访问 Field 对象, 获取字段属性
```

```
'访问 Index 对象, 获取索引属性
```

```
End If
```

```
Next
```

```
'访问 Relation 对象, 获取关系属性
```

```
Close #fileNum
```

在文本文件中, 信息是按表对象的结构特征, 并以“属性名 = 属性值”的形式存储, 与下文使用 SaveAsText 方法自动导出的窗体和报表信息表达方式保持一致。而且, 对象名都是显示在该对象其他属性的前面, 并过滤掉了没有属性值的数据项。

对于查询操作, 利用 QueryDef 对象可以获取相应的 SQL 语句。

### 2.2 窗体、报表和宏对象信息的提取

使用 Access.Application 对象的 SaveAsText 方法, 可以将数据库对象的内容转换为文本格式, 方法的声明为:

```
Application.SaveAsText <objType>,<objName>,  
<fileName>
```

<objType>表示对象类型, 取值为 acForm(窗体)、acReport(报表)和 acMacro(宏)。

<objName>表示要转换的窗体、报表和宏对象的名称。

<fileName>表示保存信息的文本文件名。

导出的信息自动按对象的结构特征, 如窗体对象按窗体页眉、主体、窗体页脚等节的顺序存储, 且属性项均以“属性名 = 属性值”的形式表示。

对于 Access 2010 中建立的 accdb 格式的数据库文件, 导出的宏操作信息以两种形式表示: 一种是“属性名=属性值”的形式, 另一种直接就是 XML 格式, XML 格式清晰地描述了 If...Else...EndIf 程序块的逻辑结构及其嵌套关系。

### 3 XML的形式化表示

#### 3.1 操作点配置文件

Access 系统为表、窗体、报表等数据库对象设置了很多属性, 利用上述方法提取的对象信息比较繁杂, 而且不是所有的数据项都要评阅, 为提高评阅效率, 可以设置一个操作点配置文件筛选出需要评阅的操作信息。

根据教学大纲或考纲的要求, 对需要考查的知识点进行分解, 将 Access 对象的各个属性映射为操作点<sup>[9]</sup>, 并将这些操作点存储在相应的配置文件中, 其表达形式为:

```
opType = {attr1,attr2,attr3,...,attrn}
```

其中, opType 代表操作点的类别, attr 代表该类别包含的各个操作点, 这些操作点代表了操作题中所有需要检查的项目。类别主要按数据库对象的组成元素来划分, 类别的名称是自定义的, 操作点的名称则使用相应的属性名标识。

##### (1)表对象的配置文件

该文件包含 tbl、fld、idx、rel 四个类别, 分别代表表对象中所有需要检查的表属性、字段属性、索引属性和关系属性。如, fld 类别表示为:

```
fld = Type,Size,DefaultValue,Caption,...
```

其中, Type、Size、DefaultValue、Caption 等均为字段的内部属性标识符, 分别表示字段的数据类型、大小、默认值、标题等属性。

##### (2)窗体和报表对象的配置文件

该文件包含 frm、label、textBox 等类别, 分别代表窗体及各类控件。如, label 类别表示为:

```
label = Left,Top,Caption,...
```

报表对象的配置文件的构成与窗体类似。

查询对象和宏对象不需要设置操作点配置文件。

#### 3.2 XML 文档结构

XML 的结构表现为语义单元相互嵌套的层次关系, 通过 XML DOM 节点的层次关系可以反映数据库对象各元素之间的关系, 通过 XML DOM 节点及其属

性可以反映数据库对象的属性和状态, 可见 XML 文档能够清晰地表达数据库对象的结构和属性信息。各数据库对象的 XML 形式化表达如下:

##### (1)表

一个 <Table> 节点代表一个表对象, <Attr>、<Field>和<Index>为子节点, 分别代表表级属性、字段及索引等表的组成元素;<Relation>节点代表表间关系。各节点的 Name 属性分别标识表对象的名称、字段名、索引名和关系名。最底层的元素节点描述各个属性项, 并以 Access 系统内部的对象属性标识符作为节点的标记名, 文本节点记录属性值。用 DTD 描述如下:

```
<!ELEMENT Tables (Table+,Relation*)>
<!ELEMENT Table (Attr,Field+,Index*)>
<!ELEMENT Attr (RecordCount,ValidationRule?,...)>
<!ELEMENT RecordCount (#PCDATA)>
...
<!ELEMENT Field (Type,Size,DefaultValue?,...)>
<!ELEMENT Type (#PCDATA)>
...
<!ELEMENT Index (Index-Key,Primary,Unique,
    Order)>
...
<!ELEMENT Relation (Table,ForeignTable,Foreign)>
...
<!ATTLIST Table Name CDATA #REQUIRED>
<!ATTLIST Field Name CDATA #REQUIRED>
<!ATTLIST Index Name CDATA #REQUIRED>
<!ATTLIST Relation Name CDATA #REQUIRED>
```

##### (2)查询

一个 <SQL> 节点代表一个查询对象, 节点的名称属性标识查询对象的名称, 每个子节点代表一个 SQL 子句, 子命令名作为子节点的标记名, 文本节点记录子句内容。例如, 代表查询条件的 Where 节点表示为:

```
<Where>系部="CS"</Where>
```

##### (3)窗体和报表

其结构与组成窗体和报表的各节对应。

一个 <Form> 节点代表一个窗体对象, <Attr>、<FormHeader>、<PageHeader>、<Section>、<PageFooter>、<FormFooter>作为 <Form> 节点的子节点, 分别代表窗体属性以及窗体页眉、页面页眉、主

体、页面页脚、窗体页脚等组成元素, 这些节点还可以再包含<Label>、<CommandButton>、<TextBox>等代表控件的子节点. 最底层的元素节点描述各个属性项, 并以 Access 系统内部的对象属性标识符作为节点的标记名, 文本节点记录属性值.

报表与窗体类似, 一个<Report>节点代表一个报表对象, 此外, 还可以有<BreakHeader>和<BreakFooter>代表组页眉和组页脚的节点.

<Form>和<Report>节点用 Name 属性标识窗体或报表对象的名称, 各控件节点用 Name 属性标识相应的控件对象的名称.

#### (4)宏

直接从提取的宏对象信息中截取 XML 表达部分.<Sub>节点代表一个子宏, <ConditionalBlock>节点代表一个 If 程序块, <If>和<Else>是其子节点, <Condition>节点表示 If 条件, <Action>和<Argument>节点分别表示宏操作命令和相应的操作参数, 文本节点记录操作信息. 图 2 所示是宏的一个 XML 表达片段.

```

- <Sub Name="login">
- <Statements>
- <ConditionalBlock>
- <If>
  <Condition>[Forms]![登录]![txtAccount]="123abc"</Condition>
- <Statements>
- <Action Name="MessageBox">
  <Argument Name="Message">登录成功, 欢迎使用! </Argument>
  <Argument Name="Type">Information</Argument>
  <Argument Name="Title">密码判断</Argument>
</Action>
+ <Action Name="CloseWindow">
</Statements>
</If>
+ <Else>
</ConditionalBlock>
</Statements>
</Sub>

```

图2 宏对象的XML形式化表达

按照配置文件中的设置, 将提取到文本文件中的操作信息用上述 XML 形式化表达.

对于系统自动建立的样文 XML 文档, 有时还需要根据具体题目的评阅要求适当修改.

① 删除操作点配置文件中包含的但本题不需要考查的项目. 如, 配置文件中设置了标签控件的 Top 和 Left 两个考查点, 如果在设计报表的题目中不需要检查某个标签控件的位置信息, 则可以从 XML 文档中删除这个<Lable>节点下的<Top>和<Left>子节点.

② 如果某项设置有多解, 可以将节点值修改为答案组的形式. 例如, 查询“学号”的前 2 位为“09”的学生记录, 则<Where>节点可表示为:

<Where>(Left([学号],2)="09")|(Mid([学号],1,2)="09")</Where>, “|”表示多选一.

## 4 试题评阅与错误反馈

将数据库信息存储为 XML 格式后, 数据库内部各个对象及其属性通过 XML 技术被表示为元素节点或者文本节点, 对数据库文档的评阅就转换为对 XML 文档的解析, 而解析 XML 文档的方法就是使用 DOM<sup>[10-12]</sup>. 通过对 DOM 树的节点进行深度遍历完成评阅.

评阅过程是按照样文 XML 文档中的对象顺序进行的, 遍历样文 XML 文档的各个节点, 根据节点的 Name 属性获得要评阅的对象, 然后根据对象名查找作业文档, 若作业文档中包含该对象, 则按上述方法继续定位该对象的子对象直到取得属性值.

### 4.1 评判策略

数据库对象的属性值并非都是惟一的, 凡是使用表达式设置的属性, 如字段的有效性规则、查询条件、窗体或报表中用计算表达式作为控件来源等, 属性值就可能有多答案.

另外, 对象的属性值还可能由多个数据项构成, 从而使得操作顺序的不同产生不同的设置结果. 这种情况在查询对象的设计中最常见. 例如, 添加字段的顺序不同会导致 SELECT 子句中输出项的顺序不同; 在涉及 3 个表的查询中, 添加表的顺序不同会导致 2 个联接条件在 FROM 子句中出现的顺序不同, 但这种排列顺序上的不同并不会改变查询结果的信息内容与记录顺序. 而对于排序项, 虽然也可以由多个数据项构成, 但排列顺序的不同却会影响查询结果的信息显示顺序.

由于操作题的灵活性和操作结果的多样性, 需要根据不同的情况采用不同的评判策略.

① 多个答案值的评判 根据样文 XML 文档中的设定, 采用正则表达式或拆分答案组的方式检查学生文档中的操作结果.

② 顺序相关的多数据项的评判 将各数据项作为一个整体, 直接比对. 若要区分错误情况, 可以拆分为数组后再比对.

③ 顺序无关的多数据项的评判 取得节点值后, 先排序, 再比对.

## 4.2 错误反馈

对于错误操作,应该生成明晰的错误报告,这在实验作业的评阅中尤为重要,可以帮助学生及时改正错误.上述建立的XML文档具有清晰的语义并保持了数据库对象的结构特征,能够很方便地按树状层次报告错误的位置;而且,存储属性项的节点的标记名都是以Access对象属性标识符表示的,能够直接根据标记名指出错误源.

为了使错误报告更易于阅读,还需要将报告中的所有属性标识符“翻译”为中文,方法是:建立一个文本格式的外部配置文件,以“名-值”对的形式为操作点配置文件中的每个操作名标识符(即属性标识符)提供相应的中文描述.生成错误报告时就利用这个配置文件进行翻译.

## 4.3 成绩评定

对Access操作题的评阅主要就是检查每个操作点的完成情况.如果各操作点的分值是平均分配的,可以设定每个操作点的基准分值为1;若设置错误,则分值为0.5;若没有设置,则分值为0.实际完成情况与应完成情况的比值乘以100,就是最终的操作成绩,计算公式如下:

$$s = \frac{\sum_{j=1}^m V_{a_j}}{\sum_{i=1}^n V_{A_i}} * 100$$

式中, $A_i$ 表示样文XML DOM中的一个节点, $V_{A_i}$ 表示该节点的成绩分值, $V_{A_i}=1(i=1,2,\dots,n)$ . $a_j$ 表示学生文档中匹配的操作点, $V_{a_j}$ 表示该操作点的分值,若操作正确,则 $V_{a_j}=1$ ;否则 $V_{a_j}=0.5$ .

例如,题目中有10个操作点,学生作答的结果中有6个正确、3个错误,1个没有设置,则:

$$s = (6*1 + 3*0.5) / 10 * 100 = 75.$$

若要区分不同操作点的难易度,可以在样文XML的相应节点增加一个权重属性Weight.例如,字段有效性规则对应的节点可表示为:

```
<ValidationRule Weight="1.5">...
</ValidationRule>
```

如果操作点的分值不是平均分配的,可以在样文XML文档的相关节点增加一个分值属性Score.例如,可以为“学生”表中的“学号”字段分配5分,为“出生日期”索引分配3分:

```
<Field Name="学号" Score="5">...</Field>
<Index Name="出生日期" Score="3">...</Index>
```

## 5 结语

本文针对Access数据库对象的设计特点和结构特征,将存储在Access数据库中的操作信息统一表示为XML形式,通过对XML文档的解析实现机器自动评阅.

操作点配置文件对数据库信息进行了有效的筛选.并且,当需要增加或删除操作点时,只要修改相应的配置文件,而不必修改代码.使用本策略建立的评阅系统先后应用于Access2003和Access2010数据库操作的评阅,Access2010新增了“计算”类型的字段,为此只需要在表对象的操作点配置文件中增加Expression和ResultType两个属性项.

XML文档清晰地表达了数据库对象的结构和内容,通过将XML标记与数据库对象属性相关联,不仅提高了系统的评阅效率,而且易于错误的定位和反馈,也易于应用的扩展.在评阅过程中,还充分考虑到操作题的灵活性,尤其是查询语句的多样性表达,增强了系统的评阅能力.

评阅系统已在实际教学中得到有效应用,特别是平时实验作业的评阅,可以大大简化出题过程,不必为每一题一一设置考查点.同时,能够快速而准确地反馈评阅结果,既便于学生及时改正错误,也利于教师及时了解教学情况,从而把教师从人工评阅的繁重工作中解放出来,将更多时间和精力投入教学研究中.

## 参考文献

- 1 杨芳菊,李梅.Access机试系统的评分方法探讨.计算机应用与软件,2011,28(9):182-184.
- 2 李富星,牛永洁.Access自动评分方法研究.信息技术,2010,(8):170-171.
- 3 胡海峰.基于XML的Access自动阅卷技术研究与实现.甘肃联合大学学报(自然科学版),2009,23(7):37-42.
- 4 王卫国.基于VB的ACCESS无纸化考试系统的设计与实现.微计算机应用,2007,28(1):108-112.
- 5 陈尧妃,陈焕通,倪应华.基于XML的SQL Server数据库技能操作自动阅卷.计算机系统应用,2010,19(3):158-161.
- 6 赵晓峰,肖丹.基于XML与ACTIVEX技术的Office操作自动评分系统.计算机与数字工程,2006,34(12):178-180.
- 7 丁俊辉,龚沛曾.基于XML的HTML自动阅卷算法设计与实现.计算机技术与发展,2012,22(1):1-4.
- 8 梁妍.基于XML的考试评分系统的设计与实现.中国教育技术装备,2007(3):57-59.
- 9 汤克明,陈峻.Word自动阅卷系统的设计与实现.计算机工程与应用,2008,44(35):69-72.
- 10 孔令波,唐世渭,杨冬青等.XML数据的查询技术.软件学报,2007,18(6):1400-1418.
- 11 朱艳红,何东彬.基于XML文档检索的存储研究与实现.计算机系统应用,2007,16(5):49-52.
- 12 廖述梅,万常选,徐升华.XML信息检索探究.情报学报,2007,26(2):229-234.