

标准化试卷的文法的构建

一、背景描述

试题库自动构建系统是一个非常实用的应用，可以帮助教育机构快速构建试题库、组织试卷、提高试题的质量和效率。其中，试题库自动构建系统的核心技术是利用计算机程序从现有的试卷中抽取试题，按照不同题型构建试题库或自动组卷。在试题库自动构建系统中，标准化试卷的文法设计是一个关键的环节，它可以影响系统的准确性、效率和可扩展性。

本研究项目的目标是实现一个试题库自动构建程序，该程序将对现有的试卷进行词法分析、语法分析以及语义分析，并将其中的不同题型的试题进行识别与提取，最后依题型构建试题库。本文给出了一个含有填空题、判断题、选择题的标准化试卷文法设计，为试题库自动构建程序明确了理论基础。

二、文法定义

试题库自动构建系统所处理的对象是现有的试卷，试卷的结构可以被表示为以下的文法定义：

$$G[< \text{试卷} >] = (V_N, V_T, P, < \text{试卷} >) \quad (1)$$

其中非终结符号集 V_N 如下：

$$V_N = \{< \text{试卷} >, < \text{试卷头} >, < \text{考试时间} >, < \text{考试科目} >, < \text{题型列表} >, \dots\} \quad (2)$$

终结符号集 V_T 如下：

$$V_T = \{A, B, C, D, \text{年}, \text{试卷}, \dots\} \quad (3)$$

P 中产生式如下：

```
<试卷> ::= <试卷头> <题型列表>
<试卷头> ::= <考试时间> <字符串> "试卷"
<考试时间> ::= <数字> "年"
<题型列表> ::= <题型> | <题型> <题型列表>
<题型> ::= <题型号> <题目类型>
<题型号> ::= "一、" | "二、" | "三、"
<题目类型> ::= <填空题头> | <判断题头> | <选择题头>
```

```

<填空题头> ::= "填空题：" <填空题列表>
<填空题列表> ::= <填空题> | <填空题> <填空题列表>
<填空题> ::= <题号> <字符串> <填空题答案>
<填空题答案> ::= "答案：" <字符串>

<判断题头> ::= "判断题：" <判断题列表>
<判断题列表> ::= <判断题> | <判断题> <判断题列表>
<判断题> ::= <题号> <字符串> <判断题答案>
<判断题答案> ::= "答案：" ("正确" | "错误" | "对" | "错" | "T" | "F")

<选择题头> ::= "选择题：" <选择题列表>
<选择题列表> ::= <选择题> | <选择题> <选择题列表>
<选择题> ::= <题号> <字符串> <选项列表> <选择题答案>
<选项列表> ::= <选项> | <选项> <选项列表>
<选项> ::= ("A" | "B" | ... | "Z") "." <字符串>
<选择题答案> ::= "答案：" <答案列表>
<答案列表> ::= ("A" | ... | "Z") | ("A" | ... | "Z") <答案列表>

<题号> ::= <数字> "."
<数字> ::= <数字> <数字> | "0" | "1" | "2" | ... | "9"
<字符串> ::= "" | <字符> <字符串>
<字符> ::= "a" | "b" | "c" | ... | "z" | "A" | "B" | "C" | ... | "Z" | "0" | "1" | "2" | ... |
    "9" | " " | "-" | "_" | "+" | "=" | "." | "@" | "#" | "$" | "%" | "^" | "&" | "*" | "(" |
    ")" | "[" | "]" | "{" | "}" | ";" | ":" | "'" | "," | "<" | ">" | "/" | "?" | "!" | "请" |
    "下" | "阅" | ... | "答"

```

三、单词列表

试卷文法的单词列表如表1所示。

表1 试卷文法单词列表

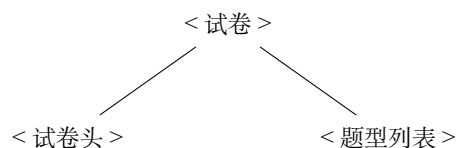
单词名称	单词值
< 试卷头 >	由考试时间和字符串组成，并以“试卷”结尾
< 考试时间 >	由数字和“年”组成
< 题型号 >	由“一、”、“二、”或“三、”之一组成
< 填空题头 >	由“填空题：”和填空题列表组成
< 填空题答案 >	由“答案：”和字符串组成
< 判断题头 >	，由“判断题：”和判断题列表组成
< 判断题答案 >	由“答案：”和“正确”、“错误”、“对”、“错”、“T”或“F”中的一个组成

< 选择题头 >	由“选择题:”和选择题列表组成
< 选项 >	由“A”、“B”...“Z”和字符串组成
< 选择题答案 >	由“答案:”和“A”、“B”...“Z”中的一个或多个组成
< 题号 >	由数字和“.”组成
< 字符串 >	由任意字符组成
< 数字 >	由“0”到“9”中的一个或多个数字组成
试卷	-
年	-
填空题:	-
答案:	-
判断题:	-
正确	-
错误	-
对	-
错	-
T	-
F	-
选择题:	-
A	-
B	-
C	-
D	-
一、	-
二、	-
三、	-

四、 结构举例

下面将以图1中所示试卷为例，给出被处理对象的结构。

根据图1中所示试卷，可以绘制如下语法树：

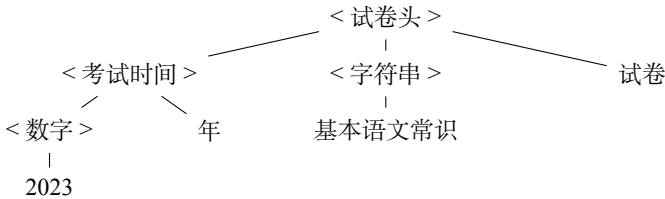


2023 年基本语文常识试卷

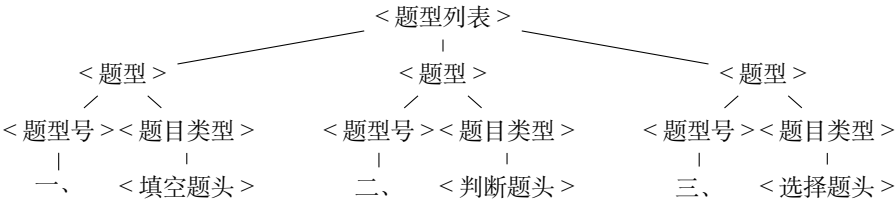
- 一、填空题：
1. 昨天我去了一家_____店买手表。答案：钟表
- 二、判断题：
2. 中国的首都是北京。答案：正确
- 三、选择题：
3. 以下哪个国家不属于欧洲？
A. 法国 B. 英国 C. 美国 D. 德国
答案：C

图 1 用例试卷

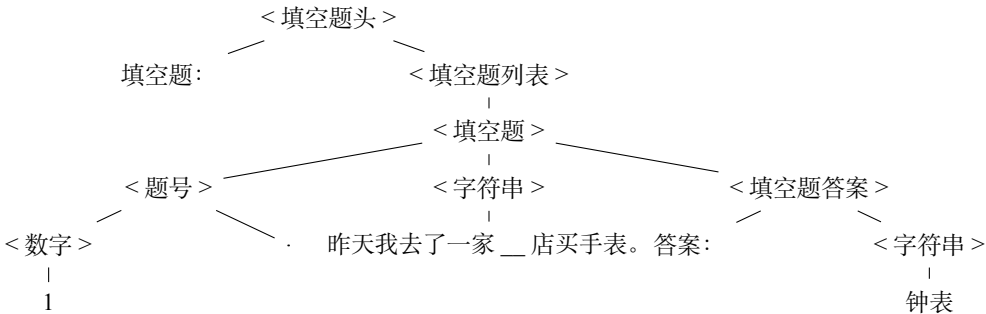
其中 < 试卷头 > 的语法树如下



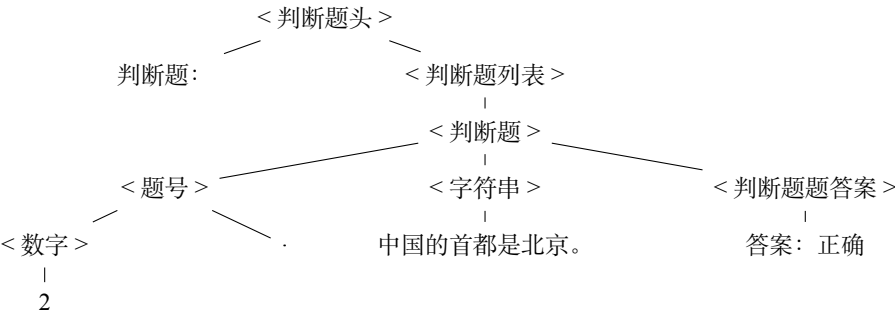
其中 < 题型列表 > 的语法树如下



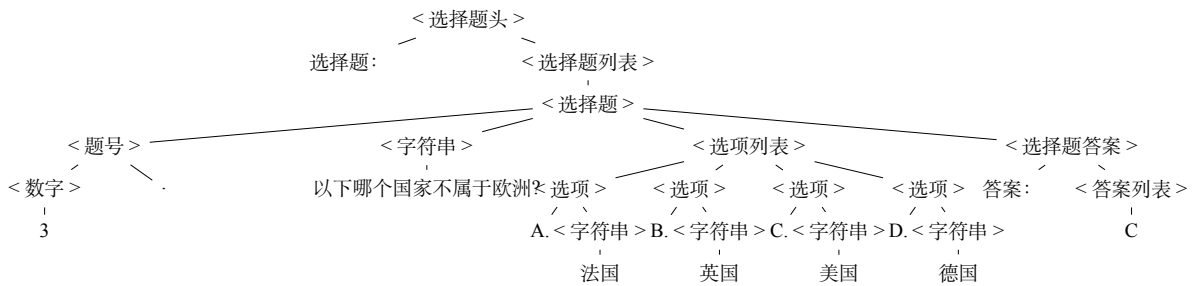
< 填空题头 > 的语法树如下



< 判断题头 > 的语法树如下



< 选择题头 > 的语法树如下



五、任务明确

5.1 词法分析

识别出试卷中的各个最基本元素，包括试卷名称、考试时间、题目、答案、选项和正确答案等单词。

5.2 语法分析

根据试卷的文法定义，识别出试卷的结构，包括试卷头、试题列表和各种题型的具体内容。

5.3 语义分析

识别出被处理对象后执行的处理/应用目的，即试题库自动构建程序从现有试卷中抽取试题的各个部分，按不同题型构建试题库，或者按给定要求自动组卷。具体而言，程序需要从试卷中识别出各种题型的题目、选项和答案等信息，存储到试题库中，然后根据用户的要求自动组卷生成新的试卷。在自动组卷的过程中，程序需要根据用户要求的试卷结构和难度等级，从试题库中抽取相应数量和类型的试题，生成符合要求的新试卷。