

基于母题的简单算术应用题出题系统

Research on the Examination System of Simple Arithmetic Word Problem Based on Motif

贺培鑫

Peixin He

江西财经大学

中国·江西 南昌 330013

Jiangxi University of Finance and Economics,

Nanchang, Jiangxi, 330013, China

【摘要】论文提出了基于母题的简单算术应用题出题系统的设计思想和实现方法。该系统利用简单神经网络和自动分词技术来实现通过对母题的分析,再经过关键词替换和算术运算公式得到多道不同子题以及答案,供出题者选择。

【Abstract】This paper puts forward the design idea and implementation method of the examination system of simple arithmetic word problem based on Motif. The system analyzes Motif by using the simple neural network and automatic word segmentation technology, then the test paper can choose the word problems from a number of different sub-questions and answers through keyword replacement and arithmetic operation formula.

【关键词】母题;题型分类;自动出题

【Keywords】motif; type classification; automatic problem setting

【DOI】10.36012/sde.v2i3.1645

1 引言

小学数学的算术应用题无论在平时作业还是考试中都是重点,教师在给学生布置作业或设计考试题目时,主要利用人工出题或题库出题的方式。然而人工出题和从题库选择出题都需要很多的准备工作,这使得出题过程耗时长而且题目的质量也得不到保障。基于母题的简单算术应用题出题系统通过对母题的分析生成类似的子题,可以生成大量与母题运算结构相同的题目,避免了人工出题和题库出题过程中的弊端。

2 题型分类

在数学算数应用题中,题型分类有很多种方法,一般情况下主要依据题型的计算方式由加、减、乘、除四则运算来区别题目类型^[1]。但是在题目中也会出现某些带有特征性或关键性的信息也作为题型分类的指标,由于在生成题目的期间要从不同的语料库中调用词语来替换母题中的可替换词,本系统首先根据题干中的带有特征性的关键词将算术应用题分为了人物类型题目,地点类型题目。

所谓人物类型题目指的是题目中主要以人物或人物活动充当题干,是关于人物本身、人物与人物之间存在某种数学关系的题目。例如:“小明有 8 个苹果,小亮有 5 个苹果,他们两个人加起来有多少个苹果?”该问题描述的就是两个人物“小明”和“小亮”之间产生的数学问题。

地点类型题目指的是题目中主要以地点、或关于某地点发生的活动充当题干,是关于地点本身、地点与地点之间存在某种数学关系的题目。例如:“电影院有 6 排座椅,每排座椅能坐 15 位观众,请问这个电影院一共能容纳多少观众?”该问题就是描述“电影院”本身产生的数学问题。

本系统在此基础上再结合算术应用题中加、减、乘、除四则运算的不同的运算公式和运算方式,将题目类型进一步的划分。例如:在简单加法算数应用题中,求解的方式为确定题干中已知的两个数求,第三个未知数,根据“加数+加数=和”这个运算公式得到相应的答案。因此,本系统将这三个参数分别作为未知数再次确定了 3 种新的题型,当运算公式中的第一个“加数”未知时,可以将题目划分为起始量未知的加法应用题,当运算公式中第二个“加数”未知时,可以将题目划分为转移量未知的加法应用题,当运算公式中“和”未知时,可以将题目划分为结果量未知的加法应用题。减法、乘法、除法同理,便得到了 12 种不同的题型,再结合人物类型和地点类型题目的分类,本系统一共将简单算术应用题分为了 12 种人物类型题型和 12 种地点类型题型,共计 24 种题型。

3 构建语料库

语料库的大小和精准性决定了系统可以生成子题的数量和质量,因此语料建设在系统中也极其重要。本系统主要将母题作为模板生成类似于母题的子题库,而且人物类型和地点

类型都有不同的词句搭配,这也是构建语料库的难点。

本系统通过替换母题中可替换的语料生成子题,但母题中的可替换词并不是由系统随意替换的。例如:母题为“小明有一些玩具,其中有 2 个小汽车,3 个变形金刚,请问小明一共有多少个玩具?”,在这道人类结果量未知的母题中可替换的词有人物名“小明”,物品词“小汽车”“变形金刚”,还有问题词“一共”。在替换过程中为了保证题目不出现歧义性,“小明”必须对应替换为人名,“小汽车,变形金刚”要替换为相应的玩具名,“一共”也要替换为相应的问题词,否则就会出现问问题,所以本系统针对不同需求以字典的形式构建了对应的语料库。例如:人名库、人物动作库、地点库、物品库、介词库、问题词库等。库与库之间也有对应的联系,所以不仅要构建语料库,还要建立各个库之间的关系,本系统利用用键值对的形式把它们联系起来。

4 分析母题生成相应的子题

首先在对母题的分析中,本系统采用了结巴分词的自然语言处理方式对母题进行标记并划分为词段,利用简单神经网络统计母题中各词的词性和出现的词频,并收集母题中的数字,最终得到母题中不可替换的题干和可替换的语料^[9]。其次在子题生成的过程中进行语料替换。具体实现思路如下:

例如:母题为“动物园里有 5 只老虎,猴子的数量是老虎的 5 倍,请问动物园里有多少只猴子?”

①根据母题的问题类型将母题划分为地点类型结果量未知的乘法应用题。

②将母题用字符串的方式写入程序中,且数字之间分别用空格隔开,方便分词和统计数字。

```
import jieba
```

```
str='动物园里有 5 只老虎,猴子的数量是老虎的 5 倍,  
请问动物园里有多少只猴子?'
```

③提取母题中的数字存放于数组中{5,5},确定题干中只出现数字的个数,并调用随机函数生成随机的数字。

```
tn = [int(s) for s in str.split() if s.isdigit()]
```

```
t[0] = tn[0]
```

```
t[1] = tn[1]
```

④对母题进行标记和分词,分词结果为:

```
str =jieba.cut(text,cut_all=False)
```

“动物园里”“有”“只”“老虎”“猴子”“的”“数量”“是”“老虎”“的”“倍”,“请问”“动物园里”“有多少”“只”“猴子”“?”

在此过程中标记出母题中每个字符所在的位置;

⑤通过语料库对比得到不可替换的题干,结果为:

“里”“有”“的”“数量”“是”“的”“倍”“请问”“?”

可替换的语料为:“动物园”“只”“老虎”“猴子”“老虎”“动物园”“有多少”“只”“猴子”

⑥确定不可替换的题干后,从语料库中选择相应的语料来替换母题中可替换的语料,并将母题中的数字用随机生成数替换;

```
if len(replacements) > 0:
```

```
replacement = replacements[randint(0, len(replacements) - 1)]
```

```
output = output + " " + replacement
```

```
else:
```

```
output = output + " " + words[i]
```

⑦根据母题类型对母题和子题列式和计算结果。

```
def answer(self):
```

```
self.k=self.t[1]+self.t[0]
```

```
return str(abs(self.k))
```

以上七步是针对分析母题生成第一道子题的具体实现方法,只需设置生成子题的个数重复循环第六第七步就可以生成大量的子题和答案来满足出题人的需求。

5 关于本系统的补充

为了更好地满足出题人的需求,本系统针对 24 种不同的题型都进行了预处理,分别针对 24 种题型提前设置了母题,出题人可以直接使用系统出题的功能生成某一类题型的题库和答案,也可以根据出题人的意愿选择不同题型和不同的题目数量生成加、减、乘、除的混合题库实现自动出题的功能。另外,本系统所使用的语料库均为人工创建,未使用外部语料接口,也是人们日常生活中经常使用或熟悉的一些语料,在学生作答时简化了他们对题目的理解,可以更好地完成作答。

6 结语

本文介绍了基于母题的简单算术应用题系统的实现方法,并通过案例展示详细说明了系统实现的思路。大多数简单算数应用题都可利用本文介绍的方法实现智能化出题,可以借助本系统帮助出题人解决切实的问题^[9]。

参考文献

[1]谭文丽.小学数学应用题阅读能力调查及对策研究[D].重庆:重庆师范大学,2019.

[2]曾小芹.基于 Python 的中文结巴分词技术实现[J].信息与电脑(理论版),2019,31(18):38-39+42.

[3]赵天彤.小学低年级学生数学阅读理解调查研究[D].天津:天津师范大学,2014.