

# C语言上机考试系统中自动评分方法的探讨

王 颖

(湖南科技大学 计算机科学与工程学院, 湖南 湘潭 411201)

**摘 要:**利用计算机考核学生的C语言编程实践能力,实现编程题的自动评分,已成为计算机辅助教学的一个热点。依据程序设计题特点,总结了目前能应用于实践的几种C语言编程题的自动评分方法,并对这几种自动评分方法进行了探讨与分析。

**关键词:**C语言;自动评分;辅助教学

**中图分类号:**TP31      **文献标志码:**A      **文章编号:**1674-5884(2014)01-0149-02

目前,计算机C语言已作为各类高校计算机专业甚至其他相关专业的必修课程。由于学习该课程的学生人数多,为了提高教学效率,避免手工阅卷时出现的费时且易出错等问题,必须采用能确保公正、客观的计算机自动评分系统。目前,国内在C语言的上机考试系统中选择题和填空题的自动阅卷技术已经很成熟,而对于评价实践能力的C语言编程题,计算机自动评分技术还处于发展与完善中。自20世纪60年代开始,国内外许多学者致力于研究程序设计题计算机自动评分方法。发展至今,目前较完善的C语言编程题自动评分技术已可总结为好几种。

## 1 基于程序理解的编程题自动评分方法

程序理解是通过分析程序,对程序进行抽象、推理来获取知识。该方法将人工阅卷的某种思维过程和程序理解的几个基本策略进行结合,通过比较模板程序与学生程序,并给学生程序进行评分<sup>[1]</sup>。其评分过程如下<sup>[2]</sup>:

- (1)首先将程序代码转换成系统依赖图(SDG)这种中间表示形式;
- (2)采用标准化处理方法转换SDG,消除程序实现形式的多样化;
- (3)通过匹配学生程序和模板程序给学生程序评分——从程序的规模、结构、深度等几个方面着手。

根据文献[1]中给出的实验结果数据可知,这种自动评分方法符合人工阅卷的思维过程,总体上优于人工评分,且能应用于实际需要,只需提供模板程序即可,不过这种方法对模板程序要求较高。

## 2 基于语义相似度的编程题自动评分模型

文献[3]提出的基于语义相似度比较的编程题自动评分模型对程序的一些标准化规则如布尔运算和算术运算进行了改进和完善,在原有模型上采用函数调用和数

组分析等一些改进方法,根据语义相似度匹配学生程序与模板程序。该模型主要依据如下几点进行自动评分<sup>[4]</sup>:

- (1)首先定义好相应的模板程序,以此为标准来判断学生程序是否正确;
- (2)只要标准答案的其中一个语义与学生程序的语义对等,则可判断该学生程序的语义正确;
- (3)程序语义也可转换并通过系统依赖图SDG这种中间表示形式来表示,如果两个系统依赖图等价,则两个程序具有等价的语义;
- (4)语义等价转换也就是通过系统依赖图标准化语义等价程序,那么,不管程序以何种形式进行计算,只要其结果不变,则称该转换是等价的。

基于语义相似度比较的编程题自动评分模型在分析词法和语法上比较严格,虽然也已能应用于实际需要,但仍有需要完善的地方,比如,对指针和结构体的处理就不是很理想。

## 3 用正则表达式来描述程序得分点的自动评分模型

该模型使用的是一种分割程序语义的处理方法。此评分模型并不是拿整个程序进行完全匹配,而是将模板程序拆分成若干得分点,并采用正则表达式来描述它,然后在学生程序中查找与得分点模块相似的代码段,并进行匹配评分<sup>[5]</sup>。通常一道编程题在判卷时可分为以下3种情况:

- (1)程序能编译运行并且结果正确;
  - (2)程序能编译运行但结果是错误的;
  - (3)程序不能编译运行的,也得不到结果。
- 对第一种情况,我们只需要分析其代码是否可信即可;对第二种程序能编译运行但结果错误的,就可通过基于得分点的评分模型进行评分;对第三种得不到结果的

情况,可先从词法或语法上判断是否出错,而词法和语法的正确性也是编程题的一个评分点,因此,必须对其词法和语法错误的数量进行统计,再结合其他得分点的情况得到程序的最后成绩。

文献[5]提出的自动评分模型与人工评分结果大致相符,然而构造正则表达式需考虑精确和全面,因为正则表达式构造的优劣会影响评分的结果。

#### 4 基于动态测试和静态分析相结合的自动评分模型

文献[6]先从程序结构、运行结果和数据完备性等方面对静态分析与动态测试这两种评分方法进行了对比分析,根据比较结果提出两者相结合的自动评分模型,如图1所示。此模型应用编译原理中的词法分析和语法分析,对仅有少量错误的考生程序,在系统监控下,先通过动态测试查找错误并修改错误,根据结果信息和错误信息进行评分;而对于有较多错误或结果不正确的程序,就需通过静态分析,先构建系统依赖图并将程序标准化,然后将学生程序与模板程序进行匹配评分。这种相结合的评分模型具有较强的互补性,从而使评分结果更接近于人工阅卷。

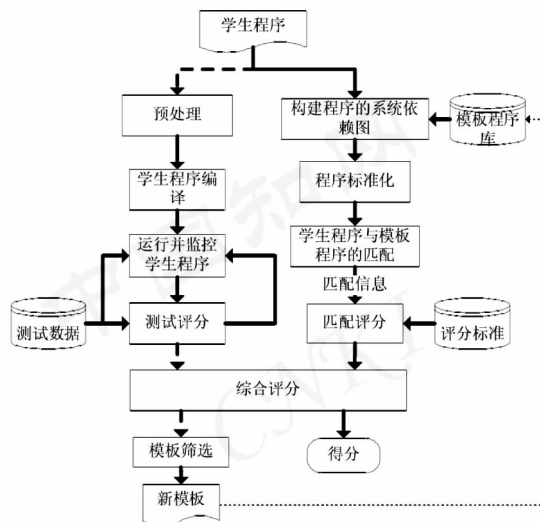


图1 静态分析与动态测试相结合的自动评分模型

#### 5 程序分析与树核相似算法相结合的自动评分模型

文献[7]提出了程序分析与树核相似算法相结合的自动评分方法,此模型与文献[6]提出的一样,其实质也是采用了动静结合的自动评分方法。区别在于,文献[7]采用语法分析树这种中间表示形式对学生程序的语句单元进行词法和语法分析,然后利用树核相似算法等关键技术计算程序的相似度,进行采分点语句匹配。

该方法采用抽象语法树描述程序结构,是对前面所提到的自动评分模型的一个必要的补充。但这种方法仍有一定的缺陷,从文献[7]给出的实验结果分析可知,它对模板程序的要求比较高。

#### 6 基于字符串相似度算法的自动评分模型

字符串相似度算法最早用于计算机基础课程中的Office文档的自动评分。文献[8]通过比较填空题与编程题的特点,在现有的几个字符串相似度算法中选用编辑

距离算法应用于计算机程序编程题的自动评分中。文献[8]对于编程题的自动评分采用3个层次,这3个层次的评分过程如图2所示。其中,对第1层能得到运行结果的采用黑盒测试法,在第2层中使用了字符串相似度算法计算程序中两个字符串的相似度,而如果学生程序在第2层中仍然得分为零,就进入第3层进行关键点匹配。

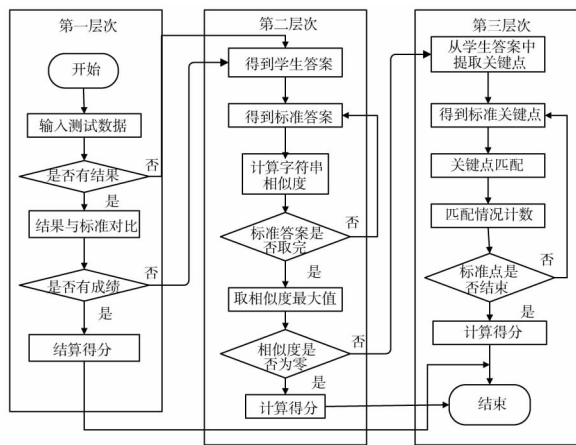


图2 基于字符串相似度算法的自动评分模型流程图

相比仅使用关键点匹配进行自动评分,这种自动评分模型经过3个层次的评测,能更客观、公正地对学生程序进行评分。

#### 7 结语

C语言编程题自动评分这个课题对理论知识和实践能力要求高,它涉及到多个学科甚至多门专业知识,比如软件工程、人工智能、编译原理、程序语义学等,具有很高的研究和实用价值。本文总结了目前可用于程序编程题自动评分的6种模型,经分析这几种评分方法的优缺点得出,要想使得自动评分系统更智能化,接近于人工阅卷,还需在自然语言分析技术和人工智能技术方面继续努力,只有不断地将编程题自动评分技术的理论与实践结合起来,并不断发展和完善它,才能使其功能更完善,更符合考试的实际需求。

#### 参考文献:

- [1] 马培军,王甜甜,苏小红. 基于程序理解的编程题自动评分方法[J]. 计算机研究与发展, 2009, 46(7): 1136-1142.
- [2] 李永浩. 基于程序理解的编程题自动评分系统研究与应用[D]. 哈尔滨: 哈尔滨工业大学, 2004.
- [3] 王甜甜. 基于语义相似度的编程题自动评分系统研究与应用[D]. 哈尔滨: 哈尔滨工业大学, 2005.
- [4] 王宁. 编程题自动评分系统中结构体的研究与实现[D]. 哈尔滨: 哈尔滨工业大学, 2006.
- [5] 余石泉. 编程题自动阅卷技术的研究与实现[D]. 长沙: 中南大学, 2007.
- [6] 赫伽宁. 结合动态测试与静态分析的编程题自动评分方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2008.
- [7] 陈纯纯. 程序设计题自动评分系统的研究与设计[J]. 济南职业学院学报, 2011(12): 75-77.
- [8] 杜利峰,牛永洁. 字符串相似度在自动评分系统中的应用[J]. 电子设计工程, 2011, 19(7): 42-44.

(责任校对 莫秀珍)