

Using Unit Test to Realize Automated Hinting and Grading in Student Programming

Zhiguo Zhang*

Sun Yat-sen University, Guangzhou, China

*Corresponding author. Email: lnszgz@sysu.edu.cn

ABSTRACT

In doing machine work, students may encounter problems or difficulties. It should be very helpful for students' learning if students can get some hints or feedback immediately. We present in this paper a method that can automatically give students helpful guidance hints and judge and grade students' work in the process of students' programming on machine. This method can be used in many classes such as programming and other related subjects. The core idea of the method is to use unit test in assigned work to students. The teacher first writes the unit test modules which will test students' work at runtime according to the correctness requirements of the training work assigned to students. The unit test modules may also contain some hints information to students at runtime when students meet difficulties. Students must use the unit test module in their programming environments as library when they are doing their training work on machine. Because these tests can be very flexible forms reflecting a variety of different teaching requirements, and can be coded with many helpful hints information for students in many failure situations, this method for automated hinting and grading has more powerful functions and more wide applications than other traditional online judging systems.

Keywords: Computer aided instruction, Programming course, Unit test, Automated hinting and grading.

在程序设计学习过程中使用单元测试实现自动提示和自动评判

张治国*

中山大学, 广州, 中国

*通讯作者. 邮箱: lnszgz@sysu.edu.cn

中文摘要

学生在机上训练过程中如果遇到问题或困难能够及时获得提示或获得结果的反馈, 对学生学习是非常有帮助的。本文给出了一个在学生编程或设计过程中能够自动提示指导信息并且对学生结果能够进行自动判定评分的方法。这种方法可以应用于很多相关的课程当中, 比如程序设计等各种设计类课程当中。方法的核心思想是在学生机上训练的过程中使用单元测试。教师首先编写单元测试模块, 其测试的内容反映了对学生训练的要求、对正确性的要求、和学生在训练过程中各种错误后的提示信息。学生在机器上的训练过程中一定要使用该测试模块以便学生可以获得有用的提示和评判打分。由于这些测试可以是非常灵活的形式, 可以反映各种不同的教学要求, 且能够在不同的层次或侧面提示有用的帮助信息, 它比其它的在线评分系统有更强大的功能和更广泛的应用环境。

关键词: 计算机辅助教学, 程序设计课程, 单元测试, 自动提示和评判

1. 导论

自动提示和评分教学辅助系统的重要性体现在当教师不在其旁边时学生对自己编写的程序运行结果可以立刻获得反馈或提示。特别是在在线教育的情况下或学生独立学习的情况下更是如此。一般都认为在很多情况下自动评判系统是有助于教学的[1]，特别是从经济性、管理、在线环境等其它方面的考虑[2,6]更是如此。很多文献讨论了自动评判的软件实现框架，如Web-CAT[3]、APOGEE[4]、和Infandango[5]。本文我们强调自动提示和评判的扩展功能及其在课程教学中的使用。

程序设计类课程中学生必须参与实际的机器上训练。传统的机上训练实验是让学生完成给定的编程任务。对于单一学生出现的各种问题，教师需要单一的指导，在学生数很多的情况下，指导很难有效的完成。另外学生课外自己机上训练遇到问题时也希望能得到立即的指导或反馈。这种情况下学生完成一个任务之后，学生一般也很难确认是否正确，所以希望能够立刻得到解答。在传统的实验室内指导教师可以给出结果。但是在自学和网络环境下就很难直接获得指导和知道自己的结果正确与否。这些问题可能阻碍了学生的继续训练，影响了学生的有效学习。

有很多的自动评判系统，它们能够给出实验结果的评判，但只是最终结果的评判。这样的评判系统很难做到交互式的指导，也很难针对细节技巧和特殊教学要求进行评判。

本文首先给出一个基于单元测试[7,8,9]的自动提示和评判方法。然后，讨论这种方法在传统课程和其他形式的课程当中的应用。

2. 基于单元测试的教学训练环境

这种教学训练环境的基本要求是有一个网络支持。它可以是一个局域网或远程网，甚至可以是最简单的支持FTP服务的无线局域网，比如教师的机器上有一个WiFi服务和一个FTP服务就可以。学生的机器上必须有网络的访问支持和编程环境和工具。

学生的训练任务是由软件包（称为任务包）形式发布。任务包包含了训练任务的题目描述（也可以任何其他的形式发布），另外包含了一个测试模块。测试模块中包含了对于题目的各种要求、各种情况下的帮助提示信息、和答案，这些均以单元测试的代码形式出现。教师事先要准备好这些任务包，这是一个创作题目和代码编写的过程。任务包可以发放在网上，或以其他任何形式交给学生。

学生接收到任务包后，要将这个任务包连接到自己的编程环境当中去，一般是作为连接库的形式加入到自己的编程环境当中。一般情况下任务包会包含开发环境的一个项目管理文件，学生只要直接打开项目管理文件就可以阅读题目、进行编程或进行训练的设

计任务了。编程过程当中测试模块会测试学生的程序，并给出错误信息和启示性提示，学生完成后可以获得是否正确反馈和成绩。

学生的训练可能有多种不同的情况。一种情况是，学生是自学过程，不需要提交任何结果给教师，此时学生的机器可以不联网。另一种情况是学生可能需要提交自己的训练结果，这种情况学生机器可以是联网的或不联网。联网的情况下会自动提交结果，不联网的情况下学生需要把结果以其它任何方式提交（比如发邮件或拷贝）即可。还有一种方式可能是现场交互方式，这种方式下学生的任何结果，甚至包括中间结果，都是可以随时上传给服务器的。服务器上可能有成绩管理、竞赛规则、实时排名等功能，教师甚至可以直接使用这些信息做交互式的指导和教学。

3. 教师准备题目内容及其提示和评判

学生要完成的任务（作业、练习题目等）可以被看成是要进行单元测试的编程模块。教师先要给出任务的描述，任务描述在代码中可以以注释的方式出现，但这种方法有一定的限制，因为不能出现图形等特殊符号。所以任务描述往往采用单独的文件形式包含在开发项目当中。任务描述也可以以任何其他的形式发布给学生。教师最主要的任务是编写和准备测试模块。测试模块完成对学生的编程模块的测试。测试模块其实是一个单元测试模块，其中根据题目的要求列出若干测试项。测试项是一些评判条件，可以根据评判结果给出评分和一些提示信息。下面是一个评测项的简单结构例子。

```
TEST(HomeWork1, test1) //一个测试项
{
    HomeWork1 hw1;
    hw1.step();
    CHECK(hw1.Result[0]==10, grade10,
    "Hint: Result[0]应该是正数， ...");
    CHECK(hw1.Result[1]==11, grade11,
    "Hint: Result[1]是递增结果， ...");
    hw1.step();
    CHECK(hw1.mid[0]==100, grade12,
    "Hint: mid[0]应该是正数，中间值...");
    UpLoad(); //上传服务器
}
TEST(HomeWork1, test2) //一个测试项
{
    HomeWork1 hw1;
    hw1.step();
```

```
CHECK(...);
```

```
...
```

```
}
```

测试模块可以非常细化或非常简单。简单的情况只是测试最终结果的正确性并且打分。一个非常细化的评测可以把学生程序可能会出现的很多情况都进行检测。所以，测试模块可以完成我们需要的许多功能，如下列：

- 识别学生程序的错误。
- 在学生程序的很多错误点上给出信息提示。
- 自动评分。
- 实现各种灵活的评判策略。
- 部分或全部信息和结果上传服务器。
- 在服务器侧可以做任何要求的处理。
- 等等

4. 各种课堂形式

基于单元测试的方法可以嵌入到各种不同形式的课堂教学当中。从离线到在线，从自学、网络课程，到传统课堂等等都可以应用这种方法。下面我们简单列举几种可能的应用场景。最主要的选项区别是服务器侧的功能设置，不同的功能设置会使系统应用于各种不同的场景。比如，服务器可能拥有如下某些功能。

- 服务器作为作业任务的发布源。
- 服务器接收学生的最终结果。
- 服务器可以接收学生的中间结果和最终结果。
- 服务器可以对以上的信息进行处理，比如排序、动态提示、实现某种竞赛规则等。

4.1. 传统实验室的实验课

传统的实验室一般会支持一个局域网。我们需要一台教师的机器作为服务器，用于发布作业任务和收集学生的结果。学生在机上训练时可以获得若干提示，如果需要也可以不给予这些提示。每个学生在训练过程中出现的编程问题和完成情况和自动评判成绩会即时出现在教师机器上。

4.2. 传统教室课堂

这种情况需要局域网的支持。如果教室没有局域网支持，教师可以搭建一个简单的WiFi无线局域网。当然这要求所有学生的机器支持WiFi访问。教师机器上可以构建一个WiFi接入点（比如用360的免费WiFi）。简单的环境并不需要教师机器安装特别的软件或服务，只要开启FTP服务就可以了。在这种环境下，教师可以即时发布给学生一些非常简短的训练题目，或

者一个简短的测试，教师很快可以得到学生完成的各种信息，据此可以直接点评指导。

4.3. 在线课程

教师可以在线发布作业题目，学生获得任务包后，可以是独立的完成。这种情况下学生不需要继续联网，但在学生编程过程当中，任务包会给出若干提示并给出评定分数。在线课程也可以是交互式教学环境，教学环境非常类似于传统教室课堂的情况，教师会及时获得学生的完成情况。

4.4. 自学课程

学生可以以任何形式获得教师发布的任务包，训练的时候，可以在自己的开发环境中使用这个任务包，从而在自己训练过程当中可以获得很多提示，并且获得最终的评定成绩。

4.5. 程序设计竞赛

在竞赛时每个参加竞赛的学生可以在竞赛开始时刻获得竞赛的任务包，然后在自己的开发环境当中直接使用任务包开始编程。这种情况下可以取消掉对于学生的任何提示和评判分数的显示，但这些信息可以直接上传到服务器，在服务器上设定竞赛规则和评判标准，可以动态给出评判的结果，也可以有动态的排名机制等等。由于在竞赛任务包当中有设计好的很多条件、测试位置，甚至还可以加入时间上的要求等等，所以竞赛规则和评判标准可以做得很复杂。这比传统的自动评分方法要灵活很多，而且能支持非常复杂的竞赛规则和评判标准，竞赛的实施功能和管理功能上也可以非常复杂。

4.6. 机上考试

机上考试非常类似于程序设计竞赛的情况。实际上可以适用于任何课程的考试，可以是自动评分而学生立即获得分数。对于不容易自动评分的考试内容，也可以是学生只提交考试的结果，这种结果由评判模块自动生成，教师后期再进行人工判卷或其他方式的自动评判。

4.7. 其他可能的教学过程

这种方法可以作为一个课程的作业发布和作业提交系统。测试模块可以允许学生完成作业时有自动提示和自动判分功能，也可以不允许提示和评分，可以将作业的结果提交给老师后再做判分。作业提交可以是自动提交或者以其他方式提交给教师。

另外由于这种方法可以获得了学生在进行编程和完成作业的过程中表现的各种不同的信息，这些信息不但包括评判的分数也包括了评判分数的依据及学生在各种情况下出错或没出错的信息，这些信息都可

以作为研究和分析学生学习情况的统计数据,也可以成为教师对于发布的作业的分析研究的依据。

另外,这种方法可以实现在云端,在云端的服务器上包含教学服务和学生的编程环境,而对于学生侧的机器就没有了任何要求。

5. 在实际课程当中的应用

我们采用这种方法,在实际的教学过程当中,对于多门课程进行过实际的应用。这些课程主要是程序设计类的课程,但也包括了像计算机组成原理这样的硬件设计课程。

在C/C++程序设计、Java程序设计、综合程序设计实验、数据结构等课程中,学生的全部实验内容都是采用这种方法完成的,学生的部分课外作业中程序实现形式的作业也是以这种方式发布的。

另外我们在计算机组成原理等课程当中使用了此方法。也是在实验环节和作业布置发布与提交当中采用这种方法。内容主要是处理器设计、逻辑设计等类练习和作业。

通过使用这种方法,实验课程极大的提高了学习上的效率,在学生进入训练的开始,我们先是关闭了自动提示,在学生进行了一两轮的试探后,我们再动态的打开自动提示功能,在前几轮学生的自己思考基础上,在有自动提示后学生都能很快理解题目和找到解决问题的思路,完成编程任务。

实验课程也极大的提高了时间上的有效性,一是学生能够很快进入状态,二是学生有自动提示后请求老师帮助的次数很少,老师可以节省时间来针对特别要关注的学生。在班级学生人数比较多的情况下,这一点显得更为有意义。

并且对于学生的训练要求和实验分数的判定更加灵活、多样、全面。

6. 总结和结论

本文给出了一个在学生进行程序设计中能够自动提示指导信息并且对学生结果能够进行自动判定评分的方法。其基于单元测试的思想,将学生正在进行设计的程序看成是单元测试的对象,而教师编写的单元测试模块要测试的内容反映了对学生训练的要求、对正确性的要求,也包含了学生在各种可能错误后的提示信息。学生在机器上进行程序设计时一定要将测试模块作为项目的链接库之一,以便学生可以动态获得编程提示和对结果的评判打分。这种方法可以应用于很多相关的课程的各种形式的课堂教学当中,比如程序设计、数字逻辑设计、处理器设计、软件设计等各种设计类课程当中。由于这些测试可以是非常灵活的形式,可以反映各种不同的教学要求,且能够在不同的层次或侧面提示有用的帮助信息,它比其它的在线评分系统有更强大的功能和更广泛的应用环境。

这种方法的一个缺点是需要教师事先要做好完整的测试模块,这需要教师花费较多的编程和题目创作时间。但也正是因为这一点,对学生的要求和题目的形式,及其成绩判定的规则,都可以更加灵活和完整。

需要继续讨论的问题是如何将这些方法能够更有效的应用于传统的教学或者是网络的教学或者其他形式的教学中,如自学过程中,使其能够发挥更大的教学促进作用。另外需要讨论的问题是在各种教学环境当中,如何去有效简洁地实现这种方法,涉及到一些实现技术。

REFERENCES

- [1] C. Douce, D. Livingstone, and J. Orwell. Automatic Test-based Assessment of Programming. *Journal Educational Resources Computing*, 5(3), Sept. 2005.
- [2] C. Wilcox, Testing Strategies for the Automated Grading of Student Programs, *Proceedings of the 47th ACM Technical Symposium on Computing Science Education (SIGCSE16)*, p56, Memphis Tennessee USA March, 2016
- [3] S. Edwards. Using Test-Driven Development in the Classroom: Providing Students with Automatic, Concrete Feedback on Performance. In *Proc. International Conf. on Education and Information Systems: Technologies and Applications (EISTA 03)*, pages 421–426, 2003.
- [4] X. Fu, B. Peltsverger, K. Qian, L. Tao, and J. Liu. APOGEE: Automated Project Grading and Instant Feedback System for Web Based Computing. In *Proc. of the 39th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '08*, pages 77–81, New York, NY, USA, 2008. ACM.
- [5] M. J. Hull, D. Powell, and E. Klein. Infandango: Automated Grading for Student Programming. In *Proc. of the 16th Annual Joint Conf. on Innovation and Technology in Computer Science Education, ITiCSE '11*, pages 330–330, New York, NY, USA, 2011. ACM.
- [6] C. Wilcox. The role of automation in undergraduate computer science education. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education, SIGCSE '15*, pages 90–95, New York, NY, USA, 2015. ACM
- [7] Nunit, <https://nunit.org/>
- [8] JUnit 4, <https://junit.org/junit4/>
- [9] PHPUnit, <https://phpunit.readthedocs.io/en/9.5/>