A Research on Online Judge Technology Based on MOOC Platform

Siwan Tang\(^1,a\), Linfang Zou\(^2,b\), Xuehua Liao\(^3,c\)

\(^1\)Sichuan Normal University, Chengdu, Sichuan, China
\(^2\)Chengdu University of Technology, Chengdu, Sichuan, China
\(^3\)Sichuan Normal University, Chengdu, Sichuan, China

\(^a\)1447909844@qq.com, \(^b\)1289351504@qq.com, \(^c\)191716312@qq.com

Keywords: Online Judge; MOOC; Automatic grading

Abstract. Programming courses in Chinese university MOOC(massive open online courses) platform use Online Judge as a real-time program automatic grading tool. One the one hand, the use of Online Judge increases the ability of programming, on the other hand it decreases the workload of teachers. So more and more programming courses are taught based on Online Judge. Although Online Judge has many advantages, we find some shortcomings during long-term use. This paper describes the technology of online judge, then analyses advantages and disadvantages. Considering these disadvantages, we put forward a method to make it more suitable for programming practice and test.

1. Introduction

MOOC platform offers numerous free, sharing, high quality courses, and makes study at any time and any place. The rise of MOOC are changing the traditional teaching mode. Many universities have offered courses in the MOOC platform. Our college follows this trend to carry out teaching reform of programming courses based on MOOC platform. Teaching and research teams have taught a course named \textit{C language programming} in MOOC platform. In addition to teaching in the classroom, students can watch corresponding study videos and are required to finish test in the MOOC platform. These test are no longer the consolidation of knowledge, but consist of small programming assignments which can make up for the lack of training in programming. Instead of grading by manual, these programming assignments are graded automatically and timely by Online Judge. The real-time grading stimulates the interest of students. The programming ability of students has been promoted.

But in the process of teaching, we find that Online Judge, the automatic grading system, is relatively rigid, and objectivity and accuracy is not enough. Accordingly, we need to find a more object, accurate assessment mechanism of Online Judge, so that is more suitable for programming assessment.
2. Online Judge

Online Judge is an automatic grading system, originally used in the ACM International
Collegiate Programming Contest. Nowadays, it is widely used in the teaching of programming
courses. Peking University's POJ, Zhejiang University's ZOJ are well-known.

2.1. Automatic grading process of Online Judge

Automatic grading process of Online Judge is as follows\cite{1}\cite{2}: Students write programs
according to the requirements, then submit it. Online Judge uses compiler to compile it. If passed,
Online Judge executes each test cases in turn, otherwise gives the reasons of failure. The judge of
test case is specifically as follows: the input of submitted program is the input provided in the test
case. Then run the program. Compared with the output of the executed program and the output
stored in test case, the students can get the corresponding score of that test case if the two outputs
are same. Otherwise students get zero point and the prompts of wrong. In the end, add scores of
each test case as the final score. The program grading process is shown as Fig. 1:

![Diagram of the automatic grading process of Online Judge](image)

Figure 1. automatic grading process of Online Judge.

2.2. Compare Online Judge with manual grading

Students usually hand in programs by electronic or paper. Teachers grade programs by
manual in the past. In the grading process, more attention is focused on knowledge involved,
appropriate notes, the conventional solution, modularity, meaningful variable names and aesthetic
of layout\cite{1}. The process is often subjective, because how the program is beautiful, what kind of
annotation is reasonable, what is enough modular is short of strict regulations. The tedious and time
consuming of the grading makes the programming assignments less. Students often have good
command of knowledge, but cannot use the learned knowledge flexibility in programming practice.
Time-consuming of the grading hinders timely feedback. Students even forget the content of
assignments and difficulties encountered during programming. So programming is only to complete
assignments while the purpose is to enhance understanding and use in practice.

Online Judge effectively improve this situation. You can use it at anytime and anyplace if you
have network. After submitting the program, students can get the assessment results immediately.
The results are as follows: pass, fail to pass, timeout etc. Test cases which cannot be passed have
wrong prompts. So students can modify their program according to the corresponding prompt.
When the modification is completed, they can submit it again until it is passed. Students can train
repeatedly according to their own need, rather than waiting for the long process of manual grading.
Online Judge can reduce the burden of programming grading for teachers. Teachers can concentrate on assignments and test cases. The selection of test cases is particularly important, because whether the test case is suitable will affect the grading of the Online Judge. Test cases should include all situations in order to instruct them to consider comprehensively.

2.3. Advantages of Online Judge

Programs that students submit can be graded by Online Judge. Teachers can save a lot of grading time, so that they can pay more attention to teaching and learning situations. Online Judge is objective. Score difference does not exist due to unbeautiful layout, unmeaningful variable name, different grading standards. Timely feedback makes students quickly know the grading result. Students can find and correct mistakes quickly based on prompts. This is benefit to the accumulation of programming experience. Online Judge makes programming at anytime and anyplace. Students can determine the number of programming training according to their own situation. Learning theory with practice ,Online Judge has good effect on program language learning.

2.4. Disadvantages of Online Judge

However, in the process of using Online Judge, we find some shortcomings.

2.4.1. Only consider the output of the committed program

Opportunists directly output the right answer of the program without any implement process. Online Judge still returns a full mark after submission. It is not reasonable to only consider the output without the implement process. At the same time, it cannot verify whether the program meets certain special requirements.

2.4.2. Cannot distinguish the degree of error

Once the program fails to compile, the student gets zero point. One student's program fails to compile because of lacks of a semicolon. Another student get zero point because of no code wrote. In this case, program grading of Online Judge cannot reflect true level of students.

2.4.3. Too many constrains

To pass test cases successfully, students need to meet a lot of constraints, such as the format constraints of input and output. Students tend to have errors at this aspect, which increases the barriers of learning.

3. Improvements of online Judge

The grading way of Online Judge is called by dynamic grading. That is, run the program and check whether each test cases pass successfully. Considering disadvantages mentioned above, this paper presents a static and dynamic combination method for program grading. Since dynamic is mentioned above ,we will concentrate on the part of the static grading. The most important thing of static grading is that it doesn't run the program. Static grading simulates manual grading in providing reference programs and setting key statements. Each of the key statement is set score according to the importance. The corresponding score is obtained based on the similarity between the key statement and the statement in the submitted program.

Static grading of program can be more accurate in reflecting the true level of programming. The specific process is shown in fig. 2:
Figure 2. the process of static grading.

3.1. Pretreatment stage
   
   purpose: initially remove the part which cannot affect the essence of the program
   
   using method: delete
   
   details: delete continuous spaces, newlines, and text; delete natural similar codes, such as the import of library code; uniform of uppercase and lowercase etc.

3.2. Standardization stage
   
   There are so many ways to write the program to realize a same function. Accordingly, how to reduce the diversity of the program will be the focus of the research.
   
   purpose: reduce the diversification caused by language itself.
   
   using method: equivalent transform.
   
   details: the standardization of conditional statements: unify the way of if..else ...; the standardization of loop control statements: the consistent use of while-cycle; the standardization of compound assignment operator: for example: x+=y to x=x+y; etc.
3.3. Intermediate representation stage

The intermediate representation unit has the tree, the graph, the string, the metric value and so on. Measure the similarity of the program based string is from perspective of structure and lexical about source code. Compared with the tree or graph, this way has better efficiency in space and time, but is less in the grammatical information. The source code is substituted by a series of metric values in the metric value method. Due to the loss of too many program structure information, the accuracy based metric value is low. Intermediate representation of source code should be able to reflect the grammatical information, and has a good time and space efficiency.

3.4. Program similarity algorithm selection stage

The premise of algorithm selection is not sensitive to the statement sequence. We can choose the GST (greedy string tiling) algorithm, and make appropriate improvements to make the time efficiency higher.

GST is a greedy string matching algorithm. This algorithm is greedy search for two strings to find the largest public substring. It needs to search repeatedly for the two strings. Each search is to find the longest common substring from the current unmarked part. And mark on the found longest common substring to avoid repeatedly use\(^6\). GST algorithm has best space-time efficiency when the length of pattern string is equal to the text string. At this point, the time complexity of GST algorithm is O\((n^2)\), and the time complexity is O\((n^3)\) in the worst case.

3.5. Calculate static grading score stage

the calculation method of static grading score is shown as Eq. 1.

\[
\text{score} = \sum_{i=1}^{n} w_i \cdot \text{sim}_i
\]

\[ (1) \]

\[
\text{total} = \sum_{i=1}^{n} w_i
\]

\[ (2) \]

w is the score of each key statement. The relation between w and the full mark of program meets Eq. 2. Simi is the similarity between statements in the submitted program and the key statement. \(n\) represents the total number of key statements.

3.6. Combine dynamic grading score with static grading score

The static grading of program can avoid getting a full mark by only outputting the right answer and get zero point because of few flaws. Even if the program contains all the key statements, it cannot guarantee to run the program successfully. So the static grading of program mentioned above cannot replace the dynamic grading method. It will make the grading of program more reasonably by considering the contents and results at the same time. So the combination of dynamic and static grading is desirable. The center of this stage is to find the appropriate proportion between dynamic grading score and static grading score.

4. Summary

Online Judge can free teachers from the grading of a lot of programming assignments. At the same time, it can provide more opportunities to practice programming for students. However, Online Judge still has some deficiencies. We propose a method to optimize the grading of Online
Judge, that is, the combination of dynamic and static grading. Based on the research in this paper, it will have an effective improvement on Online Judge in the future.

References


