

Report of TDG Project
Computer Assisted Assessment Software for Math Proofs

In this project, our goal has been the development of algorithms for “proof-type questions” based on the open source software WeBWork (see <https://en.wikipedia.org/wiki/WeBWork> for more detailed description of this software).

Background

The software WeBWork is a student oriented, actively learning tool. For us, we see the following points as being the key features, among others, of WeBWork:

1. Immediate Feedback Capability – students can get response to their homework questions immediately.
2. Symbolic Answer Checking Capability – an answer to a WeBWork question can be as complicated as an equation such as

$$2x \cos y = \frac{e^{xy}}{x^2 + y^2 + 1}$$

Most mathematics tutoring software do not possess capability to check answer of such complexity, hence the instructor has to change the question types to either Multiple Choice questions or similar questions which are more tractable.

3. Flexibility (on the instructor’s side) to (i) set up allowed number of attempts per student, (ii) to archive and data-mine students’ actions through its database, enabling in-depth analysis of students’ performances.

Motivation of Our Work

One main motivation behind our work has been the lack of “free-form writing” type questions in the existing WeBWork question bank (which comprises more than 35,000 questions). For our purpose, we are interested in developing “free-form writing” for science subjects, especially in Mathematics.

The following example may illustrate our point:

Question: For any odd natural number, show that its square is also an odd natural number.

A typical answer to this may take the following form:

Answer: Let n be such a number.
 Then $n = 2k + 1$ because it is an odd number.
 Squaring it gives $n^2 = (2k + 1)^2$.
 Hence $n^2 = 4k^2 + 4k + 1$
 $n^2 = 2(2k^2 + 2k) + 1$.
 $n^2 = 2l + 1$, where we have let $l = 2k^2 + 2k$.
 Therefore n^2 is an odd number.

This solution of a student consists of a text, in which his/her arguments are described. During the entire process of writing down the solution, the student has lots of freedom, and it is our goal to automatically/semi-automatically check to what extent the student's writing is correct/incorrect.

To achieve this, we have utilized one feature of WeBWork, viz. the Essay Box inside WeBWork. Originally, the Essay Box in WeBWork is an empty box for students to fill in their answers to be graded by a human grader subsequently. Our work consisted mainly in writing subroutines to turn this Essay Box into a "semi-auto grading box" using some pattern matching facilities and ideas from Linguistics.

Results

The following screen shot summarize our result to date.

The screenshot shows a web browser window with a WeBWork interface. On the left, a sidebar lists 'Problem 1' through 'Problem 5', with 'Problem 5' selected and marked with a checkmark. The main content area is divided into three columns: 'Entered', 'Answer Preview', and 'Message'. The 'Entered' column contains the student's handwritten-style solution. The 'Answer Preview' column shows the same solution rendered in a clean, mathematical font. The 'Message' column is highlighted in yellow and contains the following text: 'This answer will be graded at a later time. Full Match: EQUATION0 Then EQUATION0 because it is an odd number. Full Match: EQUATION0 Squaring it gives EQUATION0. Full Match: EQUATION0 EQUATION0 where we have let EQUATION1. Full Match: FORMULA0 is an odd number Therefore FORMULA0 is an odd number. INFERENCE 100%'. Below this, a green bar states 'The answer above is correct.' and a box contains the problem statement: '(1 point) Let n be an odd natural number, show that n^2 is also an odd natural number. (Variables you can use: k, l, p)'. Below the problem statement is a 'Proof:' section with a text box containing the student's solution. At the bottom, there are buttons for 'Preview My Answers' and 'Submit Answers', and a status message: 'Your score was recorded. You have attempted this problem 3 times. You received a score of 100% for this attempt.'