Design of Questions and Distracters for a Dynamic, Algorithm-Based Suite of Physics Problems for Engineering Students

Luis Neri, Víctor Robledo-Rella, Enrique Espinosa, Julieta Noguez
Tecnológico de Monterrey, Campus Ciudad de México
neri@itesm.mx, vrobledo@itesm.mx, enrique.espinosa@itesm.mx, jnoguez@itesm.mx

Abstract - We discuss a method to design problem sets for publication in an on-line tool built to coach engineering students in solving quizzes and exams during an undergraduate Physics course. This tool contributes to reinforce the student’s skills when approaching and solving problems. The coach is dynamic because values in formulas for a given problem are unique to every student. We present criteria and guidelines followed to design appropriate distracters for each problem, which may give the teacher insight on common mistakes. If the difference between the answer given by the student and the correct answer is larger than a given offset, then the coach issues feedback to guide them in finding the source of the mistake. We have undertaken a field study with first year university students. We present our results and discuss the usefulness of this tool to assist students through their own learning process.

Index Terms - Online assessment, e-learning, adaptive content, self-testing, problem solving.

INTRODUCTION

Our group has been active in the conception, design and implementation of computer-supported learning tools [1]. In addition, considerable know-how has been accumulated in the design of sound pedagogical methods for enhancing the learning process in diverse areas, contexts and educational levels [2]. In this paper, we present work in progress on undergraduate Physics courses. It is commonplace that freshmen find it hard to acquire the knowledge and skills needed to effectively reason through traditional problem-solving. Consequently, they tend to lose interest in the subject, and thus their motivation falls. It is important then to develop an institutional computational system on line to assist this need in our students.

FOCUS OF THE PROJECT: PEDAGOGICAL REINFORCEMENT

Our project deals with coaching engineering students when solving quizzes and exams. The focus of the project is to provide students with sufficient pedagogical reinforcement during the learning process by assessing them as best as possible in their own particular needs. We attempt to provide personalized and immediate feedback whenever they make common mistakes. Our purpose is to serialize (i.e. to write in electronic media) the skills and experience of our Physics professors as witnessed during class time and routine office hours. An experienced instructor knows a set of algorithms for inferring the learner’s mistakes. For example, as unknowns are identified and cleared in mathematical equations, a number of miscalculations may occur, such as missing physical variables, choosing the wrong units or not converting to the right units, dividing instead of multiplying, computing the power of 2 instead of the square root, or solving for the wrong unknown. The professor then provides feedback so to make the student able to understand the origin of the mistake and take the needed steps to find out the right answer.

AUTOMATED & ADAPTIVE FEEDBACK

The project at stake is an early attempt to replicate the feedback mentioned in the previous section. We offer an online coach capable of detecting a limited set of key mistakes, and providing appropriate feedback which will presumably provide a basic reinforcement to the student, at times when the professor is unavailable.

APPLICATION ARCHITECTURE

We have developed a prototype of a web-based coach that offers questionnaires to the students that have enrolled in the aforementioned Physics course. The questionnaires are grouped by modules that cover the main content of the course. Our current study centers on classical mechanics. Each questionnaire contains a set of dynamic questions that have precise algorithms to compute the correct answers, as well as a complementary set of algorithms to compute the “wrong choices”, hereby called distracters. These wrong choices are defined according to the instructor’s teaching experience. In the prototype version, the system compares the answer given by the student versus the right answer. If they differ by more than a given offset previously defined
by the professor, it provides the student a first feedback as
guidance.

The web tool has been programmed using the rule-based
programming paradigm. The Rete Algorithm [3] has been
used to fast index a large number of these distracters without
sacrificing performance. The user interface was programmed
using Java 2 Enterprise (J2EE) tools, such as the Model
View Controller (MVC) design pattern, plus Java Standard
Taglibs (JSTL) and Ajax on the user interface. Witness

FIGURE 1. There are two types of users: (1) the student and
(2) the instructor. Of these, the latter requests services to the
application. Such services include (a) uploading information
about the students in a group from an Excel spreadsheet, (b)
viewing a grade report based on the summary of student
answer sessions and (c) services related to the backup and
assembly of the application itself.

The layout of the components that make up the
application suggests the implementation of MVC. The
question editor allows the instructor to add, edit, or delete
individual questions to each questionnaire, along with the
correct answer and distracters algorithms. It also allows the
professor to upload a figure associated with a given question
(see Figure 2). These distracters are made up of (a) equations
that lead to common incorrect answers and (b) feedback
when the incorrect answer is reached. The students log on to
the system using their student ID and an initial password.
Once logged in, they can change their password, and select a
questionnaire within a module to solve it. Students are then
presented with a problem statement which may include a
figure, and has a limited time to solve it set by the
professor. The numerical values for some (or all) of the
variables within the statement will be randomly generated by
the computer and will be displayed with three digits within
previously selected intervals set by the professor. The
students are given a field where they must give their answer
rounded to three significant figures, in scientific notation,
and a last field where they must select the corresponding SI
units for the answer. The student then can submit each
worked out question to be evaluated by the system, one a
time and the application will come back with a report card
that includes automated feedback according to the given
answers. The application also detects and gives feedback on
other types of errors, such as using incorrect units.

FIGURE 2

DISTRACTER DESIGN

As pointed out above, an outstanding feature of our system
is the careful design of distracters in order to provide
appropriate and individualized feedback to the students. To
this aim, we rely on our large teaching experience giving
Physics courses for freshmen engineering students at our
institution since 1991. We know with enough confidence the
average Physics background that most students have at this
level, and the most common mistakes and misconceptions
they often have. This empirical knowledge has been
carefully applied in designing each of the aforementioned
distracters. For each question, we have designed design
usually from 5 to 10 distracters, according to the difficulty
level of the question; the higher the difficulty level, the
larger number of distracters.

For each problem, the system is supplied both with (a)
the algorithm for the correct answer, including its

En la figura adjunta, los bloques de masas \( m_1 \) y \( m_2 \)
son de una inclinación \( \theta \) y están conectados sin friction.
Los coeficientes de friccion entre los bloques y las
superficies inclinadas son iguales. Se obtiene que los
bloques se deslizan con aceleración constante de
magnitud a = 10 m/s\(^2\), m1 hacia arriba y m2 hacia abajo. ¿Cuál es el valor del coeficiente de friccion?

variables

\( m_1 \)

\( m_2 \)

Equation

...
distracters, the system asks the student to go back to the approach and solution of the problem.

**PROBLEM AND DISTRACTERS SELECTION EXAMPLE**

To illustrate how the system works, an example associated with the applications of Newton’s laws is given here. A typical statement will read as follow, where the squared brackets indicate three digit numerical values that will be randomly generated by the computer:

“In Figure 3, three blocks of masses \([m_1]\) kg, \([m_2]\) kg and \([m_3]\) kg are in contact between them and located on a horizontal surface. The coefficient of kinetic friction between the blocks and the surface is the same and equal to \([\mu_k]\). A constant horizontal force of \([F]\) N is applied to the left of block \(m_1\), as shown. Find the contact force between blocks \(m_2\) and \(m_3\).”

Variables and ranges:

\[ m_1 = [1.00, 2.00] \text{ kg}, \quad m_2 = [2.00, 3.00] \text{ kg}, \quad m_3 = [3.00, 4.00] \text{ kg}, \quad \mu_k = [0.200, 0.300], \quad F = [30.0, 50.0] \text{ N} \]

Unknown: contact force between blocks \(m_2\) and \(m_3\); \(F_{23}\)

Correct answer:

\[ F_{23} = \frac{m_1 m_2 F}{m_1 + m_2 + m_3} \]

or

\[ F_{23} = \frac{m_2 F}{m_1 + m_2 + m_3} \]

Correct units: N (Newton)

System feedback: Very good!

Table I shows the most common distractors for this problem, grouped by “distracter class”, the selection criteria for these distracter classes and the corresponding feedback given by the system.

<table>
<thead>
<tr>
<th>Distracter class</th>
<th>Criteria for distracter class</th>
<th>Feedback given by the system</th>
</tr>
</thead>
<tbody>
<tr>
<td>(F_{23} = F)</td>
<td>Many students think that (F) is directly applied to block (m_1) and that it is the only force applied to it, instead of analyzing the free-body diagram of block (m_2), and/or (m_3).</td>
<td>Does force (F) applies directly to (m_2)? What happened to kinectics frictional forces? Review the free-body diagrams of each block.</td>
</tr>
<tr>
<td>(F_{23} = F - \mu_k m_1 g) or (F_{23} = F - \mu_k m_2 g) or (F_{23} = F - \mu_k m_3 g)</td>
<td>Many students think that (F) is directly applied to block (m_1) and to block (m_3), instead of analyzing the free-body diagram of each block. In addition only frictional forces on (m_1), (m_2) and/or (m_3) are considered.</td>
<td>Force (F) is not applied directly to (m_2) or to (m_3). Review the free-body diagrams of the blocks.</td>
</tr>
</tbody>
</table>

Confusion between the free-body diagrams of the blocks. Draw the free-body diagram for all the system and for \(m_3\) alone. Write down the corresponding motion equations.

This is the total friction force on the system of three blocks. Draw the free-body diagram of each block.

The student writes masses instead of weights in the equations. A very common mistake. The rest of the reasoning is correct.

What is the relation between weight and mass?

The frictional force on \(m_3\) is not included or frictional forces on \(m_1\) and \(m_2\) are not included.

Review the frictional forces on the blocks.

Any other result that does not lie on any value interval for distracters.

Review carefully both the phrasing and procedure of the problem.
CASE STUDY: CLASSICAL MECHANICS

In order to evaluate the academic usefulness of the software, we have proceeded as follows: we considered three of our Physics I courses of the January – May 2008 term at the Tecnológico de Monterrey, Campus Ciudad de México. The selected topic for the present study was kinematics in one dimension and particle dynamic with friction, which covers about 1/4 of the total time assigned to the course. Two of these groups (hereafter, groups A and B) were asked to use the Dynamic Online Tool (hereafter, DOT), while the third group (hereafter, group C) was not asked to do so. The participants were university students between 1st and 2nd semester of diverse engineering majors. The total populations of the groups that participated in this study were 17, 19 and 63 students, for groups A, B and C, respectively. The groups which used the tool (\(N_A = 17\) and \(N_B = 19\); so \(N_{\text{WITH}} = 36\)) will be referred as the experimental group, while the group that did not use the tool (\(N_{\text{WITHOUT}} = 63\)) will be referred as the control group.

We designed a Pre-test and a Post-test consisting of 5 questions each, which were to be solved by the students in all three groups in the classroom in about 20 min. We then asked groups A and B to work with the DOT for about a couple of weeks. After this, we applied the Post-test to all three groups. The time span between the application of the Pre-test and the Post-test was about 3 weeks, during which the professors of the respective groups continued lecturing according to their normal procedures.

In order to measure how much each student improved his ability to solve physics problems (as measured by the Pretest and Posttest), we define a “relative gain” \([4]\):

\[
G = \frac{(\text{Posttest} - \text{Pretest})}{(100 - \text{Pretest})}
\]  

(2)

That is, \(G\) measures how much the student did actually improve his grade from the Pretest to the Posttest, relative to how much the student could have improved. A student with a Posttest grade equal to 100 would get \(G = 1.00\), while a student with equal grades for the Pretest and the Posttest would get \(G = 0\).

In Figure 4 we show the relative gain \(G\) for each student in our sample \((N = 99)\). For the sake of clarity we have ordered the students in increasing relative gain and we have labeled those students that used the DOT and those students that did not use it. The average gain of groups A and B (with DOT) was \(<G>_{\text{WITH}} = 0.41\) (STDV = 0.29), while the average gain of group C (without DOT) was \(<G>_{\text{WITHOUT}} = 0.16\) (STDV = 0.59). The average gain for the whole sample was \(<G>_{\text{ALL}} = 0.23\) (STDV = 0.53).

The statistical soundness of these results has to be improved before any strong conclusion can be drawn of these results. Among the main factors affecting these results are: (a) intrinsic variability among the students’ academic skill within the different groups, (b) reliability of the software at the time the students used it (some minor technical problems with group A), and (c) there was not a serious commitment by the students in groups A and B to use the DOT.

USABILITY QUESTIONNAIRE

In addition to the Pre-test and Post-test, we designed a questionnaire to be applied to groups A and B in order to measure the ease with which students were able to use the DOT. This usability questionnaire \([5]\) focused on four elements: (a) learnability, the capability of the tool to enable the students to learn how to use it, (b) efficiency of use, related with the performance and response time of the Web tool, (c) interface design, considering if the students easily accomplish their intended tasks, and (d) subjective students’ opinion about the tool’s usefulness. The questionnaire consisted of 10 organized questions related to these usability elements. Only 9 students answered it in Group A and 7 students in Group B.

The results for some elements of the usability questionnaire (learnability, interface design and students’ opinion) are shown in Figures 5 to 7, respectively. In each Figure the first 5 columns refer to opinions from group A and the remaining ones to those from group B.

We can observe on the graphs that different results about the usability of the DOT were found between groups A and B. In general, Group B showed better results than Group A. The main reason is that some of the exercises given to Group A were corrected while the system was running and there were some technical difficulties in having the software running properly for this group, as was mentioned earlier.
The students understand easily the use of the tool
The tool helps students recover from errors

FIGURE 5
LEARNABILITY MEASUREMENTS FOR GROUPS A AND B

The exercises were useful to improve the students' Physics knowledge
The students' general opinion of the Web tool

FIGURE 7
SUBJECTIVE STUDENTS' OPINION ABOUT THE USEFULNESS OF THE DOT

We are conscious of some of the limitations of this work that we comment next. First, we point out that the fact that a student response to a given problem is within the interval associated with a given distracter does not necessarily imply that the student incurred in the mistake expected by the professor. In fact, a student can actually follow a series of wrong procedures, different from those envisaged by the teacher for the given distracter, and nevertheless arrive to the same wrong answer. All we can say at this point is that the student made the mistake expected for that distracter only with a certain probability. In order to investigate in more detail what kind of mistake or mistakes students actually did, we need to disclose student's reasoning in more detail. For this purpose, the system could ask the student a set of simpler questions or to provide them easier problems that will lead the professor to find out exactly what type of errors they actually made. For instance, in the example of the blocks discussed above we could present to the students a problem with only two blocks in contact, say m₁ and m₂, or even present only one block. Only after they provide the right answer for this simpler problem (say one block), the system will present the two blocks problem, and if they solve correctly this problem, the system will then present them the three blocks problem. In this way, the system can go forwards and backwards providing problems of different difficulty level, according to the initial difficulty level chosen by the professor.

Note that this problem is also present even in the case when students provide the right answer for a given problem: we cannot be absolutely sure if they actually followed the correct procedure to get the right answer. It is also possible that a combination of mistakes may lead them to provide a result that lies within the given interval around the right answer.

We plan to tackle this problem in the near future by taking into account a Bayesian probability model [6], [7], [8], and [9] and a knowledge-objects theory [10].

Another limitation on the feedback given to the students is the possibility that those intervals associated with each of the distracters may overlap each other. Right now, when students enter an answer value that lies within these overlapping intervals, the system returns all the feedback statements associated with each of the distracters, which can be confusing for them. To reduce this possibility we may reduce the number of distracters associated with a given problem, only considering the most representative one for each distracter class. We can also reduce the interval width associated with each distracter and to carefully select the input variables' values so to minimize this overlapping.

CONCLUSIONS AND FURTHER WORK

The results obtained so far with our case study with freshman engineering students showed that the students that used the DOT got an average relative gain larger, \(<G>_{\text{WITH}} = 0.41\) (STDV = 0.29), than those students which did not used
the DOT, <G>WITHOUT = 0.16 (STDV = 0.59). However the statistical soundness of this result must be reinforced using larger samples of students and maintaining a robust architecture for the system before allowing the students use it.

Nevertheless, we feel confident that our system can provide students with handy feedback in order to help them to understand the source of their mistakes while solving physics problems and to find out the right procedures implied in solving the problem. This will motivate student’s self-learning when the teacher is not available.

According to our discussion above, future work is needed to investigate in more detail the origin of the student’s mistakes. We must follow more closely student’s reasoning. For this purpose, the system should have for each complex problem a set of simpler associated problems in order to isolate with certainty the true origin of the student’s mistakes. A classification of the problems according to their difficulty level is needed: complex, intermediate and basic ones. We can start with a complex problem and, if necessary, go back to intermediate or even to basic ones. We can also do this in the opposite direction. In this way, the system can go forward and backward providing problems of different difficulty levels according to student’s actual mastering level of physics concepts and their ability for problem-solving.

We plan to approach this problem in the near future by means of Bayesian probability models and also including a knowledge-objects theory.

This kind of intelligent, adaptive and dynamic software has a strong academic and administrative potential, and can easily be extrapolated to other disciplines including math, computing and business process management, among others.

REFERENCES


AUTHOR INFORMATION

Luis Neri Head Science Department, Associate Professor, Tecnológico de Monterrey, Campus Ciudad de México, neri@itesm.mx.

Víctor Robledo-Rella Assistant Professor, Science Department, Tecnológico de Monterrey, Campus Ciudad de México, vrobledo@itesm.mx.

Enrique Espinosa Assistant Professor, Computer Science Department, Tecnológico de Monterrey, Campus Ciudad de México, enrique.espinosa@itesm.mx

Julieta Noguez Associate Professor, Computer Science Department, Tecnológico de Monterrey, Campus Ciudad de México, jnoguez@itesm.mx