Reasoning About Systems of Physics Equations

Chun Wai Liew[1] and Donald E. Smith[2]

[1] Department of Computer Science, Lafayette College, Easton PA 18042
liew@cs.lafayette.edu
[2] Department of Computer Science, Rutgers University, New Brunswick NJ 08855
dsmith@cs.rutgers.edu

Abstract. Many problems in introductory Physics require the student to enter a system of algebraic equations as the answer. Tutoring systems must be able to *understand* the student's submission before they can generate useful feedback. This paper presents an approach that accepts from the student a system of equations describing the physics of the problem and checks to see if it is correct. When it is not, the student's equation set is analyzed vis-a-vis one or more correct sets of equations, known physics concepts, and algebraic transformations. During this analysis credit-blame assignment is performed to identify one of several types of errors including 1) algebraic errors, 2) one or more omitted physics concepts, 3) incorrect instances of a required physics concept, and 4) use of an inappropriate physics concept. Experimental data collected from an introductory physics class is summarized and discussed vis-a-vis other methods. Results indicate that the techniques applied are effective at localizing most errors but that more work is needed to distinguish between algebraic and conceptual errors.

1 Introduction

One of the keys to being a good tutor is the ability to identify and localize the error in a student's answer and then generate useful focused feedback. Effective feedback enables students to correct their mistakes or conceptual errors without the tutor explicitly giving them the answer. An Intelligent Tutoring System (ITS) must not only determine if an answer is correct or incorrect but when incorrect it must localize and characterize the incorrect or missing part(s) of the answer. The degree to which an ITS can (1) localize an error in the answer, (2) identify the conceptual errors that contributed to the mistake and (3) analyze the problem context to generate feedback, greatly impacts on the effectiveness of the system.

Many problems in introductory Physics require the student to enter a set of algebraic equations as the answer. Tutoring systems must be able to *understand* the student's submission before they can generate useful feedback. This paper presents an approach that accepts from the student a system of equations describing the physics of the problem and checks to see if it is correct. When it is not, the student's equation set is analyzed vis-a-vis one or more correct sets of equations, known physics concepts, and algebraic transformations. The ability to apply algebraic transformations to the system of equations results in there being many equivalent forms of the answer that a tutoring system must be able to recognize and analyze. Most credit-blame assignment algorithms would map the student's equation to an equation in a solution set; however, it is computationally infeasible to either generate or store all the equivalent answers in this context. We have developed an algorithm that uses algebraic transformations to change the equations in the answer set into a form that is similar and suitable for comparison to the student's answer. The mapped equations are then compared and the similarities and differences are used to identify the errors that have been made.

2 Separating Physics from Algebra: an Example Problem

One of the many challenges to employing ITS techniques in a quantitative domain such as physics is the role played by algebra. It is required to solve such problem but at the same time enables misused concepts to masquerade as algebraic mistake. The resulting ambiguity (i.e., is the mistake algebraic or is the student confused about how to apply the physics concepts) presents a significant challenge to tutoring systems. In a perfect world the student would provide an *untransformed* set of equations that describe the physics of the problem and then proceed to algebraically solve them. In the real world, inputs provided by students are somewhere in between the desired *untransformed* equations and the final solution. The difficulty in generating effective feedback is exarcebated when algebra and physics are interwoven. To compound the problem, giving inappropriate feedback may often lead to a correct answer and even reinforce *bad habits* a student may have acquired.

To date all approaches to this problem, and there are few of them, compare the student's set of equations to a correct set of equations. The assumption is that the differences are sufficient to generate the appropriate feedback. As we will show in this paper it is difficult to (1) ensure that the appropriate equations are compared and (2) classify the error when the differences are identified. In the context of this paper, we are ignoring the issues of dimensional correctness of the equations. An earlier paper [1] describes our work in this area.

This section outlines the issues involved in ensuring that an equation in the student's set is appropriately and correctly mapped to an equation in the answer set. The discussion is illustrated with an example problem based on pulleys. A simple mechanism that is introduced early in the Physics curriculum is an Atwoods machine, a massless pulley with two masses, m_1 and m_2 hanging at either end as is shown in Figure 1. A common problem based on the Atwoods machine is to ask the student for the equation(s) that would determine the acceleration of the mass m_1, assuming that m_1 and m_2 are not equal.

A set of equations that would solve the problem is shown below. Equations 1, 2 and 3 deal with Force while Equation 4 is concerned with Acceleration.

$$T_1 - m_1 * g = m_1 * a_1 \tag{1}$$

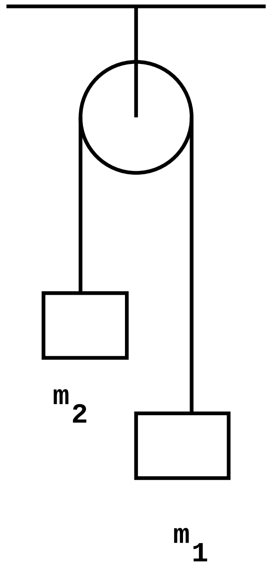

Fig. 1. Atwoods Machine

$$T_2 - m_2 * g = m_2 * a_2 \tag{2}$$
$$T_1 = T_2 \tag{3}$$
$$a_1 = -a_2 \tag{4}$$

This set of equations leads to the following solution

$$a_1 = (m_1 - m_2) * g/(m_1 + m_2) \tag{5}$$

An example of a set of equations that is different but also correct is:

$$T - m_1 * g = -m_1 * a \tag{6}$$
$$T - m_2 * g = m_2 * a \tag{7}$$

A student could enter any of the sets of equations (1 -4, 5 or 6-7 and they should be considered correct. These equations highlight some of the issues that an ITS has to deal with when it analyses a set of equations and attempts to identify the equations. The issues arise from the many equivalent syntactic forms of the equations. An ITS must be able to recognize and identify these equivalent forms of the equations. The recognition problems frequently cause an ITS to compare two equations that are not comparable. This results in the system generating feedback that is erroneous and therefore confusing to the student.

In comparing the student's equations with the answer set, an ITS must determine the mapping of the variables and subsequently equations from one set to the other. This process is complicated by several issues:

1. *variable renaming:* The student and the instructor may use different variable names to represent the same quantities. There is no restriction on the names of variables or choice of subscripts even though there are many standard variable names but there are also many commonly used variations, e.g., F, $Fnet$, F_1 usually represent a Force.
2. *simple aliasing of one variable:* Frequently, variables that have the same magnitude, coefficients and dimensions are aliased for one another. For example, the variables T_1 and T_2 in equations 1, 2 and 3 are equivalent to one another. In equations 6 and 7, there is only a single variable T_1 that is used to represent both, i.e., T_1 is an alias for T_2. When variables are aliased, then the number of equations in the set is also reduced.
3. *mapping of coefficients for a pair of variables:* Aliasing of variables also occurs when the coefficients are not the same. An example equation is:

$$a_1 = -a_2$$

   In this case, the variables are not directly equivalent because they have different coefficients (sign and possibly magnitude).
4. *elimination of a class of "equivalent" variables:* There are many ways to specify the algebraic solution to a problem. These may involve using a greater or lesser number of variables and thereby a greater or lesser number of equations. Sometimes variables that are in an equivalent set are completely eliminated and in the process the corresponding equations that specify the equivalences are removed. For example, one very different but correct solution to the example problem is:

$$m_1 * g - m_1 * a = m_2 * g + m_2 * a$$

   In this case, there is no variable representing the tension of the rope (commonly T, T_1, T_2). Instead that variable has been eliminated from the set and a corresponding equation has been eliminated.
5. *choice of coordinate axes:* For every vector, e.g., Force or Acceleration, there is a choice of coordinate axes. In the example problem, everything is aligned vertically but there is still the choice of whether the acceleration is positive in the up or down direction.

$$T_1 - m_1 * g = -m_1 * a_1 \tag{8}$$
$$T_1 - m_2 * g = -m_2 * a_2 \tag{9}$$
$$a_1 = -a_2 \tag{10}$$

   The first set of equations (Equations 8-10) is consistent with a choice of a_1 being positive in the same direction downwards and opposed to T_1. The second set of equations below (Equations 11-13 is consistent with a_1 being positive in the upwards direction. An ITS should be able to detect and recognize the choice of axes, and thus the correctness of both sets of equations.

$$T_1 - m_1 * g = m_1 * a_1 \tag{11}$$
$$T_1 - m_2 * g = m_2 * a_2 \tag{12}$$
$$a_1 = -a_2 \tag{13}$$

All the above issues must be dealt with if a system is to make sense of the student's submitted answer, even with the help of a set of correct answers. The above issues arise from the use of algebraic shortcuts by the student, a use that cannot be and should not be discouraged. Unfortunately, their use obfuscates errors and makes it much harder for an ITS to identify errors and differentiate between conceptual and algebraic errors.

3 Related Work

Most Physics tutoring systems do not support the use of equations for input. Most of the systems either ask for numeric input or use multiple choice questions[2–4]. In an earlier version of the system[5], we described a technique for performing credit-blame assignment on single equations. This technique is unable to deal with the issues and complexities that arise when the answer is a set of equations instead of a single equation.

It is fairly easy to determine if an equation is correct if all the variables have been defined. One simple way is to substitute various values for the variables and evaluate the equation. This is the approach taken in many existing systems. It is much harder to determine what is wrong about an equation once it has been labeled incorrect. One simple approach (used in an early version of the ANDES system [6]) is to generate all possible correct equations and find the closest match to the student's submission. This turns out to be computationally infeasible as the space of possible equations is very large. The current ANDES system [7] instead generates a much smaller set of equations. This set represents the most common choices of coordinate axes (issue 5). The system uses an algebraic solver to determine the correctness of a student's equation. If the equation is incorrect, the system uses a set of mal-rules [8] to perturb its set of equations to find a closest match to the student's equation and thus generate feedback based on the mal-rule selected. This obviously can lead to cases where completely erroneous and confusing feedback is generated. Note that the system still has to deal with issues such as variable aliasing(Issue 2), mapping of coefficients (Issue 3), variable elimination (Issue 4) and choice of coordinate axes (Issue 5), when trying to map a student's answer to one that it has generated. This approach also requires a great deal of effort to create a knowledge base that has sufficient Physics knowledge to generate the required solutions and the choices. ANDES also takes the approach of verifying each equation as it is entered by the student. This gives the student immediate feedback as information is entered. The system can take this approach because all the variables have been defined *a priori*. The main concern for ANDES is to map the student's equation to a member of the set that is presently in the system.

4 Approach

The first step towards "understanding" a student's submission is to classify each equation as either (1) correct, (2) almost correct and (3) incorrect. The approach described in this paper compares the student's submission to a single solution set to classify each equation. Correctness preserving algebraic transformations are used to change one set of equations (e.g., the system's) into a form that is equivalent to the other (e.g., the student's), so that individual equations may be compared. Only a single set of equations is used unlike the multiple sets that are used in the ANDES system. The technique can determine the differences between an incorrect and a correct equation but has some difficulties in differentiating between algebraic and conceptual errors.

There are two types of algebraic transformations used in our system. The first is used to transform equations into a canonical sum of products form. The second type is used to transform the set by reducing the number of variables and hence the number of equations. These two types of transformations are used to transform the student's set of equations so that each equation can be heuristically mapped to an equation from the answer. Errors are identified by comparing each equation in the student's set with the corresponding equation from the solution set. Variables are not defined *a priori* but their dimensions are inferred by the system based on their name (e.g., a_n typically stands for Acceleration). Currently the student is required to enter all the equations before the system starts the analysis.

4.1 Algebraic Transformations

The two types of correctness preserving transformations are used to:

1. *transform equations to canonical form:* These transformations are rules for rewriting equations and transforming these equations into a canonical sum of products form. An example of these rules is one that multiplies each term by the divisor of a term. The primary effect of this rule is to remove all denominators. The canonical form is then used to *validate dimensions.* The dimensions of each variable and constant in the equations is inferred, thereby determining the dimensions of the equation. This can be done using a constraint based algorithm that is described in detail in [1].
2. *rewrite set equations to eliminate variables:*
   This type of transformation eliminates variables and thereby equations from the set. In a prior step, the dimensions of each variable is determined. The variables used in the student's set are then heuristically mapped to variables in the answer set. Any variables that are unmapped are then eliminated from the set using the transformations. The transformations are used for:
   - *the elimination of equivalent variables:* Equations in the form $v_1 = k * v_2$ are removed by substituting $k * v_2$ everywhere the variable v_1 occurs.

- *the elimination of a class of variables:* The previous transformation eliminated variables that could be substituted for with an equivalent variable. There are times when a class of variables is not present in one or other set of equations. Usually, this happens when one set uses intermediate variables and the other does not. This transformation uses gaussian elimination to remove the variables from the set of equations.

4.2 Mapping Variables and Equations

Transformations are used to convert a set of equations so that it has the same number and type of variables as the set that it is being compared to. The variables from one set must then be mapped to the variables of the other set. The algorithm looks for a one-to-one match between the variables in the two sets of equations and handles simple aliasing of names. The heuristic technique uses knowledge of well-known variables (m usually represents Mass) and common use of subscripts to find a mapping.

Once all variables have been mapped, the algorithm maps equations in the student's submission to equations in the answer set. The algorithm only maps equations that have the same dimensions. The heuristic used determines how good a match is based on the terms in the equation, i.e., whether a term is present or absent in the equation that is being compared to. There is a pre-set limit on the maximum number of differences allowed. If the number of differences exceeds the limit, the equation is marked as unmapped, i.e., there was no corresponding equation in the answer set of equations.

4.3 Identifying Errors and Generating Feedback

Each of the student's equations is compared with the corresponding equation from the answer set to determine the differences between the student's equation and the correct equation. Some of the detectable differences are listed below:

- switch in sign, e.g., use of a $+$ instead of a $-$.
- missing or additional term: e.g., use of $m_1 * g + m_1 * a$ instead of $m_1 * g$.
- missing or additional trigonometric factor: e.g., use of $m_1 * g * sin\theta$ instead of $m_1 * g$.
- incorrect coefficient: e.g., use of a_2 instead of $-2 * a_2$.
- incorrect factor, e.g., use of m_2 in place of m_1.

All of these differences do not change the dimensions of the encompassing equation. Any mistakes that would change the dimensions of the equation would result in the equation being mapped to another equation. Trigonometric factors are dimensionless functions so their absence or presence does not change the dimension of the term or equation.

The switch in signs is a special case in that every other error can be identified and localized by examining only the equation in which it occurs. A switch in signs could be an error but could also represent a choice of an alternative set of coordinate axes. This can only be determined by verifying that the sign switch is consistently applied to all equations.

The final step is to generate feedback based on the critique of the equations. This is not the only place where feedback is generated. Feedback is generated immediately if any of the equations are found to be dimensionally incorrect or if the set of equations is incomplete, i.e., there are an insufficient number of equations.

The feedback that can be generated is based on the differences that are detected. This results in suggestions to consider (1) whether certain quantities are acting in concert or in opposition (incorrect signs), (2) the impact of a body on other bodies (missing or additional terms) and (3) that a body is acting (or not) at an angle with respect to another body (missing or additional trigonometric factors).

If there are unmapped equations in the student's answer, then there is a corresponding equation in the instructor's answer. The system will generate feedback suggesting that the unmapped equation does not lead to the solution and suggest considering the interactions between certain bodies in terms of Physics concepts, e.g., Force or Momentum. Earlier work in this area is described in [5]

4.4 Limitations

The errors that students make generally fall into one of the following categories:

- algebraic errors: This is detected as an incorrect equation.
- missing Physics concept error: This is detected as an missing equation.
- incorrect application of concept error: This is detected as an incorrect equation.

Problems arise when the system is unable to determine the type of error from the difference between the correct answer and the student's submission. This is caused by the use of algebraic transformations that are correctness preserving but have no corresponding Physics concept. For example, if a student chooses to (1) not use the intermediate variable T to represent tension in the rope and (2) to use a to represent acceleration, she would end up with the equation:

$$a = (m_1 - m_2) * g/(m_1 + m_2) \tag{14}$$

The use of algebraic transformations to reduce the set of equations, in this case to one equation, poses many problems to the ITS when a mistake is made. For example, if the student switches the $-$ sign for $a_1 = -a_2$, the equation above, Equation 14, is reduced to:

$$a = g \tag{15}$$

With the original set of equations (Equations 1-4), there are five places where such an algebraic error could have been made. On the other hand, the student could have made a conceptual error and not realized that the two blocks are moving in opposite directions and therefore generated the equation $a_1 = a_2$. If the student uses algebraic transformations and enters a reduced set of equations, the technique is unable to differentiate between these sources of errors. This is an example of where the student's use of algebraic transformations makes it hard for an ITS to generate useful feedback. The technique can localize and identify the difference between the student's answer and the correct answer but that is insufficient. The system needs to be able to identify the source of the difference, i.e., the error that was made.

The degree to which the student uses algebraic simplifications greatly impacts the ability of the algorithm to identify the source of an error. If the student made the above error and submitted either two or three equations, the algorithm would only have to disambiguate between an algebraic error and a conceptual error. In that case, the algorithm would assume that a conceptual error had been made and generate the appropriate feedback.

A more advanced mechanism would (1) enumerate the causes of the algebraic error, (2) identify the ones that are more likely to occur, (3) identify ways to disambiguate between the likely errors and then (3) prompt the user for information to disambiguate. For the above example, the system might prompt the user with "What is the relationship (equation) between the acceleration of the block of mass m_1 (a_1) and the acceleration of the block of mass m_2 (a_2)?". We are currently exploring this avenue of research.

5 Experimental Evaluation and Discussion

In the spring of 2001, we collected roughly 175 answers to two Physics problems from 88 different students. The two problems were (1) the Atwoods problem described in this paper and (2) a problem involving a moving pulley. The students were enrolled in an introductory Physics course for engineers and science majors.

Analysis showed that dimension errors occurred in roughly fifteen percent of the answers. These errors were detected by the algorithm and described in [1]. With the rest of the equations, the algorithm correctly maps variables in ninety percent of the equations. However, it only correctly maps seventy percent of the equations. Much of the difficulty lies in disambiguating between an equation is slightly incorrect and should be mapped from an equation that is fundamentally wrong and should not be mapped.

The ability to provide good and useful feedback is very dependent upon how well the algorithm maps the equations. There are three main categories of answers that the algorithm failed on. The first is where the student's choice of coordinate axes is not the same as the one in the answer. This is an area where the ANDES algorithm currently performs better. The second category is composed of answers where the algorithm incorrectly mapped the equations. If an equation has been correctly mapped, the algorithm performs well in determining what is incorrect about an equation. The algorithm can then correctly identify the errors and generate the appropriate feedback. Finally, there are answers where the equations were correctly mapped and the differences were identified but the algorithm could not determine the cause of the error. This was a small set because the students had been explicitly instructed to write down all the applicable equations and **not** to perform algebraic simplification. Even so, approximately twenty percent of the students used some algebraic simplifications and submitted reduced sets of equations.

6 Conclusion

Reasoning about sets of algebraic Physics equations is a difficult task that is further complicated when students use algebraic transformations to *simplify* their answer. This paper has described an approach for performing credit-blame assignment on a set of algebraic Physics equations submitted by a student. The technique is based on analysis and comparison of the student's answer (a set of equations) with a correct answer (another set of equations). The technique uses algebraic transformations to identify and map variables and equations from the student's set to the answer set. Mapped equations are compared and the differences are used to identify errors and generate feedback. The algorithm can detect and identify several types of errors including algebraic errors, omitted Physics concepts and incorrect applications of concepts. The technique has been evaluated on answers collected from students in an introductory Physics course at a small college. The results show that the algorithm successfully analyzes most of the submitted answers and generates the appropriate feedback.

References

1. Liew, C., Smith, D.: Checking for dimensional correctness in physics equations. In: To appear in Fourteenth International Florida AI Research SocietyConference. (2002)
2. Novak, G.M., Patterson, E.T.: Just-in-time teaching: Active learner pedagogy with www. In: Proceedings of IASTED International Conference on Computers and Advanced Technology in Education. (1998)
3. Hubler, A.W., Assad, A.M.: Cyberprof: An intelligent human-computer interface for asynchronous wide area training and teaching. In: Proceedings of Fourth International World Wide Web Conference. (1995)
4. Novak, G.M.: World wide web technology as a new teaching and learning environment. International Journal of Modern Physics **8, No. 1** (1997) 19–39
5. Liew, C., Shapiro, J.A., Smith, D.: What is wrong with this equation? error detection and feedback with physics equations. In: Proceedings of Thirteenth International Florida AI Research SocietyConference. (2000)
6. Gertner, A.S.: Providing feedback to equation entries in an intelligent tutoring system for physics. In: Proceedings of the 4th International Conference on Intelligent Tutoring Systems. (1998)
7. Shelby, R., Schulze, K., Treacy, D., Wintersgill, M., VanLehn, K., Weinstein, A.: An assessment of the andes tutor. In: Proceedings of the Physics Education Research Conference. (2001)
8. Cohen, P., Feigenbaum, E., eds. In: The handbook of artificial intelligence. Volume 2. William Kaufmann, Inc., Los Altos, CA (1982) 317–318