An Algebra Subsystem for Diagnosing Students' Input in a Physics Tutoring System

Joel A. Shapiro

An Algebra Subsystem for Diagnosing Students' Input in a Physics Tutoring System

Joel A. Shapiro, *Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, 15260, USA*
and
Department of Physics and Astronomy[1], Rutgers University, Piscataway, NJ 08854-8019, USA
shapiro@physics.rutgers.edu
http://www.physics.rutgers.edu/~shapiro

Abstract. To help a student in an introductory physics course do quantitative homework problems, an intelligent tutoring system must determine information of an algebraic nature. This paper describes a subsystem which resolves such questions for Andes2. The capabilities of the subsystem would be useful for any ITS which deals with problems involving complex systems of equations. This subsystem is capable of 1) solving the systems of equations at the level of introductory physics problems, 2) checking the validity of equations the students enter, 3) investigating whether an equation is independent from a set of other equations, and if not, determining on which equations it does depend, and finally 4) providing tools to help the student with algebraic manipulations, including a "solve-tool" that solves her equations. The ability to determine dependence of equations is first used by Andes during problem generation, by providing information to that component of the ITS which generates correct solutions to the problem. Later, during tutoring, it enables the help module to model which equations the student appears to know. One new feature of this algebra subsystem is that it deals with the dimensional units of physical quantities throughout. An important change from a previous approach is in the meaning of "correctness" of an equation and in the method of determining which equations it can be derived from. The theoretical differences between the two methods, and the pros and cons of each, are discussed. Then we evaluate how the capabilities of the subsystem have affected the Andes tutor's effectiveness, with a particular emphasis on the effects of the changed method.

Keywords. intelligent tutoring systems, solving algebra problems, physics problem tutoring

INTRODUCTION

An intelligent tutoring system that attempts to give guidance to a student in solving a complex problem needs to be able to distinguish which pieces of a solution the student already knows from pieces with

---

[1]Permanent address.

which she might need help. If it gives guidance incrementally, it needs to analyze each new input from the student, figure out which piece of the problem this input is supposed to address, and then determine whether it does so correctly. In addition, of course, it needs to identify incorrect steps and attribute the error to something upon which it can give help. If the set of possible correct steps is fairly simple, it may be possible to have a list of templates for all possible correct steps, together with a pattern-matcher or other simple algorithm for deciding if the student's input matches a correct step. In a field with sufficiently rich methods, however, this may not be practical. Correctness of inputs may then need to be judged by constraints (Ohlsson, 1993). It may be necessary to use discipline-specific methods to formulate these constraints, as well as in analyzing how a student has broken down a problem.

One field in which this richness of approach occurs is introductory physics at the college/university level. A major component of the educational process here is the solving by students of quantitative problems. This requires the student to 1) analyze a physical situation, often consisting of many parts, 2) extract from general physical principles relations among the physical quantities, 3) assign algebraic variables to describe those quantities, 4) translate the relations into explicit algebraic equations, and finally, when a sufficient set of equations is found, 5) use the tools of algebra [2] to solve for the required unknown quantity in terms of the given values.

For example, in the problem shown in Figure 1, the student must first recognize that the 30 kg block has three forces acting on it, and the 20 kg block two forces, that the accelerations are along known directions, and that the distance moved can be determined from the acceleration. Secondly, she must recognize that Newton's second law will determine the accelerations, that the two accelerations and the two tension forces are equal in magnitude due to the cord, and that the distance moved is given by constant acceleration kinematics. Thirdly, she must choose coordinate systems and variable names to describe all the forces, accelerations and the distance moved. Fourth, she writes down the equations, including perhaps $T_1 - m_1 g \sin 25° = m_1 a_1$ and $d = \frac{1}{2} a_2 t_2^2$, which will solve the problem, and fifth, she uses standard algebraic manipulations to solve for d. Notice these steps may be implemented in many ways. For example she might have defined the weight W_1 instead of directly using $m_1 g$. Often vector equations take different forms when components with respect to different axes are used. Had she implemented the basic physics principles in their fundamental form, she would also have defined a variable for the net force on block 1. Students rarely write down the application of physics principles in their most fundamental form, but rather they combine a number of observations and write down only a composite equation. While this is certainly to be encouraged to some extent, it greatly complicates the task of identifying the student's input. An example will be given at the end of the "Requirements..." section. In the system to be discussed, the student input is primarily in the form of defining variables and entering equations. As mentioned, in a complex problem, not only are the individual equations each describable in many forms, but there are myriad sufficient *sets* of equations, with no one-to-one mapping between individual equations of two sufficient sets. Thus modeling students' knowledge is much more complex than simply checking off equations, one at a time, as each student equation is

---

[2]"Algebra" is being used in a loose sense throughout this paper, meaning the methods of making mathematically correct manipulations on equations that one learns through the pre-calculus course, rather than in the sense of a vector space over a field with multiplication, as one meets in an abstract algebra course. In particular, these equations will involve dealing with trigonometric functions, exponentials and logarithms, as well as polynomials.

entered. Acceptable student solutions are not describable as successive steps along a predetermined path, but rather involve generating sufficient correct constraints, in the form of equations, to determine the answer.
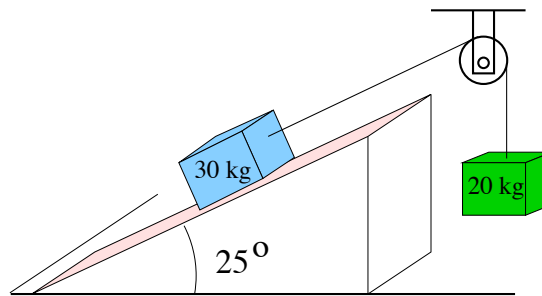
An inclined plane making an angle of 25.0 degrees with the horizontal has a pulley at its top. A 30.0 kg block on the plane is connected to a freely hanging 20.0 kg block by means of a cord passing over the pulley.
Compute the distance that the 20.0 kg block will fall in 2.00 seconds starting from rest. Neglect friction.

Fig.1. A typical Andes Physics problem.

There are a number of systems designed to deal with homework problems in introductory physics. Most (CyberProf, 2001; Hübler & Assad, 1995; Raineri et al., 1997; UTexas, 2005) are designed to give right/wrong feedback on multiple choice or numeric answer questions only. Some, like WeBWorK (Gage & Pizer, 1999) and (WebAssign, 2005), can handle symbolic expressions as well. These all consider only the student's final answer, and cannot provide help along the way. Of a more tutorial nature are the "Personal Assistants for Learning" (PALs) of the CIRCLE group (Reif & Scott, 1999), which lead the student through a tightly structured interaction with multiple choice responses. The PAL developers explicitly renounce artificial intelligence, so each problem's tutorial path is explicitly authored. (Mastering Physics, 2004), which emerged from Cybertutor (Pritchard & Morote, 2002), works similarly, though it can go beyond multiple choice. Real artificially intelligent tutorial systems have been developed for teaching algebra and other mathematics at the pre-college level (Carnegie Learning, 2005; Corbett et al., 1997). In the 1970's, a great deal of effort was made to produce an ambitious system for teaching physics, called Plato (Sherwood & Stifle, 1975; Woolley, 1994), but disappointing results, together with the huge expense of computers at that time, seem to have killed it. To our knowledge, the only current system designed for reasonably free problem solving at the level of introductory university physics is the ANDES tutor (Andes, 2002, 2005; Gertner & VanLehn, 2000; Schulze et al., 2000) developed at the University of Pittsburgh and the U. S. Naval Academy, and used at the USNA.

This paper describes a subsystem designed for integration into Andes. The subsystem provides an oracle for answering questions of an algebraic nature that the full system needs in order to determine correctness of student steps, identify the mistaken steps, and to model the understanding of the student so as to effectively provide help to the student. In Andes, the full system also needs the information the subsystem can provide about independence of equations in order to generate solution paths and a constraint network of the variables and equations in a problem. The particular system in which this

subsystem has been used, first in the Autumn of 2001, is Andes2, a revision of Andes. Andes' previous method for finding this information could not handle complex problems. The goal of the new subsystem is a strong and robust tool for providing the information without limiting the scope of physics problems presented. The issues addressed here are likely to be of use more broadly, in any tutorial system designed to deal in some generality with science or engineering problems that involve algebraic equations among physical quantities.

The next section outlines what Andes does (Schulze et al., 2000), what sort of questions it needs to ask of its algebra subsystem, and what use it makes of these answers. The following section describes in detail the capabilities required of and provided by the subsystem. Then we focus on new methods of answering two of these questions, the correctness of a student equation and upon which "canonical" equations the equation depends. We discuss two meanings of "correctness" for an equation, well-motivated correctness versus algebraic correctness, and their relative advantages. Finally we evaluate the effectiveness of the subsystem within Andes and of Andes as a pedagogical tool.

ANDES

Preparing a problem in Andes

There are two stages in the use of Andes for any given physics problem. The first is the preparatory stage. Andes accepts a formalized description of a problem, from which it develops a solution and the structures it will need in the second phase, the tutoring of students on the problem. In the preparatory stage, it uses a knowledge base of physical principles to construct a constraint network, consisting of variables and "canonical" equations in those variables, which is sufficient to find the solution to the problem. It also generates a set of solution paths (Gertner et al., 1998), or more accurately, a set of partially ordered subsets of the constraint network. Each path terminates with a subset of equations from which it is possible, algebraically, to extract the value of the required physical quantity. To generate this information, the system needs help from the algebra subsystem:

- To determine whether adding a given equation advances the solution of a problem beyond what is already specified by the previous set of equations. Roughly speaking, each new equation will determine the value of one previously unknown variable, at the possible expense of introducing new unknown variables. But it will do so only if the equation is independent of the equations already in the set.

- To determine whether the set of equations is sufficient for solving the problem. In particular, the methods used by the new algebra subsystem require an actual solution for all the variables in the problem, so the subsystem must be able to solve systems of equations.

The preparatory procedure produces a file which contains the problem solution, a list of all variables relevant to the problem, and the canonical equations, which together constitute the constraint network. It also lists the solution paths. Thus it stores all the problem-specific information necessary for the tutoring system to present a given problem to the student and to provide help to the student in working through the solution of that problem.

Andes' tutoring stage

Once this is done, Andes is able to tutor a student on the problem. It has a user interface which presents, in text and figures, the problem statement, and then allows the student to define variables, draw vectors, define axes with which to describe these vectors, and write equations. Andes is designed to let the student proceed with the solution of a problem without interference for as long as the student is on an acceptable path, making correct entries involving possibly relevant variables. It does give feedback for each student entry by turning the entered objects from black to green if correct, or red if not. Upon seeing her input turn red, a student might spontaneously correct what is wrong, or ask "what's wrong with that". As the student defines variables, Andes requires that each variable correspond to a physical quantity that could occur in trying to solve the problem, as enumerated by the preparatory solution-generating procedure. The student's equation is only accepted if it is given in terms of variables the student has already defined. Thus any student equation received by the algebra subsystem will be in terms of recognizable variables [3]. One crucial task for the tutoring system is to be able to distinguish correct equations from incorrect ones. In addition to "what's wrong" help, Andes can provide "what's next" help on request, when the student needs hints as to how to proceed. Besides tutoring help, the system can provide help of a more mechanical form — it can help the student solve the equations she has written.

Typically the solution of a problem by a student is an involved process, taking on average more than 20 minutes, with the tutor providing about 8 explanations (Shelby et al., 2001).

To provide this help, the system needs to be able to answer the questions:

- Is the equation the student wrote down correct?

- What can one conclude the student knows of the constraint network from what she has written down?

- Can a set of equations be solved for all variables, either in explicit numerical form or in terms of a few undetermined parameters? If so, it must provide the solution.

Goals for the subsystem

It is the algebra subsystem which provides specific information on the questions discussed. The subsystem is therefore designed to determine the correctness of submitted equations, based on the "canonical" equations, and to provide information about the student's knowledge of these equations. The new subsystem uses new methods, at least compared to the older version of Andes, for determining correctness and attributing knowledge to the student. These methods are based on the following observation: If any set of canonical equations has a solution space contained within the solution space of the student's submitted equation, her equation can be derived from the equations in the set. That is to say, the equation

---

[3]While the student may use her own notation for all variables, the fact that Andes2 requires a detailed definition of each defined variable eliminates the problem of variable identification. This detailed scaffolding for the student is pedagogically useful for beginning students, but is likely to get onerous for students who have made the definition clearly in their mind but resent the time-consuming process of relaying it to the tutor. An alternate approach, in which the tutor attempts to identify which physical quantity a student variable represents, is described in Liew & Smith (2002a,b).

can be derived from any set of which it is not independent. Furthermore, this can be determined without constructing such a derivation, which means that much more complex systems of equations, which arise from multistep problems and which were previously intractable, can be handled by the new algebra system.

In addition, the new algebra system incorporates physical units in all calculations, checking dimensional consistency as the first step in checking any equation. It has the capability of converting to and from SI units to others explicitly given. It also includes tools to provide algebraic help to the student in solving the equations she has written.

REQUIREMENTS AND USES OF THE ALGEBRA SUBSYSTEM

In this section, we describe what questions Andes asks of the algebra subsystem, and what use it makes of the answers. The demands made on the algebra system are:

- Given a set of correct equations, to solve, as much as possible, for all the variables in terms of "known" quantities. Known quantities may be either numerically given or described as "parameters", that is, independent undetermined quantities, in terms of which solutions may be expressed.

- To determine the correctness of an equation, given the set of correct canonical equations and their solution.

- To give reasons an incorrect equation is incorrect, such as dimensional inconsistency.

- To determine whether a correct equation might have been derived from a given set of independent correct equations, and if so, on which equations within the set it depends.

What use is made of the answers will be discussed in this section, while the methods used to determine them will be discussed in the next.

Solving a set of equations

There are several reasons why Andes needs a system which can solve the algebraic equations. One, of course, is to judge whether a presented answer is correct. Another is to be sure, when it generates what it believes is a complete solution path, that the required information can in fact be extracted from the equations.

One more use Andes makes of the solving ability of its algebra subsystem is to provide a tool for the student. The Physics professors currently using Andes regard their task as teaching the physics concepts, and not in exercising the algebra skills of their students. Thus they regard the primary task for the student in solving a problem is to write down a sufficient set of equations which follow from physics principles as applied to the problem at hand. They are happy to provide the students with a tool, even if it is a black box, which will solve the equations they have written.

The algebra system provides several "solve-tools" of varying power available to the student, to eliminate some of the drudge work of actually employing the equations to derive an answer. While it is

not clear pedagogically just how much of the work the system should take off the student's shoulders, there is no doubt that plugging numbers into equations is something the student presumably knows well enough not to need continual practice, and that she will appreciate having the algebra system do it for her at her command. Three tools are implemented:

the genie After checking that the student has entered correct equations that can determine the answer, give the answer to the student. This tool will give no explanation of how the algebra was performed (hence its name).

the simplifier The student selects an equation. This equation is then evaluated by plugging in all assignment statements the student has given, and the result then simplified.

solve-and-sub This tool asks the student to select an equation and a variable solvable within it (possibly in terms of other variables), and solves the equation for that variable. Then the student can select other equations containing the solved-for variable and have the solution substituted in and the resulting equations simplified. This would permit the student to guide the system to solve simultaneous linear equations. Thus it would probably be more suitable than the genie for engineering students, for whom the genie might be disabled.

The help system currently incorporates only the genie solver mode.

Finally, there is an internal reason for the subsystem to be able to solve for the values of all the variables in the system. The method used to determine correctness and independence of equations involves numerical evaluation about the solution point, so certainly that point must be known. As many problems involve scores of variables in addition to the one sought in the problem statement, it would be onerous to ask the creator of a problem to provide a full solution.

Solving equations is, of course, what one expects of a computational algebra system. For a discussion of why commercial computational algebra systems seemed inadequate to Andes' needs, and why the concerns of our solver are rather different from those of computational algebra systems, see the section "How the subsystem solves equations" below. But one issue is worthwhile discussing here, pertaining to the treatment of parameters.

Parameters are physical quantities that do not have known explicit values, and whose values are not determinable from the information given. When a physics problem involves such parameters, it may be asking for the value of a sought quantity as an algebraic expression depending on the parameters. There are, however, also cases in which the answer is unaffected by the value of the parameter. For example, in the elastic scattering of a cue ball off another billiard ball initially at rest, one may ask for the final velocity of the cue ball as a function of the two influencing parameters, the initial velocity and the scattering angle. The answer is unaffected by the third parameter, the common masses of the balls. Even though the answer is not affected by the mass, variables that are essential to solving the problem, namely the momenta of the balls, are affected, so that the complete solution of the set of canonical equations does involve the mass as a physical quantity.

Our algebra system would have a very hard time solving a problem such as this if forced to keep all the mentioned parameters as algebraic variables. Fortunately, Andes does not require that. For the purposes of checking that a solution exists, or that an equation is correct, or for determining dependence,

it is enough to answer those questions when the parameters are set to particular numerical values, as long as the answer to the questions does not depend on the values used. The solver assigns to each independent parameter a random "ugly" value, one which could not conceivably arise by solving an incorrect equation the student might write down for the variable in question. In the problems Andes addresses, this is enough to ensure that the answers do not depend on the values chosen. The reasons will be discussed further in the section on how the subsystem answers the question of correctness of equations. Thus this method of simplifying the solution process does not limit Andes from anything it is asked to do, although it does preclude it from giving the student the answer, if it is to be given as an expression in terms of the parameter.

Checking correctness of student equations

While a tutoring system might, under some conditions, object to a correct equation as being premature or inappropriate, it must always object if the equation is algebraically incorrect. In earlier versions of Andes the correctness of student equations was judged by whether the equation was equivalent to one on a list. The list was generated during the preparatory process by applying a set of rules for algebraic manipulation to the set of basic, or "canonical", equations produced by the knowledge base from the problem specification (Gertner, 1998). For each derived equation the generator recorded the set of canonical equations used. If the student's equation could be found as a simple algebraic manipulation of one of the derived equations on this list, it was considered correct, and which equations it depended on was given by the corresponding set. This method requires combining the full set of canonical equations in all combinations a student might generate correctly (Gertner, 1998). Unfortunately, the number of canonical equations involved, even in fairly simple physics problems, is much larger than a typical human solver would imagine. For example, in the inclined plane problem shown in Figure 1, Andes2 generates [4] 45 equations in 41 variables. The number of possible ways of combining these into a correct equation is immense. Generating such a list proved unwieldy on all but very simple problems.

The new algebra system takes a different approach. Let us call an equation "balanced" if the two sides agree on the subspace of the variable space specified by the equations which follow from the problem specification. If there are no free parameters, that means just that when the numerical values of the problem solution are inserted, the two sides of the equation match. If there are free parameters, it means the two sides are the same function for all allowable values of those parameters. Our method judges a student equation to be correct if it is balanced. For all problems the tutor addresses, the problem specification implies the solution, so the correctness of the student's equation is judged by simply plugging in the solution and evaluating the student's equation. In the inclined plane problem, where there are no free parameters, the correct numerical values of the 41 unknowns are substituted into the student equation and correctness is determined by whether the two sides balance. As correctness in Andes is indicated by turning the equation green and incorrect equations are turned red, the Andes group calls this approach "color-by-numbers".

---

[4] Andes treats all mechanics problems as at least two-dimensional, and needs to accommodate all quantities a student might refer to. In particular, many of these variables are angles which are transparently zero, but are still needed for the fundamental equations which describe this problem.

One might ask whether it would be more appropriate to define correct as derivable from the "canonical" equations, which follow from the problem statement and known physical principles, by some set of algebraic manipulations. The answer depends on exactly what is meant by derivable. If, on the one hand, it means that there exists an algebraically correct procedure for deriving the student's equation from the input, then there is a formal proof[5] that derivability is equivalent, in our context, to checking that the equation balances. If there are no parameters, one can check if the equation balances simply by plugging in the numerical values of all variables, with due attention to numerical accuracy. If there are parameters, balancing would require the two sides differ by a function which is zero for all allowable values of the parameters. The way to check correctness when there are parameters is to assign to these variables random values, and if that function evaluates to zero, to conclude that the difference is, in fact, zero[6]. Thus balancing can be judged by evaluating as green in color-by-numbers, and algebraically correct derivability is equivalent to color-by-number.

On the other hand, one might mean something else by derivability. It might mean that the student's equation could arise in a derivation starting from the input and proceeding by rationally motivated steps towards finding a solution. If her equation could never arise in that context, the student should not be writing it down. To use this definition, well-motivated correctness, would require specifying some finite set of methods by which such manipulations should proceed. For example, one could permit solving one equation for one variable in terms of the other variables and substituting the results into other equations. This, however, may not lead to a convergent procedure, so any attempt to generate all satisfactory equations will need to use a more restrictive method. The differences and limitations of these two methods will be discussed, and an example of an algebraically correct but rationally unmotivated calculation will be given, in the Evaluation section.

Help in finding errors in equations

When a student enters an incorrect equation, the tutor needs to try to identify what is wrong. As it may not know what equation the student was aiming at, a comparison of expressions may not be useful. One way to find the source of some errors is to perturb the entered equation in various ways, and ask the algebra subsystem if the resulting equation is correct. For example, signs of terms can be flipped, or sines and cosines interchanged. Because checking an equation with color-by-numbers is very computationally cheap, a large number of perturbations can be checked. Nonetheless, many near misses will not be identified by this process.

---

[5]Proof: All Andes problems can be solved by algebraically correct steps, so the solution for all variables can be derived, $v_i = f_i(\{\lambda_j\})$, where $\{v_i\}$ is the set of all variables in the problem, $\{\lambda_j\}$ is a (possibly empty) set of underdetermined parameters, and f_i are a set of functions explicitly determined by solving the canonical equations. Suppose the student has written an equation equivalent to $S(\{v_i\}) = 0$, where S is any algebraic expression in the variables $\{v_i\}$. If we can show that S is indeed 0 when we substitute $f_i(\{\lambda_j\})$ for v_i in S, then, because substituting one expression for another to which it is equal in an algebraic expression is a legitimate algebraic step, we have derived $S(\{v_i\}) = 0$.

[6]If the difference is an analytic function other than zero, the chance that the random values chosen happens to be the solution is nil. There are functions which may be zero for a range of parameters, such as $|v| - p$ given that $v^2 = p^2$, which may not vanish for all possible values of the parameter p, but Andes problems tend to specify enough information, such as "p is positive", to avoid such problems.

Dimensional Analysis

One form of mistake the algebra system can effectively detect is in locating dimensional inconsistencies. One of the basic techniques physics teachers try to impart to their students is that they should always check that their equations and values have consistent physical units (Halliday & Resnick, 1981; Serway & Beichner, 2000; Tipler, 1991). If, trying to recall the formula for the area of a circle, a student remembers $2\pi r$, she should be able to discard that because she realizes areas are measured in square inches or square meters, while the radius is in inches or meters, so it is impossible for $A = 2\pi r$. We all recall the \$125M mission to Mars lost because the required thrust was calculated in pounds, but the units left off, and only that number of newtons was applied (Pollack, 1999). So it is very poor pedagogy to have a tutorial system that ignores units. A system that is able to point to dimensionally inconsistent operations can provide important feedback on what is wrong with an incorrect student equation. Andes2 informs a student of dimensional inconsistency as the first step in checking entered equations.

When physicists or engineers use a computer to do their calculations, they have already verified their equations and chosen appropriate units, so except for oversights like the Mars disaster, it is generally sufficient to have their programs work with pure numbers. Thus most major tools for calculation do not integrate units in any essential way. But a tutoring system needs to recognize that $K = mv$ is the wrong formula for the kinetic energy ($K = \frac{1}{2}mv^2$) even in a problem giving the numerical value for v, measured in m/s, as 2. It can know this because the left hand side of the equation has units kg·m^2/s^2 while the right hand side has units kg·m/s.

In Andes1, as in many other systems, the lack of treatment of units meant that one needed to assume that all units in the problem were consistent. If you look at the problems in elementary physics books, you will find that the overwhelming majority of the ones before the modern physics sections do employ only SI standard units, but even there, there are some values for time specified in minutes. It is doubtful that even European children have a good feel for the speed of their favorite car in m/s. And there is one quantity for which the "standard unit" is quite unfamiliar to freshman — angles. Angles are dimensionless, as can be seen from the formula for the length s of an arc of angle θ and radius r: $s = r\theta$. As s and r are both measured in meters, θ is measured by a pure number. But how big is an angle of 2? It is 2 radians, not 2 degrees. Nonetheless degrees are used extensively in stating physics problems. Thus Andes1 was inconsistent in its requirement that all quantities are measured in standard units, and would have run into trouble soon, when dealing with angular velocities and momentum.

For both these reasons, but most crucially to allow degrees, the algebra system allows for quantities to be specified in non-standard units. All internal calculations are done in SI units, but a preferred set of units can be specified for each variable, and numerical values can be given together with any of a large set (though not at all a complete set) of units.

Derivation and dependence of equations

Because the tutor needs to keep track of what parts of the solution the student already used, as the student enters a new equation the tutor must try to analyze which canonical equations were used to derive that equation. One way to answer such questions is to search for a derivation using a predetermined set of

algebraically correct operators[7]. This has been used in tutors for calculus (Yibin & Jianxiang, 1992), for electric circuits (Brna & Caiger, 1992) and in multicolumn subtraction (Ohlsson & Langley, 1986). In a complex domain, the search may be prohibitively slow. The earlier Andes tried to generate all correct equations in advance, so that a submitted equation's correctness and derivation could be determined by matching, up to simple equivalences, with a predetermined list. Unfortunately only simple problems have tractable lists. The new system determines which equations could be used to derive a submitted equation by asking whether it depends on that set of equations.

The dependence-checking facility is first used by Andes in the preparatory procedure which generates the constraint network and extracts the so-called solution paths, which are really subsets of the constraint network. This process needs to see if adding an available equation increases what is known about the solution. If the equation is dependent on those already used, it is redundant with what is already known, and so provides no new information. At each stage in generating the constraint network, there is a set of variables considered not yet known, and a set of equations. If an equation proposed for addition is independent of the equations already present, it can be considered as solving for one unknown variable. It may, however, introduce new unknown variables not yet in the set. The sufficiency of constraint subnetworks is judged by having as many **independent** equations as there are unknown variables.

Dependency checking is also used in the tutoring stage. As the tutoring system wants to be able to help the student make progress on a problem when the student gets stuck, it provides "what's next" help. To do this, the system needs to have a model of what parts of the problem the student already understands. In particular, it needs a way of determining which parts of the constraint network are known. Because the student must explicitly define all variables, the variable nodes in the network are straightforward. The system must distinguish which of the canonical equations she has already used, and which others she might need to be prompted to find. This prompting should be focused on the solution path (that is, the minimal constraint subnetwork sufficient to solve the problem) which includes as much as possible of what the student has already done. The available evidence for what parts of the constraint network are known is what variables have been defined, what axis choices have been made to describe vector quantities, and, most importantly, what equations have been entered. Very often, a correct student equation will *not* correspond to any single canonical equation. The algebra subsystem can help in analyzing correct student equations to see which of the canonical equations are necessary to derive the student's equation. When a student enters a correct equation, Andes assumes that the student knows some subset of the canonical equations from which her equation can be derived. Andes deduces that subset by examining, for each solution path, on what minimal subset of the equations in the solution path the student's equation depends. It then uses heuristics based on the simplicity of the respective answers to assign credit to particular canonical equations. The new algebra subsystem provides those subsets by a different method than that previously used, which was based on the table of "all possible" derived equations, and it can occasionally produce different results.

One might have the impression that the student, not very sophisticated and entering equations with as little contemplation as possible, would be entering the basic equations with little prior calculation. It is surprising, however, how much removed from the canonical equations even a simple equation is. For

---

[7]This method has the advantage that one can also explore whether a wrong equation can result from misapplication of physical principles, or buggy rules (Brown & Burton, 1978).

example, in the problem of Figure 1 above, one step in the solution is to write Newton's second law, ($\vec{F}_{\text{net}} = m\vec{a}$) for the hanging block. In terms of the tension T in the rope, the mass m_2 of the block, the gravitational acceleration constant g and the magnitude of the downward acceleration a, a student might quite reasonably write

$$\text{S:} \qquad m_2 g - T = m_2 a$$

This, however, is not one of the basic equations of physics. For one thing, the weight force W has had its value replaced using another known law, $W = m_2 g$. But more importantly, Newton's law applies to components of forces, not their magnitudes. In fact, the closest match to the student's equation among Andes' canonical ones (that is, among the direct application of individual physical or geometrical principles to the problem at hand, in component form, plus equations which come directly from the problem description) is

$$W_y + T_y = m_2 a_y$$

To get to the student's equation requires also the canonical equations

$$a_y = a \sin \theta_a, \qquad W_y = W \sin \theta_W, \qquad T_y = T \sin \theta_T$$
$$\theta_a = 270°, \qquad \theta_W = 270°, \qquad \theta_T = 90°$$
$$W = m_2 g$$

The first three of these come from the rule for extracting components of a vector relative to some choice of axes. The next three specify the directions of these vectors. Thus the student has effectively combined eight canonical equations in writing down one fairly simple equation.

If, after writing down the one equation S, or perhaps after also including a few equations for the block on the ramp, the student is stuck and asks for help, the help system needs to know that she has correctly employed the eight equations mentioned, and not waste her time and patience tutoring her on what she already knows. With 45 equations to consider hinting at, how does the system know that these 8 are not worth looking at? The answer must be one minimal set of equations on which equation S depends.

HOW THE ALGEBRA SUBSYSTEM ANSWERS THESE QUESTIONS

How the subsystem solves equations

The first task the algebra subsystem is called upon to perform is that of solving a set of equations during the preparation of a problem for tutoring. Typically there are many equations in many variables, but they are solvable with fairly elementary techniques either as explicit numerical values or in terms of a few undetermined *parameters*. There are, of course, many very highly developed computer algebra systems with more than enough mathematical sophistication for freshman physics problems. Our first thought for handling the problem of finding the solution to the canonical equations was to use Maple (Maple, 2000) to handle the algebra manipulation. We found, however, that Maple was unable to solve automatically

what appeared to be simple equations with inequalities; for example, it failed to give an explicit answer on

$$v_x = -v_m, \ v_x^2 = 10, \ v_m = \sqrt{v_x^2}, \text{ with } v_m > 0,$$

a set of equations which occurred in a one-dimensional kinematics problem, where v_m is the magnitude of the velocity known to be moving in the negative x direction. When such problems arose in more complicated sets of equations, Maple failed to give any solution at all. The failure of Maple, even with tech support, to handle such problems encouraged us to look for alternatives. We chose to develop our own algebra system not only because this would allow us to add whatever methods we discovered were needed, but primarily because most of the known systems do not have built in support for physical units [8].

Solving a set of equations in general is not an easy task [9], but it turns out that most problems in elementary courses can be solved with a fairly simple set of techniques. Examining the 115 problems in mechanics that were already in Andes at the time we started showed that:

- The vast majority of the equations were either assignment statements, *e.g.* $m = 4$ kg, or could be reduced to assignment statements by substituting in the values of other variables already given by assignment statements. In fact, 70% of the problems could be completely solved using only this method.

- Once the variables given by assignment statements are replaced by their numerical values (with units), there will likely be simultaneous linear equations, which can be used to further reduce the number of variables. This in fact results in complete solutions of roughly half of the problems not solved by recursive substitution of assignment statements alone.

- There is no one method that handles most of what is left. Some involve nonlinear equations in a single variable, solvable by inverse functions or numerical methods. There are pairs of equations involving $\sin \theta$ and $\cos \theta$, which can be divided, and there are pairs of quadratic polynomials in two variables, which can be used together. By trying various common methods, all the problems in Andes can now be solved automatically.

It needs to be emphasized that this method which the algebra system uses to solve the equations is **not** the recommended way for students to try to solve the problem. Students are encouraged to plug in given values only at the end, the exact opposite of what the computer is doing. The major reason for the algebra system to do otherwise is that the computer deals with numbers far better than with algebraic expressions. This is not the way we want our students to work.

There is another issue that might trouble one about relying on an algebra package that desperately tries to find numerical values for all variables. Does that approach preclude the use of problems with parameters? As defined earlier, parameters are variables which can be considered known but do not

---

[8]The (June 1, 2001) version, (Maple 7, 2002), has a new package to support units.

[9]In fact, it is an impossible task. A general fifth order polynomial cannot be solved algebraically, and while that does not preclude a numerical solution if its coefficients are known, it does preclude one if the coefficients are other unknown variables. There are methods for dealing with specific classes of equations, in particular with equations that are linear, even in a large set of variables. But while the majority of equations in Andes problems are linear, not all of them are. Nor are they all polynomials.

have an determined numerical value. The most common parameter in Andes problems is the mass of an object, for example, the masses of billiard balls. Answers such as angles and velocities turn out not to depend on this parameter, while the momenta are all proportional to it. Solving for a particular value of the mass gives values for the momenta which are not general, but any generally correct formula will be consistent with this solution.

Checking correctness of equations

Let us return to the color-by-numbers method of determining if a student's equation is correct. Naturally this method requires first finding the solution to the problem. Once the system has a numerical value for all of the variables that enter a problem, it can easily check if a student equation is correct by the simple procedure of plugging in the values and seeing if the equation balances. This method for equation checking also works well with the method of assigning secret random values for parameters. A student equation that does not have the correct dependence on the parameter can be thought of as specifying the value of the parameter, and the chance that her specified value agrees with the value randomly chosen, to an accuracy better than one part in a billion, is negligible [10]. As long as the values chosen are not ones that could be stumbled upon, a student equation that is correct only for some value of the parameter has a negligible chance of being correct for ours.

This raises what is the one difficult issue in equation checking by substitution — how close do the sides need to be to balance? Our evaluations, of course, are not precise, but use standard double precision arithmetic with an accuracy of about one part in 10^{15}. If the left hand side of the equation evaluates to 10^{-7} and the right hand side to zero, does this balance? Yes for the problem with the momentum of an aircraft carrier (in kg·m/s), but no, if this problem concerns the mass difference of a grain of salt and an electron, measured in kg. In checking the balancing of equations the system also calculates maximum possible errors, though the algorithm is not perfect in estimating them.

In order to avoid marking as correct wrong equations that just stumble close to the right answer, the tolerance allowed for agreement should be held as tight as possible. This is not a serious problem for equations that do not contain numerical calculation by the students, for the computer calculations made to verify the equation are accurate enough to permit using very tight standards for agreement. But we cannot expect the students to do their calculations to 15 figures, or even to specify an answer to such accuracy. Thus the system allows final numerical answers to have a leeway reasonable for the quantity in question. Andes instructors ask the student to avoid plugging in numerical values, except for simplifying values such as 0's, 1's and 2's, until giving the final answer. So for intermediate equations the system can require machine accuracy, while perhaps asking for three significant figures on final numerical answers.

---

[10]One might worry about equations which are correct over regions, such as writing x where one should have had $|x|$. Parameters which might affect the sign of x would then have a finite chance of equating x and $|x|$ in a problem for which this is not generally true. But such problems would have bifurcated answers, depending on the parameter, and they would not be appropriate for an introductory course.

Incorporating physical units

In Andes, when a student specifies a variable, she describes the physical quantity it represents. The tutor does not at that point ask in what units the quantity is measured. However, it does know, from the physics knowledge database, what the appropriate units are for that variable. When the student writes an equation giving a numerical value for a quantity, she must include appropriate units. The algebra system, in checking that equation, checks that the units are correct for the physical quantity. In any other equation, it also checks that the units are consistent. Each expression has a units field which gives the powers of meters, kilograms, seconds, coulombs, and degrees celsius. These are the appropriate units for the fundamental International System (SI) of units. The algebraic operations have built in the correct rules for propagating these dimensions, and imposing the appropriate consistency conditions. In fact, Andes objects to dimensional inconsistency before any check on the numerical validity of the equation.

As long as all variables are expressed in SI units, ordinary algebra, including powers of the units, will be consistent. Illegal operations, such as trying to add terms with different dimensions, are a clear sign something is wrong with an equation. This should be very helpful in giving reasons that an equation is wrong.

In order to maintain flexibility of expression, Andes permits a problem specification to ask that certain variables be described in non-standard units. Thus a speed may be input in miles/hour if desired, but internally all quantities are converted to SI units.

Modeling which equations the student knows

As discussed above, the Andes help system needs to discover which of the canonical equations the student appears to know. It does this by asking for minimal sets, within each solution path, of equations from which the student's equation might have been derived. The first version of Andes tried to extract this information from its table of all possible ways of combining the basic equations, but this method breaks down on all but very simple problems. The new algebraic system is able to judge independence of equations, however, and therefore it can provide information — not always unique answers but sets of possibilities — on which canonical equations were used by the student in creating the entered correct equation.

Determination by dependency

The method Andes2 uses observes that a student's equation could have been derived from a set of other equations if it provides no independent restriction on the solution set of those equations. Equations are restrictions of the possible collection of values of the variables. If a set of equations so restrict the solution space of the variables that the student's equation provides no further restriction, then her equation is a consequence of the others. If that is not the case, then she could not have legitimately arrived at her equation from the set, for there are values of the variables for which all the equations in the set are true, but her equation is false.

Thus if Andes can determine one unique minimal set of equations with a solution space contained in the solution space of the student's equation, it can reasonably conclude that the student knows those

equations. Unfortunately there may be more than one such minimal set, in which case there are alternate sets of equations the student may have used. These can often depend on which of several possible paths to solving the problem the student has embarked upon. The algebra system cannot decide questions like this, but it can enumerate the possibilities for the help system.

Linear dependence

The method of determining the solution space of an arbitrary set of equations is again nontrivial, or impossible, as already mentioned for the special case of finding the solution of the full problem. This problem becomes much simpler if the equations are linear. Then the equations restrict solutions to hyperplanes, and the condition for *dependence* of the student's equation is that the normal to her hyperplane is a linear combination of the normals of the hyperplanes of the equations in the set. The components of the normal are simply given by the coefficients of the variables in the linear equation. Determining if a vector in N dimensions is a linear combination of a set of P other such vectors is an easy order $(P \cdot N)$ or $(P^2 \cdot N)$ calculation[11], not prohibitive.

Generalization to nonlinear equations

So judging independence would be easy if all equations were linear. Unfortunately, even elementary physics problems involve nonlinear equations, and the method just described cannot be directly applied. It is still true, generically, that each equation restricts the space to a surface of one dimension less that the full space, but that surface may be curved. It is also still true that a possible solution point on the surface is prevented, by the equation, from moving off in the direction of the normal to the surface at that point, but as the surface is curved the normal changes direction from point to point on the surface.

The methods used for linear equations may still be used, however, if we focus our attention on small deviations from the solution point, P_0 of the full problem, which the algebra system has already provided to us. We expect in all our equations, $f_i(\{v\}) = 0$, f_i to be differentiable (probably analytic) at the solution point, so everything can be expanded by Taylor series to first order in the variables. The constant term is zero, and the first order term is specified by the gradient of f_i. As each equation becomes linear to this order of approximation, the linear method discussed above can be used. The normal to the equation solution surface is the easily calculated gradient at P_0. If the student equation is independent in the linear approximation then the full equations are also independent [12]. Generically, the reverse will be

---

[11]Order $P^2 \cdot N$ for the initial setup of the set, and then order $P \cdot N$ for subsequent queries on that set. The algorithm used is to reduce the vectors to row echelon form while entering them into the set. This makes the checking of equations against that set more efficient. The help system is expected to make more queries on fixed sets than changes in the sets.

[12]Suppose there is a point P_1 for which her linearized equation has a discrepancy Δ, but the linearized canonical equations are all exactly correct. Every point $P = \lambda P_1 + (1 - \lambda)P_0$ on the line segment between P_0 and P_1 will also satisfy the linearized canonical equations and have a discrepancy $\lambda\Delta$ in the student's linearized equation. For points sufficiently close to P_0, the exact equations should differ from the linearized ones by amounts that go to zero *faster* than the first power of λ, but the linearized dependence is violated to order λ, so the full equations cannot agree. This contradicts the idea that full student equation would have no discrepancies on the solution space of the canonical equations. Thus the student equation must be independent.

true as well — if the linearized equations are dependent the full ones will usually be as well, but in this direction there are exceptions, as will be discussed below.

An explicit example of dependency determination

Let us consider a simpler problem to illustrate how the dependence calculations can help determine what the student has used. Consider this problem:

> A car starts from rest and accelerates at a constant rate to 20 m/s in a distance of 50 m. What is the acceleration of the car?

Some basic equations that deal with the kinematics of linear motion at constant acceleration are

1: $v_f^2 - v_i^2 = 2as$
2: $v_f - v_i = at$
3: $s = \frac{1}{2}at^2 + v_i t$
4: $s = \frac{1}{2}(v_i + v_f)t$

while the givens here are

5: $v_i = 0$
6: $v_f = 20$ m/s
7: $s = 50$ m

The solution point, which solves all these equations, is

$$P_0 : (t, s, a, v_i, v_f) = (5\text{ s}, 50\text{ m}, 4\text{ m/s}^2, 0, 20\text{ m/s}).$$

The first four equations are not independent, in fact no three of them are independent. Any two of them imply the other two. So there are many different complete sets of independent equations for this problem, depending on which two of the first four equations are included:

$$A = \{1, 2, 5, 6, 7\}, \qquad B = \{1, 3, 5, 6, 7\}, \qquad C = \{1, 4, 5, 6, 7\}.$$

$$D = \{2, 3, 5, 6, 7\}, \qquad E = \{2, 4, 5, 6, 7\}, \qquad F = \{3, 4, 5, 6, 7\}.$$

We will also ask about the subsets that don't include the givens,

$$\bar{A} = \{1, 2\}, \ \bar{B} = \{1, 3\}, \ \bar{C} = \{1, 4\}, \ \bar{D} = \{2, 3\}, \ \bar{E} = \{2, 4\}, \ \bar{F} = \{3, 4\}.$$

Suppose the student writes down the equation S: $s = \frac{1}{2}v_f t$. Plugging in the solution values gives 50 m = $\frac{1}{2} \cdot$ 20 m/s $\cdot$ 5 s, which is correct, so the equation is correct. From which sets could it have been derived, and which most easily?

Rewriting the equations in the form $f =$ left side $-$ right side $= 0$, and taking the gradient, gives:

function f_i	gradient $\partial f_i/\partial x$, for $x = (t, s, a, v_i, v_f)$				
	t	s	a	v_i	v_f
1: $v_f^2 - v_i^2 - 2as$	0	$-2a$	$-2s$	$-2v_i$	$2v_f$
2: $v_f - v_i - at$	$-a$	0	$-t$	-1	1
3: $s - \frac{1}{2}at^2 - v_i t$	$-at - v_i$	1	$-\frac{1}{2}t^2$	$-t$	0
4: $s - \frac{1}{2}(v_i + v_f)t$	$-\dfrac{v_i + v_f}{2}$	1	0	$-\frac{1}{2}t$	$-\frac{1}{2}t$
5: v_i	0	0	0	1	0
6: $v_f - 20$ m/s	0	0	0	0	1
7: $s - 50$ m	0	1	0	0	0
and student's equation:					
S: $s - \frac{1}{2}v_f t$	$-\frac{1}{2}v_f$	1	0	0	$-\frac{1}{2}t$

Evaluating at the solution point means plugging in the values of the variables at P_0, so, dropping units here, we have:

f_i	t	s	a	v_i	v_f
1: $v_f^2 - v_i^2 - 2as$	0	-8	-100	0	40
2: $v_f - v_i - at$	-4	0	-5	-1	1
3: $s - \frac{1}{2}at^2 - v_i t$	-20	1	-12.5	-5	0
4: $s - \frac{1}{2}(v_i + v_f)t$	-10	1	0	-2.5	-2.5
5: v_i	0	0	0	1	0
6: $v_f - 20$ m/s	0	0	0	0	1
7: $s - 50$ m	0	1	0	0	0
and student's equation:					
S:	-10	1	0	0	-2.5

First, observe the dependence of the first three equations is manifest by noting that the first line is 40 times the second minus eight times the third. Similarly the fourth line times eight, added to the first line, gives 20 times the second. This is the statement that only two of the four equations are independent. Next, observe that no linear combination of these four lines will give the student's equation; her equation is independent of the sets $\bar{A}...\bar{F}$, so she must have used one of the givens.

For each of the complete sets of independent equations, which equations are necessary to derive the student's can be determined by finding linear combinations of the gradients as above. The answers for each set are

$$A : \{1, 2, 5\} \quad B : \{1, 3, 5\} \quad C : \{4, 5\} \quad D : \{2, 3, 5\} \quad E : \{4, 5\} \quad F : \{4, 5\}$$

Thus she has definitely used Eq. 5, and it is considerably more likely that she used equation 4 than that she used two of the first three equations. She definitely knows at least one of the fundamental kinematic equations, probably Eq. 4, and she has taken note of the fact that the car started from rest, the given $v_i = 0$.

A flaw in the method

Thus we have seen how being able to judge the independence of equations can be used to help determine what the student knows, and we have also seen how to tell whether linearized equations are independent. Unfortunately there is a small hole in this argument — if the linearized equations are independent, the full equations are also necessarily independent. It is also true that *generically*, independent equations will have independent linearizations, but not always. For example, consider two equations in two unknowns, whose intersection determines a solution. In the generic case, the solution curves of the two equations will intersect, and the linearized forms of the equations, shown by the tangent lines, will be independent. But in the exceptional case that the two curves are tangent to each other at the intersection, their linearized form is shown by the single tangent line (Figure 2), so the linearized forms are not independent and do not determine the point P, even though the full equations do.
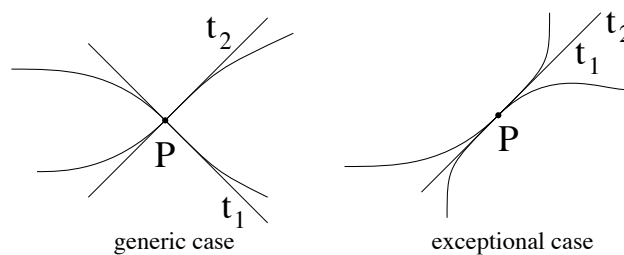

generic case   exceptional case

Fig.2. Two independent equations determining a solution point P.

As this difficulty only arises in exceptional cases, one might hope that it will not occur in the problems encountered in the introductory course. But in fact it occurs routinely in vector problems, because the solution often involves an angle of 0 or $90°$, which are critical points of the cosine and sine functions. In fact, if we look back at the equations above for the hanging block, and ask for the minimal subset of the eight equations which appear to be required to derive the student's equation, the linearized method would not include the three equations giving the angles. These equations are in fact needed to get $a_y = -a$, $W_y = -W$, and $T_y = T$ from the three equations $a_y = a \sin \theta_a$, $W_y = W \sin \theta_W$, $T_y = T \sin \theta_T$. But they do not appear to be needed in the linear approximation. Expanding $a_y = a \sin \theta_a$ in a Taylor series in θ_a about $\theta_a = 3\pi/2$ gives

$$a_y = a\left(-1 + \frac{1}{2}(\theta_a - 3\pi/2)^2 + ...\right) \sim -a + 0 \cdot (\theta_a - 3\pi/2) = -a,$$

where the $\sim$ represents the linear approximation. Thus the linear approximation might mislead the system into thinking $a_y = -a$ does not require knowledge of θ_a. The problem arises because the solution point happens to be at maximum of the expression $a_y - a \sin \theta_a$. The expression happens to have a zero value and a zero derivative at the same point.

A partial fix

How can the system deal with the fact that this situation, which is in some sense exceptional and should have little probability of ever arising by chance, actually arises often in the problems assigned to students? Examining dependence without approximations is a very complex issue, and even going to second order in the expansion[13] would make the calculations much larger. The variables in question are generally givens, and the help system may be able to deal with uncertainty in whether the student has recognized these. So the approach taken is this: When calculating the gradient of each equation's function, the system also notes which variables the full equation depends on. If a proposed dependence involves only functions with zero derivatives with respect to a given variable, but nonetheless one or more depend on that variable, the help system is warned that the equation *might* depend on some equation that gives the value of that variable, in addition to the ones it depends on in linear approximation. It can also be sure of dependence if the number of variables involved is not greater than the number of independent equations in the canonical set.

EVALUATION: EFFECT OF CHANGING METHODS

As was mentioned earlier, any student equation which is colored green by color-by-numbers has a derivation starting from the canonical equations and proceeding by algebraically correct steps. The derivation, however, might not pass muster of any instructor examining the result, because it might involve steps that have no motivation in solving the problem. A tighter definition of derivability, well-motivated correctness, would require each step to be a credible step forward in deriving an answer. The distinction is best understood with an example.

In linear kinematics, there is an equation holding if the acceleration of an object is constant:

$$\text{A:} \qquad v_f^2 - v_i^2 = 2 \cdot a \cdot s$$

where v_f and v_i are the final and initial velocities, a the acceleration, and s the distance travelled. Very often a problem will state that the object starts from rest, *i.e.*

$$\text{B:} \qquad v_i = 0$$

If the student enters the equation

$$\text{S:} \qquad v_f^2 + v_i^2 = 2 \cdot a \cdot s$$

any instructor would conclude that the student had misremembered a sign in the equation and mark the equation wrong. But equation S can be derived from A and B by squaring B and doubling the result, giving $2v_i^2 = 0$, and adding that equation to A. Thus S is derivable by legitimate algebraic steps, but the

---

[13]While it is probably true that an introductory physics problem would never run into the situation where the expression, its first derivatives, and its second derivatives all vanish at the same point, but the function is still not identically zero, there are in principle still these exceptional situations. In fact, the function $\{ f(x) = e^{-1/x^2}$ for $x \neq 0, f(0) = 0 \}$ has a minimum at $x = 0$ where it is zero and so are *all* its derivatives, and yet it depends on x.

derivation is misguided because there is no reason to take these steps if your goal is to solve for one of the unknowns — the only possible motivation is to justify your mistake. So the old Andes would have marked S wrong, which is good, while the new one will mark it correct.

How often color-by-number approves equations that should be rejected in actual use has not been systematically studied. While it is clearly a weakness in principle, the consequences need to be compared to the weakness of methods for generating "well-motivated" derivations. The possible sets of rules for such derivations need to be sufficiently strong to include all reasonable correct student equations, while being sufficiently limited to provide tractable (and certainly finite) sets. Whether this is possible is not clear, but certainly the rules used by Andes1 failed to produce tractable sets for a fairly large fraction of the problems physics professors wanted to assign. The old Andes was simply unable to generate the lists of derived equations for 27 of the 115 problems used by Andes in the Autumn of 2000. Color-by-numbers has avoided that problem. In Autumn 2001, Andes2 was used by 119 students, who completed 5766 problems (102 distinct ones). Only one incident was observed of an equation which was unmotivated, and thus pedagogically wrong, being marked correct because the incorrectly included terms evaluated to zero.

There have been evaluations of the instructional effectiveness of using Andes in place of written homework at the USNA in 1999, 2000, and 2001, during which time Andes changed dramatically. In 1999, Andes1 used the old algebra system with generated lists of algebraic equations. In 2000, in order to incorporate problems for which these lists could not be made, Andes used color-by-number evaluation, but did not yet use the independence checking features of the algebra system described here. 2001 was the first year the new help system was used, along with a full use of the capabilities described here. In all years the Andes group did significantly better than the control group, who did their homework problems on paper. The results for 1999 (Schulze et al., 2000; Gertner & VanLehn, 2000; Shelby et al., 2001) showed a 0.21σ effect on an exam which covered only the first six weeks of mechanics. In 2000 (Shelby et al., 2001), with more problems in Andes, there was a 0.92σ effect on a similar early test, but no significant effect on the final exam. An evaluation of Andes2 in Autumn 2001 (VanLehn et al., 2002) gave a 0.52σ effect on an exam covering 12 weeks of material. Because so many things changed from year to year, the differences in these results probably have as much to do with the maturity of the rules of the "help" production system (Andes1 was quite mature before being discarded) as with the change in methods. Despite the in-principle possibility of misattributing dependencies, instructors in 2001 reported that Andes2 did not suggest outlandish equations when asked "what's wrong" with a student equation. This had been a very significant problem the previous year, for Andes1, which used the previous algebraic methods. It is, however, not possible to separate the effects of changing the algebraic methods from the changes due to the complete rewrite of the help system, in determining what caused this improvement.

Since 2001, there has been a large expansion in the domain of Physics covered by Andes. There are now 311 problems for which the system, including the algebra subsystem, is being used. A summary of pedagogical effectiveness can be found in (VanLehn et al., 2005b,a). The professors involved with the course become aware of recurring problems, and while some programming bugs and poor interactions between the algebra system and the help system have been found and fixed, there has been no indication of a problem with the basic methods described here.

The method described in "a partial fix" above to handle the exceptional cases of dependency appears to correctly give the dependencies in the problems the Andes group has examined so far, although the treatment of "might depend" answers by the help system could use further work. Currently it ignores "might depends", thereby possibly undercrediting some equations. But this does not appear to have caused any problems in the use of this method at the USNA.

As was mentioned earlier, the ability of the algebra system to solve sets of equations permits Andes to provide the students with "the genie", which will give them the answer if they have supplied adequate equations to determine it. While one might question whether allowing this might diminish the student's ability to complete problem solutions when away from Andes, preliminary analysis by Robert Shelby shows (Shelby, 2004) "the least we can say is that using the solver most certainly did not seem to result in a disadvantage for the Andes students" on the final exam.

SUMMARY

A new physics tutorial system has emerged from the Andes effort. It makes very substantial use of a new powerful algebra subsystem. This subsystem has introduced new capabilities for dealing with dimensional analysis, for solving systems of equations, and for providing algebraic help to the students. In addition, it has made two significant changes compared to other systems. Prior versions of Andes, as well as many other tutoring systems involved with algebraic equations, judge the correctness of a submitted equation and which canonical equations it depends upon by seeking a derivation of the submitted equation. Because this is a slow process, previous versions of Andes tried to prederive all correct equations for each problem. The new system is based on the observation that determining that an algebraically correct derivation exists can be done by simple evaluation of the equation, without actually finding a derivation. A new method is then required for determining which canonical equations are needed for a derivation. The new system does this by examining the linear dependencies of the expansions of all equations about the solution point. This is not an infallible method, possibly failing if the solution point is a critical point of the equations. However, the most common occurrences of this problem can be handled by heuristic methods which have been incorporated into the system. The main advantage of the new method is that it is extremely efficient, compared to a system which severely limited the kinds of problems that could be handled. These new methods allow much more complex problems. In the Autumn 2001 use at the USNA, Andes2 was able to present many problems which were beyond the abilities of the previous method. The range of problems has increased greatly since then, successfully handled by the new algebra system. It also provided a very popular tool for helping the students do the required algebraic solutions.

ACKNOWLEDGEMENTS

leave to participate in this project.

REFERENCES

Andes 1 (2002). http://www.pitt.edu/~vanlehn/andes.html.

Andes 2 (2005). http://www.andes.pitt.edu/.

Brna, P., & Caiger, A. (1992). The application of cognitive diagnosis to the quantitative analysis of simple electrical circuits. In C. Frasson, G. Gauthier & G.I. McCalla (Eds.) *Intelligent Tutoring Systems; Proceedings of the Second International Conference,* (pp.405–412). Berlin: Springer-Verlag.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science,* 2(2), 155–192.

Carnegie Learning (2005). Carnegie learning. http://www.carnegielearning.com/.

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). The PUMP/PAT tutor. *Handbook of Human-Computer Interaction*, chapter 37: Intelligent tutoring systems. 2nd edition. Amsterdam, The Netherlands: Elsevier Science.

CyberProf (2001). http://www.howhy.com/home/.

Gage, M. E., & Pizer, A. K. (1999). WeBWorK – math homework on the web. In *Electronic Proceedings of the Annual International Conference on Technology in Collegiate Mathematics.* http://archives.math.utk.edu/ICTCM/EP-12/P3/html/paper.html.

Gertner, A., & VanLehn, K. (2000). Andes: A coached problem solving environment for physics. In G. Gauthier, C. Frasson & K. VanLehn (Eds.) *Proceedings of the 5th International Conference, ITS 2000,* (pp. 131–142). Berlin: Springer.

Gertner, A. S. (1998). Providing feedback to equation entries in an intelligent tutoring system in physics. In B.R. Goettl, H.M. Halff, C.L. Redfield, & V.J. Shute (Eds.) *Proceedings of the 4th International Conference on Intelligent Tutoring Systems, ITS '98,* (pp. 254–263). LNCS 1452. Berlin: Springer.

Gertner, A. S., Conati, C., & VanLehn, K. (1998). Procedural help in Andes: Generating hints using a bayesian network student model. In *Proceedings of the 15th National Conference on Artificial Intelligence* Madison, WI, USA.

Halliday, D., & Resnick, R. (1981). *Fundamentals of Physics,* (pp. 34). 2nd edition. New York: John Wiley.

Hübler, A. W., & Assad, A. M. (1995). CyberProf: An intelligent human-computer interface for asynchronous wide area training and teaching. In *Proceedings of Fourth International World Wide Web Conference.*

Liew, C. W., & Smith, D. E. (2002a). Checking for dimensional correctness in physics equations. In *Proceedings of Fifteenth International Florida AI Research Society Conference,* (pp. 299–303). AAAI Press.

Liew, C. W., & Smith, D. E. (2002b). Reasoning about systems of physics equations. In S.A. Cerri, G. Gouardères,

and F. Paraguaçu, (Eds.) *Proceedings of the Sixth International Conference on Intelligent Tutoring Systems, ITS 2002*. LNCS 2363. Berlin: Springer-Verlag.

Maple (2000). Maple©6, Waterloo Maple Inc. http://www.waterloomaple.com/.

Maple 7 (2002). Maple©7, Waterloo Maple Inc. http://www.mapleapps.com/categories/maple7/html/Units.html.

Mastering Physics (2004). http://www.masteringphysics.com/.

Ohlsson, S. (1993). Constraint-based student modeling. *Journal of Artificial Intelligence in Education,* 3(4), 429–447.

Ohlsson, S., & Langley, P. (1986). Psychological Evaluation of Path Hypotheses in Cognitive Diagnosis. In *Learning Issues for Intelligent Tutoring Systems*, chapter 3, pages 42–62. New York: Springer.

Pollack, A. (1999). Missing what didn't add up, NASA subtracted an orbiter. *New York Times*, page 1.

Pritchard, D. E., & Morote, E.-S. (2002). Reliable assessment with CyberTutor, a web-based homework tutor. In *World Conference on E-learning in Corporate, Government, HealthCare, and Higher Education*, Montréal, Canada. E-Learn.

Raineri, D. M., Mehrtens, B. G., & Hübler, A. W. (1997). Cyberprof^{tm} — an intelligent human-computer interface for interactive instruction on the world wide web. *Journ. Asynch. Learning Networks,* 1, 20.

Reif, F., & Scott, L. A. (1999). Teaching scientific thinking skills: Students and computers coaching each other. *Am. J. Phys,* 67, 819–831.

Schulze, K., Shelby, R., Treacy, D., Wintersgill, M., VanLehn, K., & Gertner, A. (2000). Andes: An intelligent tutor for classical physics. *The Journal of Electronic Publishing,* 6(1).

Serway, R. A. and Beichner, R. J. (2000). *Physics For Scientists and Engineers*. Saunders College Publishing, Fort Worth, 5th edition. pages 10–11.

Shelby, R. N. (2004). private communication.

Shelby, R. N., Schulze, K. G., Treacy, D. J., Wintersgill, M. C., VanLehn, K., & Weinstein, A. (2001). An assessment of the Andes tutor. In S. Franklyn, J. Marx, and K. Cummings (Eds.) *Proceeding of the 5th National Physics Education Research Conference,* (pp. 119–122). Rochester, NY.

Sherwood, B., & Stifle, J. (1975). The PLATO IV communications system. Technical report, University of Illinois Computer-based Education Research Laboratory, Urbana, IL.

Tipler, P. A. (1991). *Physics For Scientists and Engineers,* (pp. 5–6). 3rd edition. New York: Worth Publishers.

UTexas (2005). https://hw.utexas.edu/bur/functionality.html.

VanLehn, K., Lynch, C., Taylor, L., Weinstein, A., Shelby, R., Schulze, K., Treacy, D., & Wintersgill, M. (2002). Minimally invasive tutoring of complex physics problem solving. In S.A. Cerri, G. Gouardères, and F. Paraguaçu, (Eds.) *Proceedings of the Sixth International Conference on Intelligent Tutoring Systems ITS 2002,* (pp. 367–376). LCNS 2363. Berlin: Springer.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005a). The Andes physics tutoring system: Five years of evaluations. In *12th International Conference on Artificial Intelligence in Education (AIED-2005)*.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005b). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education,* 15(3).

WebAssign (2005). Webassign. http://www.webassign.net/.

Woolley, D. R. (1994). PLATO: The emergence of online community. http://www.thinkofit.com/plato/-dwplato.htm.

Yibin, M., & Jianxiang, L. (1992). Intelligent tutoring system for symbolic calculation. In C. Frasson, G. Gauthier & G.I. McCalla (Eds.) *Intelligent Tutoring Systems; Proceedings of the Second International Conference, ITS '92,* (pp. 132–147). Berlin: Springer-Verlag.