Algebra Subsystem for an Intelligent Tutoring System

Joel A. Shapiro
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA, 15260
and
Dept. of Physics and Astronomy*
Rutgers University
Piscataway, NJ 08854-8019
shapiro@physics.rutgers.edu

Sept. 28, 2001

Abstract

I describe a computer algebra subsystem specifically designed for use by a tutorial system for introductory physics quantitative problem solving. It is capable of solving the systems of equations involved in such problems, checking the validity of equations the students enter, investigating whether an equation is independent from a set of other equations, and if not which equations it depends on, and finally 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 used during problem generation in providing information for the solution-path generating module, and later during tutoring so the help module can model which equations the student appears to know. One important feature of our algebra system is that it deals with the dimensional units of physical quantities throughout.

An important change from a previous system is in the meaning of "correctness" of an equation and in the meaning of which equations it can be derived from. An evaluation of the theoretical differences is given here, while an evaluation of the effect on student feedback will be given in a subsequent publication.

Introduction

This paper describes an algebra system designed to be integrated into an intelligent tutoring system (ITS) for helping students do quantitative problems in introductory college/university physics courses. The particular tutoring system involved is Andes2, a revision of the Andes tutor[1, 2, 3] developed at the University of Pittsburgh and the U. S. Naval Academy. The issues addressed here are likely to be of use in any tutorial system designed to deal in some generality with science or engineering problems that involve algebraic equations among physical quantities.

Andes is designed to accept a formalized description of a problem, from which, with the help of a knowledge base of physical principles, it constructs a set of solution paths, as well as a set of equations. It has a user interface that allows the student to define variables, draw vectors, and write equations. It also allows the student to request help, and gives feedback and hints in response to student actions. To do this, the help system needs to be able to answer the questions

*Permanent address.

- Is the equation the student wrote down correct?
- What can we conclude the student knows of the solution path from what she has written down?
- Does adding a given equation advance the solution of a problem beyond what is already specified by the previous set of equations?

Andes is designed to let the student proceed with the solution of a problem, defining variables and axes, and writing down relevant equations, without interference for as long as the student is on an acceptible path and is writing down correct things. 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". In defining variables, the system requires that the variable correspond to a physical quantity that could be relevant to solving the problem, as enumerated by Andes' solution-generating module. 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 package will be in terms of recognizable variables. A crucial task for the tutoring system is to be able to distinguish correct equations from incorrect ones.

Correctness of equations

In earlier versions of Andes the correctness of student equations was judged by whether the equation was equivalent to one on a list of all possible equations that could be generated from the basic, or "canonical", equations produced by the knowledge base from the problem specification[5]. In generating the derived equations, the generator kept track of the canonical equations used. If the student's equation could be found as a simple algebraic manipulation of one of the equations on this list, it was considered correct, and which equations it depends on determined by the entry in the list. Generating such a list proved unwieldy on all but very small problems.

The new algebra system takes a different approach. We define an equation to be correct if it is true given the problem specification. The tutor may, under some conditions, object to a correct equation as being premature or inappropriate, but it must always object if the equation is incorrect. As the problem specification implies the solution, the correctness of the student's equation is judged by simply plugging in the numerical solution and evaluating the student's equation. As correctness in indicated by turning the equation green and incorrect equations are turned red, we call this approach "color-by-numbers".

How this new approach differs in its answers from the criterion of derivability will be descussed in later sections.

Physical dimensions (units)

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. I recently overheard two otherwise intelligent sounding adults trying to recall the formula for the area of a circle and coming up with $2\pi r$. It had probably been nearly a decade since their geometry class, so forgetting the formula might be understandable, but they should know that area is measured in square inches or square meters, while the radius is in inches or meters, not squared. And of course 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. So it is very poor pedagogy to have a tutorial system that ignores units.

In Andes, when a student specifies a variable, she describes the physical quantity it represents, but the tutor does not at that point ask what units the quantity is measured in. When the student writes an equation giving a numerical value for a quantity, however, she must include appropriate units. The algebra syatem, 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. In fact, it objects to dimensional inconsistency before any check on the numerical validity of the equation.

Dependence of equations

As the Andes 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 student model telling which of several possible solution paths the student has been pursuing, and how far along that path she is. The available evidence for this is what variables have been defined, what axis choices have been made, and what equations have been entered. The algebra system can help in analyzing correct student equations to see which of the canonical equations are necessary to derive the student's equation. It does that by examining whether a specific equation is algebraically dependent or independent of a set of canonical equations. The new algebra system answers this 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. These differences between this lookup-table method and our new method, described below, will be investigated in the forthcoming article[6].

The dependence-checking facility is also used at an early stage in problem development, when a problem is being readied for the tutor by the solution-path generating module. This module needs to see if adding an available equation increases what is known about the solution, or whether, because the equation is dependent on those already used, it is redundant with what is already known, and provides no new information.

The method we use for checking independence requires that the system have a solution of the equations involved. As solving an system of equations is a difficult problem, and as the solution to the full problem is a solution to any subset of the canonical equations, a solution to the full problem will suffice. That solution is also required by our method of checking student equations. Thus the first task of the algebra system for any problem is to solve the system of canonical equations.

Solve-tool

The equation-solving ability of the algebra system may also be employed to offer the student help in algebraic manipulations. Andes provides several "solve-tool"s of varying power available to the student.

Outline

In the rest of this paper, I will discuss these aspects of the algebra system:

- How it tries to solve the set of equations
- How it checks correctness
- How it handles physical units
- How it judges the independence of an equation from a set of other equations.
- How it provides algebraic manipulation help to the student.

- An evaluation of the theoretical differences in behavior of the new methods of equation checking and dependency determination.

Solving equations

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[4] 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 found we needed, but primarily because most of the known systems do not have built in support for physical units[1].

Solving a set of equations in general is not an easy task[2], as witness the fact that even very sophisticated systems can fail on very easy problems. As I was not prepared to launch a Maple sized effort, I needed to see if we could restrict our methods and still handle the full scope of problems we expect to ask freshmen to solve.

Examining the problems that were already in Andes at the time we started, 115 problems in mechanics, I found 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 the algebra system uses to solve the equations is **not** the way we want students to try to solve the problem. Students are

[1]The just-released (June 1, 2001) version, Maple7, has a new package to support units.

[2]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 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 our equations are linear, not all of them are. Nor are they all polynomials.

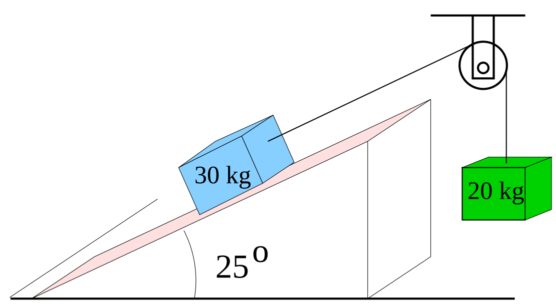
encouraged to plug in given values only at the end, the exact opposite of what the computer is doing. The major reason for the algebra system to do otherwise is that the computer deals with numbers far better than with algebraic expressions. This is not the way humans ought to work, although sadly many students cling to their preference for working with numbers, rather than with algebraic expressions.

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? 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 require knowledge of the mass.

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, we do not have to do so. We sidestep this problem by assigning each independent parameter an arbitrary “ugly” value. The student never sees this value, but as her formulae are supposed to work for all values of the parameter, they must work for our arbitrary numerical ones as well. As we will see in the next section, this method does not limit us from anything we would like to do, although it does preclude us from giving the student the answer, if it is affected by the parameter.

Checking student equations

As I mentioned, earlier versions of Andes tried to enumerate all correct equations and judge each student equation entry by whether it was equivalent to one on that list. This requires combining the set of canonical equations in all combinations[5]. 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 problem shown in Figure 1, Andes2 generates 45 equations in 41 variables. The number of possible ways of combining these into a correct equation is immense. Thus in recent versions of Andes the correctness of student equations has been judged by checking them against the correct solution of the problem — that is, the correct values of the 41 unknowns are substituted into the student equation and correctness is determined by whether the two sides balance.



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

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 we mean by derivable. If, on the one hand, we mean that there exists an algebraically correct procedure for deriving from the input the student's equation, then we can give a formal proof that derivability is equivalent, in our context, to evaluating as green in color-by-numbers. For in all our problems we can solve the problem by algebraically correct steps, so we can write the solution for all variables, $v_i = f_i(\{\lambda_j\})$, 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 explicitly determined functions. If the student has written an equation equivalent to $S(\{v_i\}) = 0$, and 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 $S(\{v_i\}) = 0$. Thus algebraically correct derivability is equivalent to color-by-number.

On the other hand, we might mean something else by derivability. We might mean that the student's equation could arise in a derivation starting from the input and proceding 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. But this definition requires that we specify some finite set of methods by which such manipulations should proceed. For example, we would permit solving one equation for one variable in terms of the other variables and substituting the results into other equations. This, however, can easily lead to a divergent procedure, so any attempt to generate all satisfactory equations will need to use a more restrictive method. I will discuss the differences and limitations of these two methods in the Evaluation section.

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 our system has a numerical value for all of the variables that enter a problem, we can easily check if a student equation is correct by the simple procedure of plugging in the values and seeing if the equation balances.

The method of solution substitution for equation checking also works well with our method of assigning secret messy values for parameters. 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 our checking of equations we also calculate maximum possible errors, though our algorithm is not perfect in estimating them.

In order to avoid marking as correct wrong equations that just stumble close to the right answer, we want to make sure the tolerance we allow for agreement is 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. We will allow final numerical answers to have a leeway reasonable for the quantity in question. We want the student to avoid plugging in numerical values, except for 0's, 1's and 2's, until giving the final answer, so for intermediate equations we can require machine accuracy, while asking for, perhaps, three significant figures on final numerical answers.

Incorporating physical units

One of the major lessons of problem solution strategy in physics is to check equations for dimensional consistency. Thus incorporating physical units as an integral part of the tutoring system in general, and into the algebra subsystem in particular, is essential for good pedagogy in physics. A system that is able to point to dimensionally inconsistent operations can provide important feedback on what is wrong with an incorrect student equation.

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 the major tools for calculation do not integrate units in any essential way. But we want a system that will 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 $\mathrm{kg{\cdot}m^2/s^2}$ while the right hand side has units kg·m/s.

Internally, our algebra system assigns to each term in an expression units as measured in terms of the fundamental International System (SI) units, meters, kilograms, seconds, coulombs, and degrees Kelvin. As long as all variables are expressed in such units, ordinary algebra, including powers of the units, will be consistent, and 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 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. I doubt if 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 Andes has been inconsistent in its requirement that all quantities are measured in standard units, and would have run into troubles 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.

Modeling which equations the student knows

Judging the correctness of the student equations, as we have seen, is straightforward once the problem solution is known. More sophisticated information is needed when the student asks for "next step help". At this point the help system has to judge which pieces of the problem the student has already correctly used. In particular it must judge which of several possible solution paths the student is pursuing. It must also distinguish which of the canonical equations she has already used, and which others she might need to be prompted to find. Here too, 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. Our 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.

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 basic equations even a simple equation is. For example, in the problem described above, one step in the solution is to write Newton's second law, $(F = ma)$ for the hanging block. In terms of the tension T in the rope, the mass m_{20} of the block, the gravitational acceleration constant g and the magnitude of the downward acceleration a, a student might quite reasonably write

$$m_{20}g - T = m_{20}a.$$

This, however, is not one of the basic equations. For one thing, the weight force W has had its value replaced using another known law, $W = m_{20}g$. But more importantly, Newton's law applies to components of forces, not their magnitudes. In fact, the closest we can come to the student's equation in the canonical ones is

$$W_y + T_y = m_{20}a_y$$

To get to the student's equation, we also need the canonical equations

$$\begin{array}{rcccl} a_y = a\sin\theta_a, & W_y & = & W\sin\theta_W, & T_y = T\sin\theta_T \\ \theta_a = 270^o, & \theta_W & = & 270^o, & \theta_T = 90^o \\ & W & = & m_{20}g & \end{array}$$

Thus the student has actually combined eight canonical equations in her head in writing down a fairly simple equation.

If, after writing down this one equation, or perhaps after 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 method we use involves finding minimal sets of canonical equations from which the student's equation could be derived. One way to judge this would be to try many algebraically correct manipulations to see if we could arrive at the student's equation, but that is a very open-ended task. We can conclude 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. 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 we can determine one unique minimal set of equations with a solution space contained in the solution space of the student's equation, we 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.

The method of determining the solution space of an arbitrary set of equations is again nontrivial, or impossible, as we mentioned for the special case of finding the solution of the full problem. This problem becomes much simpler if the equations are linear. Let us take as an example a situation with three variables, x, y, and z. A linear equation has the form

$$\text{Eq. 1:} \qquad a_1x + b_1y + c_1z + d_1 = 0,$$

where a_1, b_1, c_1, and d_1 are known constants. A linear equation reduces the space of variables by one dimension, so here Eq. 1 restricts the three dimensional space of possible (x, y, z) values to lie on a plane. If we have another equation on these variables

$$\text{Eq. 2:} \qquad a_2x + b_2y + c_2z + d_2 = 0,$$

it also restricts the values to *its* plane. To satisfy both equations, the point (x, y, z) must lie on both planes, that is, on their intersection. Usually this will be a line of intersection, when the equations are independent. But it is possible the two planes are the same, in which case the second equation adds no information, or they might be parallel, in which case there is no intersection and therefore no solution. That will not happen in our case, because we will not present the students problems that have no solution. Note that the equations are independent if the planes are not parallel, which might also be stated that their normals are not parallel.

Suppose we have two equations, which we already know to be independent, and we want to know if the student's equation

$$\text{Eq. S:} \qquad a_Sx + b_Sy + c_Sz + d_S = 0,$$

is independent of our two equations. Her equation also determines a plane on which the solution must lie (if it is a correct equation). If that plane intersects the line of intersections from Eq. 1 and Eq. 2, there are points on that line that do not satisfy the student's equation, and therefore her equation is independent and not derivable from the others. On the other hand, if that line of intersection lies within the plane solving the student's equation, her equation is not independent and can be derived.

This geometry may make the situation clear but the computation not. Fortunately, the condition for *dependence* of the student's equation is that the normal to her plane is a linear combination of the normals of the equations in the set. This is true even if we have many more equations and variables involved and would have even more trouble picturing the geometry. Fortunately, determining if a vector in N dimensions is a linear combination of a set of P other such vectors is an easy order (PN) or (P^2N) calculation[3], not prohibitive.

Life 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.

We may still use the method of the linear equations, 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 we may expand everything by Taylor expansion to first order in the variables. The constant term is zero, and the

[3]Order P^2N for the initial setup of the set, and then order PN 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. We expect the help system to make more queries on fixed sets than changes in the sets.

first order term is specified by the gradient of f_i. As each equation becomes linear to this order of approximation, we can use the method discussed above. 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 there is a point P for which her linearized equation has a discrepancy Δ, but the linearized canonical equations are all exactly correct. Every point on the line segment between P_0 and P will also satisfy the linearized canonical equations and have a discrepancy $\lambda\Delta$ in the student's linearized equation, where λ is the fraction of the distance $|P_0 P|$ that the point is away from P_0. 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, which 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. Generically, the reverse will be 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 we discuss below.

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:

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

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

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 $s = \frac{1}{2}v_f t$. Plugging in the solution values gives 50 m $=$ $\frac{1}{2}$ 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, we have

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: $s - \frac{1}{2}v_f t$	$-v_f t$	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:	-100	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, we 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.

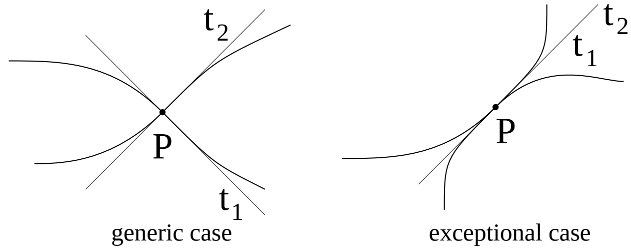
We can ask, for each of our complete sets of independent equations, which equations are necessary to derive the student's, by finding linear combinations of the gradients as above. The answers for each set are

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

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

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 we can tell whether linearized equations are independent. Unfortunately there is a small hole in this argument — if the linearized equations are independent, so are the equations themselves, always, as we saw above. 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, so the linearized forms are not independent and do not determine the point P, even though the full equations do.



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 we see 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 $\theta_a = 3\pi/2$, we have

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

where the $\sim$ represents the linear approximation. Thus the linear approximation might mislead us to thinking $a_y = -a$ does not require knowledge of θ_a. The problem is arising 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.

How do we deal with the sad fact that this situation, which is in some sense exceptional and should have little probability of ever arising by chance, actually arises all the time in the problems we assign students? Examining dependence without approximations is a very complex issue, and even going to second order in the expansion[4] 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 we have taken is this: When we calculate the gradient of each equation's function, we also note 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. If only one of the equations in the linearly dependent set involves the variable, then we can definitely say that for the full equation, this dependency is incorrect, and we need to include the equation giving the variable's value. We can also be sure of dependence it the number of variables involved is not greater than the number of independent equations in the canonical set.

The method just described to handle the exceptional cases appears to correctly give the dependencies in the problems we have examined so far, but it is not mathematically rigorous. In our forthcoming article[6] we will examine the results of this method versus the table of derived equation method of the old Andes, from logs of student entries on the subset of problems that method was able to handle.

[4]While it is probably true that we 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 $1-e^{-1/x^2}$ has a minimum at $x = 0$ where it is zero and so are *all* its derivatives, and yet it depends on x.

Providing a *solve tool* for the students

Finally, an algebra system should be able to help the student with 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. We have decided to implement three tools

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), but for humanities majors, the professors feel the emphasis should be on the physics principles only and not on teaching algebra techniques

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 equation simplified. This would permit the student to guide the system to solve simultaneous linear equations without it being done as black art. Thus it would probably be more suitable than the genie for engineering students, for whom the genie might be disabled.

Various diminished versions of the genie are also available within the algebra package but no interface for them is currently planned, so they will not be available.

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 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 derivation is fishy 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. On the other hand, the old Andes was simply unable to generate the lists of derived equations for 27 of the 115 problems used by Andes in the fall of 2000. How often color-by-number approves equations that should be rejected in actual use needs to be examined. We would like to examine the differences in responses which occur in student use between color-by-numbers and the comparison lists generated by various algorithms for "all possible" derivations. We are not yet ready to do so. We would also like to compare the imputed dependencies with those recorded by the derivation engines. The new method for evaluating the dependency of student equations is being used for the first time this fall, while the "color by number" method for determining correctness has been used for one year. We do not currently have data for examining the effectiveness in practice, but this will be discussed[6] in a forthcoming article.

Summary

A new physics tutorial system is emerging from the Andes effort, which will make very substantial use of a powerful algebra subsystem. This algebra system has introduced new capabilities for dealing with dimensional analysis, for solving systems of equations, in particular the full physics problem, and for providing algebraic help to the students. In addition, it is using a new method of judging, indeed a new definition of, the correctness of student equations (or of possible variations thereupon) and of judging the subset of canonical equations upon which a student equation depends. These new methods allow much more complex problems to be handled, but the extent to which their answers differ in practice needs to be evaluated, as will be done in a forthcoming article.

Acknowledgements

The author wishes to thank Kurt VanLehn and his Andes Group at the Learning Research and Development Center of the University of Pittsburgh, and the Office of Naval Research's Cognitive Science Division grant N00014-96-1-0260, which supported this work. He also wishes to thank Rutgers University, particularly the Sabbatical Leave Program, which has provided most of the support for his sabbatical leave to participate in this project.

He especially wants to thank Kurt VanLehn for numerous discussions on the use of substitution for equation checking and dependence for determining the student's use of equations, as well as for determining if a new equation adds new knowledge to those already written down. Professor VanLehn made suggestions that materially improved the treatment of dependence determination in the case of zero gradient components. The author would also like to thank Linwood H. Taylor, Collin Lynch, and Anders Weinstein for much valuable programming advice and assistance they provided in aid of his effort to implement code for this system.

References

[1] `http://www.pitt.edu/~vanlehn/andes.html`

[2] Gertner, A. and VanLehn, K. Andes: *A Coached Problem Solving Environment for Physics*. In Proceedings 5th International Conference, ITS 2000, Montreal Canada, June 2000.

[3] Schulze, K.G., Shelby, R.N., Treacy, D.J., Wintersgill, M.C., VanLehn, K., Gertner, A. *Andes: An intelligent tutor for classical physics.* The Journal of Electronic Publishing, University of Michigan Press, Ann Arbor, MI, 6:1 (2000).

[4] Maple©6, Waterloo Maple Inc., `http://www.waterloomaple.com/`

[5] Gertner, Abigail S *Providing feedback to equation entries in an intelligent tutoring system in Physics*, Proc. 4'th Intern. Conf. on Intelligent Tutoring Systems, August 1998.

[6] Shapiro, Joel A., VanLehn, Kurt, and Weinstein, Anders, *Meanings of correctness and derivability of equations in a tutoring system*, in preparation.