Can technology really improve teaching introductory physics

Abstract:

Modern electronic tools have many possible uses in improving physics instruction. Especially in the large courses at which most college students are first exposed to physics, interpersonal contact is rare and traditional lecturing has proven lacking. Computers, the internet, and wireless communications have all provided opportunities to find new ways of communicating with students. They provide improved means of presentation, increased interactivity, and the prospect of 'intelligent" automated tutoring.

I will just touch on presentation tools and on the use of clickers and other current computer aids to increase student involvement in large lectures, but my chief interest is in computer tutoring of student problem solving. I will discuss some of the grading schemes and homework systems, which are gradually developing some tutorial capabilities. Then I will give my views on the prospect for more serious intelligent tutoring systems (ITS), and what is needed for such systems to become effective and practical tools.

1 Introduction

I think your patience for this talk has already been used up by your reading Larry's extended announcements of it. 1. Title

Many of you who are involved in the large introductory courses have already made use of interactive technologies. Some of you, however, are probably not aware of just what is currently available. I also hope to spark some interest in what tools might be possible in the future, that would make a real improvement in what most students could learn in these courses.

So my talk will consist of two parts. After a brief discussion of what needs help, the first part will describe some of the interactive technologies which are in common use today, which can both facilitate our teaching tasks and strengthen our ability to get through to the students. In the second part, I will discuss what I imagine might be a long range goal for developing really effective tools in the future. Here I should warn you that I became a string theorist in 1968, and you see how quickly those efforts led to practical results. 2. Structure

First off, why do our introductory courses need any new attention at all? 3. Problems

Many very intelligent teachers have prepared careful presentations of the very well understood, coherent core ideas of the physics we expect our first year students to learn. Yet there is a Physics Education Research (PER) community that insists that none of this gets transmitted to most of our students, and indeed most of what I hear from the instructors of such courses is dissatisfaction with what the students are learning, and complaints that the students learn at a surface level and forget everything after the exam. Very very few students take these courses voluntarily.

Our lectures are given in large lecture halls, where we give clear presen- 4. Large
tations of the fundamental concepts and demonstrations of their effects. 5
But the students are not really thinking about these concepts. They are 6
trying to catch the tricks for doing problems they will see later on exams or
homework, or they are watching the demos as at a circus, with those with the
biggest explosions making the biggest impressions. During a derivation the 7.
students' attention will wander off, and after a while they even stop coming 8.
to lecture unless they get points for doing so.

Of course we have known for a long time what is now one of the most con-
vincing claims of the Physics Education Research community, that students can 9. active
learn physics only by being actively intellectually involved, and that getting
most students actively involved in large lecture courses is far from auto-
matic. There is a well known solution: One-on-one teaching using the So- 10.
cratic method. This is not **just** a pretty picture: There are studies that
show that one-on-one tutoring by a trained human tutor can be extremely 11 HT
effective, improving performance by two standard deviations. That means a student who would have performed at the 16th percentile is brought to the 84th percentile, or in my last class, from a low C to a middle B+, which is certainly impressive. The key, of course, is that a good tutor does not directly provide the correct information, as a lecture or textbook does, but rather solicits from the student his understanding, both correct and incorrect, and leads the student to think, and then to understand what is right or wrong about the way he is thinking.

Unfortunately we do not work at the School of Athens, and most of our students do not avail themselves of whatever Socratic opportunities we may provide. What can we do to encourage active thinking in our students in large lecture courses?

The same fundamental problem applies to the textbook as well: However 12. Text
coherently presented, the fundamental ideas do not seem to take root in the students' imagination. Textbook publishers have tried to address this

in many ways. They use lots more paper and lots more color, and loads
more examples exhaustively explained, while playing down the importance
of explicating the fundamental ideas. The number of homework problems
has exploded. I thought it would be interesting to compare the first edition
of Sears and Zemansky, which I used as an undergraduate, to the current,
12th edition, now by Young and Freedman. Except near the end, the table of
contents is almost identical. Chapter 8, which is "Impulse and Momentum"
in the first addition and "Momentum, Impulse and Collisions" in the 12th,
has grown from 11 pages $6'' \times 8\frac{7}{8}''$ to 27 pages $8\frac{3}{4} \times 10\frac{3}{4}$ without any change
of coverage. (that is more than quadrupling the area allocated). The ratio of
example to exposition has increased from 0.45 to 0.96, more than doubling.
On top of all this, they are including lots of side thingies in addition to the 13 more
traditional exposition and examples, things like explicit "Learning goals",
pictures of football players colliding, "test your understanding", "problem-
solving strategy", online "Activ Physics" and the like. But I see no sign that
this effort has encouraged students into deep analysis (of the physics).

Of course, even in the old days, we understood that presentation of ideas
alone was insufficient, and the solution was to present *homework problems*
which a student needs to work out herself. The ♡ first edition has 19 problems back to
at the end of this chapter, the 12th has 26 discussion questions, 112 "exer- 12
cises" and four "challenge problems". I doubt that the number of problems a
student actually does has grown, but the instructor certainly has more choice
now. Of course the exams include many similar problems, and in preparing
for these exams students ideally should realize that they need to actively
utilize the fundamental concepts covered to analize physical situations.

This brings us to the effectiveness of homework problems. Here is where
the student will surely have to think for him/herself, right? Unfortunately we 14. HW
have found that many students fail to adequately engage with the material.
They have two unsuitable, though sometimes effective, methods for dealing
with problems. One is "search for formula": List the given quantities and
the sought quantity (the requested answer) and find a formula that includes
those and only those variables. The second is imitation: look for an example
that is the same, except for numerical values, and copy that solution with
new values entered. I think there are even instructors who encourage these
methods, and the kinds of questions we ask often reward them, but they do
not constitute a real involvement with the physics concepts, which most of
us would claim is the primary thing we would like to teach.

The most important place for students to think and show us they are doing so is on exams, especially open-ended written solutions to complex problems. Unfortunately, we don't like to do that any more. 15 Exams

2 How can technology help (today)

Technology can help. Here is my list of the three biggest technological improvements in teaching: Of course the first, other than Archimedes stick and sand on a beach, and probably the most important, was Gutenburg's printing press (1455), followed by blackboards (1801), But to be a bit more up to date, let's talk about current technology. 16 Big tech 17 Gut 18 used today

What I will not talk about is presentation and simultation tools. I am sure you have seen pleanty of PowerPoint presentations, and can judge for yourselves how they may facilitate, or more often destroy, communication of sophisticated ideas. There are thousands of simulations available which are somehow supposed to give students an intuitive feel for underlying principles, but I am not prepared to talk about these. And as far as computer interfaces for laboratory aparatus, at my institutions theorists like me are kept far from the laboratory courses.

2.1 Clickers

One way to engage the students in a large lecture is to ask for questions from the students, or to ask them questions. Unfortunately, it is very difficult to solicit questions from a large class, or even to get responses by raising hands. But it turns out that one can get many more students to respond to multiple choice questions using "clickers", devices that record their choices without making their individual response apparent to the whole class. Students can be graded on participation and/or on correctness. Various pedagogic methods are used with such a system. I used it partly to test at the beginning of lecture whether they had done the assigned reading, but I also asked sometimes ill-defined questions to get them puzzled. One can ask for responses to a question, then ask them to discuss with their neighbors, and respond again, and one can discuss the results of these polls. Students will participate, especially if they get a point for doing so, regardless of how small a fraction of the total grade this represents. This also greatly increases attendance. 19 Click-Help

To my knowledge, the first person to incorporate clickers in a physics lecture was Raphael Littauer at Cornell in 1972. He used discarded high- 20 L&JaS

energy apparatus to build a system in which students discharged capacitors and the net charge for each answer was read. Suzanne Brahmia came to our department from Cornell very impressed with the pedagogic potential of such a system, and tapped into my weakness for fiddling with electronics to convince me to build such a system in our Physics Lecture Hall, which held 330 students. At the time, 1993, one could get clicker systems commercially, but they cost between $200 and $500 per student, while we were able to raise only about a quarter of that, just under $15,000. So for that and three years of my time, I installed a system with a telephone keypad and three LEDs in each armrest, and we used this system for about six years. Here is a photo of the MSDOS screen right after a question was asked and answered, the class told to reconsider after talking to their neighbors, and answered again. Only the upper right quarter was projected to the students, showing a histogram of their answers before and after they discussed the question with their neighbors. Because my system was hardwired, the response by seat could be shown (by color), and I could click on a seat and get the name of the student, so I could ask, by name, for the student's reason for his answer. But I don't think anyone ever used that feature. Furthermore, one day one of the boxes installed under each row of seats started smoking, my system was declared a fire hazard, and never repaired. So much for having string theorists design and build electronics.

But our faculty had gotten quite attached to having a student response 21
system, and commercial systems had gotten much cheaper, so my system Comm
was replaced with an infrared system (PRS) which the students had to buy at about $40 apiece. Commercial clicker systems have greatly improved and gotten slightly cheaper since, so we now use an RF system from iClicker. A number of other departments are also using the iClicker system, so this is no longer a serious financial issue for the students. This system allows asking 5 choice questions, displaying a histogram of responses after the period given for the students to respond, and recording their responses.

Some issues to consider in choosing a clicker system are: what is the 22
communication technology, and how much will be communicated.

- hardwired, IR or RF?
  - Hardwired not available commercially, takes lots of work. But provides seat info.
  - IR came first, but student clicks can interfere, doesn't work reli-

ably enough for grading in a large class.

  – RF seems completely reliable

- feedback: none, few LEDs, or screen?

  – Students don't trust their answer has been received without feedback. Some systems show student ID on projector, but that doesn't work well in a large class

  – Even one LED is enough to show response received. But this requires transmitter to receive info.

  – A multicharacter screen is enough for sophisticated interaction, but I don't know anyone who makes use of this.

2.2 Course Management Systems

There are a lot of very popular course management systems, though they tend to be popular with the non-science communities more than with Physics Departments. This is because they focus on things we can do ourselves without much difficulty (posting web pages, keeping a gradebook, sending and reading email, etc.) or things we don't generally value as much as some humanists (discussion boards, chat rooms, etc.) The commercial ones have grading/homework systems which are too primitive for physics, without numerical or symbolic answer understanding.

Neither I nor anyone else in my department, as far as I know, makes use of a commercial course management system, although my university has put considerable effort first into WebCT and now into Sakai, an open-source system. I have no experience with these, and so I will say no more about them.

2.3 Exams

We began using multiple choice computer graded exams in the late 70's, 24 Exams
again with a homemade system which is still the grading system used in our department. We also used a homemade gradebook, but now there are so many commercial or freeware gradebooks of commercial quality available, that there is no particular reason to use a homemade one. But our exams are still produced by homemade software, because

- the exam can be written in latex
- the exam can be automatically extracted from a "database" file
- the software produces randomized versions, answer keys for the grading program, and all the information needed to give performance results, not just for each student but for each question.
- I have been accumulating exam questions in this format for years.

I am perhaps a bit too proud of the latex capabilities I wrote for this software, so I want to show off some questions. Here is a question with a 25 fpix
figure short enough so that not all the answers need to be narrowed. Even when the answers are permuted, only those upper answers which need to be are wrapped narrowly. Next, we have questions where the answers them- 26 lmva
selves are shown on the picture, with the letters permuted as necessary for the different versions of the exam.

Probably the biggest improvement is that we have extended our grader's abilities to allow numerical answer questions. Here is the form we use. 27 Num Ans
You probably can't read the form, but you can see that there is room for the student's name, ID, course information and exam code, and for 15 numerical answer problems and 30 multiple-choice questions. Here is a 28 NumQ
blowup of one numerical answer box, with the answer to "what is the charge of the electron, in Coulombs". In the five years this feature has been available, I think it has only been used by two professors, but I do think it can be valuable, especially in the course for engineers.

People give multiple choice exams because of the painfulness of grading 29 MC OK?
written exams by hand for a large class. Many professors question whether a multiple choice exam can assess deep student understanding as well as an open-ended written problem, and it is clear at least that multiple-choice questions need to be carefully designed, especially clearly stated and with well thought out distractors, but if that is done, they may well be as accurate as written exams. They certainly help with objectivity and consistency of grading. Even if they evaluate as accurately, however, there may still be a drawback in how the students perceive how they should be preparing for the exams. I don't have any concrete information on this.

In the last 10 or 15 years, our propensity for off-loading exam grading to the computer has been extended to homework. A number of computer-based homework systems were created in physics departments, while at the 30 comp HW same time the commercial course management systems, such as WebCT and BlackBoard grew components that would grade multiple choice or short answer questions. I believe most physicists regard the latter as insufficient, but many departments now use systems such as WebAssign or Mastering Physics which grew out of physicist-created systems that accept numerical answers or symbolic expressions. There is also a system which remains open-source, the CAPA system, but CAPA has now been incorporated into LON-CAPA, on which recent development seems to be more on the course management side than on the capability of the homework grader. I found it difficult to use. These three systems have hinted at developing tutorial functions, that is, providing more help than just marking each answer right or wrong. When last I looked, Mastering Physics had more of this than the others, but its tutoring abilities are still rather limited, and the tutoring needs to be hand-authored for each problem.

I am rather concerned with the cost of these homework systems. When 31
we first used WebAssign, in the fall of 2002, it cost \$8.50/semester, but this CvsOS
semester, if you didn't get it free with the purchase of a new copy of the very expensive Wiley textbook, it cost \$32/semester. I wish someone at Rutgers had taken the trouble to learn how to use LonCapa so I could convince my colleagues this is doable. There is also an open-source system used primarily in Mathematics, called WebWork.

3 Tutorial Systems

What I do want to talk about, the last part of this talk, is what my vision is of what would be needed to build a true computer-based tutorial system.
What we have is not enough. Nearly all the professors I know teaching the 32 NG
introductory courses are appalled at how poorly the students learn the concepts and how shallow their learning is. I find the students average 95% on the homework problems but when asked to do simpler problems on exams they score around 60%, and these are not even challenging questions. In the old days students would at least write out with several steps their approach to a problem, and a grader could criticize their approach and correct components. With the current feedback, which consists mostly of right or wrong on very specific questions (that is, answers which are a single quantity or expression, rather than a full solution), students seem not to improve their general understanding.

To generate a deeper understanding, students need to attack problems in which they must decompose the situation into pieces that fit the known physics, and then apply it correctly to generate equations. In problems, the final step is to solve those equations. What sort of tutoring system could help?

At the heart of such a system would be a problem tutor which could present a complex problem to the student, accept not only a final answer but all the material we mean when we say, on written assignments, to "show all work", and be able to comment on those inputs, not only on the correctness of each piece, but what is wrong with each piece, hint at how to correct it, and hint at how to proceed if a student is stuck. With such a system, it would be possible to concentrate on problems which are not simple one-step or plug-and-chug. Without such a tutor, we cannot rely on complex problems because many students will be unable to make progress on them, except by getting the answer from somewhere else (which is a growing problem now that almost every answer is available from google).

What would such a tutor need to do? The essential features would need 33. like
to enable a student to

- draw the crucial features of a physics diagram, such as a free body diagram,
- input equations,
- define variables,
- assign numerical values to variables *with units, of course*,
- and ask for help when needed.

It would need to give feedback at all stages.

Can such tutors be made? There is very little work in the field, but one 34 possible
serious attempt in physics is the Andes tutor, and there is also a commercially available algebra tutor in widespread use in high schools which came out of Carnegie-Mellon. I will describe what features I think a tutor should have, which is pretty much what the Andes tutor, part of which I wrote, has, and then I will make some comments on where I would like to see it improved. First, let's talk a bit about Andes.

The Andes tutor was created as an attempt to answer the question asked by the Office of Naval Research of the Artificial Intelligence grantees it was supporting: "What can you do to help the mission of the Navy". It turns out every student at the Naval Academy has to take introductory physics, and there were people there, as well as at the Learning Research and Development Center at the University of Pittsburgh, for whom an ITS for physics was the perfect project.

Here is the opening of a problem in Andes. You can see that on the left, 35
top part, is the problem specification, and, as we will see, the area for the Andes
student to draw diagrams. Below that is a currently empty dialog area. On the right, above the many slots for entering equations, is a list of variables so far defined. Andes has predefined two times, when the system is let go, and $t = 2$ seconds, the time specified for the answer. (In case you can't read it, the question asks how far the block will fall in two seconds.)

The student needs to define a set of axes, draw force and acceleration 36 later
vectors, define the mass variables, and enter equations. Here is the screen after he has made quite a bit of progress.

So, more generally, what must a real tutor have? It needs to *understand* 37 need1
the physics! If the problem is sufficiently challenging, we do not know what the first step the student must take, and often he can take one of several approaches. The problem author will not be able to spell all these out, but a computer system that can actually solve the physics problems can do so. That step is actually easier than matching student inputs to the possible solutions. Even when the student input equation is correct, it may be a combination of equations of the tutor's solutions, rather than a match to a single one. The tutor must not only judge the correctness, but also which elements of a solution are entailed in that equation. Wrong entries are even harder to identify, and a tutor that can say "You need to think about the direction the normal force is pointing" will be far more effective than one that says "that formula is not correct". A tutor also needs to help when a student is stuck, by giving what the Andes group calls "What's next" help.

Andes did not have broader model of the student's knowledge, beyond the current attempt to solve the current problem, but that is clearly something which would be good to have, for an even better tutor.

In order to build a tutor with the capabilities I have just described, the
system needs to have several complex components. The first is, of course, a 38 com-
user interface. The student needs to be able to see a clear problem statement, ponents
with pictures, of course, but he also needs to be able to communicate ideas such as directions. We generally insist on free-body diagrams, so a student must be able to draw a vector, and declare what it represents, and give

it a variable name. Of course problem solutions need to have equations, generally quite a few of them. Inputting equations is fraught with questions, and I think a system ought to be able to accept them in ascii form (like in TeX) or with tools such as Word with MathType provides, but in either case it should display the equations (in addition) in typeset form, so the student can see that his expression has been understood as he intends. The system needs to be able to dialog with the student, though how will depend on how much progress there is in natural language parsing. And it should display a summary of what has been achieved so far.

Because I want the tutor to handle complex problems in flexible ways, the solution will not follow a single path that a problem author can spell out. So the system needs to understand the physics of the problem. The physics un- 39 solver
derstanding consists of production rules which incorporate both basic physics principles and the techniques for applying them. It needs to understand a problem statement, probably a formal statement that is authored along with the statement for the student. From this it keeps a list of what is known about the current problem, and as it applies the production rules, and learns more about the problem, it needs to know when all possible conclusions have been drawn, whether the problem has been solved, and what paths it has developed to solve the problem.

This might sound like the most ambitious requirement of an ITS, but in fact, the Andes system does this part quite well. What is actually more difficult is the *help system*.

The help system controls the interaction with the student. Thus it needs to keep track of where the student is in his attempt to solve the problem. The problem-solving component has provided it with solution paths, includ- 40 help
ing correct equations (with their justifications). I call these equations the *canonical equations*. It also provides a list of variables involved in the problem. As the student makes entries, the help system has to judge if they are correct and also what they indicate about where the student is along a solution path. The student's equation is often not one of the canonical equations, but it should be derivable from them. AndesI tried to list all derivable equations for each problem, but this proved unworkable except for the simplest problems. One of my major contributions to the reworked system, AndesII, was to convince the builders that derivability was equivalent to evaluating as correct on the solution space of the canonical equations, which is usually a single point in the space of the variables. This made correctness trivial to check, but it doesn't make clear which of the canonical equations he is aware

of and has used. This can be a problem if the tutor is called upon to hint at a next step. Even harder is determining what a student is trying to do and what he did wrong when he presents an incorrect equation. It could be a totally wrong concept, or he may have neglected one force, for example, in the total, or he may have made a mistake in projecting a component of a force, or confused two different quantities.

Both the problem-solving system and the help system need information 41 alg
of an algebraic nature. As the problem-solving system applies rules to give equations, it needs to know whether this new equation adds knowledge of the variables, and when it has enough equations to determine a solution, what that solution is. The help system needs to know, when a student enters an equation, whether it is correct and on which canonical equations it depends, this last so as to give the student credit for using those equations. Andes has a separate component, the algebra system, which I wrote, that answers these questions. There are some tricky issues in determining dependencies, not for linear equations, of course, but most physics equations are not linear in all the variables together. I wrote a paper on this. See refs.

Ideally a tutoring system would keep track of what each student knows 42 stud
and present problems designed to expand that knowledge. This requires a model
model for the student. The Andes group did at one time try to build a Bayesian model for the student, but this has not actually worked.

4 My dream

If Andes has so much of the features I want, what is lacking? First, Andes 43 to see
was designed to be installed on the PC the USNA gives to each student, and was not web based. This needs to be changed, and they have gone some way down this road, but it is not fully web oriented.

It has a definite idea of what a student must always do, including onerous specification of each variable. For example, to define a force, the student must draw it, tell what kind of a force it is, what time it acts, on which body, due to what, and what the name of the force is. These are all good things but inputting them for every force becomes very tedious, and the average time to do a problem is 15—30 minutes.

The help system needs lots of improvement. Students can get very frustrated with Andes lack of understanding of what they are trying to do. Sophia Gershman, who some of you know, after giving it praise for some things, says

"On the negative side, ANDES often clouds instead of illuminating the approach of starting from basic principles! The very issue that ANDES prides itself on gets destroyed by the computer limitations. Often, the student who gets scolded by ANDES for not using a "basic principle" simply didn't write the principle in the syntax or form that ANDES prefers. Students then spend an enormous amount of time trying to find just the right form of the expression instead of analyzing the phenomena and thinking about physics principles."

Other than the year I spent on sabbatical at the Learning Research and
Development Center of the University of Pittsburgh, my work has been done 44 me
together with Don Smith and Chun-Wai Liew. In suitable challenge to Andes, we call ourselves the *Watchung Group*. We have been working on techniques for recognizing the referents of student-chosen variable names, enabling a tutoring system to avoid requiring the student to identify each variable precisely. We feel a good tutoring system would start with Andes-style tight scaffolding requirements, but would relax those requirements as a student developed proficiency in solving physics problems.

Why has there been so little progress and so few people working on this? 45 to be
One reason is that short term projects are more sellable than long term ones. done
In the AI community, where a long range ambitious project might be more sellable, they are more interested in ideas with very general applicability, such as natural language tutors, rather than systems in which mathematics plays a major role, which is a much more limited audience. Many journals in AI insist that all articles have an experimental evaluation included, which precludes (or almost does) publishing developments on part of a system unless it can be incorporated into a functioning system. The same is true of the funding agencies.

One way around this problem would be to have a modular system de- 46 can
velop, so pieces could be improved and progress shown. Here are some of the do
things needed:

- an equation slot, with convenient input and better display of equations, including greek, subscripts, exponents, *and units.*
- a graphic input buffer adequate for the student to draw free body diagrams, simple trajectories, etc.
- a general API for fitting together the various components

- a more robust knowledge representation. Among the things that might be included are real vector equations (and dot and cross products), variable functions (such as $x(t)$ rather than x at specific times) 47

- recognizing variable without full specification by the student. This is what my group has been working on for many years now. We are fairly successful.

References: On my home page, you can find this talk, and also links to 48

- Andes
- Watchung Tutor Group and its publications
- my homemade clicker (SRS)
- an example and instructions for my exam system
- and some other stuff.

I am not a plagiarist. We have to be careful now (Biden, Obama, ...) 49

5 Summary

In this talk, I have 50

- explained why improvements need to be made in our teaching of introductory physics
- told you some of the technological aids you could use now
- told you what I think an intelligent tutor might do in the future, and how it would need to function.

I would love it if some of you decided to create some of these pieces. But I can't make any promises of how they will be received.