Experiences using the open-source learning content management and assessment system LON-CAPA in introductory physics courses

Gerd Kortemeyer
Lyman Briggs College and Division of Science and Mathematics Education, Michigan State University, East Lansing, Michigan 48825

Edwin Kashy
National Superconducting Cyclotron Laboratory and Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48825

Walter Benenson
Lyman Briggs College, Michigan State University, East Lansing, Michigan 48825

Wolfgang Bauer
National Superconducting Cyclotron Laboratory and Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48825

(Received 13 September 2007; accepted 21 December 2007)

We discuss the development and functionality of the LON-CAPA system with a particular focus on its homework and examination functionality. We also describe its more general approach to course management and its infrastructure for course content sharing and reuse. We then focus on measures of student learning and the effectiveness of different content types. © 2008 American Association of Physics Teachers.

[DOI: 10.1119/1.2835046]

I. INTRODUCTION

In 1992, CAPA (a Computer-Assisted Personalized Approach) was started at Michigan State University in a small introductory physics course as a way to provide randomized homework with immediate feedback.1,2 Different students would have different versions (for example, different numbers and options) of the same problem, so that they could discuss problems with each other, not simply exchange the solution. As an example, Fig. 1 shows two versions of the same homework problem, as seen in LON-CAPA today. It might be argued that different randomizations beyond simple changing of numerical values may result in problems of different difficulty, and a problem like Fig. 1 would be unfair in an exam setting. However, when used in homework settings, higher randomization (including major variation of the scenario) leads to more fruitful online discussions. (In contrast to the results for other problem characteristics, outlined in Sec. III D 3, evidence for this claim so far is only anecdotal.)

When CAPA was first introduced, students would receive a printout of the problems and would enter their solutions through a terminal. In later years, a web interface for answer input was introduced. Almost in parallel, starting in 1991, the University of Texas Homework Service was developed,3 which followed (and still follows) very much the same approach. The main difference between the systems is that CAPA generated the problem variations on demand and dynamically, and the UT Homework Services generates all randomized versions of the problems ahead of time. In either system, the problems are coded in what amounts to a mini-programming language, which makes the generation of new problems rather cumbersome.

Both the UT Homework Service and CAPA were soon adopted by other universities, but with a major difference: the UT system runs centrally as a service for other institutions, and CAPA was distributed to other institutions and run locally. The main reasons that CAPA did not adopt a service model were scalability concerns (because problems were generated on the fly, CPU power was an issue at the time), as well as privacy concerns: the developers believed that grade-relevant student information should not leave campus. In light of today’s Federal Educational Rights and Privacy Act (FERPA),3 this concern could not be resolved, as several universities are interpreting the law as prohibiting storing grade-relevant information off campus or outsourcing services handling such data.

CAPA’s distribution principle brought some challenges that the UT Service did not encounter. Because editing new problems is a time-consuming task, and because introductory physics problems are very similar, faculty at different institutions soon started to exchange problem libraries with each other. But because they had separate installations, such exchanges meant sending the associated files via FTP or exchanging floppy disks. Overcoming this infrastructural shortcoming was one of the main design principles in the next generation of CAPA, which we will discuss in Sec. II.

Other technical implementation issues of CAPA, such as (at the time) dependence on X-Windows for problem editing and course management, prompted a team at the University of Rochester to develop WeBWorK.5 The system follows the same educational philosophy as CAPA and the UT Homework Service, but uses the Web as its interface and the Perl programming language in homework editing.

In 1997, WebAssign6 was started at the University of North Carolina, and soon became a commercial spin-off. Very early on, WebAssign worked with textbook publishers to offer back-of-the-chapter problems as a centralized service. Students pay for access to these problem sets. The ability for instructors to create their own questions was added only later. The editor is template driven, with instructors filling in the blanks to create questions of certain types.

In summary, CAPA, the UT Homework Service, WeBWorK, and WebAssign offer very similar problem functionality with comparable randomization features. The systems differ in their distribution mechanisms, their technology choices (CAPA and the UT Homework Service initially were...
strongly driven by paper-based assignments and terminal input, and only later added web interfaces, while WeBWorK and WebAssign were web applications from the start, their problem editing interfaces [CAPA, the UT Homework Service, and WebWork offer programming languages (own implementation, enhanced C, and enhanced Perl, respectively), and WebAssign uses templates]. CAPA, the UT Homework Service, and WeBWorK are free, and WebAssign is a commercial product.

In addition to these more traditional homework systems, online tutorial systems, most notably CyberTutor, ANDES, and Interactive Examples, were developed. These systems attempt to guide physics students by asking Socratic questions, offering hints, and evaluating steps along the way. In contrast, typical homework systems just evaluate the final answer for the most part. The underlying algorithms in these tutorials are of differing complexity, with Interactive Examples following the most deterministic methods, and ANDES being the most flexible. In any case, editing new problems in these tutorial systems is very challenging, and most problems have to go through an extensive evaluation and refinement process before being ready for general deployment.

None of the systems we have mentioned offers full-featured course management functionality, and cannot be seen as a replacement for a course management system such as BlackBoard, WebCT, or ANGEL. Because the homework functionality in these course management systems is insufficient and not well adapted for use in science and mathematics, we consider the specialized physics homework and tutorial systems as complementary rather than competing with mainstream course management applications.

A. Features of LON-CAPA

Also in 1992, another group at Michigan State University started developing multimedia physics content, including self-contained CD-ROMs containing hyperlinked lecture and homework materials, interactive exercises, video clips of lecture demonstrations, and simulations for use in classes. This MultiMedia Physics package also included video analysis software to study motion of simple objects.

In 1997, LectureOnline was started as a rudimentary learning content management system to host the multimedia content previously included on the CD-ROM. It was completely web based and allowed for content to be used and assembled across different courses at the same university. LectureOnline had a somewhat less sophisticated problem engine than CAPA (with a template-driven editor) and supported many multimedia formats. The initial course content materials were imported from MultiMedia Physics, and MultiMedia Physics was absorbed into LectureOnline.
In 2000, the three Michigan State groups joined forces and started the LearningOnline Network with CAPA (LON-CAPA), which was designed to address some of the shortcomings of CAPA and LectureOnline. The features of LON-CAPA include the following:

- Completely web-based interface for all system functionality. X-Windows is no longer needed for problem editing, and Telnet access is discontinued.
- Cross-institutional resource management. Although LON-CAPA is still deployed at separate installations, problems and web pages are copied on demand between machines and automatically updated. Digital rights management protects sensitive content (for example, exam questions) and potential commercial interests (for example, back-of-the chapter libraries of textbook publishers). The system eliminates the problems associated with content sharing across different installations without becoming a service like WebAssign or the UT Homework Service.
- Cross-institutional load balancing. Although all permanent data storage is at the instructor’s home institution, any server in the network can host any session. Thus, sessions can be offloaded across the network in peak load situations, achieving scalability without generation of all problem variations like in the UT Homework Service.
- Access to the full sophistication of the CAPA problem engine through an XML-based problem format with optional embedded scripting, which allows for the use of template-driven editors. The system offers the same ease of generating content as does WebAssign, yet preserves the flexibility of scripted problem formats.
- Multimedia content and problems are made available through an embedded course-management system with functionality similar to BlackBoard and other commercial systems, thus eliminating the need to use two separate systems for the same course.

II. CONTENT SHARING

A. Faculty as users and authors

In fall 2007, LON-CAPA had over 120 participating institutions (about half secondary and half postsecondary) with over 275 000 shared resources, over 100 000 of which are randomizing online problems, almost 100 000 are images, and a little over 50 000 are web pages. The remaining 25 000 resources are other multimedia assets, for example, content assemblies (approximately 8500 “learning paths”), sound and movie files (approximately 780), and animations and simulations (approximately 1700). Figure 2 shows the growth of the resource pool over the years.

Faculty members have written most of the resources in the resources pool, sometimes in connection with externally funded projects, but originally mostly for use in their own courses. For example, through the MultiMedia Physics project and contributions from other authors, complete algebra- and calculus-based physics courses are available as a free textbook replacement. For years, several otherwise rather traditional physics courses at Michigan State University have been taught without a traditional textbook. We found much willingness of the faculty to make their materials available to the pool, and the vast majority of the material in the system is published “system wide.” From workshops and private conversations it is apparent that the few authors who wish to restrict their materials to their own institutions are often hesitant to submit their material to public scrutiny rather than a general unwillingness to share.

At the beginning of the project, we believed that monetary incentives would be needed to motivate faculty to share their material, and a system for charging for the use of others’ resources was included. In the meantime, we found that faculty are far less interested in earning usage fees than feeling a sense of accomplishment when they see the usage counters clicking or receive positive feedback from their peers. Proposed payment and bartering schemes were also seen as too complicated, and it was decided that they might inhibit rather than foster the expansion of the network.

It seems that sharing resources fits into the academic culture, just like research papers are shared. This conclusion might be subject area specific: most participating faculty are from the natural sciences and most resources are intended for introductory courses. Faculty members might take pride in writing a high quality homework problem regarding angular momentum conservation, but hardly base their reputation on it, or are competing with peers teaching the same topic. Also, the physics taught in these courses is far from controversial, so that authors do not have to worry about scrutiny with regards to matters of opinion.

In discussions with authors, the most important aspect appears to be good stewardship of the material: the project needs to guarantee that materials, some of them exam or grading-relevant homework problems, do not “leak” out of the pool; that is, students only have access to the material that faculty select for them. Particularly sensitive is the XML source code of the problems, because it allows for reverse engineering of the randomization and determination of the correct solution for any variation of the problem.

B. Commercial publishing companies

Although faculty generate the majority of content in the system, the network also hosts commercial content from textbook publishers. In addition to the resources shown in Fig. 2, the back-of-the-chapter problems from nearly all major introductory physics textbooks are available in the network. Publishing companies pay commercial companies to prepare these homework libraries in LON-CAPA format, to host them, and to selectively make them available to courses using their text. Faculty and students do not pay for this
service, instead, the publishers use LON-CAPA-coded libraries as an incentive to adopt their books—a model which is very different from WebAssign, where students directly pay for access to the online homework. However, it might be argued that LON-CAPA’s business model leads to an overall increase of the cost of already expensive textbooks.

Because LON-CAPA is a distributed network, it is sufficient to have these homework libraries hosted on one server in the network. The institutions using the problems can deploy them from their own servers and integrate them as homework or exam questions into their own courses, using the mechanisms we have described.

C. Content selection

In a large resource pool like LON-CAPA, it can be difficult to find appropriate resources. Several efforts have been made to facilitate the selection of materials:

- The content is sorted first by institution and then by author.
- Assemblies of content and learning paths are reusable. For example, different homework problems on momentum conservation can be assembled into a homework set, which the next faculty person can use as a whole.
- The system collects usage data, so when instructors locate an appropriate resource, they can see which other resources have been used in the same context.
- Search results can be ranked by different criteria, for example, by total number of accesses.

In an ongoing research project, we are attempting to automatically identify communities of practice in the system, and thus direct members to appropriate authors and content.

III. MEASURES OF AND INFLUENCES ON LEARNING OUTCOMES

A. Course grades

Although faculty acceptance depends on factors such as time savings, convenience, and philosophical considerations, both faculty and students care about learning outcomes. Faculty care, because it is frustrating to see bad exam scores, and students care, because they want good grades.

A simple and reliable measure of learning outcomes is—hopefully—the course grade. Figure 3 shows grade distributions in the standard introductory physics course for scientists and engineers (PHY 183) at Michigan State University. The solid graph shows the averaged distribution in the years 1992–1994 without online homework, and exhibits the classic bell shape with a maximum at around 2.5. Unfortunately, the data, which were obtained from the physics department in 1998, are available only in semester-averaged form and do not allow us to retroactively calculate standard deviations. We were also unable to reconstruct which sections of the course correspond to the data, but we know that at least six different instructors were involved.

In 1996, online homework was introduced. We were able to obtain semester data from five different instructors who taught the course between 1999 and 2007 (gray graph in Fig. 3). Only one of these five instructors may have also been included in the data of the pre-CAPA time period. We indicate standard deviations, which reflect grading differences between the five post-CAPA instructors.

The consistent observation in subsequent years with online homework is that the grade distribution is smaller, around 2.0 in favor of higher grades, and overall becomes skewed.

The average course grade was 2.4 ± 1.0 before the introduction of online homework, and 2.7 ± 1.2 afterward—the difference (0.3 grade points), although not statistically significant, is definitely not inflationary and is noticeable. The difference is not a result of an increased number of points for homework, because homework in all of these courses constituted only a small contribution of the grading criteria, but a result of higher exam grades.

B. Surveys of student attitudes

The reason for the improvement is most likely increased time on task: students are on average self-reported working 1–2 h more per week on physics. Students who on the same survey found online homework particularly helpful on the average worked 2.4 h more. There were also a few students who worked 10 h more per week on physics, some of them finding it very helpful, and others (who failed the course), finding it useless. Another frequently quoted reason is the immediate feedback and the ability to do the problem over and over (within a limited time and for a bounded number of attempts) until mastery is achieved. Although this feature is certainly perceived to be helpful by students and appreciated by faculty, there is some well-reasoned concern, confirmed by research results, that it can also “turn thinkers into guessers,” where students adopt a trial-and-error approach to problem solving.

C. Gender differences

A preliminary conclusion is that online homework most strongly helps students who are on the brink of failing the course. An interesting and unexpected pattern emerges: as Fig. 4 shows, the skewing of the standard bell-shaped course grade distribution can be attributed mostly to female students. PHY 231/232 is a two-semester course, where in a particular year the first semester was taught without online
homework (black histogram in Fig. 4), and the second semester with online homework (gray histogram).

It is striking that the overall improvement of grades (compatible with Fig. 3) is mostly due to female students. In the first semester, the average grade of males was 2.8 ± 0.8 and the average grade of females was 2.5 ± 1.1. In the second semester, it was 2.8 ± 1.1 for males and 2.8 ± 1.0 for females. Much more striking is the fact that the grade distribution of female students was significantly different from the male distribution in the first semester ($\chi^2 = 3500; p < 0.0001$), but this difference almost vanished in the second semester ($\chi^2 = 14; p = 0.05$).

Several simple reasons for this gender difference in the effectiveness of online homework were rejected:

- Enrollment changes: PHY 231 had 422 students (194 male, 228 female). Between the semesters, 20 male and 7 female students dropped the course, and 2 male and 3 female students added to it. These enrollment changes are too small to explain the change in the grade distribution. The students who dropped the course were students who had low grades, and because more male than female students dropped, the opposite effect would be expected.
- Instructor change: Although the instructor changed between the two semesters (both instructors were male), the effectiveness of online homework were rejected:
- Population effect: Although PHY 231/232 (Fig. 4) is a non-majors course, PHY183 (Fig. 3) is a course for scientists and engineers. The same effect was observed at Central Michigan University.

In spite of considerable effort, the reason for this gender difference has not yet been explained satisfactorily, although in the study at Central Michigan University, it was found that females in semesters in which they outperformed the males usually did their online homework earlier, that is, not as close to the due date.

D. Analyzing online interaction between students

A side effect of teaching online (completely, or in our case as an online component to a traditional lecture course) is the increased potential for peer interaction. Within or outside of LON-CAPA, students will congregate online in an effort to solve their homework most efficiently. Although peer discussion of physics problems is a powerful vehicle for learning and discovery among expert physicists, its effectiveness among novices strongly depends on the manner in which such discussion is conducted.

1. Effect of sanctioned and non-sanctioned online discussion boards

Students might have some understanding of the public epistemology of physics, but their personal epistemology may be quite different. In other words, what they believe about the generally accepted nature of knowledge and learning in physics might be very different from what they believe works best for them. A typical statement might be that “yeah, I know I should start with the concepts and derive a symbolic solution, but for me, it’s more efficient to just search for a formula with the right letters in it and plug in my numbers.” And although students might make some effort to follow expert-established expectations in a realm that they know will be visited by their instructors (that is, the discussion features within the system), they often tend to trust their own devices outside of that. Michigan State University, like other universities, has independent discussion sites, for example, AllMSU, where access by faculty is forbidden. It turns out that the self-reported extent of the students’ usage of the third party site was negatively correlated to all aspects of the course (final exam, midterms, and quizzes), as well as the Force Concept Inventory gain, and only slightly positively correlated with homework percentage. In contrast, posting to the internal discussion board was positively correlated with all of these aspects.

2. Types of discussion as correlated with grades

Similar correlations can be observed within LON-CAPA’s internal discussion board. In a study analyzing several thousand student discussion contributions in introductory physics courses, it was found that the prominence of purely solution-oriented contributions (“To get the right number, do …”) is on the average 75% for a student with a grade of 2.0, and is only 45% for a student with a grade of 4.0 (see Fig. 5). In contrast, contributions that discuss the underlying physics and concepts of the problem go from only 18% and 6%, respectively, for a 2.0 student to 38% and 15%, respectively,
for a 4.0 student. Details on these classifications can be found in Ref. 26. The discussion patterns of students who do well in the course give evidence of a much more expert-like epistemology than those of weaker students. However, the direction of the causality, that is, they do better because they discuss better or vice versa, could not be conclusively determined.

3. Types of discussion as influenced by problem

If more expert-like discussion behavior leads to better learning, then it is interesting to find which online problem properties lead to such behavior. Simple multiple choice and numerical problems lead to the worst discussion behavior, while ranking and click-on-image problems are the most promising. An example of a ranking problem would be cylinders with different mass distributions rolling down an incline where the student needs to rank them in order of arrival at the bottom, while a click-on-image problem might be a circuit diagram where the student needs to click on the diagram to “cut” wires to make a light bulb brighter. Another even more important problem property is its difficulty. Although we might naively expect that more difficult problems necessarily lead to more expert-like discussions, this is not the case. In Fig. 6 we show the prominence of discussions that involve the underlying concepts in the problem, as well as the prominence of procedural discussion (“first do this, than that, …”), as a function of the degree of difficulty of the problem. For problem difficulties above 6 (out of 10) on our scale, the error bars of conceptual discussions are compatible with a horizontal line. Also for very easy problems, the prominence of conceptual discussions goes up, and procedural discussions go down. In addition, we classified student emotional discussion contributions, where positive remarks were counted positively, and negative remarks (for example, complaints) negatively. Above a difficulty of seven, the majority of student emotional discussion contributions were negative. The somewhat flippant conclusion is that above a certain average problem difficulty, there is more pain for no (significant) gain.

IV. CONCLUSIONS

We have introduced LON-CAPA as an open-source tool to develop, use, and share online teaching and learning resources. Initially, this system was developed and used for introductory physics courses. Although the subject coverage has broadened to include chemistry, biology, mathematics, statistics, accounting, psychology, food science, and several other subjects, LON-CAPA is still driven by innovation from science faculty, mostly physicists. During the past 15 years we have used this tool for teaching physics classes of various sizes and for various target audiences and for performing
research on the effectiveness of teaching methods and approaches. We find that individualized, online homework can be an effective learning aid. Sharing of such resources across institutional boundaries is a reality.

ACKNOWLEDGMENTS

The authors thank the National Science Foundation (Grant No. ITR-0085921) for support of this project. Support in earlier years was also received from the Howard Hughes Medical Institute, the Alfred P. Sloan, and Andrew W. Mellon Foundations, as well as other grants and grant supplements from the National Science Foundation. The authors are grateful to Michigan State University and to its administrators for over a decade of encouragement and support, as well as our partner institutions. The authors would like to thank two anonymous reviewers for their detailed and constructive comments which helped to improve the quality of this paper.

Author to whom correspondence should be addressed. Electronic mail: korte@lite.msu.edu