Using Handheld Devices for Tests in Classes

Franklin Chen Brad Myers David Yaron

July 2000

CMU-CS-00-152
CMU-HCII-00-101

Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

An important method of evaluating students' progress in courses is the administration of tests. The increasing availability of handheld computers, such as Palm and Windows CE devices, in conjunction with wireless networks, allows the automating of aspects of giving tests in the classroom. During the Spring 2000 academic semester, we experimented with using Windows CE devices in a chemistry course to allow the instructor to intersperse, with lecturing, the administration of a form of "concept tests", in order to determine whether material just covered was understood, thereby enabling the instructor to modify the content or presentation of the rest of the lecture. We found that most students preferred the use of handhelds for this purpose to the use of a show of hands or holding up of flashcards.

This research is supported by grants from Microsoft, Hewlett Packard, Lucent and DARPA. This research was performed in part in connection with contract number DAAD17-99-C-0061 with the U.S. Army Research Laboratory. The views and conclusions contained in this document are those of the authors and should not be interpreted as presenting the official policies or position, either expressed or implied, of the U.S. Army Research Laboratory or the U.S. Government unless so designated by other authorized documents. Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

1 Introduction

An important method of evaluating students' progress in courses is the administration of tests. Tests come in many different forms, from simple true-false tests, and multiple-choice and fill-in-the-blank tets, to more free-form tests involving essays. The simplest forms of tests are easy to grade mechanically, so as technology becomes more and more used in education, it seems reasonable to use it to aid in the creation of tests, and grading and analyzing students' results.

The most obvious way to immediately make use of technology is to simply treat it as an improved medium for doing what was done before. In this scenario, let's call it "ordinary" testing, a student logs into a course's test Web site and is presented with a simple test, e.g., multiple-choice, and after the student submits a completed form, the test is graded and statistics gathered, and feedback is given to the student. This scenario amounts to a more efficient method of giving the old paper-and-pencil test.

It is interesting to consider, however, ways of using technology to do new things, rather than just do old things better. With the increasing availability and use of handheld computers, such as Palm and Windows CE (now Pocket PC) devices, and wireless networks, it has now become possible to take advantage of in-class use of computers and immediate feedback to the instructor and students. So here is another scenario for testing: a instructor with a server PC creates some kind of test, e.g., multiple-choice or fill-in-the-blank, and loads it into a program that will cause students with PDAs to receive a representation of the test and submit responses. The server then collects all the data, tabulates the results, computes statistics, and displays them to the instructor, who then makes use of the information to continue the lecture. The main point of this accelerated form of pop quiz, a version of what is known as a "concept test" [1], is to enable the instructor to determine whether material just covered in lecture was understood; so, for example, if it turns out that it was not understood, the instructor could elaborate on the material and slow down the lecture as appropriate.

In summary, the work dealt with two kinds of tests:

Ordinary test A test designed to work as a one-time quiz that can be taken and scored just once.

Concept test A test designed to be taken as part of a lecture, with immediate feedback displayed to the students and instructor, in which a student may repeatedly submit new responses to the same question.

2 Approach

We decided that instead of using the Pebbles[1] PDA/PC technology, we would use off-the-shelf Web technology. The main advantage of doing so is a kind of *portability*. Instead of supporting just Windows CE, our first intended platform, we can support any combination of platforms that involves running a Web server and running Web clients. Example clients other than handheld Windows CE devices would be desktop PCs, laptops, and in the future, Palm devices.

[1] `http://www.cs.cmu.edu/~pebbles/`

We originally began by developing software to aid in the creation of custom online "ordinary" tests, with arbitrary questions, answers, and accompanying content, then worked on modifying the software in light of collaborating with an undergraduate chemistry course to demonstrate possible innovative uses of handhelds in the classroom given the infrastructure set in place of the Wireless Andrew project[2] .

The course *Chemistry 09-106: Modern Chemistry II*[3] was taught by David Yaron and Garry Warnock. It turned out that Prof. David Yaron, of the CMU chemistry department, for the purposes of his course, only needed and wanted specific kinds of "concept" tests, so we specialized the just-developed concept test software to fulfill his needs. This specialized concept test, call it the "generic concept test", is a variant of the concept test in which there is only one question, and no question/answer uploading by the instructor is necessary, because questions and choices are given in lecture through paper handouts distributed before class. The instructor can reset for the next question immediately.

3 Hardware

The handheld we used was the HP Jornada 680, since HP had donated many units. But actually, the software is platform-independent, since it uses standard Web CGI. The only platform-dependent issues that came up were the following:

- strange behavior in Microsoft Pocket Internet Explorer having to do with choice CGI field name (this did not occur with the older handheld HP 620LX, used before the new ones arrived), and
- the desire to lay out the GUI elements in such a way as to comfortably fit in the limited screen size of the Jornada handheld.

The server software was all developed and run on a Linux PC, running CMU facilitized Red Hat 4.2.

4 Administration

The distribution of handhelds to the students of the chemistry course were handled by the CMU Computer Store.

Unfortunately, a mapping of handheld unique IP address to student was not available, so we had to have students "register" their handhelds in class. We did this by providing a Web page in which students entered their Andrew IDs through their handhelds, and building our own table of which student had which handheld.

A Linux PC, `dahmer.pscico.cs.cmu.edu`, used as the Web server, was aliased to `peb.cs.cmu.edu`, to make it easier for students to remember and access. The Web server on this machine is Apache[4], which comes installed and ready to go on Red Hat Linux.

[2] `http://www.cmu.edu/computing/wireless/`
[3] `http://ir.chem.cmu.edu/chem106/`
[4] `http://www.apache.org/`

5 Classroom use

There were 98 students present in each lecture, in the earlier stages of the course.

Each day on which concept tests were given in the class, approximately six tests were presented to the students.

The number of students who responded during the concept tests varied: sometimes there were as few as 20, sometimes as many as 50, depending on the day. Many students either did not have their handhelds with them; did not have them out of their backpacks; or did have them, but chose not to participate in the concept testing.

During the first trial uses of the handhelds for concept testing, a few students played around and deliberately changed answers repeatedly, or chose "None". Eventually this behavior disappeared.

Eventually, in order to collect statistics for each student, we had the students "register" their handhelds, so that it could be determined who was taking concept tests, and what their answer patterns were (Figure 1). It might have been better if registration could have been handled as the handhelds were being assigned to students.

It was interesting to see how Prof. Yaron worked the concept tests into the lecture. He would pose and explain a question, and students would start answering. After perusing the results for a while, he would then start to explain the answer, or remark that, for example, most of the people were settling on an answer that was incorrect, with the result that the distribution of newly obtained student responses would change.

During two of the lectures, Prof. Yaron saved out score data for a completed test before moving on to the next test. The total number of tests logged during these two lectures was 11. There were several days of tests from previous lectures that were not logged, because the logging feature had not yet been implemented or had not yet been tried.

5.1 Concept tests in conjunction with a demonstration

Several times concept tests were given interleaved with a chemistry demonstration. The chemistry demonstration worked as follows: Prof. Warnock came prepared with test tubes, chemicals, torches, etc., and when Prof. Yaron turned over control of the classroom to Prof. Warnock, the projector was switched to take input from a camera directed at the demonstration area rather than from Prof. Yaron's laptop.

This seemed to work well. The demonstrations were enjoyable, and involved highly visible consequences of chemical reactions, e.g., soda bottle rockets being launched, light bulbs turning on, gases forming, liquids changing color. Prof. Warnock would pause his demonstration at strategic points, e.g., preparing substances but not yet mixing them, whereupon control would pass to Prof. Yaron, who would give a concept test asking the students to predict the consequence of the next step. After enough submissions, he would then pass control over to Prof. Warnock, who would then continue the demonstration.

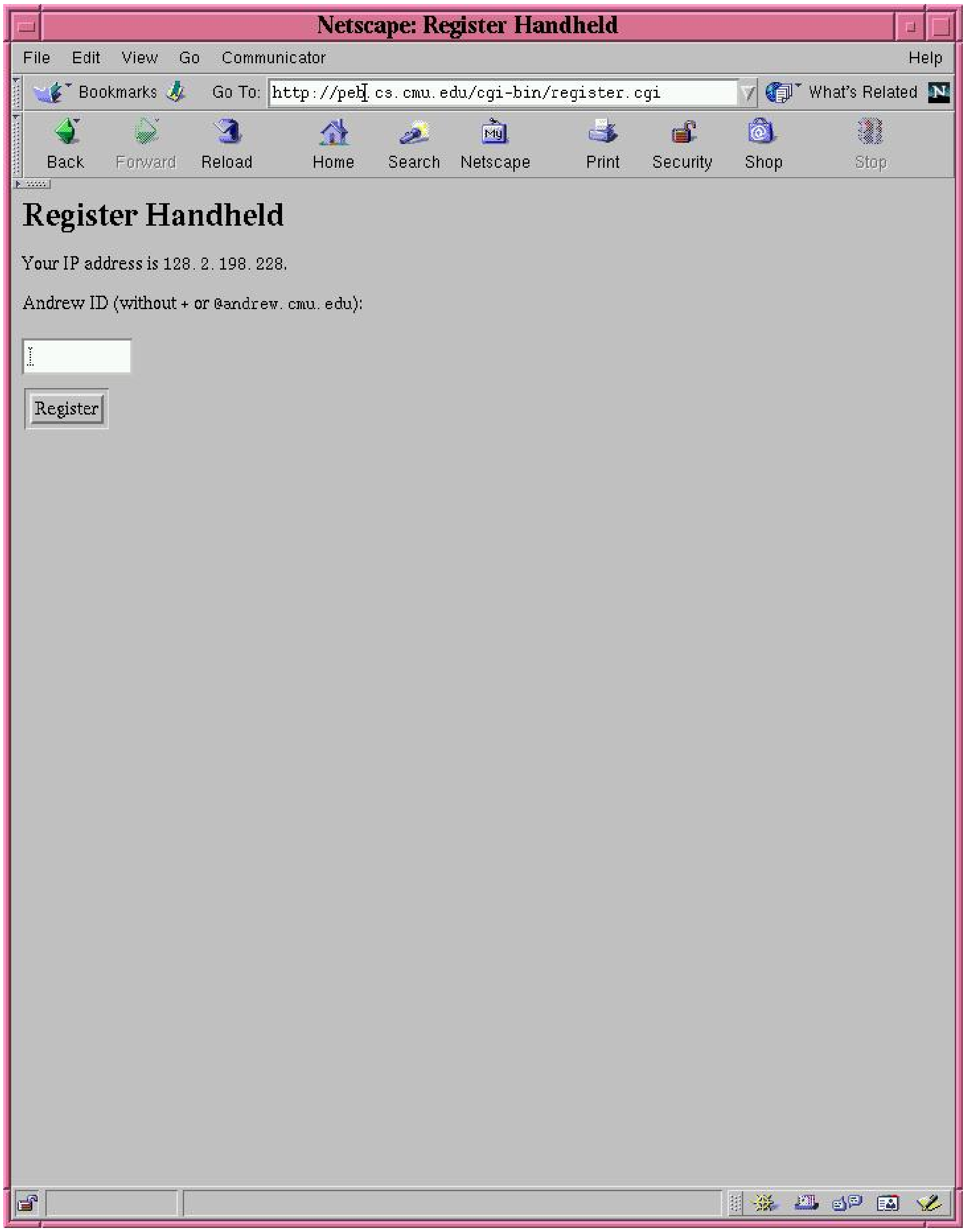

Figure 1: Page for students to register their handhelds

6 Software

There are two classes of users:

- The student, who sees a test pop up and submits responses.

- The instructor, who creates a test and loads it into the system and initiates a testing session.

Correspondingly, there are two separate ways of accessing the test data:

- `http://peb.cs.cmu.edu/` (Figure 2) is for students. From this page students may

  – take the generic concept test;

  – take a particular concept test;

  – take a particular ordinary test.

- `http://peb.cs.cmu.edu/admin/` (Figure 3) is for the instructor. The instructor may

  – list registered handhelds;

  – monitor the generic concept test;

  – display generic concept test scores;

  – upload a test;

  – monitor a concept test;

  – monitor a ordinary test.

For both the student and the instructor, the user interface consisted of Web forms with buttons, text fields, radio buttons, tables, etc.

We were not sure whether the Web server and CGI programs would perform sufficiently efficiently, or whether 80 handhelds in a room would overwhelm the network, but there seemed to be no problems.

The organization of the course information is that each course has its own directory, and each test has its own subdirectory under a course. Another organization (some kind of graph rather than a tree) would be needed if one wished to share a test or related data among multiple courses, for example.

Some access control was provided to prevent students from using the Web site intended for instructors: Apache's `htaccess` feature was used so that only those knowing the correct login and password, e.g., the instructor, would be able to access the Web page.

Figure 2: Main Web page for students

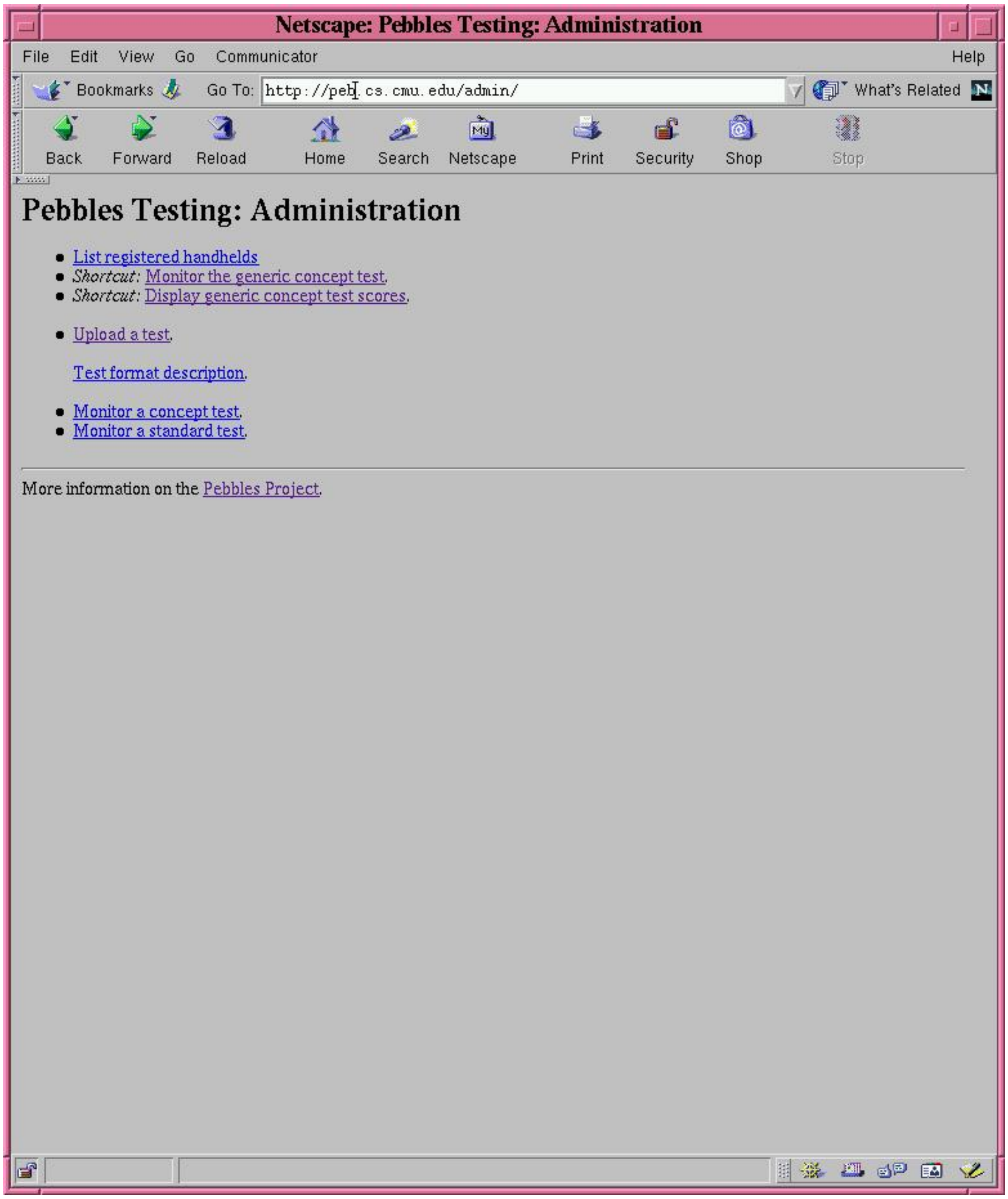

Figure 3: Main Web page for instructor

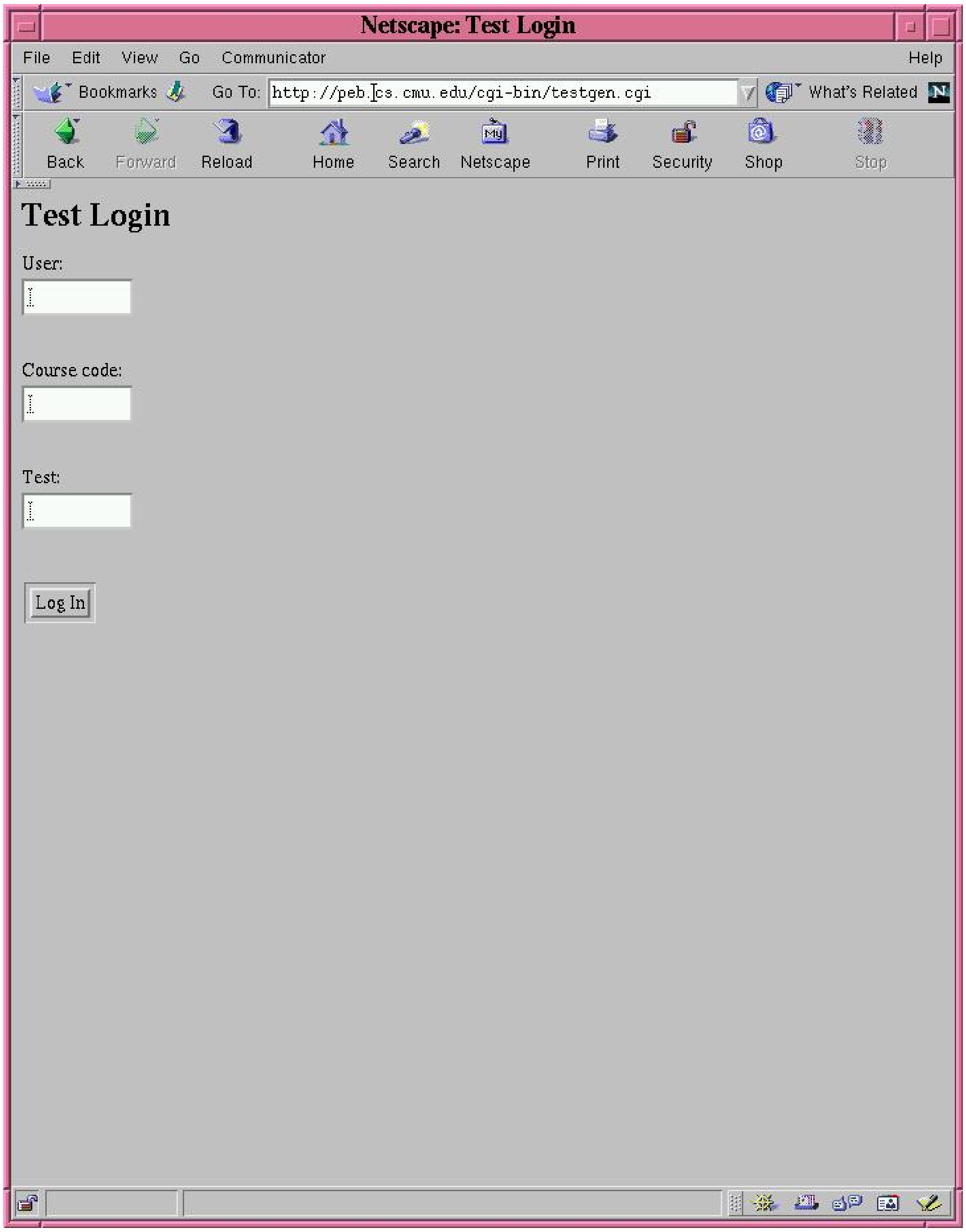

Figure 4: Login page for an ordinary test

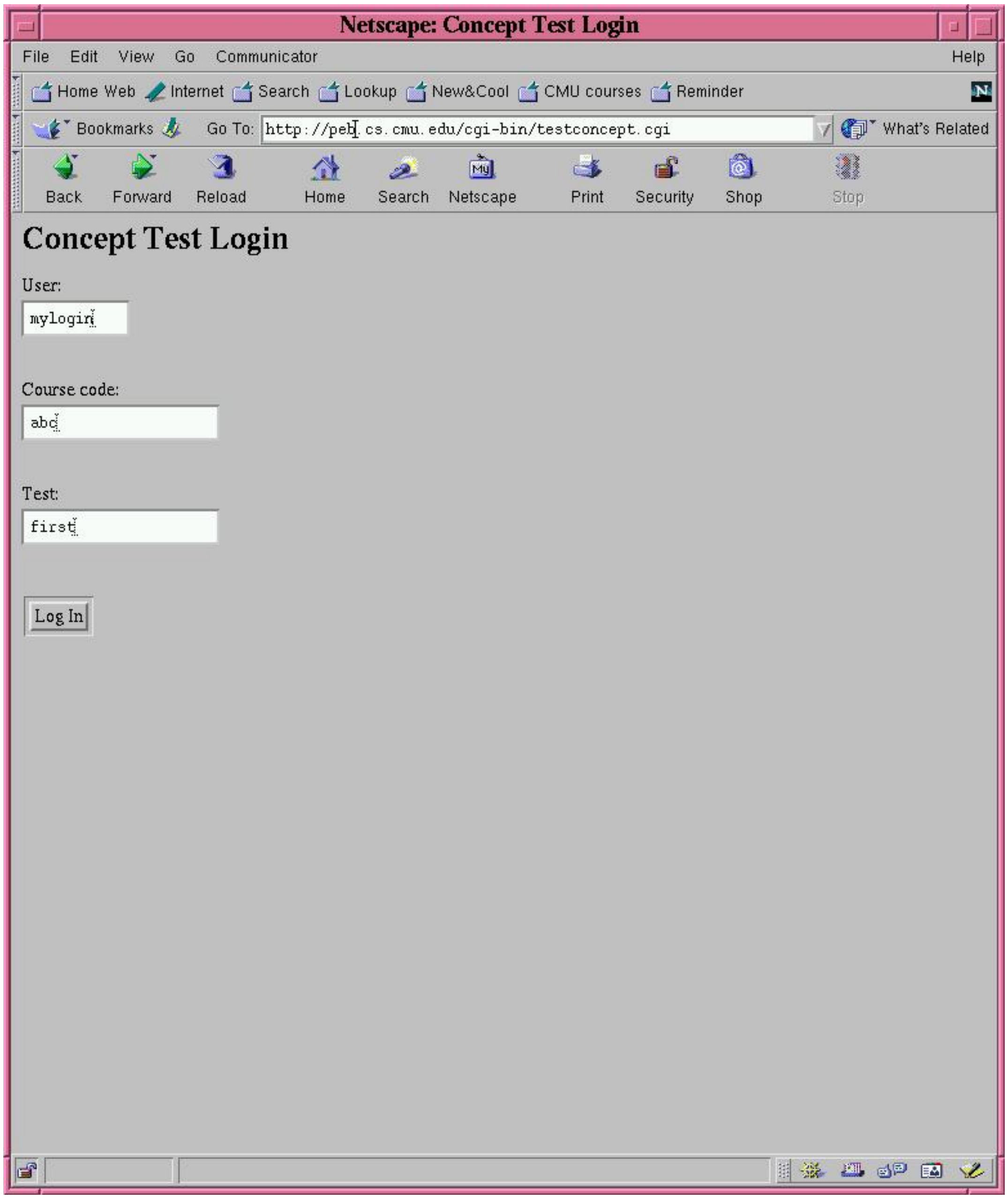

Figure 5: Login page for a concept test

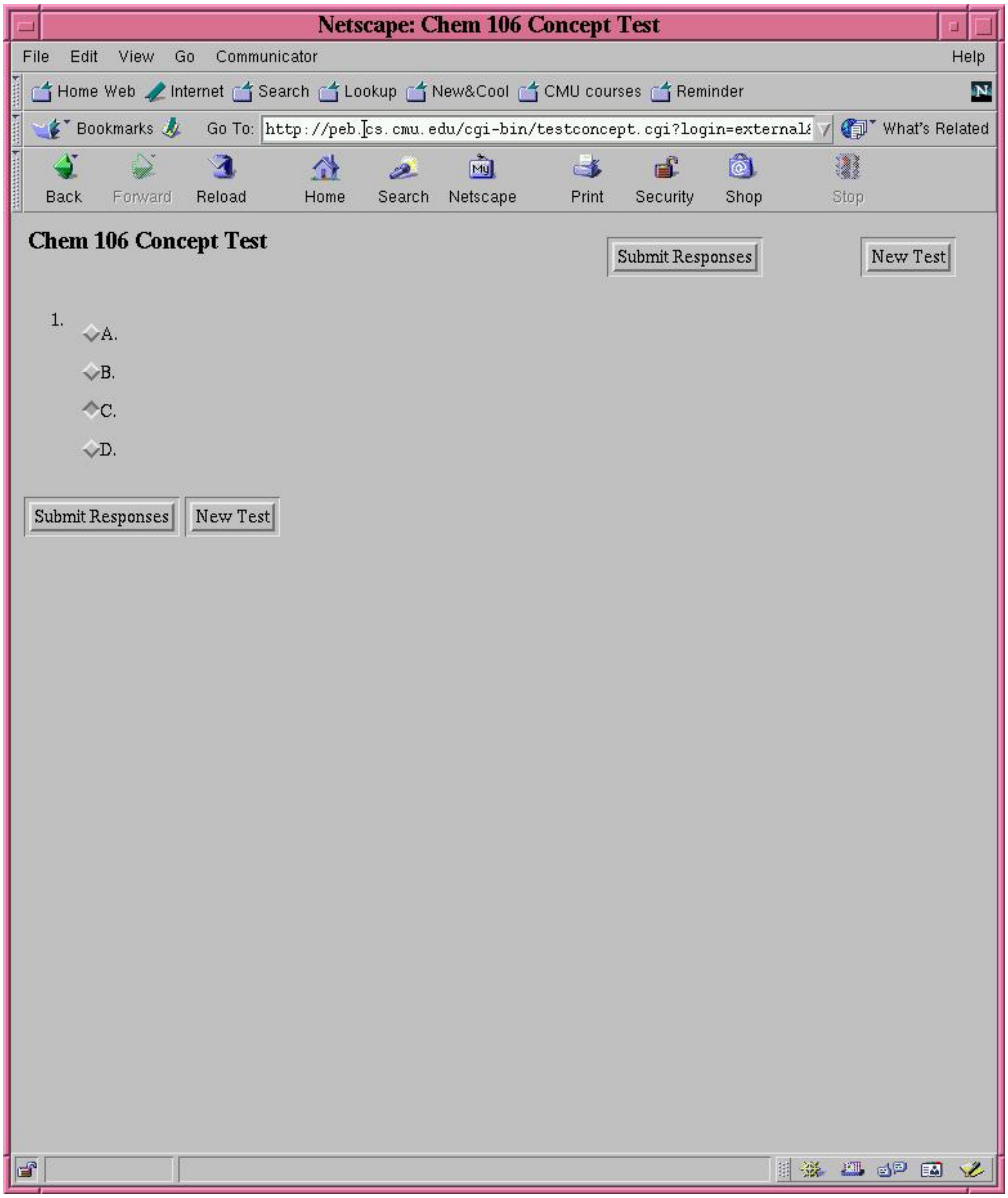

Figure 6: Initial page for the generic test

6.1 Take test

The main Web page for students is a login form. There are two different pages, one for ordinary tests (Figure 4) and one for concept tests (Figure 5). The generic concept test is just a particular kind of concept test that bypasses the login form (Figure 6).

The student, upon logging into a particular course and test, sees formatted content that is an interleaving of the HTML that the instructor designed and uploaded, and a form where the question/answer content can be specified. For example, in Figure 7, the student has successfully logged into a concept test.

The form consists of a numbered list of items, each of which has question text (which has been specified by the instructor with HTML) and a lettered list of possible answers (also specified with HTML), with a radio button preceding each answer. There is a "submit" button at the bottom of the page, and for usability reasons, a duplicate submit button at the top of the page (to make it immediately visible on the handhelds, which have very limited vertical space). There is also a "reload" button at the top of the page, which must be pressed when a new test begins, because it may have different content from the old one; an attempt to submit a form for an old test fails because each test is tagged by an ID that is kept track of.

If the student does not answer a question, this fact is detected, and shows up as the response "None".

After the student has submitted responses for the current test, the result differs depending on whether the test is a concept test or an ordinary test. If it is a concept test, the new responses are logged in the test score database, without overriding old responses, and a page appears with the same concept test being displayed, except this time there is a note indicating the last selected response, in case the student wishes to choose a different one before the instructor stops the test (Figure 8). On the other hand, for an ordinary test (Figure 9), the submission process halts, and a page appears that indicates for each question the student's response and the correct response (Figure 10).

6.2 Submit test

XML was used as a repository of test information, to avoid the need to needlessly invent special syntax for learning and parsing. We designed an XML DTD `test.dtd` to specify the syntax of the annotations.

However, it was not practical for complete online content to be written in XML, e.g., an online test will typically have an introduction, images, and interspersed formatted content, which would be created in a word processor or HTML editor, so it was necessary to use a "neutral" text format for representing test information. In order to make the source file for the test compatible with HTML editors, e.g., Microsoft FrontPage, the question/answer portion of the test was specified as an encoding as ordinary text in HTML, where < and > are replaced by { and }. This way, an ordinary HTML document with an embedded test is still valid HTML and at the same time its test content would be visible (see example in Appendix B). A subset of the XML DTD is used for the embedded text in HTML: `<title>`, `<prologue>`, and `<epilogue>` are inferred from the HTML, rather than specified by the test creator. The test format is defined in Section A.

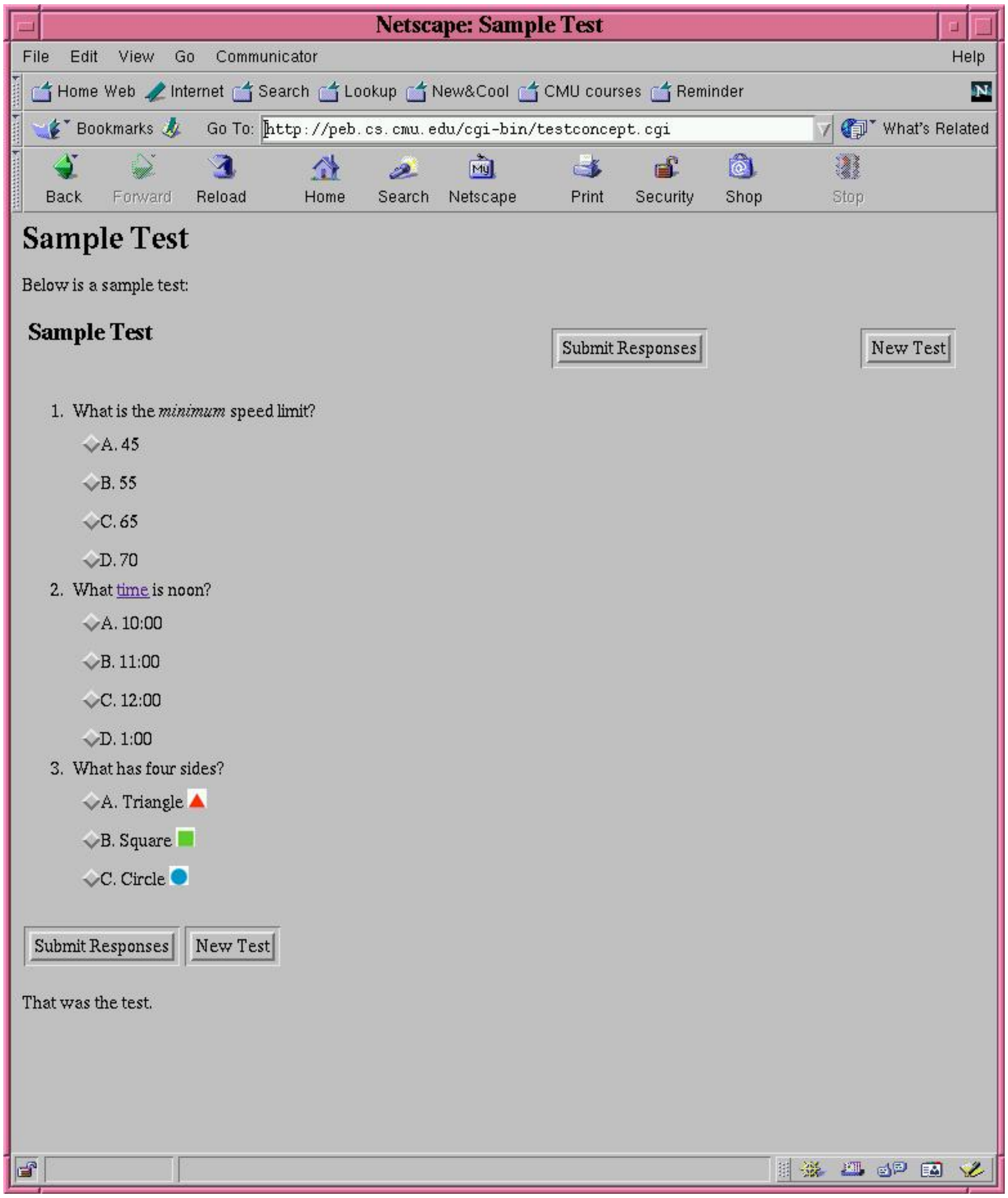

Figure 7: Initial page for a sample concept test

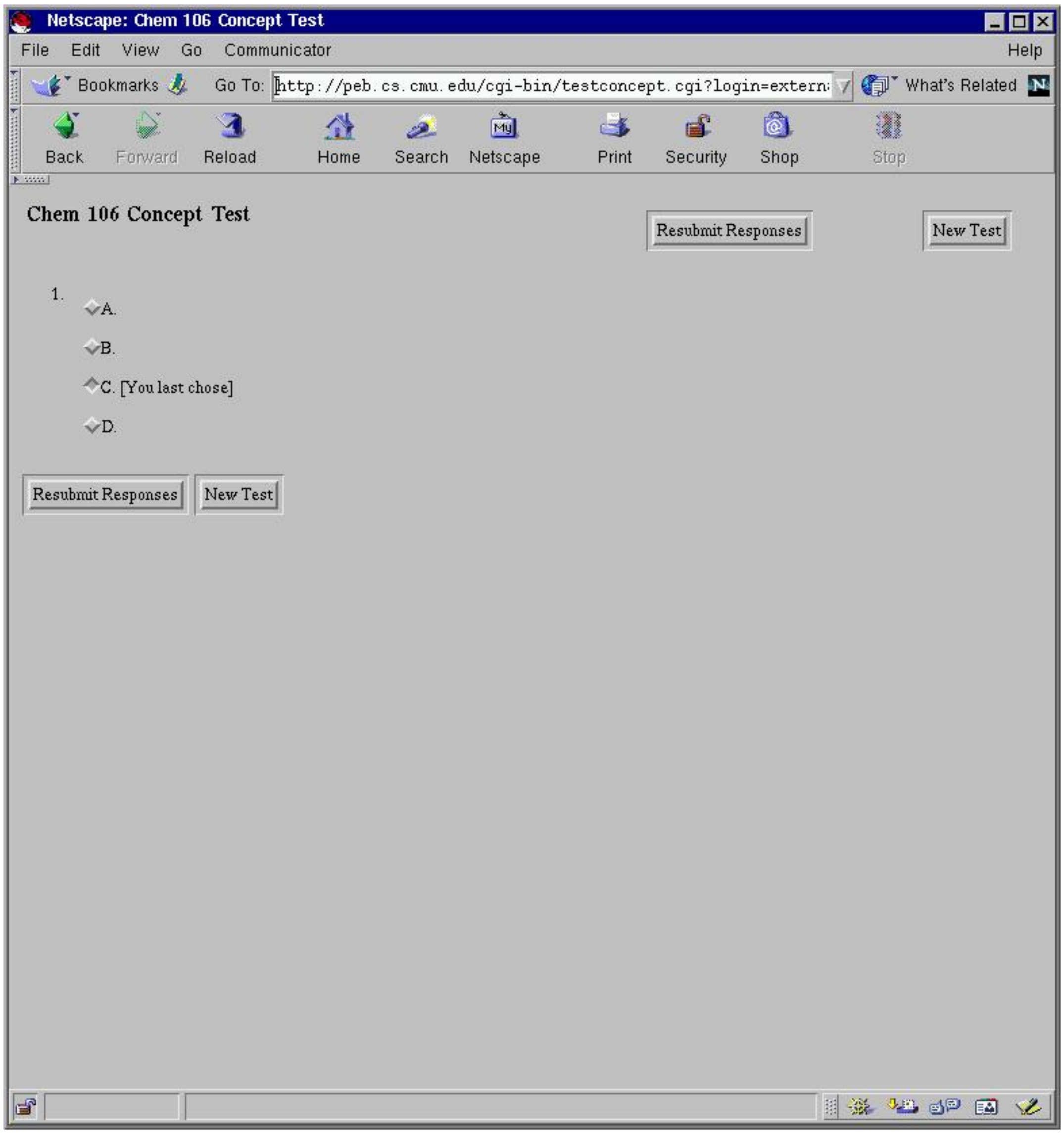

Figure 8: Page for the generic test, after a response has been submitted

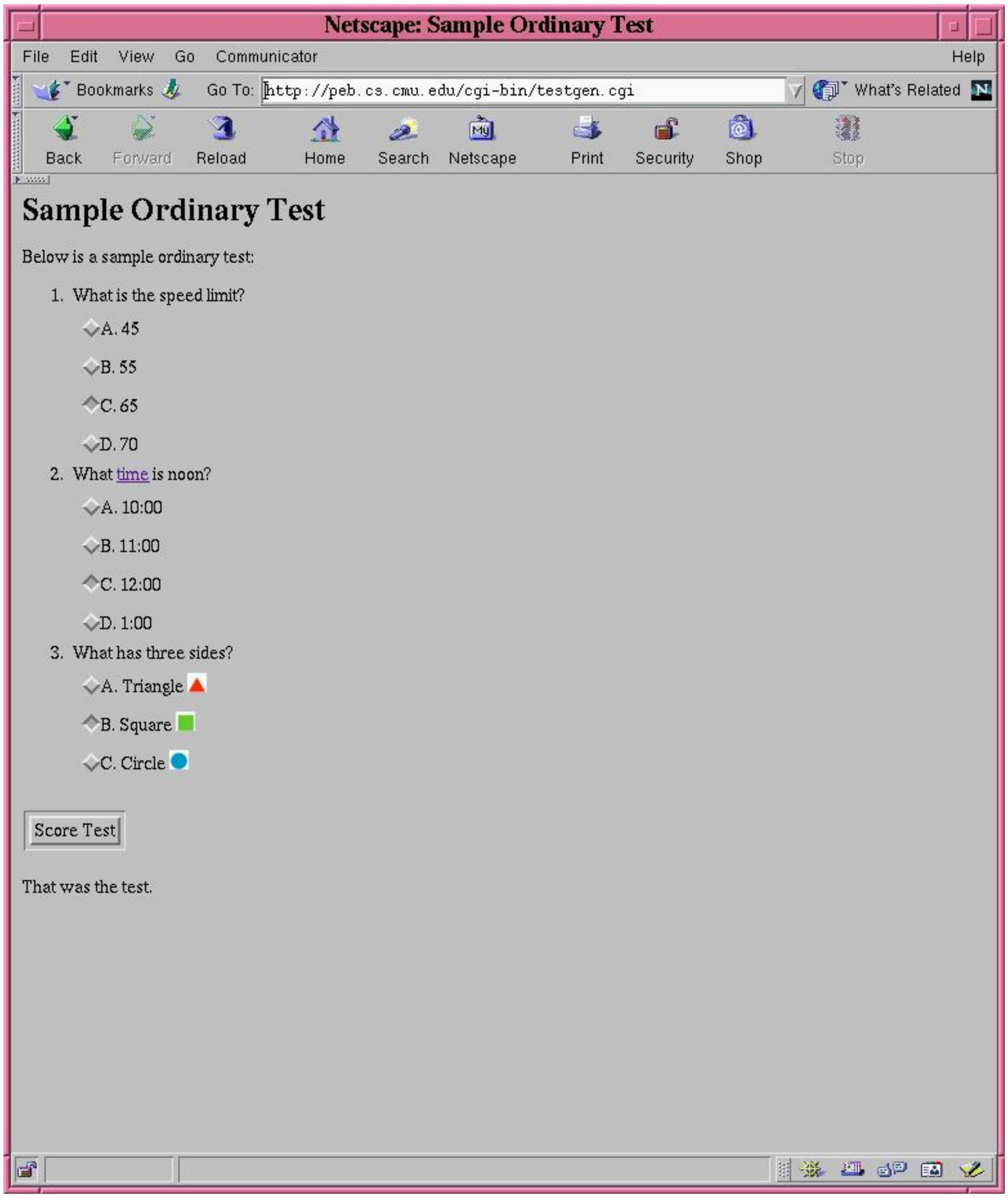

Figure 9: Page for an ordinary test, after a successful login

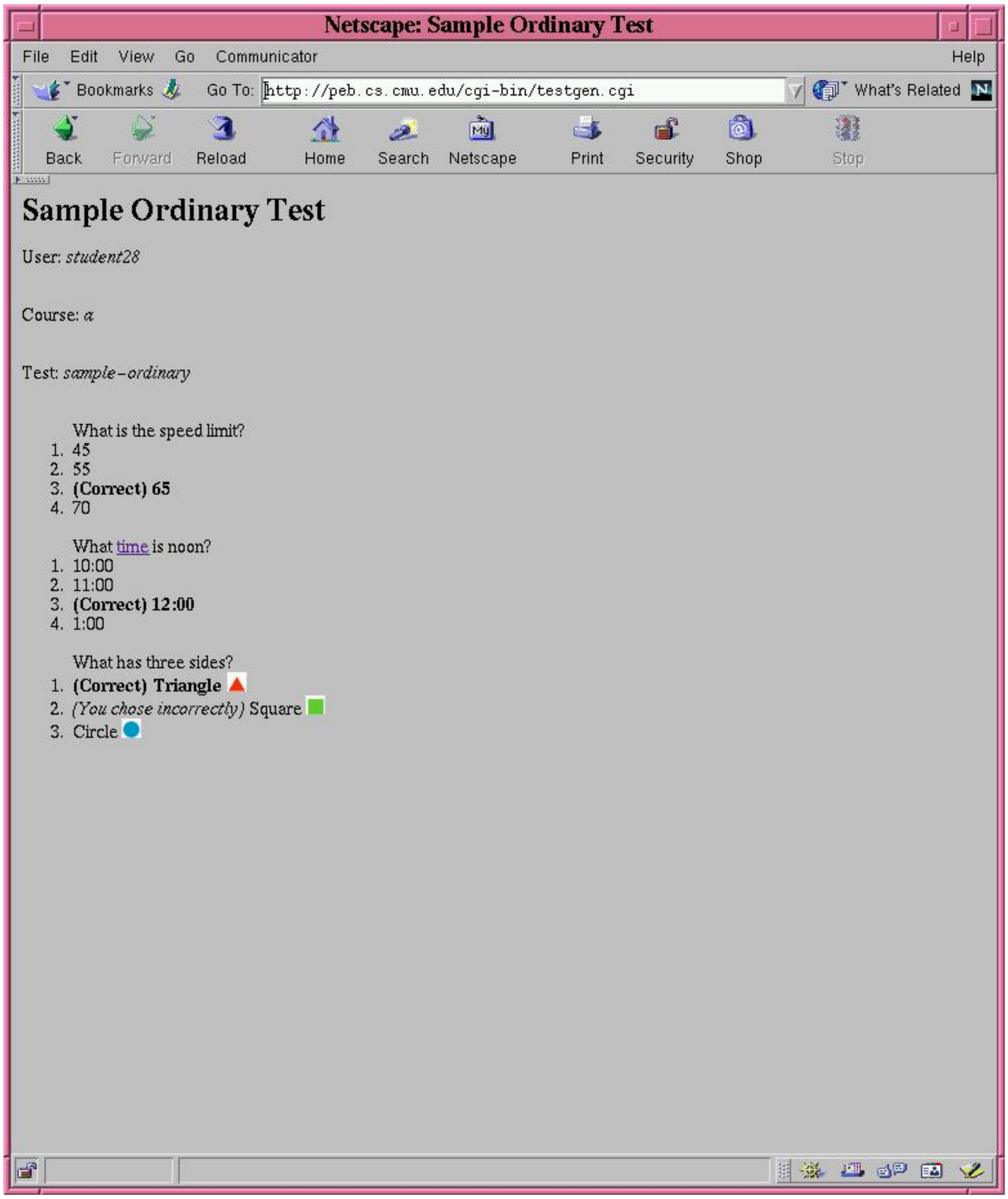

Figure 10: Page for an ordinary test, after a response has been submitted

To upload a new test, the instructor goes to the Web page for uploading tests, and specifies a local HTML file to upload (see Figure 11). The software extracts the XML portion (encoded as HTML annotations) and parses it with a validating parser, before saving out a binary representation of the complete HTML test (which may include links, images, etc.), and then indicate that uploading was successful (Figure 12).

Note: with Prof. Yaron's class, test submission was not used, because of the simplified nature of the "generic" concept test, which does not have question or answer text.

6.3 Monitor test

The "monitor test" user interface consists of the following components:

- instructions to the student,
- a bar graph of current student responses, indicating how many responses there were for each choice, and
- buttons and text fields for the instructor to control what is displayed and the score database.

The page for starting the monitoring of an "ordinary test" is in Figure 15. The page for starting the monitoring of a "concept test" is in Figure 13. The page showing the bar chart used for monitoring the "generic concept test" is in Figure 14); note that this page is arrived at directly from a link on the main adminstrative page (Figure 3).

Instructions to the student were included on this page as a useful reminder of the URL that the student needs to load in order to take a test, and a brief summary of how to submit responses. The instructions were put in after the first trial because some students came in late, or didn't hear the instructor describe the procedure, or were unfamiliar with what was expected.

When "auto refresh" is turned on, the Web page is updated every 3 seconds, in order for the students and instructor to see an updated graph of the latest student responses. There is a button to turn on or off auto refresh, because some students said it was distracting to see the page refresh at the beginning when either no or few students had yet submitted responses. Originally, the default behavior was to have a test begin with auto refresh, but this was changed so that the instructor could set up a test to begin without having to immediately turn off auto refresh.

There is a "start/stop" button to allow the instructor to either start a test, or once it is in progress, stop it, meaning that when the test is stopped, no more submissions are accepted from students, and refresh is turned off. This was to prevent a problem that seemed to be occurring in the first trials of the software: some students would continue submitting responses after Prof. Yaron had already explained the answer, and some students would begin submitting responses for the next question before he had reset the test, resulting in those answers being lost.

For the generic test, there is a text field for setting the number of answers for the next test, and also a text field for specifying the name of the score info to be saved out upon resetting the score database to begin a new test, an operation performed by the "save and reset" button.

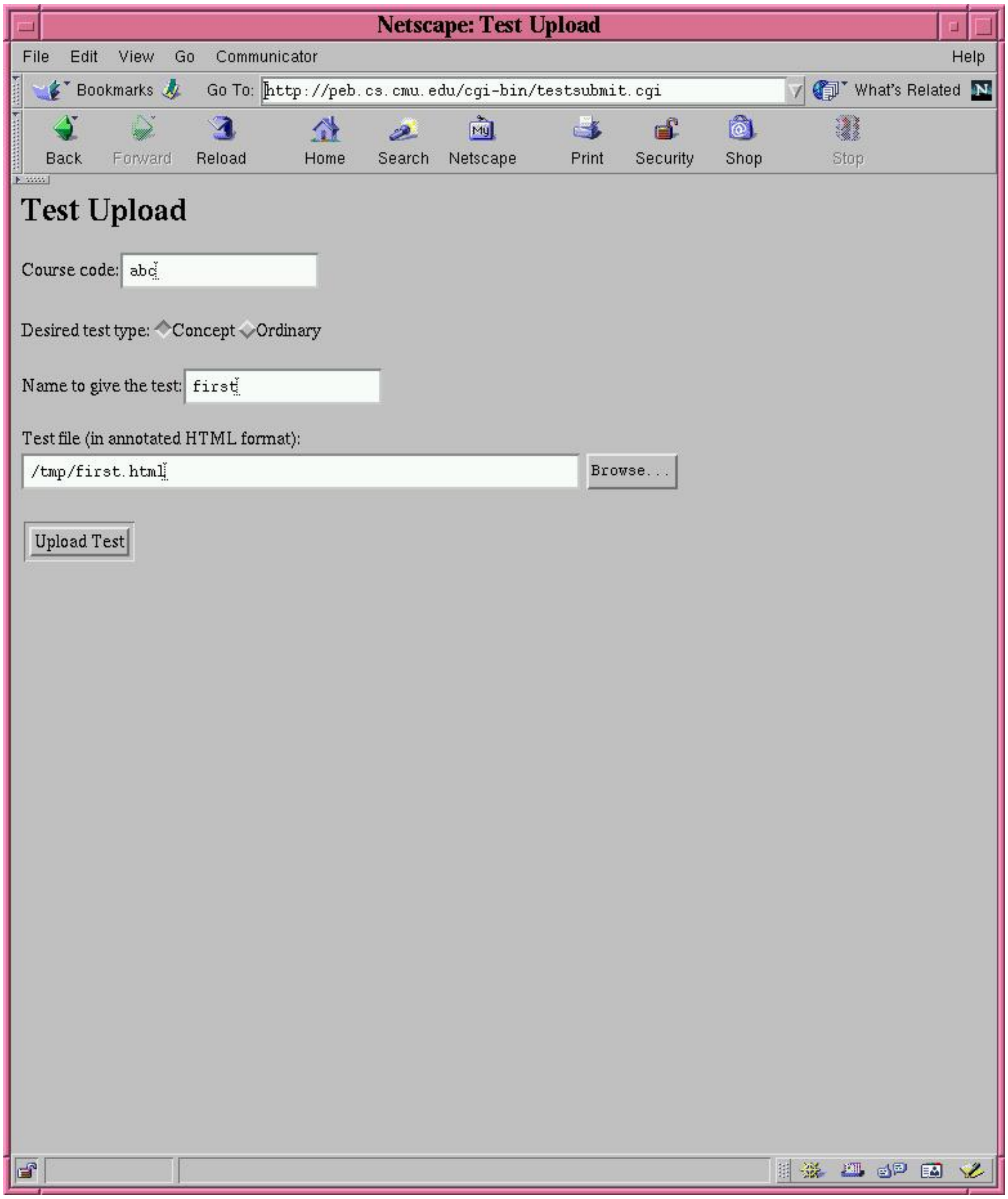

Figure 11: Page for instructor to upload a test

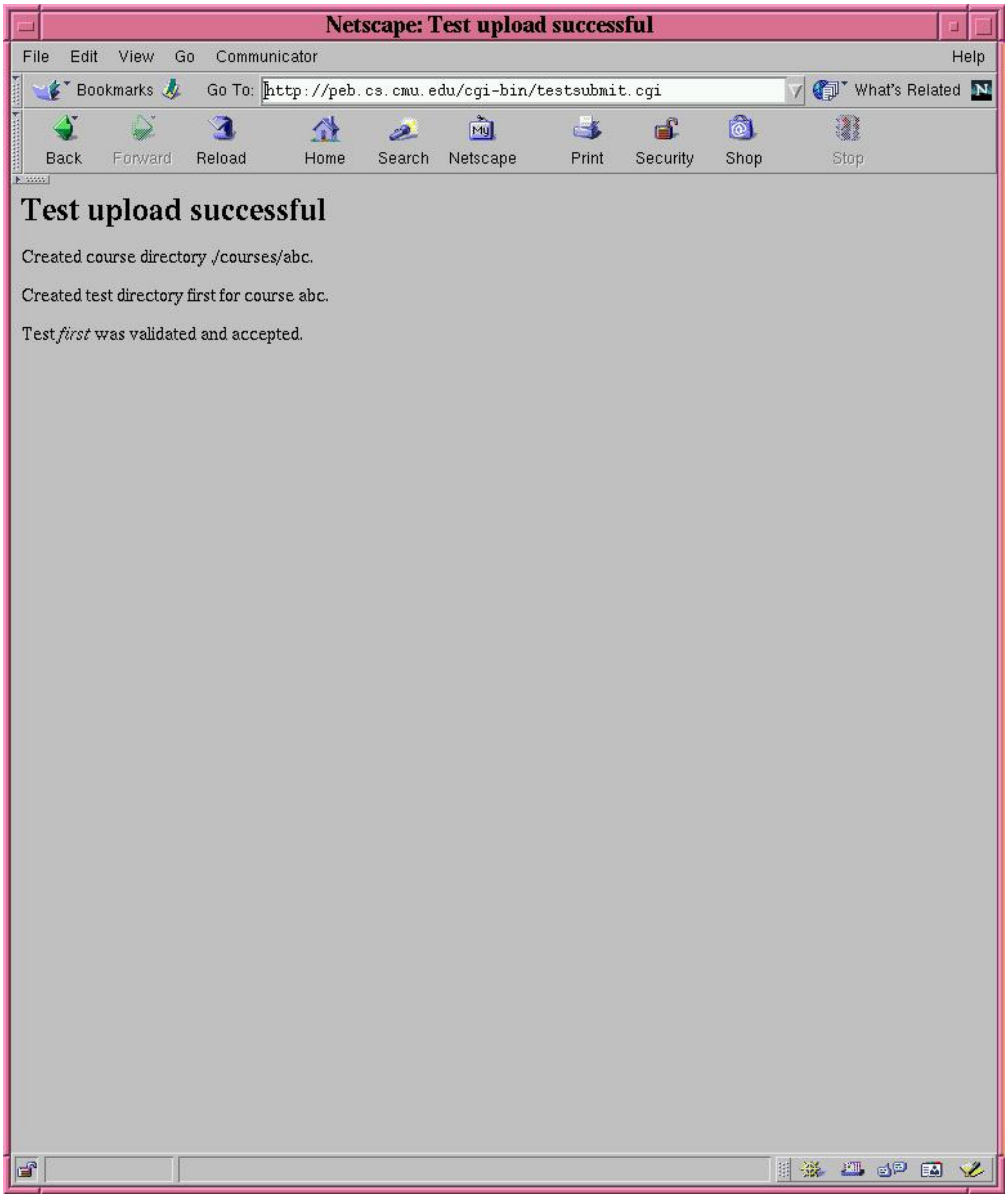

Figure 12: Page after instructor has successfully uploaded the test

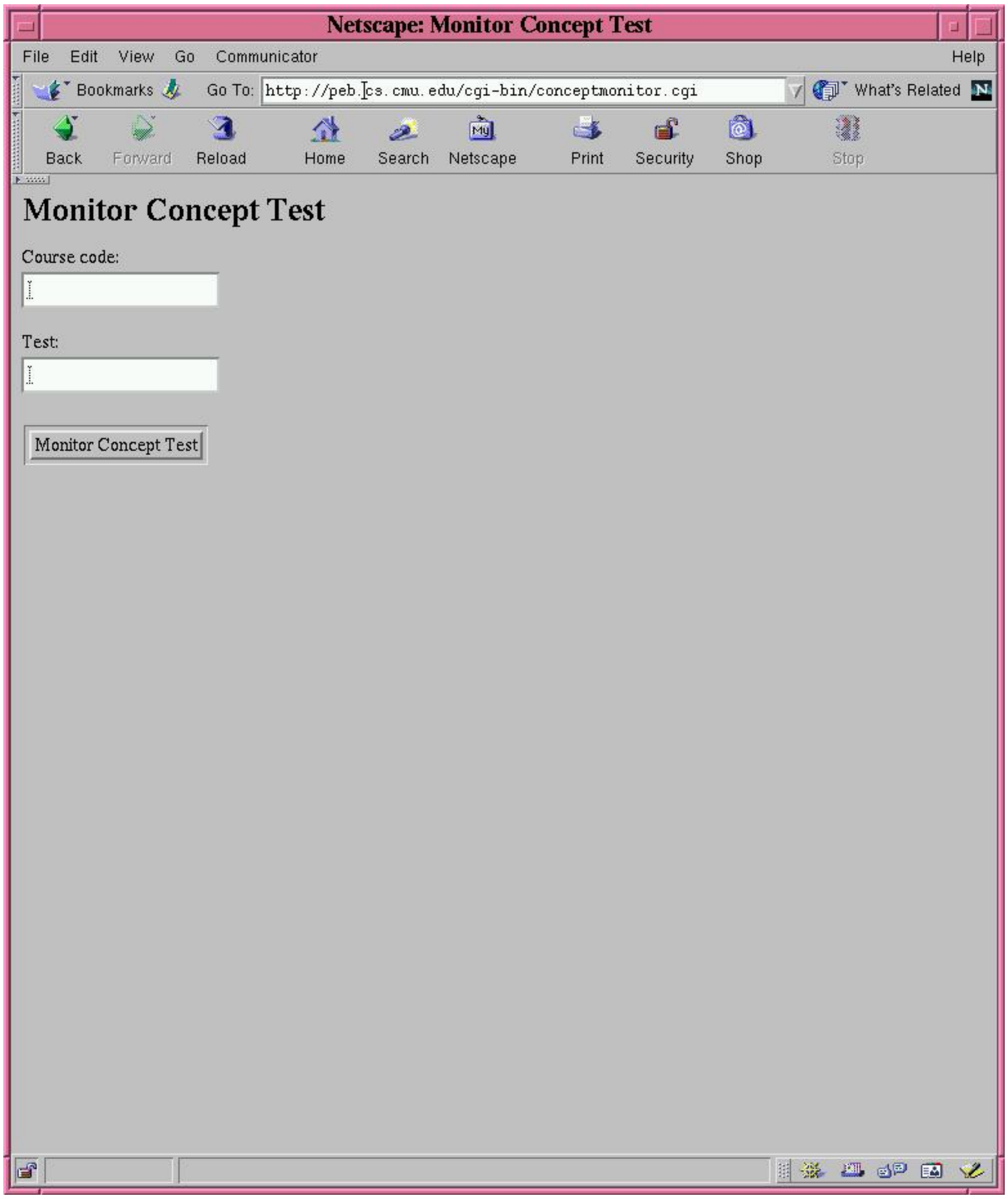

Figure 13: Page for instructor to specify concept test to monitor

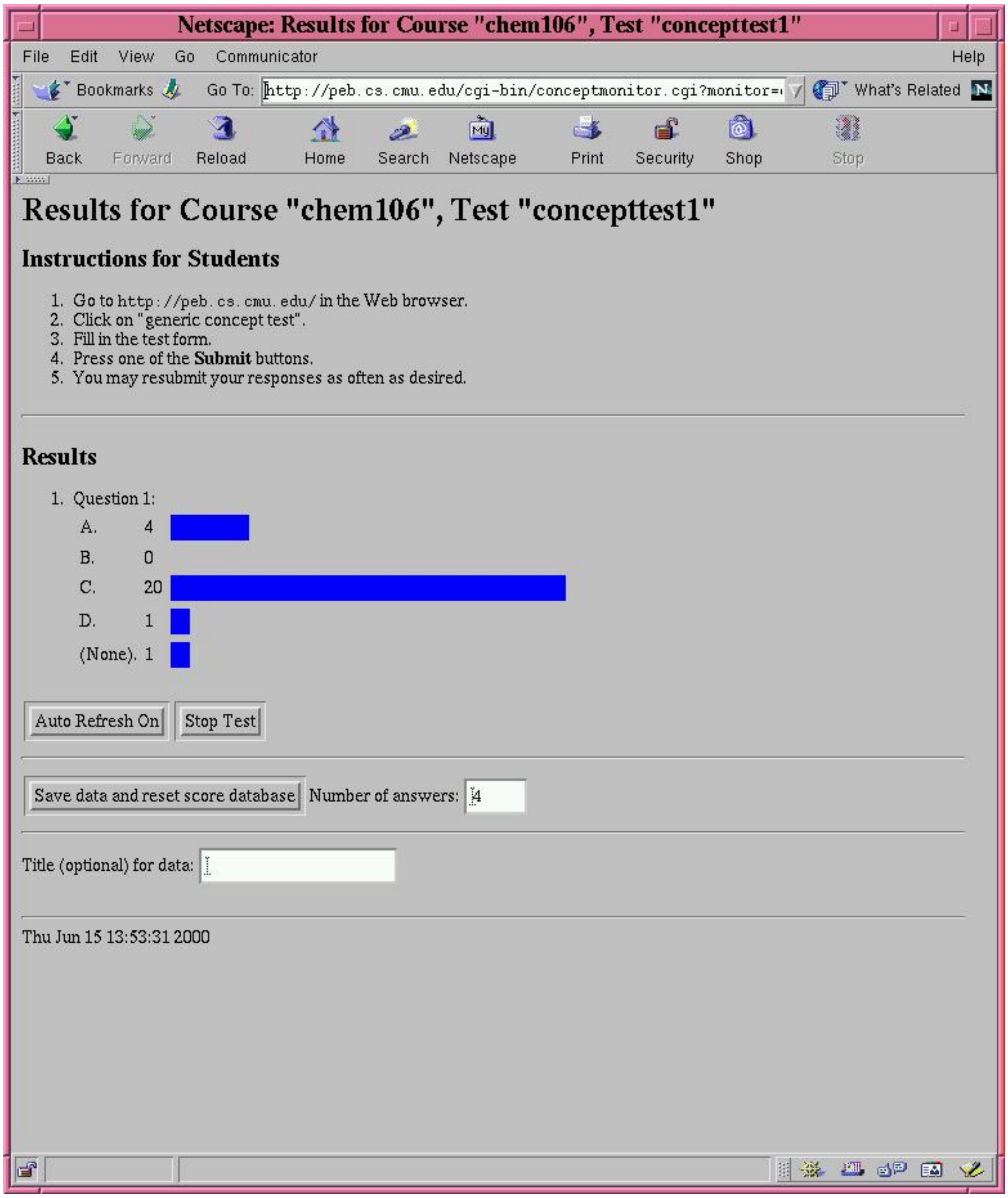

Figure 14: Page monitoring generic concept test

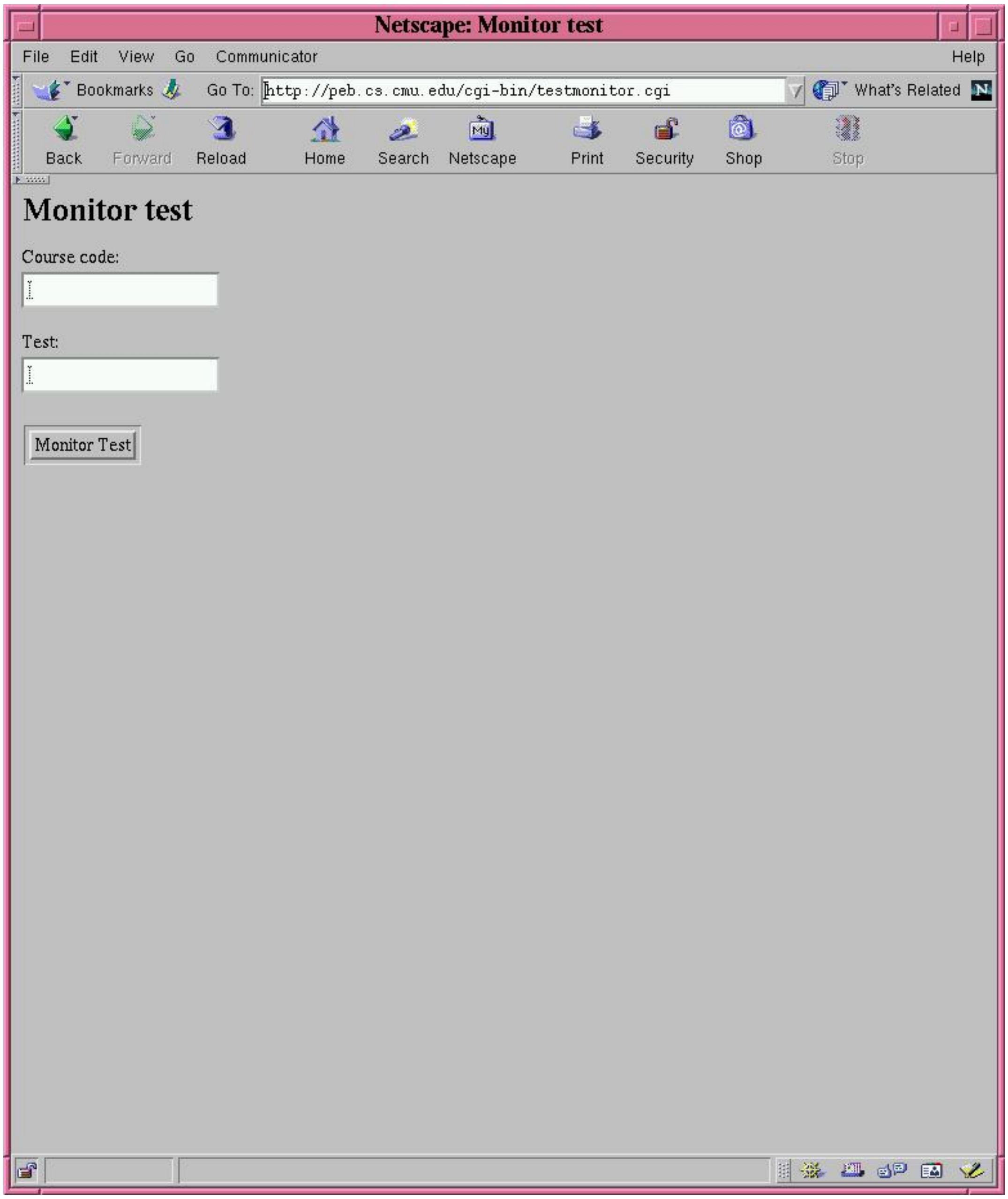

Figure 15: Page for instructor to specify ordinary test to monitor

The name of the score data file saved out consists of the current date and time followed by the name specified by the instructor (so that if the text field is left blank, by default the full name will simply be the date and time).

The saved data is accessible by means of a separate Web page (Figure 16) that pops up a list of score files and prints the scores in a table suitable for post-processing, and also recreates, for easy visual summary, the (anonymous) score histogram for the particular concept test in question.

For example, Figure 17 shows a portion of the data for a single test. In the end, most students had come up with a final answer of D. The complete history of each student's responses is saved, so we can see from the visible portion of this Web page that some students switched to the 4th answer, D, from another answer (and one switched from A to D).

7 Logged results

The test databases show that 58 students registered their handhelds.

Figure 18 shows how many students responded to a given number of tests.

It is encouraging that 14 out of 58 students responded to all the tests that were given. The second largest bin occurred at 5 tests out of 11. It would be hard to obtain this kind of data about the distribution of student participation without the use of handhelds, because the most casual assessment of class participation if human memory and hands or flashcards were used, would involve noting for each test how many hands or flashcards go up, rather than nothing for each hand whether it goes up for a given test.

8 Student reactions

Prof. Yaron at the end of the course gave out a survey for students to fill out, which had questions involving the use of handhelds in the course as well as outside it. There were 49 responses to the survey, a number less than the 58 who registered their handhelds.

39 students indicated that they reponded to concept tests "frequently (i.e. whenever possible)". The data collected suggests that this is not quite true. Unfortunately, the next choice on the questionnaire was "infrequently", so those who answered half the time probably had to choose "frequently". If the question had provided more bins, e.g., "50% of the time", "75% of the time", then it might have been interesting to gauge students' own perceptions or memories of their participation against reality.

9 Prof. Yaron's Interpretations

Prof. Yaron's interpretations of the survey results and assessment of the use of handhelds overall:

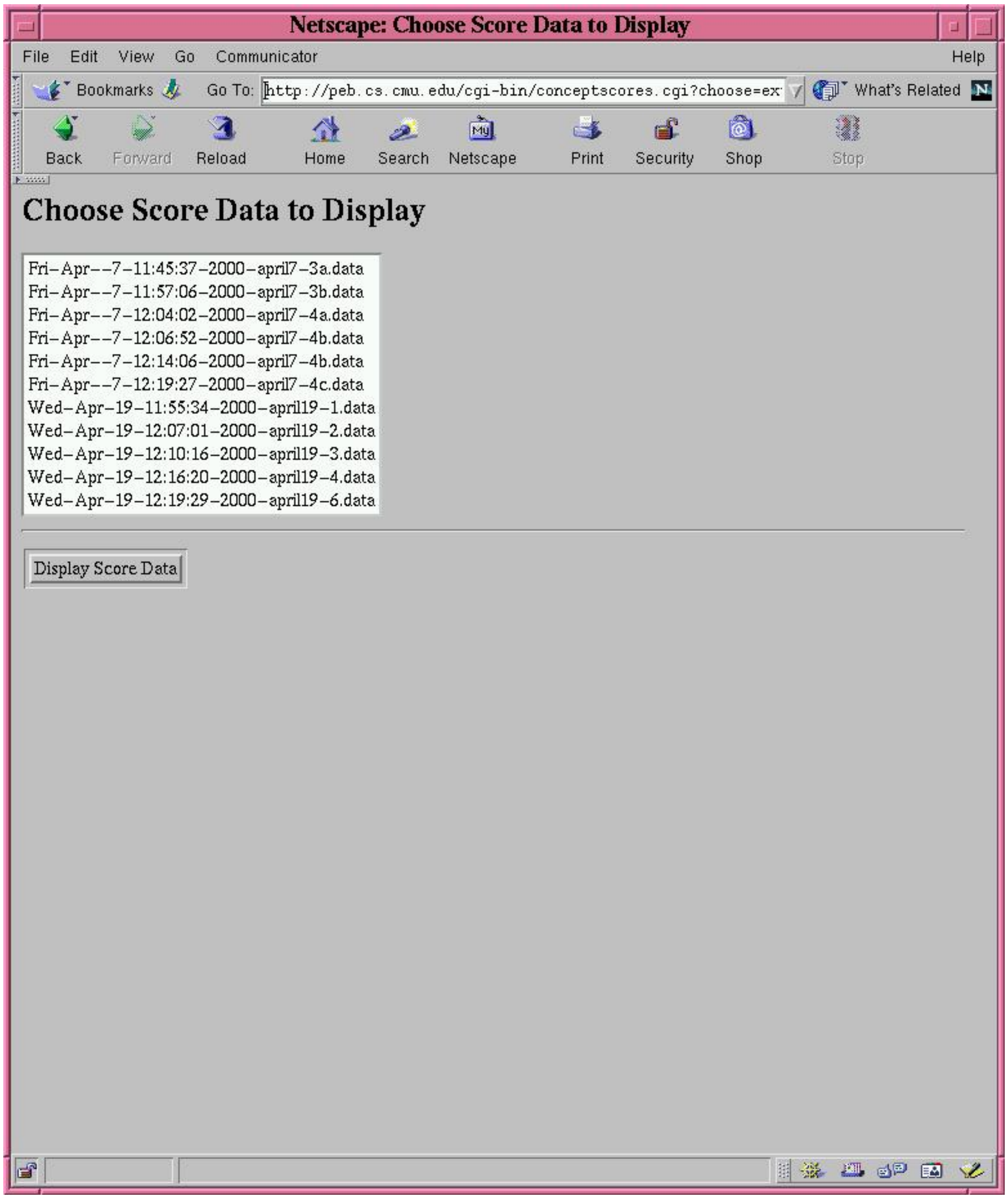

Figure 16: Web page listing saved generic concept test data to display.

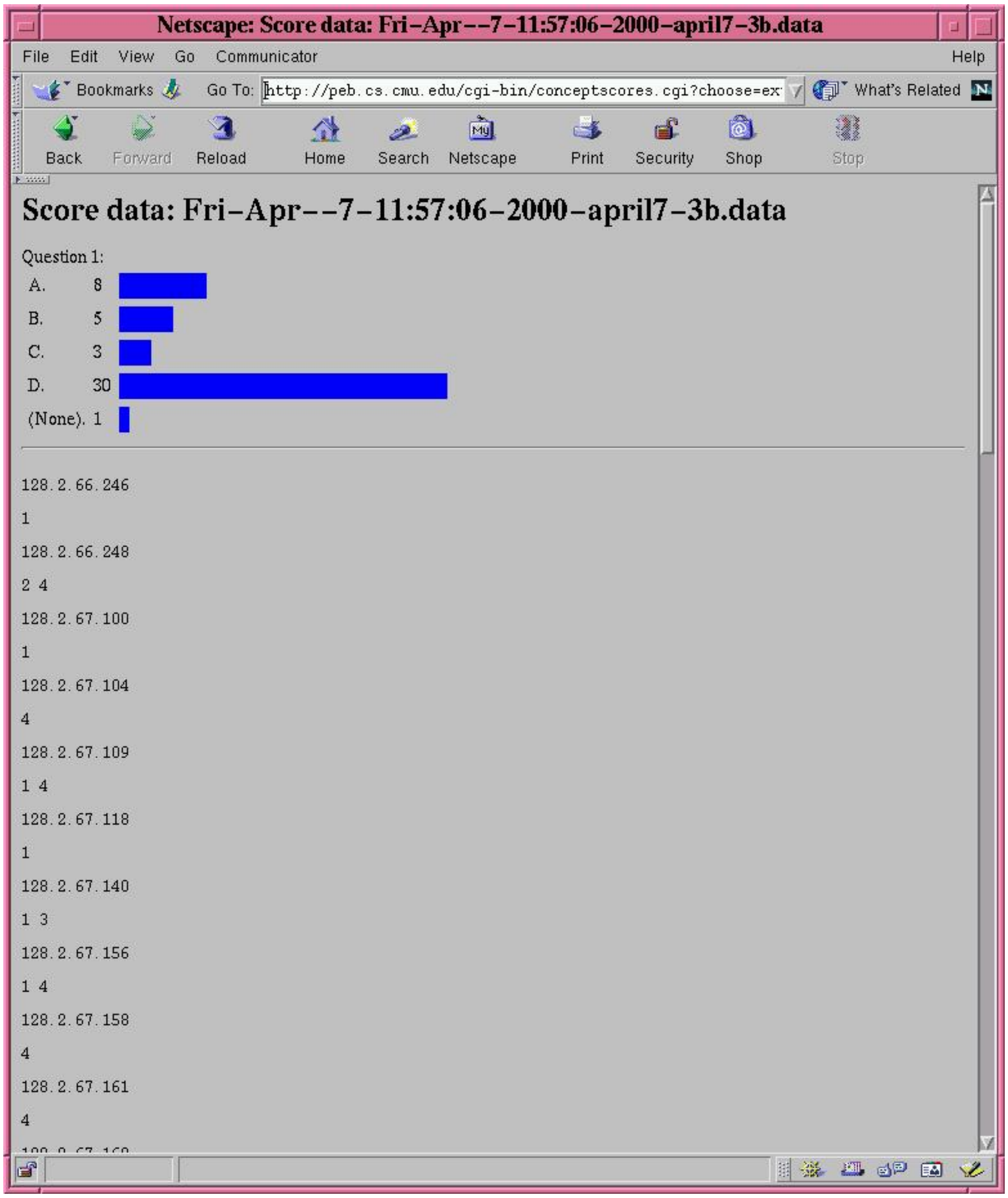

Figure 17: Example of display of saved concept test data.

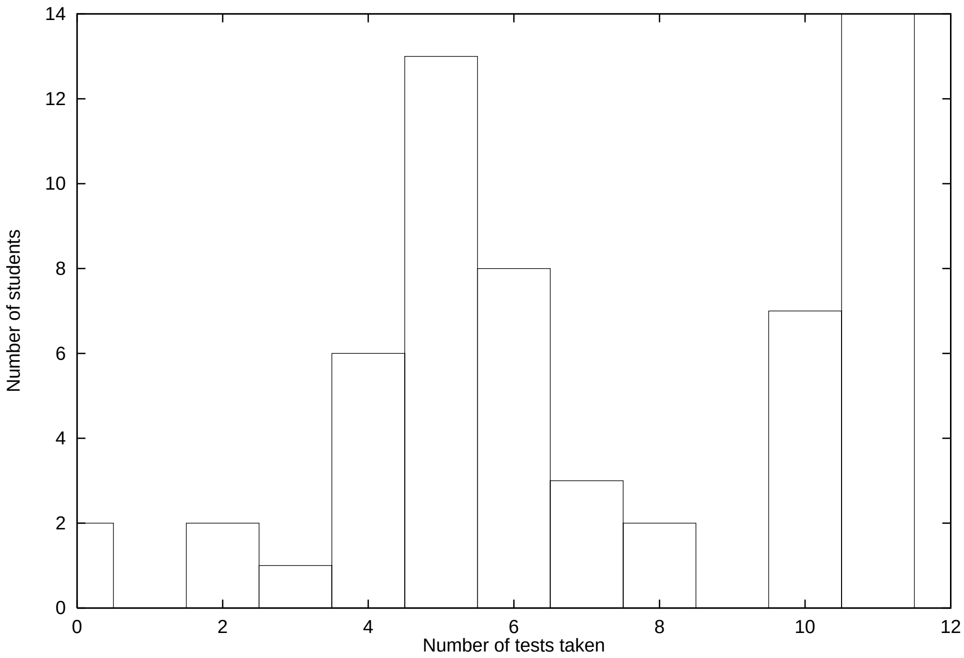

Figure 18: Number of students (total 58) taking a given number of tests (total 11).

9.1 What we did

We used the handhelds for concept tests (short multiple choice questions posed in lectures). We tried three different approaches:

- cards with letters A B C D that students hold up
- show of hands (how many people think A etc.)
- handhelds

The handheld software was done by Brad Myers and Franklin Chen, and allowed students to select a choice, and allowed me to show the results as a bar chart on the computer projector. It is all web based. The student page is at http://peb.cs.cmu.edu/ (click on take the generic concept test) and the faculty version is at http://peb.cs.cmu.edu/admin. The students register their own handhelds (identified by ip #) and, after each questions, I can save a file with all student names and responses.

We have concept tests almost every lecture, but I only used the handhelds on days where we would be doing alot of them. This was about 6 or 7 times in the semester. The last couple of times, we did these on days with chemistry demos. Demos are really popular with the students (things exploding, colors etc), and we had student answer questions predicting what would happen. I think this worked pretty well. For an example see Leture Notes EE on `http://ir.chem.cmu.edu/chem106/notes/`.

On many of the days, participation with the handhelds was about 25 or 30 students out of 60-70 attending the lecture. One day, we gave credit in the form of a guaranteed 8/10 on the next quiz if they participated (correct answers not required). To handle people who forgot their handheld or had trouble, we also provided a sign-up sheet and those people could hold up their hands. 53 people used the handhelds and got credit that way. If we were going to do grading routinely, we would have to do a better job of telling students when to bring their handhelds, and making sure all students had working handhelds and could connect to the network (there were usually about 5 people with serious problems at any given time).

9.2 Evaluation

I've attached a word file with the questionnaire we did on Wednesday of the last week of classes (Appendix D). Attendance was a bit low that week, so we have 50 responses (of a class of around 100). I think the survey gives a pretty detailed view of what students did with the machines, and what they thought. The file is exactly what we gave to the students, but with the responses summarized in blue. We haven't yet had time to do more than just count responses.

The results of the question on concept tests surprised me a bit. From my perspective, both cards and the handhelds have the advantage of better timing: I can wait until the answers start coming in. How long it takes them to vote lets me know how confused they are etc. With hands, it is much more difficult to know when to start the "how many people think A... " process (I have to just watch their faces and guess). Cards are easier to manage, so that is probably my first choice. But only about 4 students agree with me. 17 students prefer show of hands and 28 prefer the handhelds.

My reading is that the students were neutral to positive on the experience. I suspect this would be quite different if they had to buy the devices.

9.3 Future

I think the use of handhelds for concept tests is somewhat useful, but not worth too much effort. I'll do it again if it is easy and convenient, but won't go out of my way for it. I see much more potential for the handhelds in laboratories, if students can use them for data collection and analysis. Also, I'd see alot more uses if we could run our virtual lab and other Java software on them.

10 Implementation

10.1 Tests

Each test is, upon being uploaded as annotated HTML, stored both as the original HTML and as a binary format through Perl (to avoid having to reparse the original source each time). A validation step is currently plugged in, by means of calling an external validating parser. There exist XML validators on the Web and in free source or binary form; RXP[5] was chosen simply because it is freely available in source form. After validation, the Perl library *XML::Parser* (which is not validating, hence the need for an external validator) is used to construct a Perl tree representation of the file.

The Perl library `Storable` is used to store the Perl object for the parsed test file.

Each test has a test score database containing both "permanent" and "session" information. The kind of information stored depends on what kind of test the test is. We considered using `MySQL` for use as a database, but used Berkeley `dbm` supplied with Linux instead, for initial simplicity. In the end, we did not switch to `MySQL` after all, because it turned out that `MySQL` was known and criticized for its lack of support for efficient concurrent read and write access, so it did not seem there would be an advantage in using it[6].

Each test score database is saved out using Perl's libraries `MLDBM` (for `dbm`), `DB_File`, and `Storable`. It was later discovered only when moving the software to a Solaris SPARCstation, unfortunately, that the resulting database file is not portable across platforms, because of byte order issues. The test score database contains the test type (concept or ordinary), an ID, to distinguish it from other tests, a flag to indicate whether the test is "stopped" (i.e. will accept no more responses from students), the number of answers, and a table of score information for each student. In the case of a concept test, a complete history of responses of each student is maintained, in order

[5] `http://www.cogsci.ed.ac.uk/~richard/rxp.html`

[6] `http://openacs.org/why-not-mysql.html`

for possible interesting analysis of when and how students changed their minds; in the case of an ordinary test, just the single response is kept.

10.2 User interface

CGI using Perl (module `CGI`) to generate HTML is used to handle both the student's and instructor's interfaces, i.e., forms with radio buttons, text entry fields, a submit button, in a Web browser window.

Each major task is implemented in a single CGI script, e.g., monitoring a concept test, monitoring an ordinary test, displaying concept test data, taking a concept test,

Session state information was maintained through hidden fields. This is insecure, as someone could create phony data modifying various fields, but we were not emphasizing security issues at this time. A better way to handle state information would be to save state information on the server, and only expose a "session key" to the client.

Each of the programs uses a certain way to handle state transitions, e.g., when a "submit" button is pressed or a "stop test" button is pressed. Each GUI element is named, and the state of each is carried over after each form submission using hidden fields. Each program is reinvoked by the server upon a submission, so at the beginning of each program, tests of the values of parameters corresponding to GUI elements are made to determine what form configuration has just been exited and what has been entered. Data is loaded from on-disk databases on the server as appropriate, then processed in conjunction with the user-submitted data.

In order to avoid confusion from out of date forms, a `pragma no-cache` is generated in the HTTP header of each CGI-returned Web page to prevent undesired caching of old Web pages that would otherwise be reachable, for example, through using the browser's "Back" button.

11 Transferring of software

The software was all developed on a Linux PC, running CMU facilitized Red Hat 4.2. Rob Miller transferred the software to `basalt.amulet.cs.cmu.edu`, a Solaris machine. This involved determining which files were needed, checking them into CVS, and installing the required 3rd-party software on `basalt`.

The following are the required software:

- Apache Web server;
- Perl 5.00503 (already installed standard by CMU facilities);
- Perl modules:
  - `XML::Parser`,
  - `MLDBM`,
  - `Storable`,
  - `CGI` (newer version than that which came with the version of Perl installed)

- `rxp` XML parser and validator.

Course directory access permissions also needed to be set up, and the protect Web access for the `admin` page for instructors.

12 Future work

One possibility for improving the concept test experience for students is to allow for a voting of whether students are confused about a question. Several times Prof. Yaron asked the class for a show of hands on who was confused: this information might be useful to display and record online, as feedback for the lecture and for the future.

One idea that was brought up by Prof. Yaron was to provide an “explorer” type of user interface for creating quizzes from some large database of questions, i.e., there would be a “folder” for each quiz, and questions could be dragged and dropped into such folders. This is particularly useful when there is a large body of question material and one wishes to easily reuse them to make different tests from them without having to physically copy and embed question text into each desired test. And if the database of questions kept track of which questions were relevant to which topics, creation of tests would be made more convenient.

Security concerns were hardly addressed during this initial implementation. They would be particularly important if online tests became used for significant grading.

If more and more features are added to the software, it will probably be wise to use a “real” database as a repository of all course information, rather than use a combination of different file formats and file system hierarchies.

13 Conclusions

The students and the instructor in the experimental use of handhelds in the classroom appeared to be comfortable with the technology, and the administration of tests and submission of responses of tests went smoothly, after the initial phase of adjustment to a new process.

Performance of the Web server based application appeared to be sufficient; at the given size of the class and style of use, we did not run into performance bottlenecks with the Web server or the wireless network.

Students appeared to favor the use of handhelds for concept tests to the older methods of using a show of hands or flashcards.

14 Acknowledgements

For help with this research, we would like to thank Rob Miller, Doug Blair, Garry Warnock, and the students in 09-106.

References

[1] MAZUR, E. *Peer Instruction: A User's Manual.* Prentice Hall, 1997.

A Test format

The test format is below. Currently, only multiple-choice questions are implemented. It would be simple to add fill-in-the-blank support as well: XML makes it easy to change grammars and updating validation and parsing when doing so.

An optional "prologue" and "epilogue" of arbitrary text have been provided to be inserted into the HTML output if desired.

```
<!-- test.dtd -->

<!ELEMENT test ( title, prologue?, item+, epilogue? ) >

<!ELEMENT concepttest ( title, prologue?, item+, epilogue? ) >

<!ELEMENT title ( #PCDATA ) >

<!ELEMENT item ( question, answer+ ) >

<!ELEMENT question ( #PCDATA ) >

<!ELEMENT answer ( #PCDATA ) >

<!ATTLIST answer correct ( true | false ) "false" >

<!ELEMENT prologue ( #PCDATA ) >

<!ELEMENT epilogue ( #PCDATA ) >
```

B Sample test

This is a sample test in HTML form, illustrating the ability to use arbitrary HTML.

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML//EN">
<html>
  <head>
    <title>Sample Test</title>
  </head>

  <body>
    <h1>Sample Test</h1>

    Below is a sample test:

    <p>
    {test}
      {item}
        {question}
          What is the <em>minimum</em> speed limit?
        {/question}<p>

        {answer correct="true"}
          45
        {/answer}<p>

        {answer}
          55
        {/answer}<p>

        {answer}
          65
        {/answer}<p>

        {answer}
          70
        {/answer}<p>
      {/item}<p>

      {item}
        {question}
          What <a href="http://www.worldtime.org">time</a>
          is noon?
        {/question}<p>

        {answer}
```

```
          10:00
        {/answer}<p>

        {answer}
          11:00
        {/answer}<p>

        {answer correct="true"}
          12:00
        {/answer}<p>

        {answer}
          1:00
        {/answer}<p>
      {/item}<p>

      {item}
        {question}
          What has four sides?
        {/question}<p>

        {answer}
          Triangle <img alt="Triangle"
                src="http://peb.cs.cmu.edu/icons/TRIANGLE.GIF">
        {/answer}<p>

        {answer correct="true"}
          Square <img alt="Square"
                  src="http://peb.cs.cmu.edu/icons/SQUARE.GIF">
        {/answer}<p>

        {answer}
          Circle <img alt="Circle"
                  src="http://peb.cs.cmu.edu/icons/CIRCLE.GIF">
        {/answer}<p>
      {/item}<p>
    {/test}<p>

  That was the test.
  </body>
</html>
```

See Figure 7 for a rendition of the test in a browser.

C Chem 106 Concept Test

The initial 4-question Chem 106 concept test: note empty question and answer text.

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML//EN">
<html>
  <head>
    <title>Chem 106 Concept Test</title>
  </head>

  <body>
    {test}
  {item}
    {question}
    {/question}

    {answer}
    {/answer}

    {answer}
    {/answer}

    {answer}
    {/answer}

    {answer}
    {/answer}

    {answer}
    {/answer}
  {/item}

    {/test}
  </body>
</html>
```

D Chem 106 Survey and Results

Starting on the next page is a tabulation of the student responses to the end-of-course survey in Chem 106. There was a total of 49 student responses.

Questionnaire on Handheld Computers

Did you use your handheld computer to:

Read email:	*never*	*infrequently*	*frequently*
	13	15	21
Browse the web:	*never*	*infrequently*	*frequently*
	5	22	22
Respond to concept tests:	*never*	*infrequently*	*frequently (i.e. whenever possible)*
	4	9	39
Maintain contact list (phone # etc.):	*never*	*infrequently*	*frequently*
	41	7	1
Maintain calendar/schedule:	*never*	*infrequently*	*frequently*
	39	12	1
Use Pocket Word:	*never*	*infrequently*	*frequently*
	30	14	4
Use Pocket Excel:	*never*	*infrequently*	*frequently*
	43	5	2
Play games:	*never*	*infrequently*	*frequently*
	7	23	20

Please list any other uses you have found for your handheld:

Please state your level of agreement or disagreement with the following statements:

I find the concept tests to be a useful part of the lecture

strongly disagree	*disagree*	*agree*	*strongly agree*
	2	30	17

I usually participate in the concept tests

strongly disagree	*disagree*	*agree*	*strongly agree*
1	2	22	24

I liked using the handhelds for concept tests because the rest of the class could not see my response

strongly disagree	*disagree*	*agree*	*strongly agree*
4	*13*	*23*	*8*

I thought the use of cards was “not cool”, and so prefer either raising hands or using the handhelds

strongly disagree	*disagree*	*agree*	*strongly agree*
4	21	18	6

I thought the use of handhelds for concept tests took too much time

strongly disagree	*disagree*	*agree*	*strongly agree*
5 37	21	19	4

I thought the presence of handhelds, such that students could use email/browse the web, in the classroom was distracting

strongly disagree	*disagree*	*agree*	*strongly agree*
4	19	19	7

Please rank the following methods for concepts tests, based on your experiences this semester (1 –best 3- worst)

	____ *show of hands*	____ *cards*	____ *handhelds*
(first choice)	17	4	28

Please rank the following methods for concepts tests, assuming you did not have any network or technical difficutites with the handhelds (1 –best 3- worst)

	____ *show of hands*	____ *cards*	____ *handhelds*
(first choice)	13	2	33

Did you regularly sync your handheld computer with a desktop computer (using active sync)? *yes no*
9 34

How many times during the semester were you unable to connect to the network with your handheld?
___8__ *it always connected*
it failed _13__ *10% of the time* __13____ *25% of the time* _6_ *50% of the time* _3_ *90% of the time*
__2___ *it never connected*

How many times during the semester did you need to hit the reset button on your handheld?
____2__ *never* __12____ *1-2 times* ___8__ *3-5 times* ___22__ *more than 5 times*

What would you consider the biggest drawback of your handheld:
6____ size (i.e. it is too large)
14____ stability (it crashes too often)
19____ lack of power (memory etc.)
12____ other (please specify ____)
5____ no major drawbacks

I found the keyboard:
22_____ Adequate.
16_____ Too small. I would prefer a bigger device with a regular size keyboard.
7_____ I would prefer a smaller device with no keyboard.

I found the screen size:
34_____ Adequate.
12_____ Too small. I would prefer a bigger device with a bigger screen.
1_____ I would prefer a smaller device with a smaller screen.

A color screen (instead of black and white) is:

Very important	*Somewhat important*	*Not important*
28	13	4

How was the battery life?
23____ *Adequate.*
18____ *Too short.*
3____ *Didn't use device enough to tell*

Overall, I have found my handheld computer to be:
8___ *More useful than I anticipated*
18___ *About as useful as I anticipated*
18___ *Less useful than I anticipated*

How much would you be willing to pay to keep your Jornada? $_________
<50 9 50 5 100 11 150 1 200 6 250 1 300 2 400 4 500 3 600 2

Check one of the following:
19_____ *I believe that handheld computers will eventually be used widely in courses of this type*
17_____ *I believe that handheld computers are useful outside of class, but not in class*
8_____ *I don't see any real use for handheld computers in universities*

Should we use handhelds in the course next year? 27_____ *yes* 17 _____ *no*

What advice would you give to an instructor who wanted to use handhelds in their course?