DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.033—Computer System Engineering Handout #9 —March 9, 1995

Design Project #1: CLOSURES IN THE WORLD-WIDE WEB

Introduction. One of the features of 6.033 is a series of case studies discussed in recitation. Some of these case studies are special in that you are asked to go into them in more depth and to do a design project based on them. The first of these design projects relates to naming and closure.

This design project is based on a naming problem from the real world, one that the 6.033 staff actually bumped into when the Student Information Processing (SIPB) first installed a Web link to the 6.033 home page in their list of subjects with Web pages. To gain insight into the problem, the first thing you should do is encounter the problem. We have arranged such an encounter by setting up a World-Wide Web area that causes the problem to appear. After you have explored the Web area and encountered the symptom of the naming failure, your job is to come up with a design proposal that makes things work better.

Before the encounter, you might find it helpful to read section A of Saltzer's Chapter 5 (reading #23, on naming) even though it won't be assigned as regular reading until Wednesday of next week, and also the paper by Berners-Lee, et al. (reading #25), giving an overview of the Web. That way you will have a better idea of what to look for during the encounter, and the vocabulary to discuss it.

The encounter. Using your favorite Web browser, go to the 6.033 home page and follow the button labeled "6.033 Design Project 1 ENCOUNTER". On the page that appears you will find five different Web links back to the 6.033 home page, two labeled "Good", and three labeled "Bad". If you follow one labeled "Good", everything should work normally. They are there for comparison with what goes wrong when you follow the "Bad" Web links . Try any of the Web links labeled "Bad". At first,that Web link will appear to work normally, too.

From the 6.033 home page that you reach via a "Bad" Web link, try exploring the various Web links found there. The Staff, Student, and FAQ pages should appear normal, and the Web links leading out of those pages should also work normally. But the Handout page has a problem. Try following a Web link to any of the specific handouts and note what error you encounter.

Now go back to one of the Web links labeled "Good" and repeat the experiment. This time you can reach any of the handouts without encountering the error.

You might think that we have set up two distinct home pages with some subtle difference between them, but we haven't. You are actually looking at the same HTML file in each case, but each Web link on the encounter page is using a different name for the home page. If you are skeptical, you should go into the 6.033 locker ("`Attach 6.033; cd /mit/6.033/www`") and see for yourself.

To get to the bottom of the problem, here are some suggestions:

- Use a text editor or displayer to examine the HTML form of various pages in the 6.033 locker, including the handout page, which contains the Web links that sometimes don't work right. (This examination can be done either by logging in on Athena and looking in the 6.033 locker, or by invoking the browser feature that retrieves the source HTML form.)
- Make sure that you have your Web browser set up so that it shows you the URL of the page it is displaying and also the URL of the page it is about to reach for. Compare the

URL's your browser reports with the syntax of the Web links you found in the original HTML form.

- While logged in on Athena and attached to the 6.033 locker, examine the contents of the "mystery" directory to see what additional clues it can provide.
- Compare the URL's of the various Web links on the Encounter page to see what the Good links have in common and what the Bad links have in common.
- Try to construct in your own Athena locker a collection of directories, subdirectories, UNIX links, and a Web page that contains Web links that work only sometimes in this same way. To create a UNIX symbolic (soft) link on Athena use the `-s` flag of the `ln` command. (For more details see `man ln`). Once you have a working (or better yet, failing) example in your hands, you can experiment with various things that might shed light on appropriate fixes.
- Check the discussion of the BASE element, below, and see if you can get it to help make naming work right in the page you constructed in your locker. What are its limitations?

What is going wrong when you get the error? What naming system has failed? Can you explain what happened in terms of the naming concepts introduced in 6.033?

The design project. Propose a redesign of something (you get to decide what) that gets around this problem, while at the same time introducing a minimum of new problems. You do *NOT* have to implement the redesign.

Before jumping to the design stage, you should do some more reading:

- Section C of reading #23, for a discussion of naming in file systems and naming networks.
- Tanenbaum's book, *Modern Operating Systems*, pp. 287-288, to learn how absolute path names, relative path names, and symbolic links work in the UNIX file system.
- The discussion of relative addressing in Mark Andreessen's HTML primer. (You can find links to this and the next item, as well as other possibly relevant documentation about the Web, on the Encounter page.)
- The description of the BASE element in the HTML specification, to see if that offers any help in solving this problem. The BASE element is an HTML construction that looks like this:

  <BASE HREF = "some URL or local directory name">

  A BASE element may appear in any HTML page.

Make sure you don't get confused by the distinction between World-Wide Web links and UNIX links. Despite sharing the label "link" and having some common features, they also have some differences. When you use the word "link" in your writeup or in conversations with other students or the teaching staff, make sure you precede it with one of the adjectives, "UNIX" or "Web", or the quality of the conversation will deteriorate pretty quickly.

Your solution should have the following four (presumably) desirable properties:

1. A (possibly redesigned) Web browser starting with any of the URL's found in the "Good" and "Bad" Web links should not run into a naming error.

2. It is currently possible to use all of the 6.033 Web pages, including their cross-references, even if the official MIT server, Web.mit.edu, is down, by constructing a URL for the home page that invokes the alternate server. (For example, examine and try the second Good link on the Encounter page.) It would be nice if that option could continue to work.

3. Users who are logged into the Athena system can currently use the 6.033 Web pages even if both servers are down, by loading the home page directly as a file. It would be nice if that option also could continue to work.

4. It is currently possible to rename the 6.033 locker or move it to a different place in the Athena file system hierarchy without having to revise the contents of any of the 6.033 Web pages. (We would have to advertise the new URL of the home page, of course.) It would be nice to preserve that flexibility.

Your report. Your paper should be 5 to 10 pages in length, explaining carefully the background of the problem and exactly what is going wrong. Give an example, perhaps based on the failing Web links that you constructed. Next, explore alternative solutions and describe their trade–offs by explaining how well they achieve the four desirable properties, and any other properties that you notice are worth providing. (For example, if one of the solutions would be perfect except that the cost would be infinite, you should point that out.) Finally, recommend one of the alternatives, and explain why you have chosen it.

Write for an audience consisting of colleagues who took 6.033 five years ago, when naming was discussed with the help of reading #23, but before the World-Wide Web came into existence. That is, they understand naming, contexts, and closures, but they are not familiar with the particular way that naming is done in the Web and they haven't encountered the particular problem. Assume that your paper will also be used to convince whoever is responsible for changing the system to use your preferred alternative.

6.033 writing considerations: When reading your report, your instructor will be looking for things such as the following:

- Is the report easily comprehended?
  - Is it well organized?
  - Does it explain the approach or architecture before delving into details or components?
  - Does it use diagrams where appropriate? (A frequent problem when people use word processors is that they try to express everything in words, either because the word processor doesn't make it easy to include diagrams, or they haven't ever learned how to use those features. Pictures can communicate some ideas far better.)
  - Does it use the concepts, models, and terminology introduced in 6.033? If not, does it have a good reason for using a different universe of discourse?
  - Does it address the intended audience?
- Does the report provide rationales for your design decisions? Does it describe alternatives you considered and how you made choices?
- Does the design work? Does it address the issues raised in the assignment?

Special writing lecture: Note that at 2 p.m. on Friday, March 17, one of the teaching staff of the M.I.T. Writing Program will give a lecture in 34-101, the usual lecture room, on how to write a report on a design project for 6.033.

Phase Two writing considerations: If you are enrolled in the 6.033 writing practicum, you don't need to do anything special; your practicum instructor will explain how the report will get you credit for the Phase II writing requirement. If you are not enrolled in the practicum, AND you want us to forward your design project report to the writing program as your phase II writing project, please say so on the first page, and make sure that your report is at least 8 pages long. Note also that the writing program has a rule that they accept only reports that earn a B or better from the class in which they originate. Finally, be aware that the second case study will probably be a team project, and thus much more difficult to tailor to the needs of the writing program than this one.

Schedule: Your report is due in recitation ***Thursday, March 23, 1995.***