

LBSC 690 Midterm Exam**Fall 2005**

Name: Answers

You have 120 minutes to complete this exam. Time will begin as soon as you start reading the first question.

- You may use any material, including the text book, lecture slides, and notes. You may also use anything found on the Internet that existed before the exam started.
- You may NOT communicate with any other person during this exam, either in person or using electronic means.

As strategies for completing the exam, keep the following in mind:

- If you find a question to be ambiguous, make assumptions as you see fit, but write down your assumptions.
- You are more likely to get partial credit for a wrong answer if you show your work.
- Be careful not to get carried away and run over the time limit by spending too much time on one question. Plan ahead, and don't devote more time to a question than it is worth.

Please write your answers in the space provided.

Score Summary (for use by grader)

Question	Possible points	Actual points
1	20	20
2	15	15
3	10	10
4	10	10
5	20	20
6	10	10
TOTAL	85	85

1. Short answers (no more than a couple of sentences).

I just downloaded a document from

<http://www.umiacs.umd.edu/~jimmylin/LBSC690-2005-Fall/>

a. [5 points] How did my computer know what IP address should be used for that URL?

The browser contacts a series of domain name servers (DNS), which converts the URL into an IP address. Each of these servers covers part of the URL hierarchy, from “top level”—.edu, .com, etc.—down to the most specific, until the URL is converted into the correct IP address.

b. [5 points] About how long would you expect it to take the Web server to simply read the file from the hard drive? (Only the disk speed is being asked about here. Disk speeds vary, but your answer to this question should be within a factor of two of the actual number.)

A typical access (“seek”) time for a hard drive is 10 ms. Total time to read the file is dominated by the seek time; once the first byte is accessed, it takes a negligible amount of time to read subsequent bytes (assuming that the Web page is relatively small).

c. [5 points] How did the router that the Web server is connected to know where to send the packets?

Each packet contains part of the data, as well as information about where it came from and where it’s going. The router uses a routing table to look up and determine where the packet should be sent next. The router doesn’t actually know the route to the final destination, but it knows what the next hop should be, depending on traffic conditions. Packets may be sent through different routes to the same destination.

d. [5 points] What is the role of HTTP in the process of delivering this content?

HTTP is the “language of the Web.” It is the protocol used to send the actual request to the Web server, and the protocol the server uses to respond to the request.

2. Short answers (please show your work so that we can give partial credit for wrong answers).

OIT at Maryland is planning a pilot project to digitize 1,000 hours of lectures next semester (roughly 20 courses). Each lecture will be recorded using RealVideo encoded at a rate of 30MB per hour.

a. [5 points] How much hard drive space (in GB) is needed store all the digitized video?

If each hour of lecture requires 30 MB of storage, 1,000 hours will require 30,000 MB of storage, or (30,000 x 1,048,576 bytes per MB) 31,457,280,000 bytes of storage. Divide that by 1,073,741,824 bytes in a GB to get 29.29 GB of space required.

b. [5 points] OIT is planning to use DVD-R optical disks as a backup medium in case of hard drive failure. How many 4.7 GB DVDs will be required to back up all the video?

29.29 GB required divided by 4.7 GB per DVD-R is 6.23, for a total of 7 DVD-R disks required.

c. [5 points] How long will it take to download one hour of lecture over a 56 kilobits per second modem? (For this questions, assume the modem really runs continuously at 56 kilobits per second even though a real modem may not)

A 56 kilobit modem downloads 56,000 bits per second, or (dividing by 8 bits per byte), 7,000 bytes per second. We need to download 30 MB. 30MB x 1,048,576 bytes per MB is 31,457,280 bytes in each hour of lecture. Dividing that figure by our speed of 7,000 bytes per second, we get 4,493 seconds, or 1.24 hours, required.

3. Short answers (no more than a couple of sentences).

The following is a very simple XML file that encodes a business card:

```
<card type="simple">
  <name>John Doe</name>
  <title>CEO, Widget Inc.</title>
  <email>john.doe@widget.com</email>
  <phone>(202) 456-1414</phone>
</card>
```

a. [5 points] When I open the file up in Internet Explorer or FireFox, I just see the XML in a collapsible tree. Why isn't the browser able to display the information in a prettier format, e.g., put the name in bold and create a link for the email?

XML defines the structure of the data; it does not specify how the data should be displayed. When the browser sees an XML file, it doesn't know what to do with it. XHTML, on the other hand, is an XML dialect that browsers actually know how to render (as Web pages).

Dr. Oard also uses XML, but he stores the same information in a different format:

```
<contact>
  <name>John Doe</name>
  <info>
```

```
<title>CEO, Widget Inc.</title>
<email>john.doe@widget.com</email>
<tel>(202) 456-1414</tel>
</info>
</contact>
```

b. [5 points] Name one technology that can convert information stored in my format into the format that Dr. Oard uses. Briefly, how does it work?

XSLT is a technology that can convert information stored in one XML format into another XML format. In this case, it does so by mapping XML tags in one schema to tags of the same meaning in another schema. For example, it would map <phone> to <tel>, because they both have the same semantic meaning.

4. I have the following files and folders on my computer.

- A folder called “690stuff”
- A folder called “2005” inside “690stuff”
- A folder called “2004” inside “690stuff”
- The HTML file “index.html” is inside the folder “690stuff”
- The HTML file “2005.html” is inside the folder “2005”
- The HTML file “2004.html” is inside the folder “2004”

a. [5 points] I want to create links in “index.html” that points to “2004.html” and “2005.html”. What relative path expression should I use? Write two path expressions below:

2004/2004.html
2005/2005.html

b. [5 points] I want to create a link in “2005.html” that points to “2004.html”. What relative path expression should I use? Write one path expression below:

../2004/2004.html
“..” means “go up one directory”

5. [20 points] Longer answer (no longer than a few paragraphs). I own a winery in Northern Virginia. Name two ways that XML could benefit me (assuming I am able to convince others to adopt XML as well).

[Here is a sample answer that received full credit:]

There are two main ways XML coding could benefit my winery: through rationalizing my ordering of inputs I need for my business, and through

rationalizing the way distributors purchase wine from us. I need a regular supply of materials to keep my large winery running. We buy plant stocks, fertilizers, and pesticides frequently. If I could apply standard XML coding to my computer database records of purchases and use of materials, I could program software to automatically formulate resupply orders from suppliers that use XML coding as well. As soon as I've nearly used up a certain pesticide, for example, my computer could draw on our XML database to create a purchase order and send it automatically to a computer (with an XML database) at our supplier's company. This could happen with little or no intervention by my staff.

The second way XML could help is by standardizing the way outside distributors purchase wine from us. We could use XML to encode a database of all the varieties and years of wine that we have available for sale. A purchasing format could be created that would draw on the XML-coded data. Our distributors could access this database when creating purchase orders for our wine, and we could draw on this database for filling and billing orders, and for a continuously updated inventory. XML would greatly decrease the amount of human intervention required for this process.

6. [10 points] Longer answer (no longer than a paragraph). Identify one well-established usability principle that the CLIS Web site violates to a substantial degree, and suggest how that aspect of the site could be improved.

[Here are sample answers that received full credit:]

One usability principle that the CLIS Web site violates is that of alignment. Alignment has been shown to help the viewer visually connect the elements on a page, and it creates a visual flow, often a vertical flow from top to bottom, down a list of main topics. On the CLIS home page, the headings that appear in red ARE aligned vertically in two columns, but the subheadings under each red heading are not. Many lines in the subheadings contain two separate links, but the only separation factor is a vertical line and a bit of white space. Alignment and usability would be much improved if each sublink simply appeared flush left, one to a line, like this:

Master of Library Science
Master of Information Management
Doctor of Philosophy
WISE
Non-Degree
Professional Development.

A similar problem of confused alignment may be seen on the page titled Courses and Schedules. Here you have more of an attempt at alignment of subheadings, but there are two different schemes at work. In Part I Class Schedules, under the LBSC

and INFM headings, the semesters are all nicely indented, bulleted and aligned, but towards the bottom of the page, under Schedule Archives, the bullets do not conform to either of the two indents above. This lack of conformity of vertical alignment of indented material confuses the eye and should be revised.

The CLIS web site has a visual affordance problem, which is part of an overall inconsistency in link style. Along the top of the screen is the logo for the University of Maryland and the name of the College of Information Studies. Just above the school's name are what appear to be links to the various programs within the school: library science, archival studies, etc. As an archives student looking to quickly pull up information particular to my course of study, I have frequently been fooled into trying to click on these headings, which are not links at all. Since most of the links on the page are not a particular color or style, I often think these are links too. After all, it would make sense for them to be – the University of Maryland logo right next door leads to university's main page. The navigation tabs down the left side, which are a similar size, font, and style, are actual links. But these are not. The inconsistent links all over the page are a “repetition” issue that is part of the CARP principles of user-friendly graphical design.

The CLIS website violates the principle of repetition, mainly in the way navigation is laid out. On the front page, we see links to CLIS information in three separate formats: visually appealing “buttons” at the top, then bolded white links, then unbolded links. This makes the navigation bar look sloppy, especially because the links are not aligned with each other. This could be improved by presenting all navigation links in the same format—preferably that at the top, used for “Information For:” links. Separation between topics could be achieved either by changing the background color in link buttons (as has been done on other parts of the website) or by maintaining the spatial separation between them.