

LBSC 690: Information Technology Final Exam

William Webber
CIS, University of Maryland

Spring semester, 2012

Exam Duration: Two hours.

Length: This paper has 6 pages including this cover page.

Authorised Materials: Examinees may use any online or offline resource they find themselves, including lecture notes, but may not exchange materials online during the conduct of the exam.

Submission Format: Write all of your answers into a text (or Word, or similar) file and email to wew@umd.edu at the conclusion of the exam. If you use a Word or similar non-text format, please save as PDF before sending. You may write programming files as separate files; please attach them to your email. *It is your responsibility to ensure all files are correctly sent.*

Clarifications: Examinees may seek clarifications on exam questions from the instructor. Please raise your hand and wait for the instructor to come to you. Please ask the question quietly; the instructor will determine whether the question and its answer should be shared with the examinees as a group.

Academic Dishonesty: Sharing materials on- or offline during the conduct of the exam constitutes academic dishonesty.

Advice to Examinees: Do not get bogged down in a single question; allow time to attempt every question. A partial answer is better than no answer.

Question 1:

Answer *four* of the following six questions. If you answer more than four, only the first four will be examined.

a). *If you enter the search “click here” into Google, you get the Adobe Acrobat PDF Reader download page as the first result. Why?*

Many web sites have links to the Acrobat Reader download page with the anchor text (text between the tags) of “click here”. Search engines use anchor text of links to a page to help determine what the page is about, so they think that the Acrobat Reader download page is about “click here”.

[10 marks]

b). *There is much talk these days of “digital convergence”, the trend for all varieties of media, and the activities that rest upon them, to converge on the one platform (the, often hand-held, “digital computer”). What is it about digital representation that allows this convergence (compared, say, to analogue representation)?*

Every form of media has the same digital representation, as a number (or series of numbers), so they can all be processed by the one machine, a digital computer. In contrast, different media have different and incompatible analogue representations (e.g. vinyl for sound, film for pictures).

[10 marks]

c). *What are the three tiers of a three-tier webapp architecture? Assign each of the programming languages PHP, SQL, and Javascript to the tier it is most closely associated with.*

The three tiers are the data tier, the logic tier, and the presentation tier. SQL is most commonly used in the data tier; PHP in the logic tier; and Javascript in the presentation tier.

[10 marks]

d). *For each of these file types, say whether lossy compression can be usefully applied:*

- *A photograph*
- *A music file*
- *A text email*
- *Programming source code*

Briefly explain why.

Lossy compression can be used with photographs and music, since these are essentially continuous media, and the human eye (ear, brain) can compensate for, or even not detect, some loss in fidelity. Text and source code cannot be lossily compressed, since they are essentially discrete, and any change corrupts their meaning.

[10 marks]

continued . . .

e). A common rule for image files is that if they start in a vector-graphics format (e.g. SVG), you should keep them in a vector-graphics format for as long as possible before converting to a raster format. Why?

Vector formats can be arbitrarily scaled (flipped, stretched etc.) without loss of fidelity. If a raster image is scaled, it loses fidelity, e.g. becomes pixelated.
[10 marks]

f). A common term weighting scheme in text search is $tf * idf$. What do tf and idf stand for? What is the intuition behind idf ?

“TF” stands for “term frequency”; “IDF” stands for “inverse document frequency”. The intuition behind “IDF” is that the fewer documents a term appears in, the more discriminative it is.
[10 marks]

[40 marks in total]

Question 2:

Answer both of the following questions.

a). Your library is setting up a new website. It will have five main sections, with multiple sub-sections and pages. Different librarians will write content for different sections, and new content and sub-sections will be added over time. A graphic designer will be hired to produce a design for the site. The library director proposes to have people write each page in HTML using notepad (to save money on software licensing). Is it a good idea to build such a site in static HTML? Why or why not? What other technologies could be used? Is the director right to be worried about software licensing costs? Will any special hosting support be required?

Having everyone write their own static HTML file will make it difficult to maintain a consistent look and feel across the site; the navigation and templates will have to be repeated on every page; as new sections are added, every navigation will have to change, making the site difficult to maintain; and all authors will have to understand HTML. A better solution would be to install a content management system (CMS), such as Drupal. There are many open-source CMS systems, so there will be no software licensing cost. However, it will be necessary for the web server to support server-side programming, and provide a database.
[20 marks]

continued . . .

b). A colleague mentions that they've heard on the radio that the internet is "running out of addresses". She wonders if this is because the big internet companies have bought all of the domain names. Is this why the internet is running out of addresses, and if not, then what is? How many public addresses (roughly) can the internet support, and what is the fundamental technical limitation that sets this limit? What steps are being taken to solve this problem?

The internet is not running out of domain names. Instead, it is running out of IP addresses, which are used as the "phone numbers" of domains. In IPv4, each IP address is represented as a 32-bit number, which means there can be no more than 4 billion IP addresses. In the new IPv6 standard, IP addresses are being extended to 128 bits, which will support a virtually unlimited number of IP addresses. [20 marks]

[40 marks in total]

Question 3:

Complete *one* of the following two questions:

continued . . .

a). Develop a web page in HTML with a stylesheet (internal or external) in CSS. The web page should do the following:

- Have the title “Exam-Related Causes of Panic”
- Place the title in the browser title bar
- Place the title in large letters at the top of the page
- Display the image at the following URL: `http://upload.wikimedia.org/wikipedia/en/0/07/Neuro_Panic_Disorder_Norepinephrine.png`
- Display the explanatory text “Norepinephrine pathways in the brain” for the image for the visually disabled
- List the following possible reasons for exam-related panic:
 - Forgetting the exam time
 - Forgetting the exam location
 - Forgetting your own name
- Set the page background color to red in CSS
- Set the title at the top of the page to italics in CSS

40

```

<html>
  <head>
    <title>Exam-Related Causes of Panic</title>
    <style type="text/css">
      body {background-color: red;}
      h1 {font-style: italic; }
    </style>
  </head>

  <body>
    <h1>Exam-Related Causes of Panic</h1>

    
      function set_name() {
        /* Code goes here */
      }
    </script>
  </head>

  <body onload="set_name()">
    <h2>Hello , my name is <span id="name"></span>!</h2>
  </body>
</html>
```

Write the content of the Javascript function, so that when the page loads, the following happens:

- The page asks the user "What is your name?"
- When the user types their name in, the page displays the user's name directly before the exclamation mark
- The user's name is displayed in italics

No other part of the page should be modified. (Hint: you don't have to use CSS to display the name in italics if you don't want to.)

```
function set_name() {
  var name = prompt("What is your name?");
  document.getElementById("name").innerHTML = "<i>" + name + "</i>";
}
```

[40 marks]

[40 marks in total]

[120 MARKS FOR EXAM]

Make sure you email your answers to wew@umd.edu.

end of exam