

Midterm 1 Review Sheet **INFO/CSE 100 Spring 2006**

This is a list of topics that we have covered prior to midterm exam 1. This is not all inclusive of every detail and there may be items on the exam that are not explicitly listed here, but these are the primary topics of interest.

Basic terminology *Hardware*. Parts of the computer like monitor, display, Cathode Ray Tube (CRT), Liquid Crystal Display (LCD), Red-Green-Blue (RGB), pixel, processor box (or case), motherboard (or system board), daughter board (or card), processor, microprocessor, memory, Random Access Memory (RAM), kilo (1 thousand), mega (1 million), giga (1 billion), sequential access vs random access, hard disk.

***Software*.** Operating system, program, algorithm, Graphical User Interface (GUI), command line, boot.

Analytical thinking Abstract, generalize, factor of improvement, percent improvement.

Software operation Basic metaphors like buttons, sliders, close boxes, menus, keyboard shortcuts, ellipsis (...) on a menu means “more input required”. Common program operations like file operations New, Open, Close, Save, Save As, print operations Page Setup, Print, Print Preview, program control operation Exit or Quit. Common editing operations like Undo, Repeat, Cut, Copy, Paste, Clear and Select All. An “instance” of a document. Feedback methods like input dialog boxes, message dialog boxes. The ability to think about what you want the software to do, and list the operations that might be provided to accomplish that. In other words, have a strategy for exploring new applications or new capabilities of a familiar application.

Networking Descriptive names like The Internet (the networking equipment and capabilities that connects a worldwide collection of computers), Local Area Network (LAN), Wide Area Network (WAN). The concept of asynchronous communication (sending and receiving can be separated in time by an arbitrary period). Network packets and routing. Protocols are the rules by which packets are routed and interpreted, Transmission Control Protocol / Internet Protocol (TCP/IP), File Transfer Protocol (FTP), Ethernet Protocol. Network addresses (or Internet Protocol addresses), dotted quad (four numeric values separated by “.”), domain names (strings separated by “.”), how to read a domain name (right to left moving down the hierarchy), translation from domain name to dotted quad by the Domain Name System (DNS) and the DNS servers that implement it.

World Wide Web All the machines that implement HyperText Transport Protocol (HTTP) server software and the client browsers that access them. Not all the machines on the Internet participate in the Web, the Web is just one of the ways that information is passed around. Elements of a Uniform Resource Locator (URL) are: protocol specifier (eg, http://), server computer name (specified as computer name plus domain name or dotted quad), and pathname to the file on the server. Know how to construct and read a pathname as we discussed in lab and in class. Client / server interaction to request and provide a particular page or service.

HyperText Markup Language (HTML) The language in which web pages are written. Plain text files with special structure defined by a limited set of tags. The file name extension is generally html or htm. The basic structure of an html page is

```
<!DOCTYPE definition ...>
<html>
  <head>
    <title>Title text</title>
  </head>

  <body>
    Body definition including text, images, tables, etc

  </body>
</html>
```

You should be able to read a straightforward html page and understand the purpose of the tags listed above as well as bold, <p> paragraph, <hr> horizontal rule,
 line break, <h1>, <h2>, <h3> header format, unordered list, ordered (numbered) list, list item, <table> table, <tr> table row, <td> table data, anchor reference (a link to another URL), image reference. Note that some tags have required or optional attributes. In particular make sure you understand how the anchor and img tags are constructed.

Understand how pathnames are constructed and how they relate to the directory structure on disk. The “root” of the file system is specified with a single “/”. Absolute pathnames start at the root and list each directory in the path to a particular subdirectory or file. Relative pathnames start at the current location (eg, the location of the web page containing the relative pathname) and list each directory in the path from there to the particular subdirectory or file. Know that “.” means the parent directory, and that “/” is used to separate directories and filenames in a pathname.

Understand the purpose of the Document Type Definition (DTD) and the use of the Validator service from the World Wide Web Consortium to check the structure of your

web page against the associated definition. I do not expect you to be able to read or comment on an actual DTD, just know that it exists and that it defines the syntax of the tags that you are using.

Using the Web and searching for information Understand the idea of hierarchies as the traditional method of organizing information. Hierarchies are very valuable when the information is well understood and can be categorized definitively. Individual web sites are often organized in a hierarchical fashion, starting with a home page, then providing links to pages within the site for various categories of information. The linked pages then provide links to further refinements and so on. Understand the terminology of a hierarchy including the root (often drawn at the top), the leaves (the last entry in a series of increasingly detailed entries), drilling down (going from the general to the specific), moving up the hierarchy (going from the specific to the general).

Search engines like Google and Yahoo give us another way to access large piles of information such as the unorganized information on the larger web, by providing a giant index that is referred to whenever you make a search request. The index is generated by “crawling” the web, which is done by software at the search company. Not all pages that are available on the web are indexed, not all pages in the index still exist. Effective use of a search engine entails thinking about specific words that characterize the information you are looking for, understanding the flavor of the results that you get, and modifying the search to home in on the desired source. Changing the search terms or more advanced combinations of terms and qualifiers are used to control the results.

Another important aspect of searching is making sure that the information you get is reliable. Anyone can publish a web page. The data that is presented has not been fact checked by anybody at all and may be a complete fabrication, a well meaning but incorrect statement, or the best answer possible. It’s up to you to evaluate what you are looking at. Starting with known information sources is a good way to work when you have specific requirements. The UW library system has a rich set of resources to draw on and skilled librarians to help you with your search. The library web site provides access to many of their information sources, but not all. Finally, remember that the information on the web is only a small fraction of what is available in the real world. Get up and go talk to people when you want to explore a subject in depth. The world of information is deep and wide, and exploring it is a lot of fun!

UW Computing Resources You should be familiar with the material in the labs related to dante, the primary student mainframe machine available through the UW. Recognize Pine, WebPine, SSH Secure Shell, SSH Secure File Transfer, pico and what they are used for. Know that the operating system on dante is a version of Unix. Commands on dante are entered using a command shell like bash. You should be able to read a series of simple commands entered on dante and describe their effects. You will not be expected to have

memorized the details of the arguments lists for every command, but you should know that `cd` changes the current working directory to the named directory, `ls` displays a list of files in the current or named directory, and `pwd` shows the present working directory. Dante is connected to the UW network and you can easily move files from desktop machines to dante for file backup or serving on the web. The file structure on dante is a hierarchy that you should be able to navigate. For example, given a drawing representing several file directories, you should be able to write down pathnames that will identify various files or directories within that structure.

The operating system on the desktop machines available to us is usually Windows XP or Mac OS X. You should know that they both have graphical user interfaces as well as command line interfaces that you can use to enter commands. Both of these operating systems have hierarchical file systems like the one on dante, although they differ in some of the details. There are text editor programs, word processing programs, and a large variety of other capabilities available on these systems. You will not be asked about specific commands of these programs, but there may be questions of a general nature similar to the information on the first page of this discussion about common program operations. You will not be asked specific questions about the operation of Photoshop.

Information Representation Digitization (or quantization): representing information by any fixed set of symbols. Using positional notation and a fixed set of symbols, any number of different states can be identified. Different encodings can be used to represent the same set of states. *Any* phenomenon that can be set and measured can be used to encode state information. One very common encoding is to use presence and absence of some physical characteristic to encode binary bits. The word *bit* is a contraction of “*binary digit*”. A byte is a set of eight bits. 256 different states can be encoded in 8 bits because $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^8 = 256$. One way that bytes are used is to encode characters. Each different value is interpreted as a different character code. Additional levels of encoding such as XML tags are used to describe structure or other higher-level (or meta-level) information about a data set.