CSC 1051 – Data Structures and Algorithms I

Dr. Mary-Angela Papalaskari
Department of Computing Sciences
Mendel 162C

Course website:
www.csc.villanova.edu/~map/1051/
What is this course about?

- Computer Science
- Problem solving
- Algorithmic thinking
- Data representation
- Software engineering
Our textbook

Java Software Solutions
Foundations of Program Design
Seventh Edition

John Lewis
William Loftus
Overview of today’s class

• Go over syllabus/course information
  – www.csc.villanova.edu/~map/1051
• Introduction to the course – reverse history of computing
• Take the online survey
Reverse History of computing

Dig deeper into what we already know, travel backwards…

1. What we see now all around us – a connected world of computing

2. Focus on a single “traditional” computer

3. Dig deeper – data and processing
Networks

• A *network* is two or more computers that are connected so that data and resources can be shared

• Most computers are connected to some kind of network

• Each computer has its own *network address*, which uniquely identifies it among the others

• A *file server* is a network computer dedicated to storing programs and data that are shared among network users
Network Connections

- Each computer in a network could be directly connected to every other computer in the network.
- These are called *point-to-point* connections.

Adding a computer requires a new communication line for each computer already in the network.

This technique is not practical for more than a few close machines.
Network Connections

• Most networks share a single communication line
• Adding a new computer to the network is relatively easy

Network traffic must take turns using the line, which introduces delays

Often information is broken down in parts, called packets, which are sent to the receiving machine and then reassembled
A Computer Network
A Local-Area Network (LAN) covers a small distance and a small number of computers.

A LAN often connects the machines in a single room or building.
Wide-Area Networks

A *Wide-Area Network* (WAN) connects two or more LANs, often over long distances.
The Internet

• The *Internet* is a WAN which spans the planet

• The word Internet comes from the term *internetworking*

• It started as a United States government project, sponsored by the Advanced Research Projects Agency (ARPA)
  
  – originally it was called the ARPANET

• The Internet grew quickly throughout the 1980s and 90s
TCP/IP

• A protocol is a set of rules that determine how things communicate with each other

• The software that manages Internet communication follows a suite of protocols called TCP/IP

• The *Internet Protocol* (IP) determines the format of the information as it is transferred

• The *Transmission Control Protocol* (TCP) dictates how messages are reassembled and handles lost information
IP and Internet Addresses

• Each computer on the Internet has a unique IP address, such as:
  
  204.192.116.2

• Most computers also have a unique Internet name, which also is referred to as an Internet address:
  
  hector.vt.edu
  kant.gestalt-llc.com

• The first part indicates a particular computer (hector)

• The rest is the domain name, indicating the organization (vt.edu)
Domain Names

- The last part of a domain name, called a top-level domain (TLD), supposedly indicates the type of organization:

  - **edu**  educational institution
  - **com**  commercial entity
  - **org**  non-profit organization
  - **net**  network-based organization

Sometimes the suffix indicates the country:

- **uk**  United Kingdom
- **au**  Australia
- **ca**  Canada
- **se**  Sweden

Additional TLDs have been added:

- **biz**, **info**, **tv**, **name**
Domain Names

• A domain name can have several parts

• Unique domain names mean that multiple sites can have individual computers with the same local name

• When used, an Internet address is translated to an IP address by software called the Domain Name System (DNS)

• There is no one-to-one correspondence between the sections of an IP address and the sections of an Internet address
The World Wide Web

• The *World Wide Web* allows many different types of information to be accessed using a common interface

• A *browser* is a program which accesses network resources and presents them
  – Popular browsers: Internet Explorer, Safari, Firefox

• Resources presented include:
  – text, graphics, video, sound, audio, executable programs

• A Web document usually contains *links* to other Web documents, creating a *hypermedia* environment

• The term Web comes from the fact that information is not organized in a linear fashion
The World Wide Web

• Web documents are often defined using the *HyperText Markup Language* (HTML)

• Information on the Web is found using a *Uniform Resource Locator* (URL):

  http://www.cnn.com

  http://www.vt.edu/student_life/index.html

  ftp://java.sun.com/applets/animation.zip

• A URL specifies a protocol (http), a domain, and possibly specific documents
Reverse History of computing

Dig deeper into what we already know, travel backwards…

1. What we see now all around us – a connected world of computing

2. Focus on a single “traditional” computer

3. Dig deeper – data and processing
A Computer Specification

- Consider the following specification for a personal computer:
  - 3.07 GHz Intel Core i7 processor
  - 4 GB RAM
  - 750 GB Hard Disk
  - 16x Blu-ray / HD DVD-ROM & 16x DVD+R DVD Burner
  - 17” Flat Screen Video Display with 1280 x 1024 resolution
  - Network Card
Computer Architecture

- Central processing unit
- Main memory
- Bus
- Disk controller
- Video controller
- Controller
- Controller
- Other peripheral devices
Memory

Main memory is divided into many memory locations (or cells)

Each memory cell has a numeric address, which uniquely identifies it
Large values are stored in consecutive memory locations.

Each memory cell stores a set number of bits (usually 8 bits, or one byte).
Storage Capacity

• Every memory device has a *storage capacity*, indicating the number of bytes it can hold.

• Capacities are expressed in various units:

<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Number of Bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>kilobyte</td>
<td>KB</td>
<td>$2^{10} = 1024$</td>
</tr>
<tr>
<td>megabyte</td>
<td>MB</td>
<td>$2^{20}$ (over one million)</td>
</tr>
<tr>
<td>gigabyte</td>
<td>GB</td>
<td>$2^{30}$ (over one billion)</td>
</tr>
<tr>
<td>terabyte</td>
<td>TB</td>
<td>$2^{40}$ (over one trillion)</td>
</tr>
<tr>
<td>petabyte</td>
<td>PB</td>
<td>$2^{50}$ (a whole bunch)</td>
</tr>
</tbody>
</table>
Memory

• Main memory is *volatile* - stored information is lost if the electric power is removed

• Secondary memory devices are *nonvolatile*

• Main memory and disks are *direct access* devices - information can be reached directly

• The terms *direct access* and *random access* often are used interchangeably

• A magnetic tape is a *sequential access* device since its data is arranged in a linear order - you must get by the intervening data in order to access other information
Hard Disk Drive

Disks

Read/write head
RAM vs. ROM

- **RAM** - Random Access Memory (direct access)
- **ROM** - Read-Only Memory

The terms RAM and main memory are basically interchangeable

ROM could be a set of memory chips, or a separate device, such as a CD ROM

Both RAM and ROM are random (direct) access devices!

RAM probably should be called Read-Write Memory
Compact Discs

- A CD-ROM is portable read-only memory
- A microscopic pit on a CD represents a binary 1 and a smooth area represents a binary 0
- A low-intensity laser reflects strongly from a smooth area and weakly from a pit
- A CD-Recordable (CD-R) drive can be used to write information to a CD once
- A CD-Rewritable (CD-RW) can be erased and reused
- The speed of a CD drive indicates how fast (max) it can read and write information to a CD
**DVDs**

- A DVD is the same physical size as a CD, but can store much more information.

- The format of a DVD stores more bits per square inch.

- A CD can store 650 MB, while a standard DVD can store 4.7 GB.
  - A double-sided DVD can store 9.4 GB.
  - Other advanced techniques can bring the capacity up to 17.0 GB.

- Like CDs, there are DVD-R and DVD-RW discs.
The Central Processing Unit

• A CPU is on a chip called a *microprocessor*

• It continuously follows the *fetch-decode-execute cycle*:

  - **fetch**: Retrieve an instruction from main memory
  - **decode**: Determine what the instruction is
  - **execute**: Carry out the instruction

Copyright © 2012 Pearson Education, Inc.
The Central Processing Unit

- **Arithmetic / Logic Unit**: Performs calculations and makes decisions.
- **Control Unit**: Coordinates processing steps.
- **Registers**: Small storage areas.
The Central Processing Unit

- The speed of a CPU is controlled by the system clock.
- The system clock generates an electronic pulse at regular intervals.
- The pulses coordinate the activities of the CPU.
- The speed is usually measured in gigahertz (GHz).
Monitor

• The size of a monitor (17") is measured diagonally, like a television screen

• A monitor has a certain maximum *resolution*, indicating the number of picture elements, called *pixels*, that it can display (such as 1280 by 1024)

• High resolution (more pixels) produces sharper pictures
Reverse History of computing

Dig deeper into what we already know, travel backwards…

1. What we see now all around us – a connected world of computing

2. Focus on a single “traditional” computer

3. Dig deeper – data and processing
Hardware and Software

• Hardware
  – the physical, tangible parts of a computer
  – keyboard, monitor, disks, wires, chips, etc.

• Software
  – programs and data
  – a program is a series of instructions

• A computer requires both hardware and software

• Each is essentially useless without the other
CPU and Main Memory

Central Processing Unit

Chip that executes program commands

Main Memory

Primary storage area for programs and data that are in active use

Synonymous with RAM
Input / Output Devices

I/O devices facilitate user interaction

Central Processing Unit

Main Memory

Monitor screen
Keyboard
Mouse
Touch screen
Secondary Memory Devices

Secondary memory devices provide long-term storage.

Information is moved between main and secondary memory as needed.

Central Processing Unit

Main Memory

Hard Disk

USB Flash Drive

Copyright © 2012 Pearson Education, Inc.
Software Categories

• Operating System
  – controls all machine activities
  – provides the user interface to the computer
  – manages resources such as the CPU and memory
  – Windows, Mac OS, Unix, Linux,

• Application program
  – generic term for any other kind of software
  – word processors, missile control systems, games

• Most operating systems and application programs have a *graphical user interface* (GUI)
Analog vs. Digital

• There are two basic ways to store and manage data:

• **Analog**
  – continuous, in direct proportion to the data represented
  – music on a record album - a needle rides on ridges in the grooves that are directly proportional to the voltages sent to the speaker

• **Digital**
  – the information is broken down into pieces, and each piece is represented separately
  – *sampling* – record discrete values of the analog representation
  – music on a compact disc - the disc stores numbers representing specific voltage levels sampled at specific times
Analog Information

Sound wave

Analog signal of the sound wave
Sampling

Information can be lost between samples

Analog signal

Sampling process

Sampled values

12 11 39 40 7 14 47

Copyright © 2012 Pearson Education, Inc.
Digital Information

• Computers store all information digitally:
  – numbers
  – text
  – graphics and images
  – audio
  – video
  – program instructions

• In some way, all information is *digitized* - broken down into pieces and represented as numbers
Representing Text Digitally

• For example, every character is stored as a number, including spaces, digits, and punctuation.

• Corresponding upper and lower case letters are separate characters.

Hi, Heather.
Binary Numbers

- Once information has been digitized, it is represented and stored in memory using the *binary number system*.

- A single binary digit (0 or 1) is called a *bit*.

- Devices that store and move information are cheaper and more reliable if they have to represent only two states.

- A single bit can represent two possible states, like a light bulb that is either on (1) or off (0).

- Permutations of bits are used to store values.
## Bit Permutations

<table>
<thead>
<tr>
<th>1 bit</th>
<th>2 bits</th>
<th>3 bits</th>
<th>4 bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>00</td>
<td>000</td>
<td>0000</td>
</tr>
<tr>
<td>1</td>
<td>01</td>
<td>001</td>
<td>0001</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>010</td>
<td>0010</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>011</td>
<td>0011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>101</td>
<td>0101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>110</td>
<td>0110</td>
</tr>
<tr>
<td></td>
<td></td>
<td>111</td>
<td>0111</td>
</tr>
</tbody>
</table>

Each additional bit doubles the number of possible permutations
Bit Permutations

• Each permutation can represent a particular item
• There are $2^N$ permutations of N bits
• Therefore, N bits are needed to represent $2^N$ unique items

<table>
<thead>
<tr>
<th>How many items can be represented by</th>
<th>1 bit ?</th>
<th>$2^1 = 2$ items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 bits ?</td>
<td>$2^2 = 4$ items</td>
</tr>
<tr>
<td></td>
<td>3 bits ?</td>
<td>$2^3 = 8$ items</td>
</tr>
<tr>
<td></td>
<td>4 bits ?</td>
<td>$2^4 = 16$ items</td>
</tr>
<tr>
<td></td>
<td>5 bits ?</td>
<td>$2^5 = 32$ items</td>
</tr>
</tbody>
</table>
Quick Check

How many bits would you need to represent each of the 50 United States using a unique permutation of bits?
Quick Check

How many bits would you need to represent each of the 50 United States using a unique permutation of bits?

Five bits wouldn't be enough, because $2^5$ is 32.

**Six bits** would give us 64 permutations, and some wouldn't be used.

000000  Alabama
000001  Alaska
000010  Arizona
000011  Arkansas
000100  California
000101  Colorado
etc.
Homework

• Read Sections 1.1-1.3 to review material
• **Always** do all self-review exercises when you review material
• Read Sections 1.4-1.6 to prepare for next class
• Do Exercises EX 1.1-1.13
• Take the [online survey](#)