INTRODUCTION

COMS W1001
Introduction to Information Science

Boyi Xie
Today’s Topics

- General information

- Introduction to information science
  - Computers
  - Computer programs and algorithms
  - Abstraction
  - Social impacts

- What we will cover this semester
General Information

• Instructor: Boyi Xie (xie@cs.columbia.edu)
• Day/Time: Monday and Wednesday 10:10am-11:25am
• Classroom: 503 Hamilton Hall
• Textbook: Pearson Custom Computer Science COMS W1001 Introduction to Information Science
• Office hour: TBA
• TA: TBA
• Evaluation
  • Five bi-weekly homework assignments (50%)
  • Two in-class examinations (25%+25%)
• Grading
  • Top 35% some kind of A (A+, A, A-)
  • Next 40-50% some kind of B
  • The rest C, D, or F depending on circumstances
Course Objectives

• To provide an introduction to using information technology and the World Wide Web

• To explore some of the fundamental concepts involved with information storage and retrieval

• To introduce the fundamentals of algorithmic problem solving

• To provide a basic introduction to computer programming in Python
Today’s Topics

• General information

• Introduction to information science
  • Computers
  • Computer programs and algorithms
  • Abstraction
  • Social impacts

• What we will cover this semester
Computer Science

- Information science
  - An interdisciplinary area
  - Our focus is from the perspectives of computer science

- Computer science is the discipline that seeks to build a scientific foundation for
  - Computer design
  - Computer programming
  - Information processing
  - Algorithmic solutions of problems
  - Algorithmic process itself
Brief History of Computing Machines

Abacus, 2nd century BC

Pascaline, 1642

Charles Babbage’s Analytical Engine

Charles Babbage (1791-1871)
Pioneer of programmable computers

Augusta Ada Byron (1815-1852)
The first programmer
Brief History of Computing Machines

• Based on mechanical relays
  • 1940: Stibitz at Bell Laboratories
  • 1944: Mark I: Howard Aiken and IBM at Harvard

• Based on vacuum tubes
  • 1937-1941: Atanasoff-Berry at Iowa State
  • 1940s: Colossus: secret German code-breaker
  • 1940s: ENIAC (electronic numerical integrator and calculator): Mauchly & Eckert at Uni. of Penn.
Brief History of Computing Machines

• Mainframes (1960s, room size)
  • Multi-users share the computers
  • Offline preparation of tasks (punched cards)
  • No direct interaction btw the programmer and machine

• Minicomputers (~1970s, refrigerator size)
  • Graphical user interface, mouse, special software

An IBM System z9 mainframe

IBM 3279 Colour Display Terminal (1979)

Punched card from a Fortran program

A PDP-11, model 40, an early member of DECs 16-bit minicomputer family, on display at the Vienna Technical Museum
Brief History of Computing Machines

• Microcomputers (~1980s - now)
  • Computer hobbyists Steve Jobs and Stephen Wozniak built a commercially viable home computer and, in 1976, established Apple Computer, Inc.
  • IBM introduced its first desktop computer, called the personal computer, or PC, in 1981, whose underlying software was developed by a newly formed company known as Microsoft.

Apple III, 1980
IBM 5150 PC with IBM 5151 monitor, 1981
Brief History of Computing Machines

- Miniaturization of computers and their expanding capabilities have brought computer technology to the forefront of today’s society.

- Tablets and smart phones

![iPhone 6](image1.png) ![Galaxy Note 4](image2.png) ![Surface Pro 3](image3.png)
Today’s Topics

• General information

• Introduction to information science
  • Computers
  • Computer programs and algorithms
  • Abstraction
  • Social impacts

• What we will cover this semester
Computer Programs and Algorithms

• An **algorithm** is a set of steps that defines how a task is performed
  • For cooking: recipes
  • For finding your way through a city: directions
  • For playing music: sheet music

• A representation of an algorithm is called a **program**.

• The process of developing a program, encoding it in machine-compatible form, and inserting it into a machine is called **programming**.

• Programs, and the algorithms they represent, are collectively referred to as **software**, in contrast to the machinery itself, which is known as **hardware**.
**Computer Programs and Algorithms**

- “**Time and space**” conditions **restricted** the complexity of the algorithms utilized in early computing machines.
- More research efforts began to tax the **abilities of the human mind**, and were directed toward the study of algorithms and the programming process.

  - It is through this ability to capture and convey **intelligence** by means of **algorithms** that we are able to build machines that perform useful tasks.
  - The level of **intelligence** displayed by machines is limited by the intelligence that can be conveyed through **algorithms**.
The Central Role of Algorithms in Computer Science

- Limitations of
- Application of
- Execution of
- Analysis of
- Communication of
- Discovery of
- Representation of
Identifying Some Questions that Provide a Focus for its Study

• Which problems can be solved by algorithmic processes?
• How can the discovery algorithms be made easier?
• How can the techniques of representing and communicating algorithms be improved?
• How can the characteristics of different algorithms be analyzed and compared?
• How can algorithms be used to manipulate information?
• How can algorithms be applied to produce intelligent behavior?
• How does the application of algorithms affect society?
Today’s Topics

• General information

• Introduction to information science
  • Computers
  • Computer programs and algorithms
    • Abstraction
  • Social impacts

• What we will cover this semester
Abstraction

• The concept of abstraction permeates the study of computer science
• Abstraction refers to the distinction between the external properties of an entity and the details of the entity’s internal composition
• At each level of abstraction, we view the system in terms of components, called abstraction tools, whose internal composition we ignore
• It allows us to concentrate on how each component interacts with other components at the same level and how the collection as a whole forms a higher-level component
Today’s Topics

• General information

• Introduction to information science
  • Computers
  • Computer programs and algorithms
  • Abstraction
  • Social impacts

• What we will cover this semester
Social Impacts

- Computing technology has altered the ability of governments to exert control
- Had enormous impact on global economics
- Led to startling advances in scientific research
- Revolutionized the role of data collection, storage, and applications
Today’s Topics

• General information

• Introduction to information science
  • Computers
  • Computer programs and algorithms
  • Abstraction
  • Social impacts

• What we will cover this semester
What We Will Cover

• Operating System Basics
• Networks and the Internet
• Web Design Basics - HTML
• Data Storage and Manipulation
• Spreadsheets
• Database
• Data Structure & Algorithm
• Programming in Python
Operating System Basics

- OS in software classification
Operating System Basics

- Interaction of OS and users
Operating System Basics

• Processes handling in OS
Operating System Basics

- Memory and disk storage
Networks

- Network topologies
  - a. Bus
    - Computer
    - Computer
    - Computer
    - Computer
    - Computer
  - b. Star
    - Computer
    - Computer
    - Computer
    - Computer
    - Computer

- To form an internet using routers
Web Design

- HTML

```html
<html>
<head>
<title>demonstration page</title>
</head>
<body>
<h1>My Web Page</h1>
<p>Click here for another page.</p>
</body>
</html>
```
Data Storage and Manipulation

- Representing text

- Representing numeric values

<table>
<thead>
<tr>
<th>0000</th>
<th>0001</th>
<th>0010</th>
<th>0011</th>
<th>0100</th>
<th>0101</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Data Storage and Manipulation

• The binary system

**a. Base ten system**

<table>
<thead>
<tr>
<th>Hundred</th>
<th>Ten</th>
<th>One</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>

**Representation**

**Position’s quantity**

**b. Base two system**

<table>
<thead>
<tr>
<th>Eight</th>
<th>Four</th>
<th>Two</th>
<th>One</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Representation**

**Position’s quantity**

\[
\begin{array}{c}
3 \\
+ 2 \\
\hline
0111 \\
\end{array}
\]

\[
\begin{array}{c}
0011 \\
+ 0010 \\
\hline
0101 \\
\end{array}
\]

5
Spreadsheet

• Excel
Database

- Table

<table>
<thead>
<tr>
<th>ID</th>
<th>name</th>
<th>dept_name</th>
<th>salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>22222</td>
<td>Einstein</td>
<td>Physics</td>
<td>95000</td>
</tr>
<tr>
<td>12121</td>
<td>Wu</td>
<td>Finance</td>
<td>90000</td>
</tr>
<tr>
<td>32343</td>
<td>El Said</td>
<td>History</td>
<td>60000</td>
</tr>
<tr>
<td>45565</td>
<td>Katz</td>
<td>Comp. Sci.</td>
<td>75000</td>
</tr>
<tr>
<td>98345</td>
<td>Kim</td>
<td>Elec. Eng.</td>
<td>80000</td>
</tr>
<tr>
<td>76766</td>
<td>Crick</td>
<td>Biology</td>
<td>72000</td>
</tr>
<tr>
<td>10101</td>
<td>Srinivasan</td>
<td>Comp. Sci.</td>
<td>65000</td>
</tr>
<tr>
<td>58583</td>
<td>Calieri</td>
<td>History</td>
<td>62000</td>
</tr>
<tr>
<td>83821</td>
<td>Brandt</td>
<td>Comp. Sci.</td>
<td>92000</td>
</tr>
<tr>
<td>15151</td>
<td>Mozart</td>
<td>Music</td>
<td>40000</td>
</tr>
<tr>
<td>33456</td>
<td>Gold</td>
<td>Physics</td>
<td>87000</td>
</tr>
<tr>
<td>76543</td>
<td>Singh</td>
<td>Finance</td>
<td>80000</td>
</tr>
</tbody>
</table>

(a) The instructor table

<table>
<thead>
<tr>
<th>dept_name</th>
<th>building</th>
<th>budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comp. Sci.</td>
<td>Taylor</td>
<td>100000</td>
</tr>
<tr>
<td>Biology</td>
<td>Watson</td>
<td>90000</td>
</tr>
<tr>
<td>Elec. Eng.</td>
<td>Taylor</td>
<td>85000</td>
</tr>
<tr>
<td>Music</td>
<td>Packard</td>
<td>80000</td>
</tr>
<tr>
<td>Finance</td>
<td>Painter</td>
<td>120000</td>
</tr>
<tr>
<td>History</td>
<td>Painter</td>
<td>50000</td>
</tr>
<tr>
<td>Physics</td>
<td>Watson</td>
<td>70000</td>
</tr>
</tbody>
</table>

(b) The department table

- SQL - a standard language for accessing databases

Example: Find the name of the instructor with ID 22222

```
select name
from instructor
where instructor.ID = '22222'
```
Data Structure & Algorithm

- **Data structure Basics**
  - Data type and data abstraction
  - Basic structures: linear, tree, graph
  - Useful data types: array, list, set, map

- **Algorithm**
  - Time and space complexity
  - Searching
  - Sorting
Programming in Python

• Why Python
  • Easy to learn and use
  • Versatile and popular
  • Lots of online resources and support

• Learning Python
  • Basic data types
  • Control flows
  • Functions
  • File Input/Output
  • Object-Oriented Programming
Today’s Topics

• General information

• Introduction to information science
  • Computers
  • Computer programs and algorithms
  • Abstraction
  • Social impacts

• What we will cover this semester
For the Next Class

• Survey (use one or two sentences)
  1. Tell me about you: name, school, major, year
  2. Why do you want to take this course?
  3. What do you expect to learn?
  4. How do you think this course can be relevant to your current major, future study, or career?
  5. Do you have any existing knowledge in computer science, e.g. programming language, web design, database, etc?

• Email to xie@cs.columbia.edu

• Read Chapter 2 Operating Systems by Brookshear
• Read CUNIX tutorial: http://www.columbia.edu/~lgw23/cs1004/
References & Photo Credits

- http://en.wikipedia.org/wiki/Mainframe_computer
- http://en.wikipedia.org/wiki/Minicomputer