SFU SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

CMPT 120: Introduction to Computing Science and Programming 1

Binary Encoding / Representation

python™

Reminders

One-Stop Access To Course Information

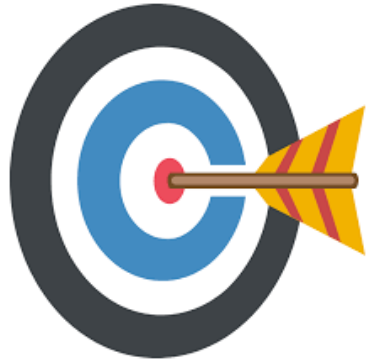
- **Course website: One-stop access** to all course information.

  http://www2.cs.sfu.ca/CourseCentral/120/liaqata/WebSite/index.html

- Course Outline
- Exam Schedule
- Python Info
- CourSys/Canvas link
- Learning Outcomes
- Office Hours
- Textbook links
- and more…
- Grading Scheme
- Lab/Tutorial Info
- Assignments

- **Canvas**: Discussions forum - https://canvas.sfu.ca/courses/39187
- **CourSys**: Assignments submission, grades - www.coursys.sfu.ca

How to Learn in This Course?



A **Attend** Lectures & Labs

R **Read** / review Textbook/Slides/Notes

R **Reflect** and ask Questions

O **Organize** – your learning activities on weekly basis,

and finally…

W **Write** Code, Write Code, and Write Code.

Deliverables

1. Deliverables are due by the given date and time.
2. For the course, we are using IDLE to write and run our Python code.
3. You can use the CSIL lab computers outside your lab hours.
4. Plan ahead your assignments and other deliverables. Computer crash, network problems etc. are not acceptable excuses for delays in deliverables.
5. You may use online Python interpreters for running and testing your codes, such as:

   https://repl.it/languages/Python3

Labs

1. Each lab has an assigned TA.
2. Attend your assigned lab and show your work to your TA for the participation marks.
3. Class enrolments and lab swaps are closed now.

Course Topics

1. General introduction
2. Algorithms, flow charts and pseudocode
3. Procedural programming in Python
4. Data types and Control Structures
5. Fundamental algorithms
6. Binary encodings
7. Basics of computability and complexity
8. Basics of Recursion
9. Subject to time availability:
   - Basics of Data File management

Today's Topics

1. Data Representation (Binary Encoding)
   - ASCII
   - Unicode
2. Data Structures:
   - int, float, Boolean
   - Lists: Methods & Functions:
   - Tuples: Methods & Functions
3. User Defined Functions

1

Data Representation

Binary Data Representation

- Data inside computer is **not represented** the same way as we represent numbers and letters in English or native language. **For example**:
  - We represent quantities using symbols (digits) **0, 1, 3,... and 9.**
  - We can write names using English letters **A, B, C,…Z** or **a, b, c,…z**
    - So, we represent a quantity **six** by using the symbol **6**.
    - Using English alphabets, we can represent a street name as: **Dawson Street.**
- **Problem!!!**
- Computer don't use (recognize) the symbols 0,1,2..9 or alphabets a, b, c,…z
- Because, computers use a completely **different** language to represent numbers or letters (or data).
- We call it machine language. (Or, binary language or representation.)

Binary Data Representation - 2

- The **binary language** consists of two symbols only: **0** and **1**
- That means, every thing in computer **MUST** be represented using the symbols **0** and **1,** only
- So, the quantity **six** must be represented using a commination of **0s** and **1s**. (Binary code)
- The name **Dawson Street** must also be represented using a commination of 0s and 1s.
- Let's create **our own binary codes** to represent letters A, B, C, …Z using a combination of **0s and 1s**.

Binary Codes (Our Own Codes)

Letter	Binary Code
A	
B	
C	
D	
E	
F	
G	
H	
I	
J	
K	
L	

Letter	Binary Code
a	
b	
c	
d	
e	
f	
g	
h	
i	
j	
k	
l	

Binary Codes: ASCII

Letter	ASCII Binary Code
A	
B	
C	
D	0 1 0 0 0 1 0 0
E	0 1 0 0 0 1 0 1
F	0 1 0 0 0 1 1 0
…	

- **ASCII:** American Standard Code for Information Interchange. (256 codes.)
- Used in computers to represent characters since 1963.
- ASCII uses 8-bits to represent one character of English language.

Letter	ASCII Binary Code
a	
b	
c	
d	0 1 1 0 0 1 0 0
e	0 1 1 0 0 1 0 1
f	0 1 1 0 0 1 1 0
…	

- Space required to represent a single binary 0 or 1 is called _____.
- Space required to represent 8-bits is called a _____.
- See a complete list of ASCII codes here: www.ascii-code.com

Number Systems

- **Binary Number System**: Uses **two** unique symbols to represents numbers or data. (0 and 1).
- **Decimal system**: Use **ten** unique symbols to represent numbers. (0, 1, 2, 3, 4, 5, 6, 7, 8, and 9).
- **Octal system**: Use **eight** unique symbols to represent numbers. (0, 1, 2, 3, 4, 5, 6, and 7).
- **Hexa-decimal system**: Use **sixteen** unique symbols to represent numbers. (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A,B,C,D,E and F).

- We can convert between number systems.
  - We can convert a decimal number representation into a binary, octal, or hexa-decimal representation.
  - We can convert a binary number representation into a decimal representation, or others.
    - A binary 1 at right most position means 1.
    - A binary 1 at next positions means 2, 4, 8, 16, 32, 64 and 128 respectively.

1	1	1	1	1	1	1	1
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
128	64	32	16	8	4	2	1

Examples

1	0	1	0	1	0	1	1
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
128	64	32	16	8	4	2	1

= 171

0	0	1	0	0	0	1	1
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
128	64	32	16	8	4	2	1

= 35

Converting from Decimal to binary

- 111
  - 128 too large from 111,
    - so there are **zero** 128 in 111.
  - 111 - 64 = 47
    - There is **one** 64 in 111, remainder 47.)
  -
  -
  -
  -
  -

1	1	1	1	1	1	1	1
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
128	64	32	16	8	4	2	1

ASCII: Decimal Equivalent

Letter	ASCII Binary Code							
	128	64	32	16	8	4	2	1
A	0	1	0	0	0	0	0	1
B	0	1	0	0	0	0	1	0
C	0	1	0	0	0	0	1	1
D	0	1	0	0	0	1	0	0
E	0	1	0	0	0	1	0	1
F	0	1	0	0	0	1	1	0
…								

- When we use Boolean expression (‘a’ < ‘A’), computer would compare the ASCII value of a (which is 97) with the value of ASCII value of A (which is 65).
-

Letter	ASCII Binary Code							
	128	64	32	16	8	4	2	1
a	0	1	1	0	0	0	0	1
b	0	1	1	0	0	0	1	0
c	0	1	1	0	0	0	1	1
d	0	1	1	0	0	1	0	0
e	0	1	1	0	0	1	0	1
f	0	1	1	0	0	1	1	0
…								

- ‘B’ <= ‘b’
- ‘cd’ <= ‘ab’
- ‘xyz’ > ‘XYZ’

Unicode Data Representation

- With ASCII we represent at the most of 256 unique characters.
- **Problem!!!**
- Unicode coding scheme was devised to encode more characters – such as characters from other world languages.
- Unicode now contains over 137,000 unique characters covering 146 modern and historic scripts. (Wikipedia)

Questions?