DS286 | 2016-08-05

L1: Data Structures and Programming

Yogesh Simmhan

simmhan@cds.iisc.ac.in

CDS
The Department of Computational and Data Science

Who is this Course For?

- CDS MTech (Computational Science) 1st Year Students
- Those new to programming, data structures, algorithms
- Those who have not had a strong programming course in your UG
- Those from other departments who will be writing programs for your research
- Prerequisites for other advanced courses

What is this Course About?

- Basics of Programing and data structures...*at the senior undergraduate level*

- Lecture covers basics of...
  ‣ Data Structures: Lists, matrices, trees, searching, sorting, graphs
  ‣ Algorithms, complexity analysis, strategies
  ‣ Concurrent data structures

- Lab covers basics of...
  ‣ Object Oriented Programming
  ‣ C++ Programming, compiling, IDEs
  ‣ Modular composition of applications, DS libraries in C++, STL, API documentation
  ‣ Concurrent programming
  ‣ Performance benchmarking

- *Learning to enjoy programming!*

What is this Course NOT About?

- Advanced algorithms, data structures
  - ‣ E0 251: Data Structures and Algorithms
  - ‣ E0 225: Design and Analysis of Algorithms
- Advanced C++ programming
- Learning other programming languages
- High Performance Computing
  - ‣ DS 292: High Performance Computing

Teaching Team

- Yogesh Simmhan
    - Assistant Professor, CDS, IISc
    - www.dream-lab.in

- Prateeksha Varshney
    - MSc(Res) Student, CDS, IISc

When are the classes?

- **Lectures (YS)**
  - ‣ Wed & Fri 10-11AM CDS 202
  - ‣ Exceptions on Institute Holidays, other schedule conflicts. May swap with Lab Session.
    - Check online, mailing list for changes
- *Days & Times non negotiable!*

- **Lab Session (PV)**
  - ‣ Mon 10-11AM CDS 202
  - ‣ Bring your laptops if available

Class Resources

- Website
  - ‣ Schedule, Lectures, Assignments, Additional Reading
  - ‣ http://cds.iisc.ac.in/faculty/simmhan/DS286

- Textbook
  - ‣ **Lectures:** Data Structures, Algorithms, and Applications in C++, 2nd Edition, Sartaj Sahni*,**
    - • http://www.cise.ufl.edu/~sahni/dsaac/
  - ‣ **Lab:** The C++ Programming Language, 3rd Edition, Bjarne Stroustrup

- Other Books
  - ‣ THE ART OF COMPUTER PROGRAMMING (Volume 1 / Fundamental Algorithms), Donald Knuth
  - ‣ Introduction to Algorithms, Cormen, Leiserson, Rivest and Stein

- Online resources
  - ‣ www.geeksforgeeks.org/data-structures/

*http://www.tatabookhouse.com/data-structures-algorithms-and-applications-in-c-plus-plus--9788173715228?ver=7519259641
**http://www.flipkart.com/data-structures-algorithms-applications-c-english-2nd/p/itmeyf6jvka3kzdu

Grading

- 10%: Online coding problems on CodeChef

- 20%: Midterm Exam
  - ‣ Written, closed book, half the syllabus

- 30%: Final Exam
  - ‣ Written, closed book, whole syllabus

- 40% Programming Assignments
  - ‣ ~7 assignments with 5-10% weightage each
  - ‣ Independent work
  - ‣ Submission using **TurnItIn**

Grading: CodeChef ☺

- Number of problems, weighted by difficulty (or) performance in competitions
- Exponential grading, e.g.
  ‣ 2%: Solving 5+ 'beginner' practise problems
  ‣ 4%: Solving 5+ 'easy' practise problems
  ‣ 6%: Solving 5+ 'medium' practise problems
  ‣ 8%: Solving 5+ 'hard' practise problems (or) top 500 rank in CodeChef (lunch, snack, cookoff, challenge) a contest
  ‣ 10%: Solving 3+ 'challenge' problems (or) top 100 rank in a contest
- Submit your handle (IISc affiliation) by Aug 15. Verify by solving assigned problem.
- Performance between Aug 15 and Nov 30 will be used.

IISc Code of Conduct

Course Guidelines

- Students must uphold IISc's Code of Conduct.
  - ‣ *Review them!* Failure to follow them **will** lead to sanctions and penalties: reduced or failing grade … **Zero Tolerance!**
  - ‣ Severe cases of academic violations will be reported to the Institute and may lead to an expulsion.
- Learning takes place both within and outside the class
  - ‣ More outside than inside ☺
- Discussions between students and reference to online material is highly encouraged
- However, you must form your own ideas and **complete problems and assignments by yourself**.
- All works submitted by the student as part of their academic assessment must be their own!

- 6.2 Violations of this policy include, but are not limited to:

- (i) Plagiarism means the use of material, ideas, figures, code or data as one's own, without appropriately acknowledging the original source.

- This may involve submission of material, verbatim or paraphrased, that is authored by another person or published earlier by oneself.

- Examples of plagiarism include:
  - (a) Reproducing, in whole or part, text/sentences from a report, book, thesis, publication or the internet.
  - (b) Reproducing one's own previously published data, illustrations, figures, images, or someone else's data, etc.

‣ (c) Taking material from class-notes or incorporating material from the internet graphs, drawings, photographs, diagrams, tables, spreadsheets, computer programs, or other non-textual material from other sources into one's class reports, presentations, manuscripts, research papers or thesis without proper attribution.

‣ (d) Self plagiarism which constitutes copying verbatim from one's own earlier published work in a journal or conference proceedings without appropriate citations.

‣ e) Submitting a purchased or downloaded term paper or other materials to satisfy a course requirement.

‣ f) Paraphrasing or changing an author's words or style without citation.

- (ii) Cheating
- Cheating includes, but is not limited to:
    ‣ (a) Copying during examinations, and copying of homework assignments, term papers, theses or manuscripts.
    ‣ (b) **Allowing or facilitating copying, or writing a report or taking examination for someone else.**
    ‣ (c) Using unauthorized material, copying, collaborating when not authorized, and purchasing or borrowing papers or material from various sources.
    ‣ (d) Fabricating (making up) or falsifying (manipulating) data and reporting them in thesis and publications.
    ‣ (e) Creating sources, or citations that do not exist
    ‣ (f) Altering previously evaluated and re-submitting the work for re-evaluation
    ‣ (g) Signing another student's name on an assignment, report, research paper, thesis or attendance sheet

IISc Penalties

- A breach of academic integrity is a serious offence with long lasting consequences for both the *individual* and the *Institute,* and this can lead to various sanctions.

- In the case of a student the first violation of academic breach will lead to a warning and/or an "F" course grade.

- A repeat offence, if deemed sufficiently serious, could lead to expulsion.

Introduction

Concepts

- Algorithm: Outline, the essence of a computational procedure, step-by-step instructions

- Program: an implementations of an algorithm in some programming language

- Data structure: Organization of data needed to solve the problem (array, list)

Problem Solving

- Data Structure
  - ‣ Organization for a collection of data items
  - ‣ Any data representation
- Eg, An integer
- Programming
  - ‣ Problem Solving
  - ‣ Programming methodology

Solving Problem on Computer

- Issues in Problem Solving
  - Problems are posed in natural language
  - Machine understand only a restricted form of language
  - No clues are given to bridge the GAP between machine language and natural language

Problem Solving

- Converting the problem to machine language so that it works efficiently

- Technique of problem solving require
  - ‣ Convert it to a sequence of steps

- Algorithm Design, Program methodology, Data structuring
  - ‣ Steps executable by computer
  - ‣ No automated way, only humans can do (partial solutions to problem solving is achieved by AI. But still very primitive)

How do we solve?

- Intuition
- Mixture of techniques
- Experience (Body of knowledge)

- Understand and use the above in order to solve the problem.
- Develop new techniques using the above

Final Solution → Program

- Needs vehicle language

- Sophisticated programming language provides high level language for implementing the steps.

- Machine language : difficult to program

Data Structures

- Data Types and Operations

- Data-Procedure Encapsulation

- Dynamic Data structuring : dynamic data alloation, de-allocation

- Algorithm Expression
  ‣ Control contructs (If, for, while …)
  ‣ Functions

- Recursive (Fortran does not provide)

Course Objectives

- Understanding Programming methodology
- Understanding Algorithms
- Understanding data structuring
- Mutual relation between the both in solving problem

- Learn C++ language (features and data structure)

To Be Continued...