CMSC 420 (Section – 0101) “Data Structures”

Course Hand-Out

http://www.cs.umd.edu/class/spring2013/cmsc420/

GOAL. The goal of this course is to ensure that students can design/select the “right” data structures to store data for a problem or application in such a way that the performance goals that an end user or application developer has for certain operations are met. In short, there is no “good” or “bad” data structure in which to store data. Data structures are “good” or “bad”, depending on the intended operations that must be performed on the data, and the performance expectations that an end user or application has.

TOPICS TO BE COVERED (TENTATIVE). The list of topics given below is a tentative list. Some topics may be added or deleted as needed. Some sample URLs are given below – you might find other useful ones on the Internet as well. But note that different authors use slightly different definitions of terms and data structures, sometimes leading to subtle variations from one presentation to another.

- Searching uni-dimensional data
  - Binary search trees and B-trees
    - http://www.cse.unt.edu/~rada/CSCE3110/Lectures/BinarySearchTrees.ppt
    - www.cs.sjsu.edu/~lee/cs146/24spxinlinB-Trees.ppt
    - www.cs.laurentian.ca/julia/courses/cosc2007/lectures/B-Trees.ppt
  - Range queries and nearest neighbor queries

- Multi-dimensional data structures
  - K-d trees
    - www.cse.lehigh.edu/~munoz/CSE335/classes/Retrieval2.ppt (also covers point quadtrees)
  - Point Quadtrees
  - Region Quadtrees MX-Quadtrees
  - PR-Quadtrees [small variant of MX-quadtrees]
  - Unit Segment Trees & Segment Trees
    - cimic.rutgers.edu/~adam/mmis03/MMIS/video.ppt
  - PM1- and PM2- Quadtrees
  - Plane Sweep and Rectangle Intersection Problems
    - www.cs.tulane.edu/~carola/teaching/cs6463/fall06/slides/sweepline.ppt
    - ima.udg.edu/~sellares/ComGeo/Intersections.ppt
  - R-trees
    - sysnet.ucsd.edu/~cfeizac/cse262/R-Trees.ppt
    - csis.bits-pilani.ac.in/faculty/goel/R-TREES/SearchTrees.ppt

- Searching text data
  - Inverted indexes
    - www.cs.northwestern.edu/~ddowney/courses/395_Spring2011/lec2.pdf
    - www.stanford.edu/class/cs276a/handouts/lecture1.ppt
  - Similarity search and TV-trees
    - coitweb.uncc.edu/~ras/KDD-02/TV-Trees.ppt
• Searching graph and social network data
  o Shortest path algorithms
    ▪ [link](webhome.csc.uvic.ca/~ruskey/classes/326/slides/Chpt7ShortestPath.ppt)
    ▪ [link](www.cs.princeton.edu/~wayne/kleinberg-tardos/04demo-dijkstra.ppt)
  o Minimal Spanning Trees
    ▪ [link](http://www.cs.sjsu.edu/~lee/cs157b/IanGravesMinimum%20Spanning%20Trees.ppt)
    ▪ [link](www-users.cs.umn.edu/~karypis/parbook/Lectures/AG/chap10_slides.ppt)
  o Graph partitioning algorithms [slides will be provided]
  o Subgraph matching algorithms [slides will be provided]

• Other topics as needed

COURSE PROJECT. A semester long project will be assigned at the very beginning of the semester. The project will have an associated time-table and interim deliverables. Each deliverable will count towards a percentage of the grade you will receive for the project, which, in turn, will count toward your overall course grade. The final course project will be due on May 2, 2013. Late project submissions will not be accepted without a note from a doctor certifying that you were too sick to do the project on time.

HOMEWORK ASSIGNMENTS. There will be somewhere between 3 and 6 homework assignments. Unless otherwise stated, assignments will be due within a week of the date on which they are handed out. Late homework will not be accepted without a note from a doctor certifying that you were too sick to do the assignment.

MID-TERM EXAM. A mid-term exam will be held during regular class hours on Thursday, March 7, 2013. The topics to be covered in the mid-term will be announced approximately one week before the mid-term.

FINAL EXAM. The final exam will be held from 8-10 a.m. on May 13, 2013 from 8-10 a.m.

COURSE GRADING. Your course grade at the end of the semester will be determined by your performance on the course programming project (30% of grade), the mid-term (30% of grade), the final exam (30% of grade) and the homework (10% of grade).

TEXTBOOK. As virtually all course material is available on the web for free today (and as textbooks are vastly over-priced), there will be no official textbook for the course. However, when discussing individual topics, a list of web sites with appropriate material will be made available to the class.

TEACHING ASSISTANT. Manish Purohit can be reached at [manishp@cs.umd.edu](mailto:manishp@cs.umd.edu) and he will provide advice on your programming project and homework assignments. His office hours will be from 2:00-3:00 p.m. on Mondays and 11:00-12:00 noon on Wednesdays in room 1112 AV Williams (AVW).

CONTACT INFORMATION & OFFICE HOURS. I can be reached at [vs@cs.umd.edu](mailto:vs@cs.umd.edu), 301-405-6724. I will be available immediately after class from 10:45-11:30 a.m. (2113 AVW) on Tu-Th, or by appointment. Please contact my secretary, Barbara Lewis ([blewis@umiacs.umd.edu](mailto:blewis@umiacs.umd.edu)) to schedule an appointment, to pick up homework/projects etc.

Good Luck!