CSE 202

Basic Information: Spring, 2008

Instructor: Russell Impagliazzo

email: russell@cs.ucsd.edu

webpage: www-cse.ucsd.edu/classes/sp08/cse202

Class: Tuesdays and Thursdays, 11-12:20, PCYNH 120;

TA: Kirill Levchenko

Russell's CSE 202 Office Hours: TBA.

TA office hours: TBA

Prerequisites: We assume some undergraduate exposure to discrete mathematics, and to algorithms and their analysis, and the ability to read, recognize and write a valid proof. We will very quickly review topics from your undergraduate algorithms courses, and then look at related advanced material. your background in these areas is weak, you may need to do some extra work to catch up. Please talk to me about this. A helpful text for proofs is Solow: How to Read and Do Proofs.

Text Book: Kleinberg and Tardos, Algorithm Design

Optional Supplementary Texts: Neapolitan and Naimipour, Foundations of Algorithms; Jeff Edmonds, How to Think About Algorithms (See webpage).

Assignments There will be one "calibration" homework, three homework assignments, a project, and a take-home final exam. Each homework assignment will have four theoretical problems and one implementation problem. You should budget 10-20 hours for each homework assignment, including time spent in office hours, and 20-40 hours for the project. Homework and the project can be done in groups up to four. Most people stay with the same group for the quarter, but you are free to switch groups (or kick people out of your group if they aren't pulling their weight.) If you have trouble finding a group, or need another person for your group, send me email and I'll put an announcement on the class website.

Evaluation: Homework will acount for 30 % of the grade, the project, 20%, and the final will account for the remaining 50 % of the grade. The calibration homework is optional and does not count towards the grade.

Standards for evaluation: Most problems will be theoretical in nature. Answers to these problems must include, implicitly or explicitly, a proof of all claims. For algorithm design questions, that involves a proof of correctness (that your algorithm actually solves the given problem) and of the time analysis. Grading of all problems (homework and exam) will be both on the basis of correctness and on logical consistency and completeness, i.e., "mathematical style". It is your obligation to provide a compelling argument that forces the reader to believe the result, not just notes from which a proof could be reconstructed. Note that I will mark off for any unsubstantiated and non-trivial claims in a proof, even correct claims.

Don't assume knowledge or sophistication on the part of the reader. A good rule of thumb is to mentally address your answer to a bright undergraduate just starting an algorithms course; don't think of your reader as a trained algorithms professor who already knows the answer. I have to evaluate your answer based on WHAT YOU WRITE not WHAT I KNOW.

More particularly, many questions will be of the form , "Provide an efficient algorithm for the following computational problem". Your answer will be graded on the following criteria:

1. Your algorithm must be clearly and unambiguously described. This can be in (well-documented, clear) pseudocode (with explanations in English) or in (precise, mathematical, well-defined) English. There should be no room for interpretation in the steps carried out by your algorithm. Any mathematically and computer literate individual should be able to follow the steps presented, or to implement the algorithm as a computer program. Such implementations by different people should still be essentially identical.
2. Your algorithm needs to solve the problem given. To certify this, you need to give a correctness proof for your algorithm. My rule is: an algorithm is incorrect until PROVED correct. I will use this rule in grading even if I KNOW your algorithm is correct. If I think you don't KNOW why your algorithm works, I will grade it as if your algorithm does NOT work. In some cases, correctness is easy or trivial; in this case, your correctness argument can be a short English explanation of Other times, correctness is highly non-trivial and requires a medium-sized mathematical argument. It's your job to distinguish these two cases.
3. Your algorithm must be efficient. This means, usually, polynomial-time, but it could mean sub-linear time in the case of data-base operations, or almost linear time when the problem is trivially poly-time. Again, unless your answer includes a well-reasoned time analysis, the presumption is that it is NOT efficient. Logic is as important as cal-

culations in a time-analysis. At a minimum, a time analysis requires an explanation of where the calculations come from. If the analysis is "easy" (e.g., with a simple nested loop algorithm), these explanations can be brief (e.g., "The outside loop goes from 1 to n, and each iteration, the inside loop iterates m times, so the overall time is $O(nm)$."). Other times, the time analysis is a tricky, mathematical proof. If you give just calculations or just a short explanation, and I think the time bound is NOT easy and clear from what you wrote, you will lose points even if you give the correct time.

4. Your algorithm must be relatively efficient. That means that, even if your algorithm is correct and reasonably fast, you may lose points if there is a FASTER algorithm. Nothing you said is false, but I'll still deduct some points because you COULD have been cleverer. How can you avoid this? There's no guaranteed way, but if the correct algorithm seems easy, you should still try to think of improvements or other approaches. Is this fair? Not really, but it is the facts of life. Your algorithm won't be used if there is a faster equivalent algorithm.

There will also be some implementation problems on the homework. For these problems, you should present an outline of the basic algorithm you used, a discussion of implementation problems or where you didn't use the fastest asymptotic algorithm and why, results from the challenge instances described in the problems, and timing information about various problem instances. I WILL NOT READ ACTUAL CODE SO DON'T BOTHER HANDING IT IN. If the actual times seemed different from the asymptotic analysis, give a short discussion of possible reasons. Give the programming language used and your computer's speed rating so that I can normalize. Your grade will be based on your answer's completeness, and your success and time for challenge problems.

Project The project for the class is to find a well-specified algorithmic problem that is important for some computer science area, preferably related to your research or job. You should 1. define the problem rigorously, 2. present and analyze algorithms in the literature to solve the problem, and 3. discuss how well the theoretical analysis captures the real-life performance of these algorithms. Your project should have at least one non-trivial mathematical proof in it (correctness proof, approximation ratio, non-trivial time analysis), and at least a page of non-mathematical discussion of the application. If you have trouble thinking of a topic, talk to me (and we can have trouble thinking of a topic together.) You may work in groups up to 4, but a team project should be somewhat longer, and everyone in the group must write up part of the project. Individual projects should be 5 (word-processed) pages or so. Tentative due date: May 26.

Academic Honesty Students will be allowed to solve and write up all homework assignments and the project in groups of size up to 4. All names should appear on the assignment. Students are responsible for the correctness and the honesty of all problems handed in with their names attached. If a group member did not participate in discussion for one of the problems handed in by the group, the group must write a note on the front page of the solution set to that effect.

Students should not look for answers to homework problems in other texts or on the internet. Yes, it is interesting to find the best known solutions. **However, do the literature search after you hand in your own solution.** Students may use other texts as a general study tool, and may accidentally see solutions to homework problems. In this case, the student should write up the final solution without consulting the text, and should give an acknowledgement of the text on the first page of their solutions. Such a solution may be given partial or no credit if it too closely follows the text. This includes all material found on the web (except on this year's class webpage), discussion with others who are either students or not (except the instructor or TA, or other students as part of office hours), or written notes from others, whether students or not, (except class notes).

Be sure to follow the following guidelines:

1. Do not discuss problems with people outside your group whether students or not (except the TA and me, of course).
2. Do not share written solutions or partial solutions with other groups.
3. Prepare your final written solution without consulting any written material except class notes and the class text.
4. Acknowledge all supplementary texts or other sources that had solutions to homework problems. Acknowledge anyone who helped with assignments, except the TA and myself.
5. Do not discuss the final exam with anyone except the instructor and TA.

Lateness Policy Late homework will be accepted until I give out an answer key and no later. So you have to be no later than me.

Reading Schedule You are expected to do a **lot** of independent reading for this course. I will be presenting selected topics from the text, but much of what I'll be going over in class will require background not covered explicitly in class, either from an undergraduate course or from independent reading. The text tends to go over material in immaculate detail, and can frequently be skimmed so the reading assignments are not quite as heavy as they may seem at first glance. However, they are substantial and you are responsible for all material included.

To help you plan, I will give a tentative schedule of general topics to be covered in class, and the corresponding sections of the text to be read:

Background: Chapters 1-3. Also, recurrence relations from 5.1, 5.2. Try some exercises, and the calibration homework. If you have **any** problems, go back and read the chapter thoroughly. All of this material is important and none of it will be covered in class.

Graph search Breadth-first and depth-first search, applications. Dijkstra's algorithm. Mainly Chapter 4. 2 lectures.

Greedy algorithms Chapter 4. Interval scheduling (4.1), Minimum Spanning Tree(4.5); (4.4), greedy approximation algorithms (11.1-11.3). 3 lectures.

Applying Data Structures to Algorithm Design (Chapters 2.5, 3, 4.6) Refers to all algorithms mentioned above, in particular greedy algorithms. 2 lectures.

Divide-and-Conquer : Chapter 5. Mergesort (5.1, in passing), Integer Multiplication (5.5), Closest Pair of points (5.4). Median Finding and selection (13.5). 2 lectures.

Search and Optimization problems, and the class NP See section 8.3 (1/2 lecture)

Back-tracking, Depth-first Search, Breadth-first Search, Branch-and-Bound Not in text, but see Chapter 10. Example problems: Maximum Independent set (p. 16), Graph Coloring (8.7), Hamiltonian cycle(8.5)addition chains (not in text). (2 lectures).

Dynamic Programming: Chapter 6. Weighted interval scheduling (6.1), edit distance (6.6), All-pairs shortest paths (6.8); Subset sum (6.5) and DP-based approximation algorithms (11.8). 3 lectures.

Network Flows, Hill-climbing and Reductions: 5 lectures.

- Network flows and the Ford-Fulkerson algorithm. (Chapter 7.1-7.3) (1 lecture).
- Hill-climbing as an algorithmic technique (not in text, but see Chapter 1, stable marriage, and Chapter 12, local search)
- Reductions between problems and applications of network flow (1 lecture, Chapter 8.1, 7.5,7.6)
- Reduction as an algorithmic technique
- Reductions and NP-completeness (Chapter 8.2, 8.4-8.7). (1 lecture)
- Coping with intractibility: Heuristics, average-case analysis, and approximations (Chapter 11, 12). 1 lecture.

Assignment Schedule To help you plan, here is a tentative list of when assignments will be due:

1. Thursday, April 10: Calibration homework due.
2. Tuesday, April 29: First homework (graph search, greedy algorithms, data structures) due.
3. Thursday, May 22: Second homework (divide and conquer, backtracking, dynamic programming) due.
4. Tuesday, May 29 : Project due.
5. Thursday, June 5: Third homework due (reductions, network flow, hill-climbing). Final exam available.
6. Thurs., June 12: Final exam due.