CSE 312 Foundations II

I. Introduction

Spring 2015
W.L. Ruzzo
CSE 312, Au '13: Foundations of Computing II

Lecture: MGH 241  MWF 1:30-2:20
Section A: MGH 242  Th 1:30-2:20 Sonya Alexandrova
Section B: MGH 228  Th 2:30-3:20 Scott Lundberg
Section C: MEB 243  Th 12:30-1:20 Yanling He
Office Hours Location Phone
Instructor: Larry Ruzzo, ruzzo@cs  F 2:30-3:20 CSE 554 543-6298
TAs: Sonya Alexandrova, sonya@cs
Scott Lundberg, slund1@cs
Yanling He, heyl@cs
M 4:30-5:30 CSE 216
Monday's
Alexandrova, sonya@cs
Scott Lundberg, slund1@cs
Yanling He, heyl@cs
3:30-4:30 CSE 2xx
Course Email: cse312aau13@uw.edu. Staff announcements and general interest student/staff Q&A about homework, lectures, etc. The instructor and TA are subscribed to this list. Enrolled students are as well, but probably should change their default subscription options. Messages are automatically archived.

Discussion Board: Also feel free to use Catalyst GoPost to discuss homework, etc.

Catalog Description: Examines fundamentals of enumeration and discrete probability; applications of randomness to computing; polynomial-time versus NP; and NP-completeness.

Prerequisites: CSE 311; CSE 332, which may be taken concurrently.

Credits: 4

Learning Objectives: Course goals include an appreciation and introductory understanding of (1) methods of counting and basic combinatorics, (2) the language of probability for expressing and analyzing randomness and uncertainty, (3) properties of randomness and their use in designing and analyzing computational systems, and (4) some basic methods of statistics and their use in a computer science & engineering context.

Grading: Homework, Midterm, Final. Possibly some quizzes, small programming assignments. Overall weights 55%, 15%, 30%, roughly.

Late Policy: Assignments are due at the start of lecture on the due date, either on paper or electronically. Late papers/turnin will be accepted (but penalized 25%) up to the start of the next lecture; not accepted thereafter, barring major emergencies.

Extra Credit: Assignments may include "extra credit" sections. These will enrich your understanding of the material, but at a low points per hour ratio. Do them for the glory, not the points, and don't start extra credit until the basics are complete.

Collaboration: Homeworks are all individual, not group, exercises. Discussing them with others is fine, even encouraged, but you must produce your own homework solutions. Follow the "Gilligan's Island Rule": if you discuss the assignment with someone else, don't keep any notes (paper or electronic) from the discussion, then go watch 30+ minutes of TV (Gilligan's Island reruns especially recommended) before you continue work on the homework by yourself. You may not look at other people's written solutions to these problems, not in your notes, not in the dorm files, not on the internet. Don't ask about whether your activities tell us before, not after, you've done the UW CSE Academic Integrity and Honor policy.

Textbooks: http://courses.cs.washington.edu/cse312
Empiricism:

1. Relying on observation and experiment, esp. in the natural sciences

2. A former school of medical practice founded on experience *without the aid of science or theory*

Synonym: Quackery, Charlatanry
Study Probability!

“Life is uncertain. Eat dessert first.”

-- Ernestine Ulmer
Counting & Binomial Coeffs: (1 wk)
• Sum and product rules, product trees, Permutations and Combinations, Inclusion-Exclusion, Binomial Theorem, Pigeonhole Principle

Probability (5 wks)
• Basics: Sample spaces, events, (e.g. coins, dice, cards, program bugs?)
• Conditional probability & Bayes theorem, ex: false positive/negative, spam detection
• Random variables: independence, expectation, linearity of expectation, variance
• Bernoulli trials, binomial, multinomial? distributions; Poisson approximation
• Tail bounds (Markov, Chebyshev, Chernoff)
• Continuous random variables; exponential and normal, central limit theorem
• Applications: average case vs random algs, hashing, fingerprinting, load balancing, entropy and data compression

Statistics (3 wks)
• Parameter estimation: confidence intervals, bias; maximum likelihood: binomial, normal, EM
• Hypothesis Testing: likelihood ratio, t-test, contingency tables & chi-squared test?
• Monte-Carlo simulation, polling and sampling?
• Bayesian estimation, Bayes classifier, machine learning
• How to lie with statistics
some example CSE applications

- Performance analysis: “events” happen randomly: unpredictable failures, unpredictable arrival of data, varying workloads, ...

- “Knowledge discovery,” data mining, AI, ...
  statistical descriptions of patterns in data

- Scientific data analysis: measurement errors and artifacts

- Uncertainty: navigation and control with noisy sensors, ...

- Algorithm design and analysis: sometimes a randomized approach is simpler or better than any known deterministic one.
Read the paper, listen to the news, surf the web. You’ll be bombarded with probability and statistics – most phrased to bias the conclusion they hope you will draw.

Defend yourself!