UMBC

AN HONORS UNIVERSITY IN MARYLAND

IS 709/809: Computational Methods in IS Research Spring 2016

Course Logistics

Nirmalya Roy

Department of Information Systems

University of Maryland Baltimore County

Welcome to IS 709/809

- Timings: Wednesday; 4:30pm to 7:00pm
- Location: Sherman Hall 210
- Instructor: Nirmalya Roy

  Faculty in IS,

  MS in CSE: UT-Arlington, 2004

  PhD in CSE: UT-Arlington, 2008

  Postdoc in ECE: UT-Austin, 2010

  Faculty at Washington State University, 2013

  Research Interests: Mobile, Pervasive and Ubiquitous Computing (MPSC)

- Office hours: Monday, 1:30 - 3pm or by appointment

  Email: nroy@umbc.edu

  Office: ITE 421

Welcome to IS 709/809

- Course website
  - http://mpsc.umbc.edu/is-809compmethods/
  - Course related information will be posted on the website
  - Please check the course website frequently
- Prerequisite:
  - IS 650 or IS 733
- Make up classes
  - Will be occasionally necessary due to travel

Welcome to IS 709/809

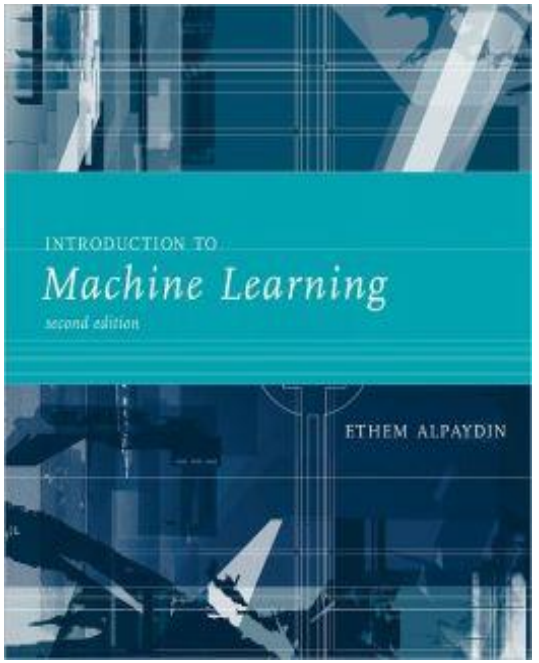
- Grading:
  - Participation/Presentation: 10%
  - Homeworks, Paper reviews, Quizzes, & Programming Assignments: 30%
  - 1 Mid-term exam: 30%
  - Research & Development Project: 30%

Course Expectations

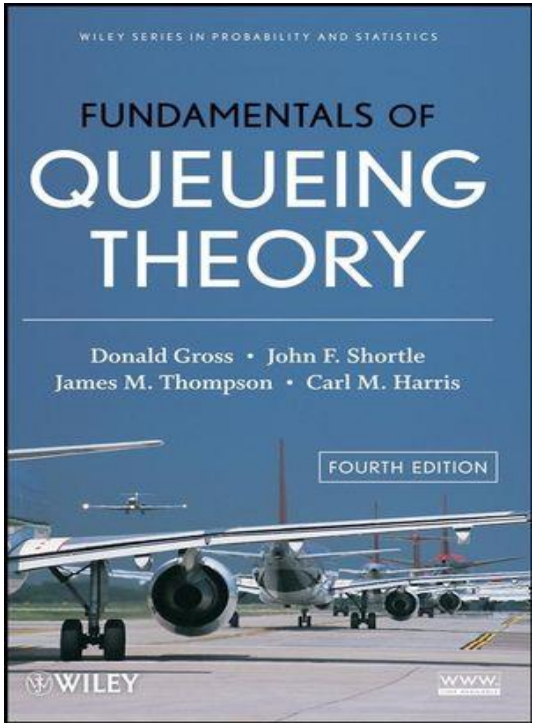
- Attendance
  - You should attend class
  - Lecture notes will be made available, but they should not be considered a substitution for attending class
- Collaboration
  - Collaboration is encouraged in general but do not copy from each other

Course Information

- Course materials:
  - Textbooks (Optional):
    Introduction to Machine Learning, Second Edition, by Ethem Alpaydin, MIT Press, 2010

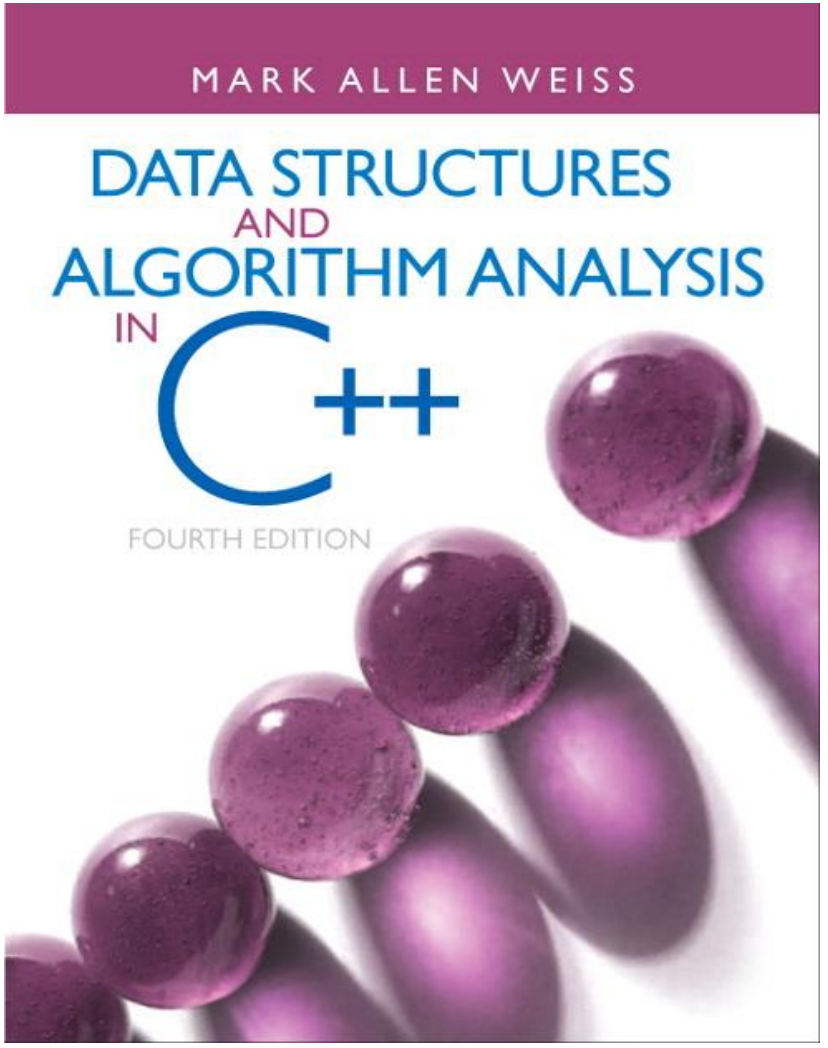


Fundamentals of Queueing Theory, 4th Ed., by Donald Gross & John F. Shortle & James M. Thompson & Carl M. Harris. John Wiley & Sons, Inc, 2008



Course Information

- Textbooks (Optional):
  - *Data Structures and Algorithm Analysis in C++, 4th Edition* by Mark Allen Weiss, 2013
- Class notes/slides
- Supplementary reading materials
  - Research Papers



What is this course about?

- Graduate level course in computational methods
  - MS and PhD students

- Learn principles of algorithm analysis
- Learn fundamentals of computational complexity
- Learn about data science
- Machine learning algorithms (e.g., supervised and unsupervised)
- Basic statistical modeling (e.g., linear and non-linear regression)
- Learn performance metrics of system
- Information systems performance & evaluation as case study
- Basic techniques of systems performance modeling
- Learn how to find an interesting IS research problem

Course Offerings

- At the end of the course
  - You understand variety of concepts …….
  - Mathematical modeling and analyzing a system
  - Exploratory and objective data analysis methods
  - Statistical methods and their computational implementation
  - Sensor networks based data mining and machine learning techniques
  - Building real systems
  - Tackling a research problem
  - …

Paper Presentation

- Class presentation
  - Choose a paper related to your tentative research project
  - You present one paper (30 minutes)
    - Deadline for selecting a paper is February 17, 2016
  - We will schedule at least one/two research paper presentation at the end of the lecture on each day

- Email me the title of the paper, authors list and the venue where it has been published
  - Don’t worry about not knowing the topic
    - Read the paper and you will understand the main concepts eventually!

Selecting the Papers

- Select a paper from a top computing or system or machine learning or data mining conference (pervasive, ubiquitous, mobile or data computing, green computing)
- Name of the good conferences and workshops in broad area of computing:
  - IEEE ICDM, ICDE, ICML
  - ACM Ubicomp, MobiSys, CHI, PerCom
  - ACM SenSys, IPSN, BuildSys
  - International Symposium on Wearable Computers (ISWC)

Course Research Projects

- Examples of projects
  - Internet of Things
    - RESTful API, IoTivity project
    - Service-Oriented Architecture (web service, OSGi, DPWS etc.)
  - Collaborative Opportunistic Sensing
    - Opportunity Project - Activity and Context Recognition with Opportunistic Sensor Configurations
  - Mobile phone based heath (mHealth)
  - Data to knowledge to decision
  - Wireless sensor networks
    - Human activity recognition (PerCom, MobiSys, Ubicomp etc.)
    - Building energy analytics (ACM BuildSys, SenSys etc.)
    - Wireless health (http://www.wirelesshealth2015.org/, ISWC etc.)
    - Smart health and big data (IEEE CHASE, ICDM, ICDE etc.)

Course Research Projects

- Projects consist of 3 parts:
  - Choosing an interesting topic
    - Identifying what new you can do
  - Proposing your novel ideas
    - Designing or modeling the solution
  - Performance evaluation
    - Testbed development and data collection
    - Prototype Implementation

Possible Research Project Ideas

- Mobile Phone and Wearable Sensor based Collaborative Activity Recognition Framework
- SmartQueue: Collaborative Opportunistic Sensing
- SenseTalk: Mobile Phone and Ambient Sensor based Conversation Detection for People Centric Applications
- StayFit: Group based Exercising using Sensor and Mobile Phones
- Control Diet: Keeping an eye on your diet
- StressSense: Measuring stress level using smart wristband

Possible Research Project Ideas

- Am I old? Mobile phone based Virtual Age Recognition
- Driving behavior and Potholes detection using smartphone and wristband
- Gesture Sense: Controlling Smart Phone from the air using depth sensor map: Is this natural interaction better than touch?
- Are you a Chain Smoker? Smart wristband based detection and intervention
- ……

Testbed Development Projects & Platforms

- Smart Home in a Box (SHiB)
  - ADL monitoring toolkit in smart homes at large scale
- Tweet-a-Watt
  - Build a wireless home-power monitoring system
- Microsoft HomeOS and openHAB
  - Enabling Smarter Homes for Community
- Shimmer GSR/Optical Pulse Development Kit
  - Stress Detection and Analysis
  - Affective Computing and Cognitive Factors Research
  - Emotional Engagement; Agitation Detection
  - Psychological Arousal like Mental Effort, Excitement, Shock etc.

Testbed Development Projects & Platforms

- ActiGraph
  - Sleep and Wellness Assessment
- Shimmer Sensing Platform
  - Wearable Sensing: Kinematics, motion, biophysical like cardio, EMG, GSR, Strain Gauge
- Fitbit
  - Fitness, ADL monitoring
- Microsoft Lab of Things
  - Connect devices home & beyond
  - Support HomeOS

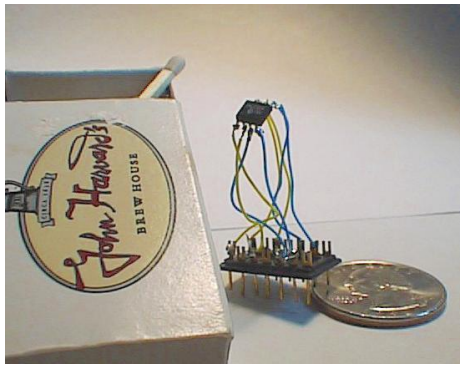
Testbed Development Projects & Platforms

- Waspmote from Libelium
  - Smart Cities (Smart parking, Smart lighting, Traffic congestion, Waste management etc.)
  - Smart Environments (Air pollution, Forest fire detection, Snow level monitoring etc.)
  - Smart Utility (Water or Gas leakages, Water quality, Energy consumption monitoring etc.)
  - eHealth (Fall detection, ADL monitoring etc.)
  - Internet of Things

“Cool” Internet Appliances

IP picture frame
http://www.ceiva.com/

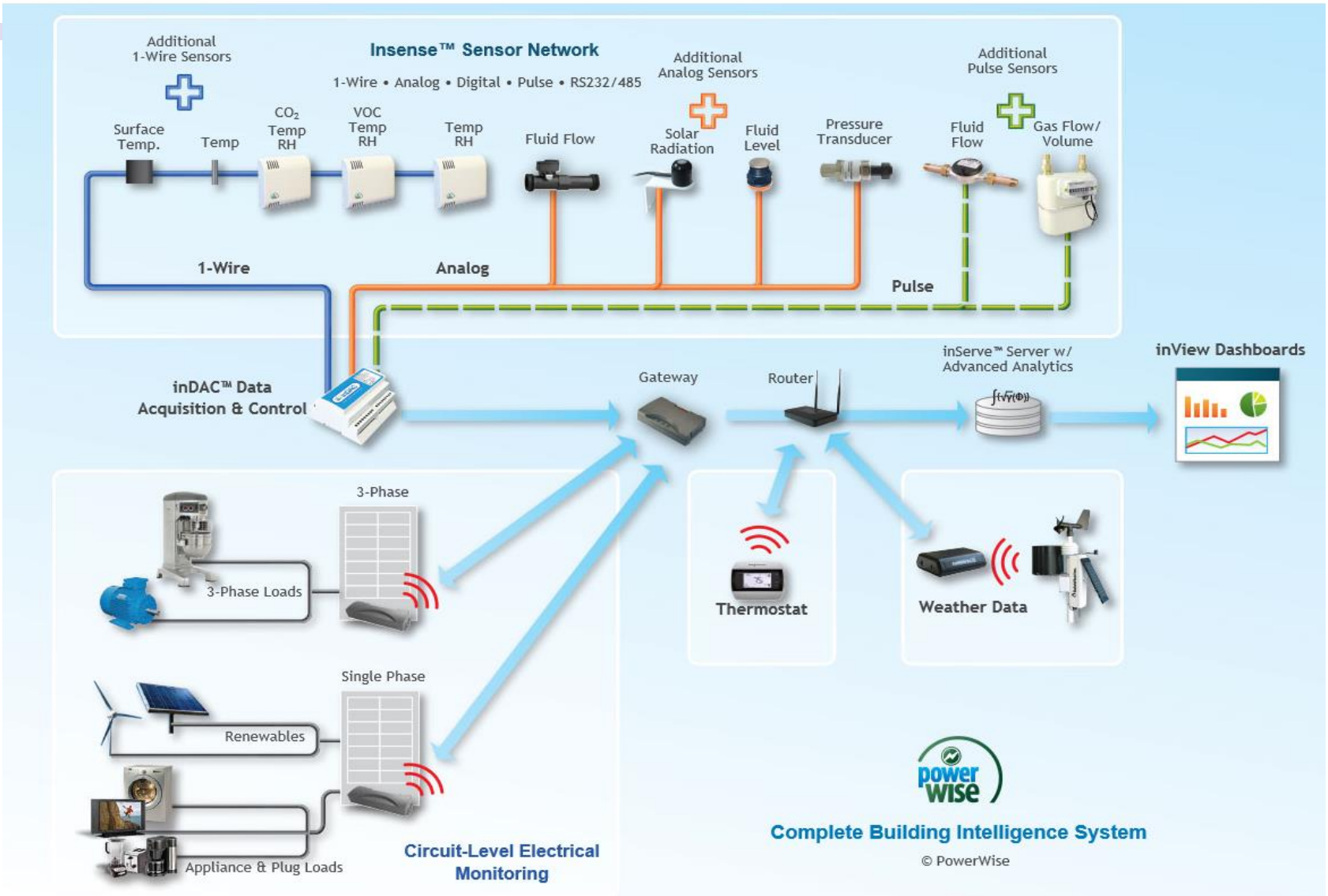
Internet phones



World’s smallest web server
http://www-ccs.cs.umass.edu/~shri/iPic.html

Web-enabled toaster +
weather forecaster

Energy Data Analytics: eMonitor



http://mpsc.umbc.edu/softwaredata/

Energy Analytics

- Appliance Energy Profiling Database Creation
  - building a dynamic catalog of the types and number of devices connected by a consumer
- Development of a Web-based Energy Management Dashboard
  - energy consumer can create goals
  - select and commit to energy savings tasks and habit changes
  - link to utility efficiency and rebate programs
  - share experiences with others and track their success and reward potential

Smart Plugs

PowerLinc Modem - INSTEON USB Interface (Dual-Band)

iMeter Solo - INSTEON Power Meter (Plug-In)

Your Take on Testbed Development Project

- Come talk to me during my office hours
- Finalize the research idea and then look for the equipment you need to do a great development project!
- Testbed development project pitch on 02/24
  - Identify the platform and devices
  - List those devices with their price
  - Let the class know what you are proposing, why you need those devices, how you will integrate all and deploy to get data, and what's the final results you are expecting
- Feel proud of your testbed development idea
- Use this testbed for your research project

Wearable Sensor

- Samsung Gear: Smart Wristwatch

http://www.samsung.com/global/microsite/gear/

- Features:
  - Accelerometer, Gyroscope, Compass, Heart Rate monitor, Ambient Light sensor, UV sensor and Barometer.
  - Watch is able to connect directly to the internet, make phone calls and send SMS's without needing a phone
  - first wearable device to include Wi-Fi, Bluetooth and 3G connectivity

Wearable Sensor

- Intel Basis http://www.mybasis.com/
  - ultimate fitness and sleep tracker with up to 4 day battery life and water resistance
  - Sensors
    - Optical Heart Rate Sensor
    - Galvanic Skin Response
    - Skin Temperature
    - 3-Axis Accelerometer

Wearable Sensor

- Infra
  https://www.facebook.com/InfraVtechnology/
- Blood Glucose Monitoring
  - Pulse monitoring
  - Blood Pressure Monitor
  - ECG - Professional Grade Heart Health Monitoring
  - EEG - Nervous System Measuring
  - Kidney Function
  - Alerts and Triggers
  - Live Data Portal
  - Integration with Peripheral Sensors

Ambient Sensors

- iBeacon http://estimote.com/
  - tiny wireless sensors
  - attach to any location or object
  - broadcast tiny radio signals
  - smartphone can receive and interpret
  - location and context awareness applications

- Texas Instrument Sensor Tags
  - supports Bluetooth Smart, 6LoWPAN and ZigBee
  - low-power sensors such as light, microphone and magnetic sensors
  - http://www.ti.com/ww/en/wireless_connectivity/sensortag2015/?INTC=SensorTag&HQS=sensortag

Energy Devices

- *Z-wave Smart Metering and Communication:*
  - *Z-Wave Smart Energy Power Strip*
- *Insteon Energy Metering and Communication:*
  - *iMeter Solo - INSTEON Power Meter (Plug-In)*
  - *PowerLinc Modem - INSTEON USB Interface (Dual-Band)*
- *Enmetric System for Intelligent Plug load Management and Power Telemetry Communication*
  - *Enmetric PowerPort*
  - *Enmetric Wireless Bridge*

Energy/Green Devices

- *SiteStage (previously was known as eMonitor)*
  - *Powerhouse Dynamics: Energy Management System*
  - *SiteSage for Homes M-24h Energy Monitor (formerly eMonitor 4-24)*

- *The Energy Detective Electricity Monitor http://www.theenergydetective.com/*

- *Energy Hub http://www.energyhub.com/*

- *PeoplePower http://www.peoplepowerco.com/*

- *Nest Lab https://nest.com/*
  - *Nest Thermostat*
  - *Nest CO2 monitoring device*

More Devices

- DrinkMate: http://www.getdrinkmate.com/
- Amazon Echo: www.amazon.com/echo
- Nike Sensor
- Jawbone UP 3
- Microsoft Band

More Devices

- Actron CP9599 U-Scan
  - Be Safe: Monitoring the Driving Behavior and Road Condition
- Myo Armband
  - Voice through Motion

Disclaimer: All the devices will be provided by the Mobile, Pervasive and Sensor Computing (MPSC) Lab in the Information Systems department at UMBC. We heartily acknowledge our sponsors Constellation Energy to make the research project activity possible in Spring 2016.

Let’s get back to our business!!

Computational Methods

Computational Methods

Introduction to Algorithm Analysis,

Introduction to System Modeling and

Introduction to Machine Learning

Algorithms

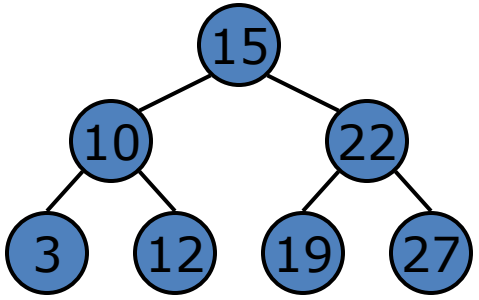
- Computational Complexity
- Runtime computation of several sorting algorithms
- Graph Algorithms
  - Shortest paths; Network flow; Minimum spanning tree etc.

Why do we need Algorithms and Data Structures?

- "Why not just use a big array?"
- Example problem
  - Search for a number k in a set of N numbers
- Solution # 1: Linear Search
  - Store numbers in an array of size N
  - Iterate through array until find k
  - Number of checks
    - Best case: 1 (k=15)
    - Worst case: N (k=27)
    - Average case: N/2

Data Structures

- Solution # 2: Binary Search Tree (BST)
  - Store numbers in a binary search tree
    - Requires: Elements to be sorted
  - Properties:
    - The left subtree of a node contains only nodes with keys less than the node's key
    - The right subtree of a node contains only nodes with keys greater than the node's key
    - Both the left and right subtrees must also be binary search trees
  - Search tree until find k
  - Number of checks
    - Best case: 1 (k=15)
    - Worst case: $\log_2 N$ (k=27)
    - Average case: $(\log_2 N) / 2$

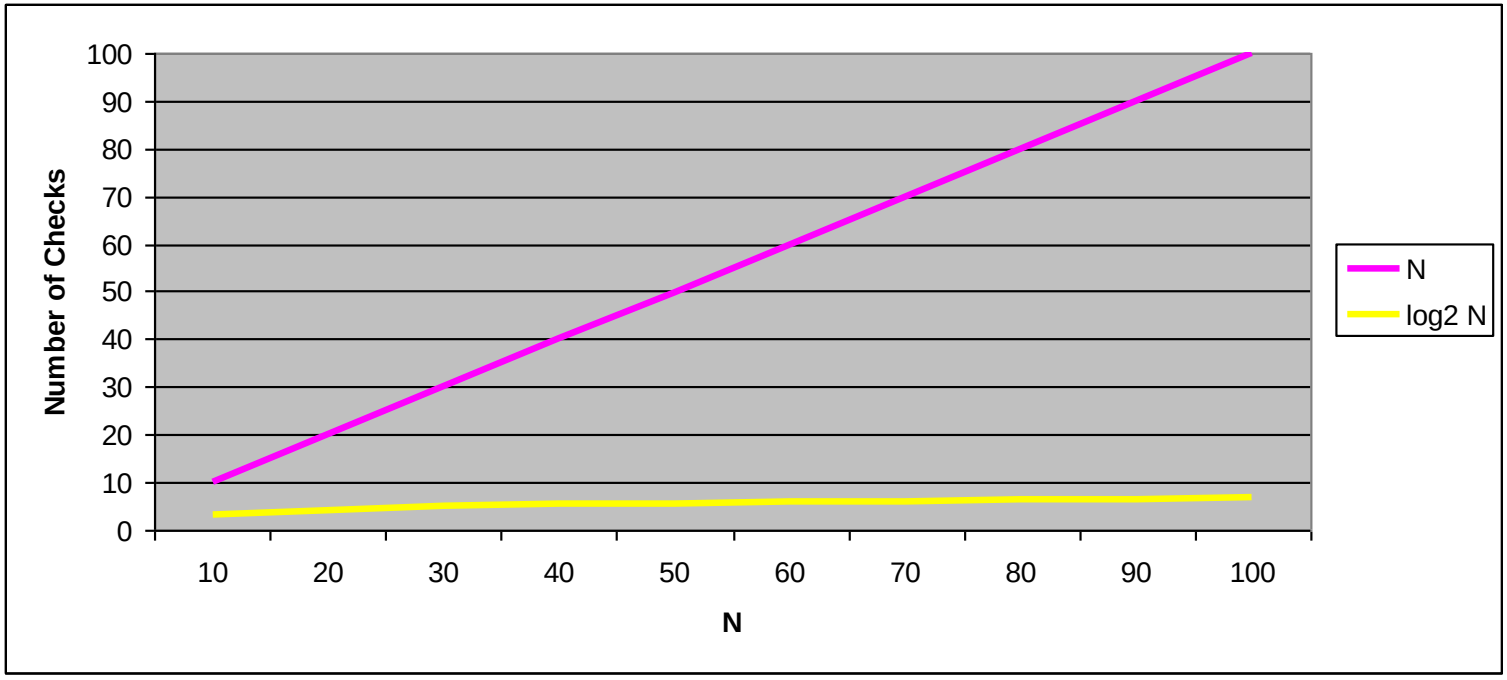


Example

- Does it matter?
- Problem Artifacts
  - N = 1,000,000,000
    - 1 billion (Walmart transactions in 100 days)
  - 1 Ghz processor = 10^9 cycles per second
- Solution #1 (assume 10 cycles per check)
  - Worst case: 1 billion checks = 10 seconds
- Solution #2 (assume 10 cycles per check)
  - Worst case: 30 checks = 0.0000003 seconds

Computational Complexity & Analysis

- Does it matter?
  - N vs. ($\log_2$ N)



Insights

- Moral
  - Appropriate data structures and algorithms ease design and improve performance
- Challenge
  - Design appropriate data structure and associated algorithms for a problem
  - Analyze to show improved performance

Why do we need Math for algorithm analysis?

- Analyzing data structures and algorithms
  - Deriving formulae for time and memory requirements
  - Will the solution scale?
- Proving algorithm correctness

Computational Methods (contd.)

System Performance Measurements

Systems

- Performance evaluation and modeling
- Concepts and techniques needed to plan the capacity of computer/information systems
  - predict their future performance under different configurations
  - design new applications that meet performance requirements
  - analytic queuing network models of computer systems
  - study the performance of centralized, distributed, parallel, client/server systems, web server and e-commerce site performance
- Database systems, mobile systems, networked systems
  - Telecommunication network design

Queueing Theory

- Waiting in lines
  - in the grocery store, on the telephone, at the airport, on the road
- Queueing theory is the mathematical study of lines
  - What are the stochastic characteristics of delay?
  - For example, what is the average delay?
  - What is the probability that delay exceeds some threshold?
  - What fraction of customers are turned away?
  - What system capacity (e.g., what number of servers) is needed to achieve a specified quality of service?
- Provide decision makers a way to efficiently allocate resources to reduce delay

Applications of Queueing Theory

- Applications to operations research, management science and industrial engineering
- Examples are
  - Traffic flow (vehicle, aircraft, people communication)
  - Scheduling (patients in hospital, jobs on machines, programs on a computer)
- Facility design
  - Banks, post offices, amusement parks, fast-food restaurants

Applications of Queueing Theory

- Applications to Networks
- Study of the performance of systems composed of
  - Waiting lines
  - Processing units
- Allows to estimate
  - Time spent in waiting
  - Expected number of waiting requests
  - Probability of being in certain states
- Useful for the design of systems such as telecommunication networks
  - Delay, blocking probability, links, bandwidth, number of processors, buffers size

In this course

- Survey the quantitative models used to analyze queueing systems

- Focus will be both on mathematical analyses of such models as well as practical issues in using such models to represent real pervasive and networking problems

Computational Methods (contd.)

Data Science/
Machine Learning/
Statistics

Questions

?