Design and Analysis of Algorithms

Course Introduction

Si Wu

School of CSE, SCUT

cswusi@scut.edu.cn

TA: 1684350406@qq.com

Outline

- Part I: About the course
- Part II: About algorithms
  - What are algorithms?
  - Why are they important to study?

Part I: About the course

Course Information

- Lecturer:

  Prof. Si Wu (吴斯)

  E-mail: cswusi@scut.edu.cn

  Office: B3-302

- Teaching Assistants：

  Mr. Wen Xue(薛文)

  E-mail: 1684350406@qq.com

  Office: B3-535

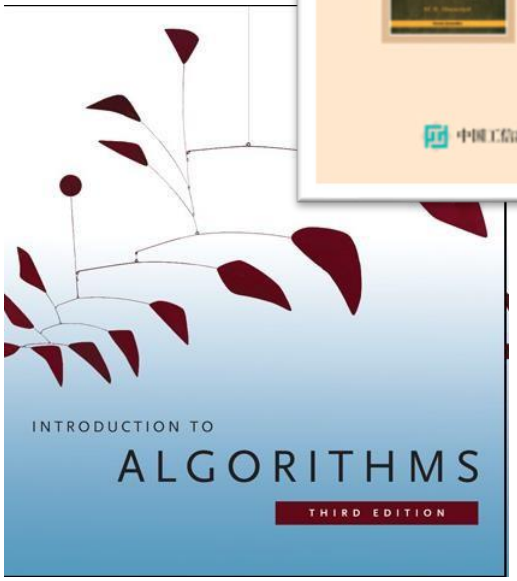
If you have any questions, please feel free to contact me by email. Please list your name and student ID when you send me an email….

Course Information

- Reference

Algorithms Design Techniques and Analysis.
(Saudi Arabia) M. H. Alsuwaiyel.
Publishing House of Electronic Industry.

Introduction to Algorithms, 3rd ed.
T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein, *MIT Press*.



Main Topics

- Algorithm Analysis
- Sorting algorithms
- Recurrence
- Divide and Conquer
- Dynamic Programming
- Greedy Algorithms
- Linear Programming
- Network Flow
- Approximation
- P & NP Problems

Course Information

- **Couse Time**: Weeks 3-14, 64 Teaching hours

  Lecture: Weeks 3-14
  (Tuesday 8:50am-10:25am), **A1306**

  Lab: TBD

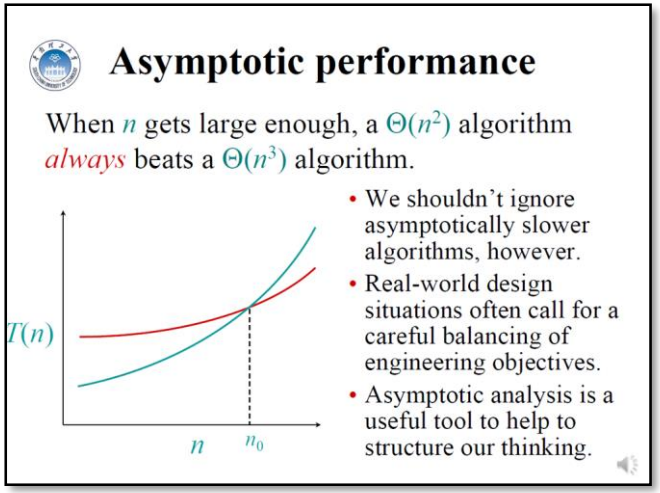
- **Final Grade**:

  Performance + Experiments (30%)

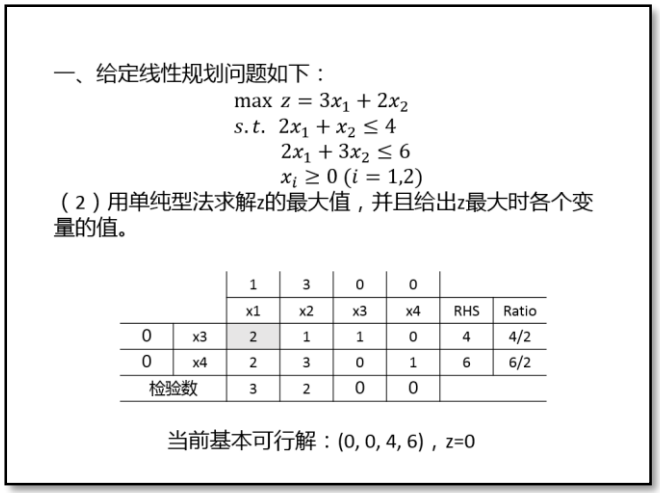
  Final Examination (70%)

Lecture

- **Teaching Session**



- **Tutorial Session:**



		1	3	0	0		
		x1	x2	x3	x4	RHS	Ratio
0	x3	2	1	1	0	4	4/2
0	x4	2	3	0	1	6	6/2
检验数		3	2	0	0		

Course Information

- Online Judge:
  - http://www.scut.edu.cn/ACM/
    (South China University of Technology)
  - http://acm.zju.edu.cn/onlinejudge/
    (Zhejiang University)
  - http://poj.org (Peking University)

About the Flavor

- It's more of a *math flavor* than a *programming one*.
- You will need to write pseudo-code, and implement it using C/C++...
- You will design and analyze, think and prove (rather than coding)

Prerequisites

- Officially:
  - Discrete Mathematics
  - Programming
  - Data Structures
- Effectively: Basic mathematical maturity
  - functions, polynomial, exponential;
  - proof by induction;
  - basic data structure operations (stack, queue, …);
  - basic math manipulations…

Experiment Policy

- Discussions and searching on the web are allowed in general
- But you have to implement the solution by yourself
- And you should fully understand your codes.

Zero Tolerance for Cheating/Plagiarism

- You may get 0 score for this course
- Will check your codes by software; scores of both the code provider and the copier will be 0 once the cheating/plagiarism behavior is confirmed

Suggestions

- In class:
  - Try to come on time.
  - Try your best to get more involved in the class.
  - Please don't chitchat.
  - Treat experiments seriously

Suggestions

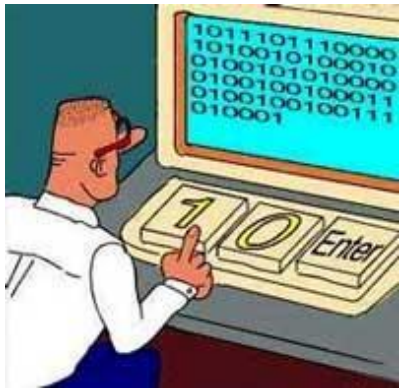
- Your suggestions will be highly appreciated.
  - Please send me/TA an e-mail

- Any questions about the course?
- My questions:
  - What are your goals?
  - What do you like to learn from this course?
  - What excite you the most in general?

Part II: About algorithms

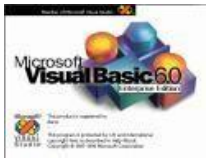
Factors of Programming

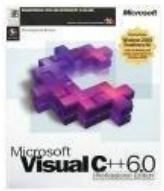
- **Programming Languages?**



BASIC PASCAL C 程序设计语言

C++ JAVA C#





Algorithms

- **Algorithm. (webster.com)**

  -A well-defined computational procedure that takes some value, or set of values, as input and produces some value, or set of values, as output.

  -Broadly: a step-step procedure for solving a problem or accomplishing some end especially by a computer.

  -Issues: correctness, efficiency (amount of work done and space used), storage, simplicity, clarity, optimality, etc.

Why study algorithms?

Importance of Algorithms

- **Problem: sorting 10,000,000 integers**
  - Case 1: Computer A executes one billion instructions per second(1GHz), an algorithm taking time roughly equal to $2n^2$ to sort n integers.
  - Case 2: Computer B executes one hundred million instructions per second(100MHz), an algorithm taking time roughly equal to 50nlogn to sort n integers.

- Case 1:

$$\frac{2\times(10^7)^2\,\text{instructions}}{10^9\,\text{instructions/second}} = 200000\text{seconds} \approx 55\text{hours}$$

- Case 2:

$$\frac{50\times10^7\times\log10^7\,\text{instructions}}{10^8\,\text{instructions/second}} = 105\text{seconds}$$

Importance of Algorithms

		$1.3\ N^3$	$10\ N^2$	$47\ N \log_2 N$	$48\ N$
Time to solve a problem of size	1000	1.3 seconds	10 msec	0.4 msec	0.048 msec
	10,000	22 minutes	1 second	6 msec	0.48 msec
	100,000	15 days	1.7 minutes	78 msec	4.8 msec
	million	41 years	2.8 hours	0.94 seconds	48 msec
	10 million	41 millennia	1.7 weeks	11 seconds	0.48 seconds
Max size problem solved in one	second	920	10,000	1 million	21 million
	minute	3,600	77,000	49 million	1.3 billion
	hour	14,000	600,000	2.4 billion	76 billion
	day	41,000	2.9 million	50 billion	1,800 billion
N multiplied by 10, time multiplied by		1,000	100	10+	10

Information Explosion

- 988EB (1EB = 1024PB) data will be produced in 2010 (IDC) ⇔ 18 million times of all info in books
- IT
  - 850 million photos & 8 million videos every day (Facebook)
  - 50PB web pages, 500PB log (Baidu)
- Public Utilities
  - Health care (medical images - photos)
  - Public traffic (surveillance - videos)
- …

Research Frontier and Hot

- 《Science》: Special Online Collection: Dealing with Data
  - In this, *Science* joins with colleagues from *Science Signaling*, *Science Translational Medicine*, and *Science Careers* to provide a broad look at the issues surrounding the increasingly huge influx of research data. This collection of articles highlights both the challenges posed by the data deluge and the opportunities that can be realized if we can better organize and access the data.

11 FEBRUARY 2011 VOL 331
SCIENCE
www.sciencemagazine.org

- 《Nature》:

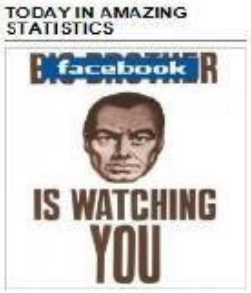
Big data, but are we ready?

Oswaldo Trelles, Pjotr Prins, Marc Snir and Ritsert C. Jansen

Data!! Data!! Data!!

- 人类共拍了35000亿张照片
- 其中1400亿在Facebook上，占据4%
- Facebook已经成为世界上最大的图像数据库

4% of All Photos Ever Taken Are On Facebook



So this is why people are intensely creeped out by Facebook's facial recognition database.

In a historical overview of consumer photography from the Kodak Brownie to Instagram, the photo-sharing site 1000memories' company blog estimated that humans have taken 3.5 trillion pictures so far -- and Facebook's servers host 140 billion of them. That's 4% of all photos taken ... *ever*.

Pretty incredible for a social network that's only been around since 2003. But it makes sense when you consider just how central Facebook is to online life, and the sheer volume of photos people take today.

With smartphone adoption growing like gangbusters since the iPhone's introduction in 2007, it's increasingly the norm to have a decent camera in your pocket at all times. And with flash memory prices at an all-time low (remember when you had to watch to make sure you wouldn't fill up the 128-megabyte card on your dedicated digital camera?), trends like taking photos of all one's meals have become convenient, if cringe-inducing.

Sure, Instagram's growth has been impressive, and new photo-sharing mobile apps launch seemingly every week. But as Facebook continues to dominate social networking -- recent estimates put its membership at 750 million, up 50% from the summer of 2010 -- it's hard to imagine any of the startups as more than a speed bump in Facebook's race to control the world's personal data.

As this infographic from 1000memories shows, compared to its nearest rivals in the photo-hoarding universe -- Flickr, Instagram, and the downright puny Library of Congress -- Facebook is the Borg unicomplex, and resistance is futile:

THE WORLD'S LARGEST PHOTO LIBRARIES

Various Problems

- Human Genome Project
  - 100,000 genes, sequences of the 3 billion chemical base pairs
- Internet
  - Finding good routes on which the data will travel
  - Search engine
- Electronic commerce
  - Public-key cryptography and digital signatures
- Manufacturing
  - Allocate scarce resources in the most beneficial way
- …

About the Course

- **Design and Analysis**
  - How can I propose an algorithm for a specific problem?
  - Is the algorithm good enough?

Analysis of Algorithms

- **The theoretical study of computer-program performance and resource usage.**

- **What's more important than performance?**

  - modularity
  - correctness
  - maintainability
  - functionality
  - robustness
  - user-friendliness
  - programmer time
  - simplicity
  - extensibility
  - reliability