

BBM 202 - ALGORITHMS



HACETTEPE UNIVERSITY

DEPT. OF COMPUTER ENGINEERING

ERKUT ERDEM

INTRODUCTION, ANALYSIS OF ALGORITHMS

Mar. 7, 2013

Acknowledgement: The course slides are adapted from the slides prepared by R. Sedgwick and K. Wayne of Princeton University.

TODAY

- ▶ **Introduction**

- ▶ Why study algorithms?
- ▶ Coursework
- ▶ Resources
- ▶ Outline

- ▶ **Analysis of Algorithms**

- ▶ observations
- ▶ mathematical models
- ▶ order-of-growth classifications
- ▶ dependencies on inputs
- ▶ memory

INTRODUCTION

- ▶ **Introduction**
- ▶ Why study algorithms?
- ▶ Coursework
- ▶ Resources
- ▶ Outline

Instructor and Course Schedule

- Dr. Erkut ERDEM
- erkut@cs.hacettepe.edu.tr
- Office: 114
- Tel: 297 7500 / 149

- Lectures: Thursday, 13:00-15:45@D2 (Section 2)
- Practicum (BBM204): Friday, 15:00-16:45@D4 (Section 3-5)

About BBM202-204

- This course concerns programming and problem solving, with applications.
- The aim is to teach student how to develop algorithms in order to solve the complex problems in the most efficient way.
- The students are expected to develop a foundational understanding and knowledge of key concepts that underly important algorithms in use on computers today.
- The students are also be expected to gain hand-on experience via a set of programming assignments supplied in the complementary BBM 204 Software Practicum.

Why study algorithms?

Their impact is broad and far-reaching.

Internet. Web search, packet routing, distributed file sharing, ...

Biology. Human genome project, protein folding, ...

Computers. Circuit layout, file system, compilers, ...

Computer graphics. Movies, video games, virtual reality, ...

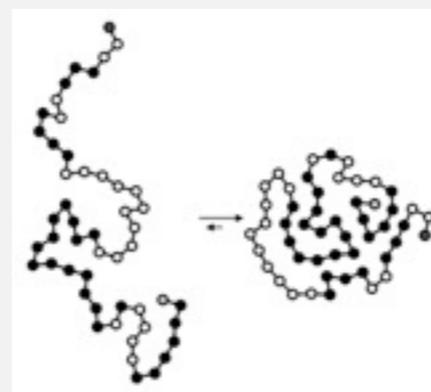
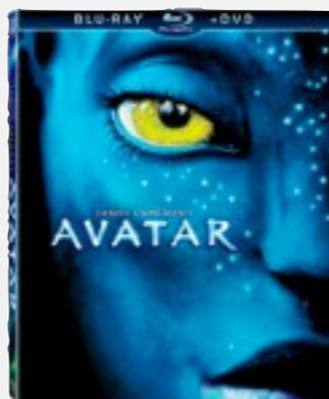
Security. Cell phones, e-commerce, voting machines, ...

Multimedia. MP3, JPG, DivX, HDTV, face recognition, ...

Social networks. Recommendations, news feeds, advertisements, ...

Physics. N-body simulation, particle collision simulation, ...

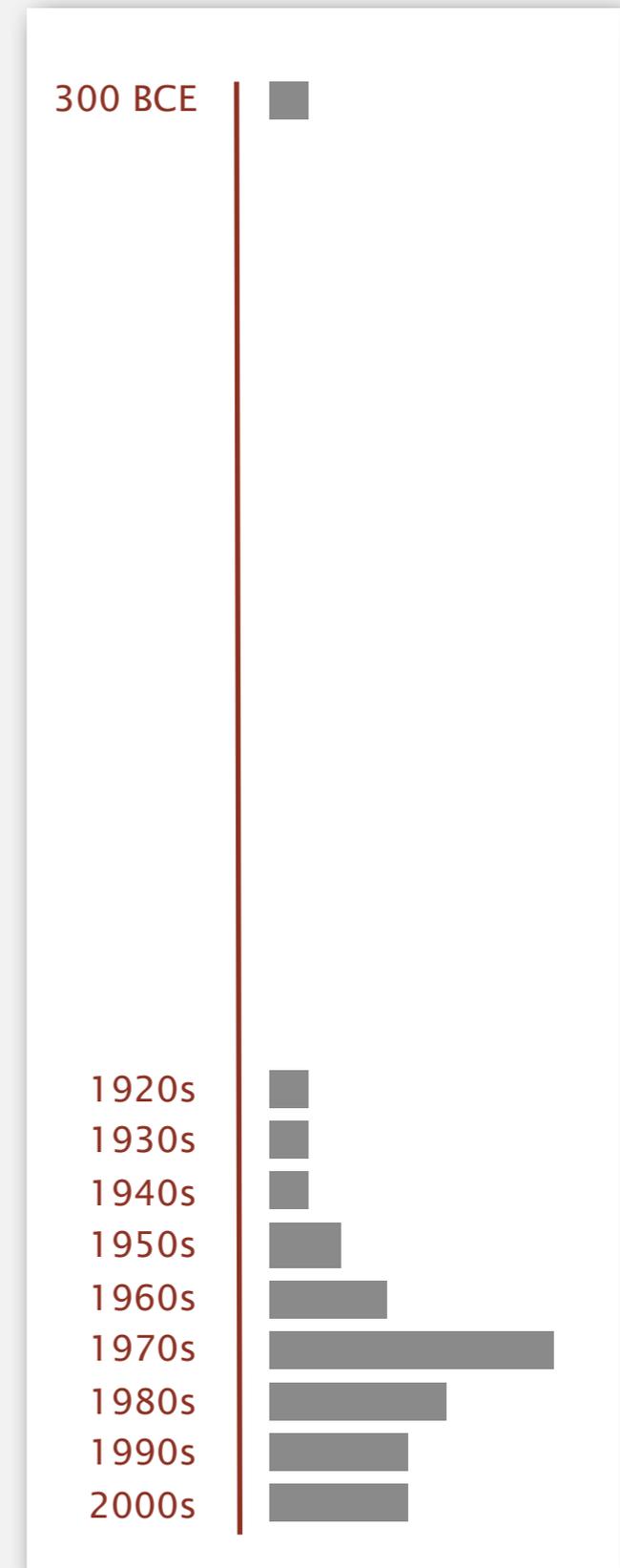
⋮



Why study algorithms?

Old roots, new opportunities.

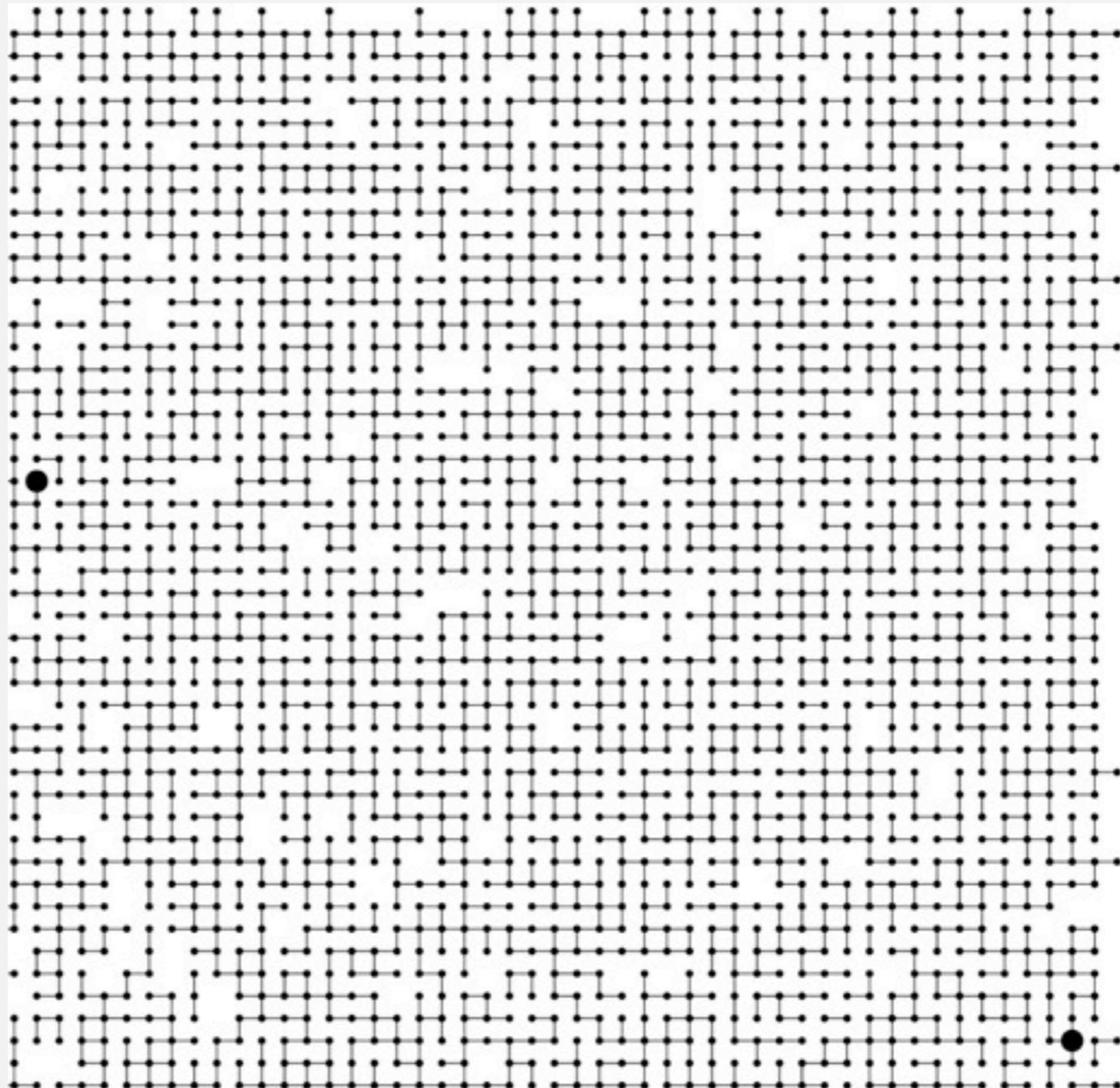
- Study of algorithms dates at least to Euclid.
- Formalized by Church and Turing in 1930s.
- Some important algorithms were discovered by undergraduates in a course like this!



Why study algorithms?

To solve problems that could not otherwise be addressed.

Ex. Network connectivity.

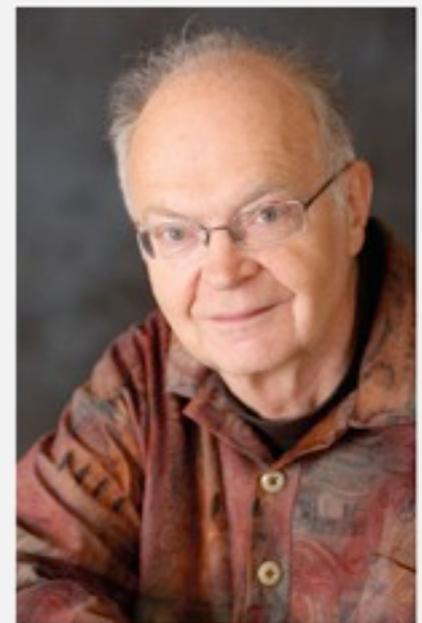
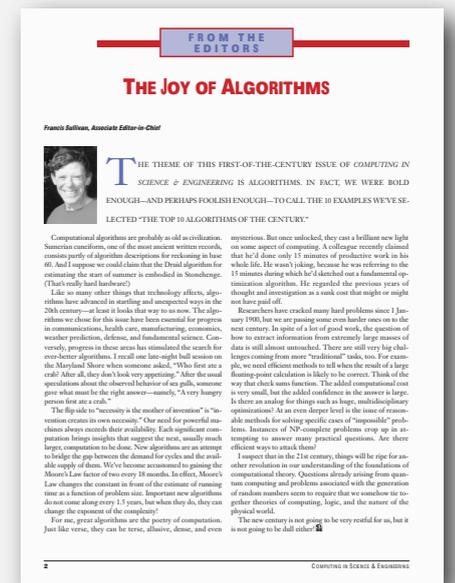


Why study algorithms?

For intellectual stimulation.

“ For me, great algorithms are the poetry of computation. Just like verse, they can be terse, allusive, dense, and even mysterious. But once unlocked, they cast a brilliant new light on some aspect of computing. ” — Francis Sullivan

“ It has often been said that a person does not really understand something until he teaches it to someone else. Actually a person does not really understand something until he can teach it to a computer, i.e. express it as an algorithm. The attempt to formalise things as algorithms lead to a much deeper understanding than if we simply try to comprehend things in the traditional way. algorithm must be seen to be believed. ” — Donald Knuth

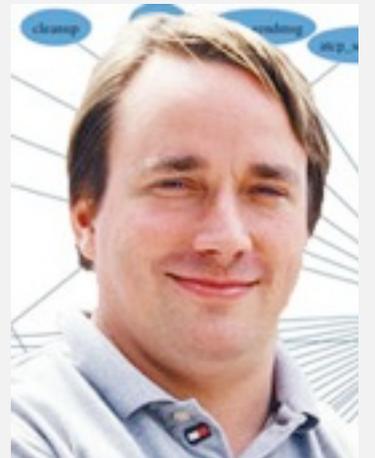


Why study algorithms?

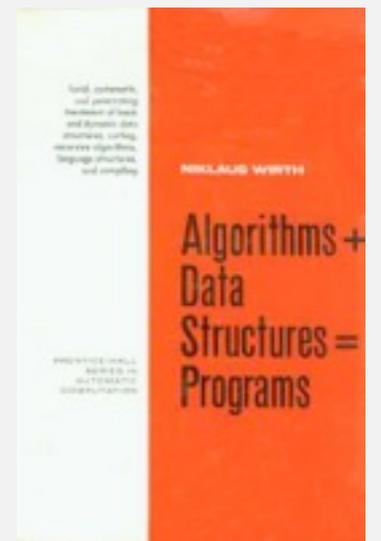
To become a proficient programmer.

“ I will, in fact, claim that the difference between a bad programmer and a good one is whether he considers his code or his data structures more important. Bad programmers worry about the code. Good programmers worry about data structures and their relationships. ”

— Linus Torvalds (creator of Linux)



“ Algorithms + Data Structures = Programs. ” — Niklaus Wirth



Why study algorithms?

They may unlock the secrets of life and of the universe.

Computational models are replacing mathematical models in scientific inquiry.

$$\begin{aligned} E &= mc^2 \\ F &= ma \end{aligned} \quad F = \frac{Gm_1m_2}{r^2}$$
$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(r) \right] \Psi(r) = E \Psi(r)$$

20th century science
(formula based)

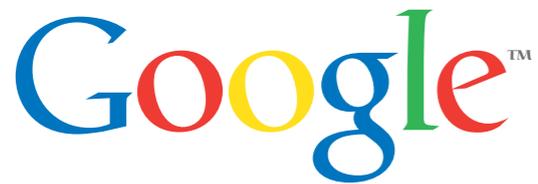
```
for (double t = 0.0; true; t = t + dt)
  for (int i = 0; i < N; i++)
  {
    bodies[i].resetForce();
    for (int j = 0; j < N; j++)
      if (i != j)
        bodies[i].addForce(bodies[j]);
  }
```

21st century science
(algorithm based)

“ Algorithms: a common language for nature, human, and computer. ” — Avi Wigderson

Why study algorithms?

For fun and profit.



Apple Computer

facebook

CISCO SYSTEMS



Nintendo

IBM



Morgan Stanley

NETFLIX



DE Shaw & Co

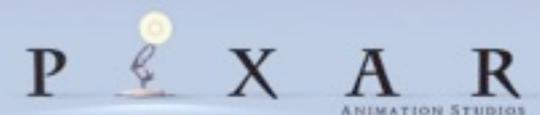
ORACLE



YAHOO!

amazon.com

Microsoft



Why study algorithms?

- Their impact is broad and far-reaching.
- Old roots, new opportunities.
- To solve problems that could not otherwise be addressed.
- For intellectual stimulation.
- To become a proficient programmer.
- They may unlock the secrets of life and of the universe.
- For fun and profit.

Why study anything else?



Teaching Staff



Dr. Erkut Erdem

erkut@cs.hacettepe.edu.tr

Tel: 297 7500 / 149



Dr. Burcu Can

burcucan@cs.hacettepe.edu.tr

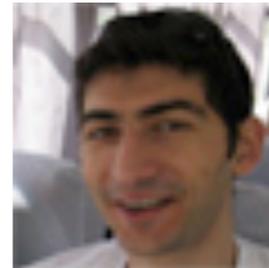
Tel: 297 7500 / 121



Pelin Aktaş

pelinaktas@cs.hacettepe.edu.tr

Tel: 297 7500



Yasin Şahin

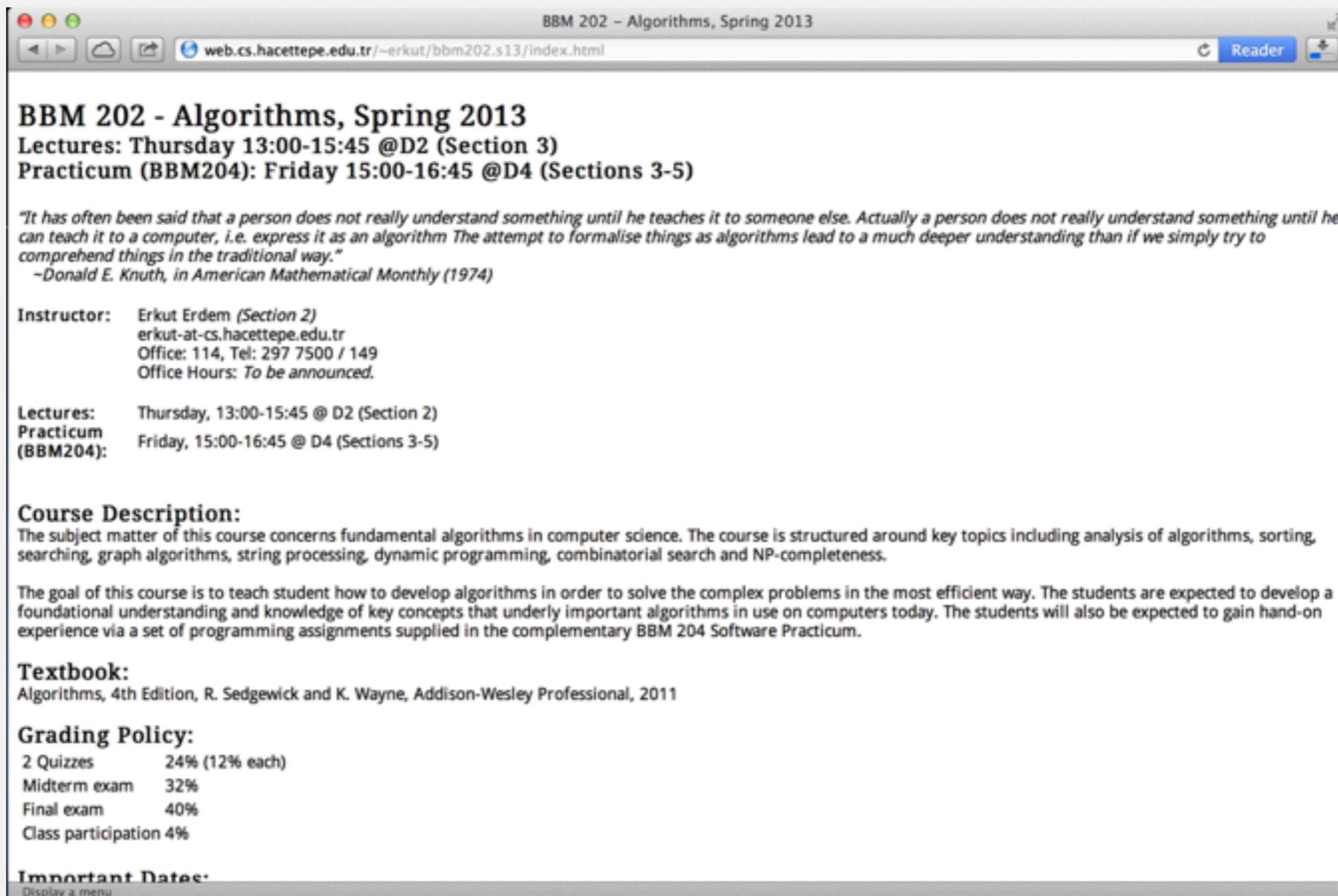
yasin@cs.hacettepe.edu.tr

Tel: 297 7500 / 123

Communication

- The course webpage will be updated regularly throughout the semester with lecture notes, programming assignments and important deadlines.

<http://web.cs.hacettepe.edu.tr/~erkut/bbm202.s13>



The screenshot shows a web browser window with the title "BBM 202 - Algorithms, Spring 2013". The address bar contains the URL "web.cs.hacettepe.edu.tr/~erkut/bbm202.s13/index.html". The page content includes the course title, lecture and practicum schedules, a quote by Donald E. Knuth, instructor information, course description, textbook information, and a grading policy table.

BBM 202 - Algorithms, Spring 2013
Lectures: Thursday 13:00-15:45 @D2 (Section 3)
Practicum (BBM204): Friday 15:00-16:45 @D4 (Sections 3-5)

"It has often been said that a person does not really understand something until he teaches it to someone else. Actually a person does not really understand something until he can teach it to a computer, i.e. express it as an algorithm. The attempt to formalise things as algorithms lead to a much deeper understanding than if we simply try to comprehend things in the traditional way."
-Donald E. Knuth, in American Mathematical Monthly (1974)

Instructor: Erkut Erdem (Section 2)
erkut-at-cs.hacettepe.edu.tr
Office: 114, Tel: 297 7500 / 149
Office Hours: To be announced.

Lectures: Thursday, 13:00-15:45 @ D2 (Section 2)
Practicum (BBM204): Friday, 15:00-16:45 @ D4 (Sections 3-5)

Course Description:
The subject matter of this course concerns fundamental algorithms in computer science. The course is structured around key topics including analysis of algorithms, sorting, searching, graph algorithms, string processing, dynamic programming, combinatorial search and NP-completeness.

The goal of this course is to teach student how to develop algorithms in order to solve the complex problems in the most efficient way. The students are expected to develop a foundational understanding and knowledge of key concepts that underly important algorithms in use on computers today. The students will also be expected to gain hand-on experience via a set of programming assignments supplied in the complementary BBM 204 Software Practicum.

Textbook:
Algorithms, 4th Edition, R. Sedgewick and K. Wayne, Addison-Wesley Professional, 2011

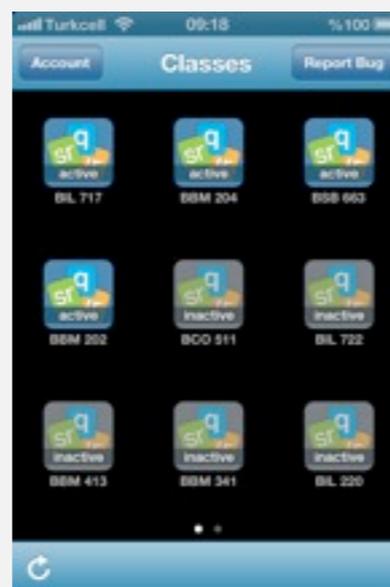
Grading Policy:

2 Quizzes	24% (12% each)
Midterm exam	32%
Final exam	40%
Class participation	4%

Important Dates:
Display a menu

Getting help

- Office Hours
 - See webpage for the schedule
- BBM204 Software Practicum
 - Course related recitations, practice with algorithms, etc.
- Communication
 - Announcements and course related discussions
 - through [PIAZZA](#):
BBM 202: <https://piazza.com/hacettepe.edu.tr/spring2013/bbm202>
BBM 204: <https://piazza.com/hacettepe.edu.tr/spring2013/bbm204>



Coursework and grading

Quizzes. 12% + 12%

- Two closed-book quizzes
- in class on Thursday, April 4 and Thursday, May 23

Midterm exam. 32%

- Closed-book
- in class on Thursday, April 25

Final exam. 40%

- Closed-book
- Scheduled by Registrar.

Class participation. 4%

- Contribute to Piazza discussions.
- Attend and participate in lecture.

BBM204 Software Practicum

Programming assignments (PAs)

- Five assignments throughout the semester.
- Each assignment has a well-defined goal such as solving a specific problem.
- You must work alone on all assignments stated unless otherwise.

Important Dates

- 25 Mar. 2013, PA1
- 08 Apr. 2013, PA2
- 29 Apr. 2013, PA3
- 13 May 2013, PA4
- 27 May 2013, PA5

Cheating

What is cheating?

- Sharing code: by copying, retyping, looking at, or supplying a file
- Coaching: helping your friend to write a programming assignment, line by line
- Copying code from previous course or from elsewhere on WWW

What is NOT cheating?

- Explaining how to use systems or tools
- Helping others with high-level design issues

Penalty for cheating:

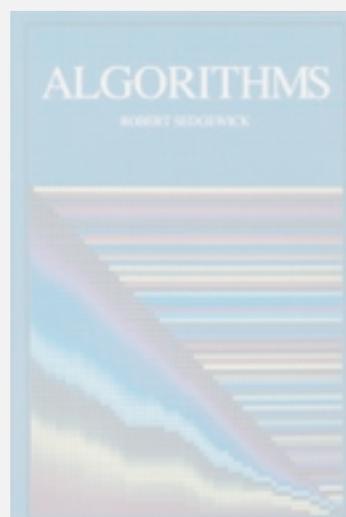
- Helping others with high-level design issues
- Removal from course with failing grade

Detection of cheating:

- We do check
- Our tools for doing this are much better than most cheaters think!

Resources (textbook)

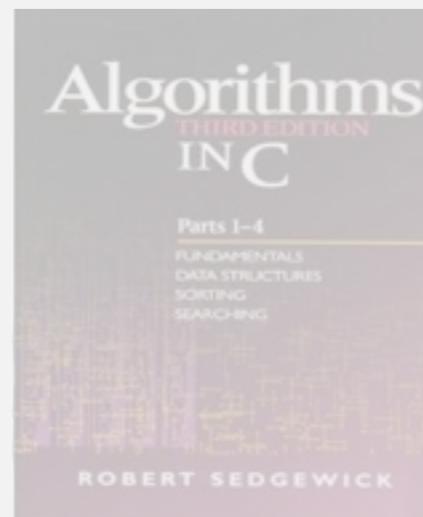
Required reading. Algorithms 4th edition by R. Sedgwick and K. Wayne, Addison-Wesley Professional, 2011, ISBN 0-321-57351-X.



1st edition (1982)



2nd edition (1988)



3rd edition (1997)



Booksite.

- Brief summary of content.
- Download code from book.

ALGORITHMS, 4TH EDITION

essential information that every serious programmer needs to know about algorithms and data structures

Textbook. The textbook *Algorithms, 4th Edition* by Robert Sedgwick and Kevin Wayne [[Amazon](#) · [Addison-Wesley](#)] surveys the most important algorithms and data structures in use today. The textbook is organized into six chapters:

- **Chapter 1: Fundamentals** introduces a scientific and engineering basis for comparing algorithms and making predictions. It also includes our programming model.
- **Chapter 2: Sorting** considers several classic sorting algorithms, including insertion sort, mergesort, and quicksort. It also includes a binary heap implementation of a priority queue.
- **Chapter 3: Searching** describes several classic symbol table implementations, including binary search trees, red-black trees, and hash tables.

Algorithms, 4th Edition

1. Fundamentals
2. Sorting
3. Searching
4. Graphs
5. Strings
6. Context

<http://www.algs4.princeton.edu>

Course outline

Introduction

Analysis of Algorithms (2 weeks)

- Computational Complexity
- Recursion, Recurrence Relations

Sorting (1 week)

- Review
- HeapSort

Searching (3 weeks)

- Sequential Search
- Binary Search Tree
- Balanced Trees
- Hashing,
- Search Applications

Course outline

Graphs (2 weeks)

- Undirected Graphs
- Directed Graphs
- Minimum Spanning Trees,
- Shortest Path

Strings (3 weeks)

- String Sorts, Tries
- Substring Search
- Regular Expressions
- Data Compression

Advanced Topics (2 weeks)

- Dynamic Programming
- Combinatorial Search, NP Completeness

ANALYSIS OF ALGORITHMS

- ▶ Observations
- ▶ Mathematical models
- ▶ Order-of-growth classifications
- ▶ Dependencies on inputs
- ▶ Memory

Cast of characters



Programmer needs to develop a working solution.



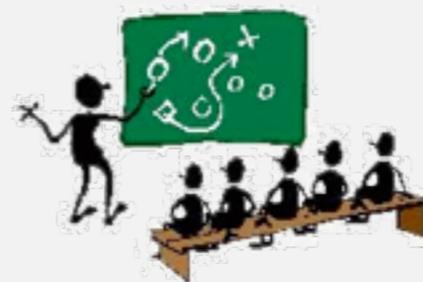
Student might play any or all of these roles someday.



Client wants to solve problem efficiently.



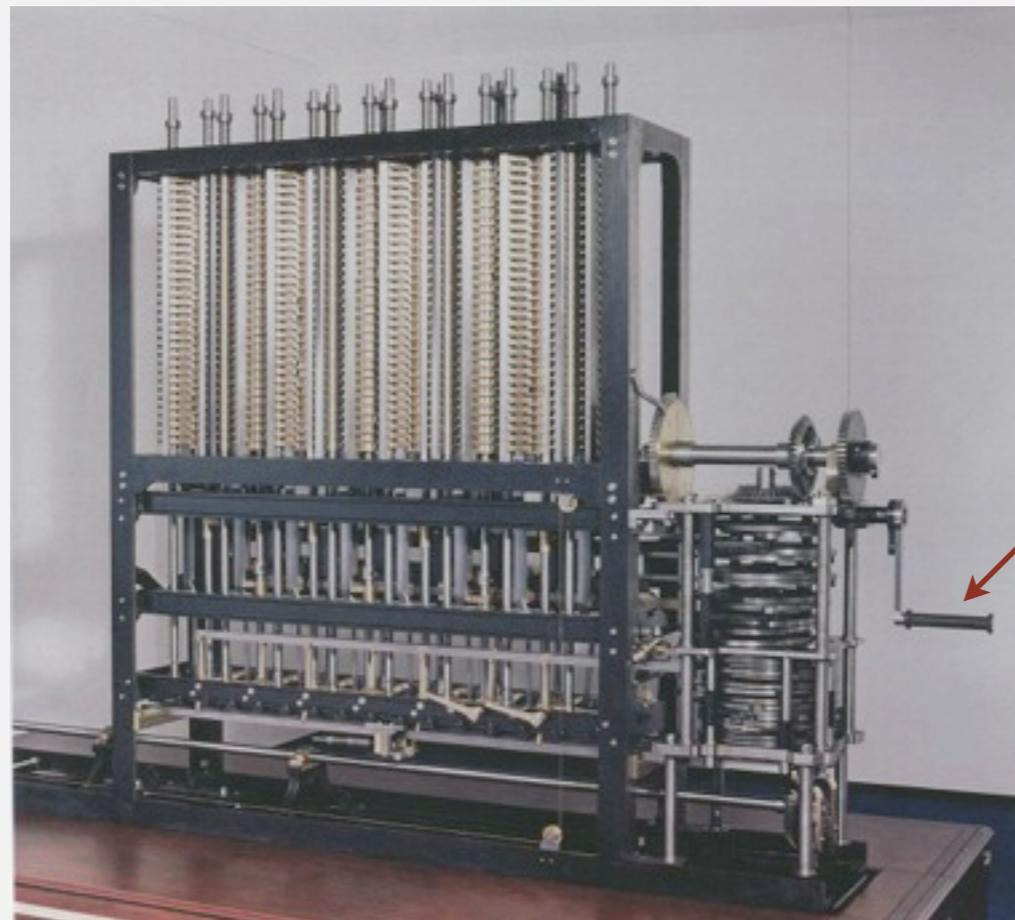
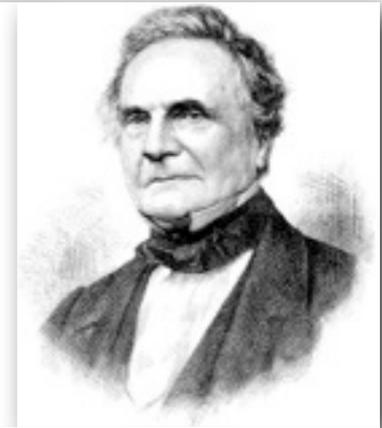
Theoretician wants to understand.



Basic **blocking and tackling** is sometimes necessary.
[this lecture]

Running time

“ As soon as an Analytic Engine exists, it will necessarily guide the future course of the science. Whenever any result is sought by its aid, the question will arise—By what course of calculation can these results be arrived at by the machine in the shortest time? ” — Charles Babbage (1864)



how many times do you have to turn the crank?

Analytic Engine

Reasons to analyze algorithms

Predict performance.

Compare algorithms.

Provide guarantees.

Understand theoretical basis.

this course (BBM 202)

Analysis of algorithms (BBM 408)

Primary practical reason: avoid performance bugs.



client gets poor performance because programmer did not understand performance characteristics



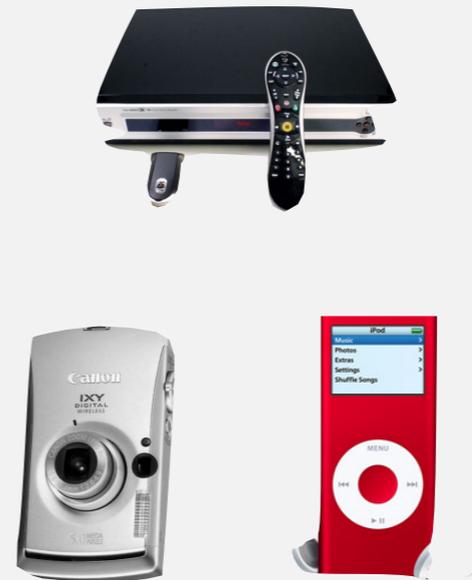
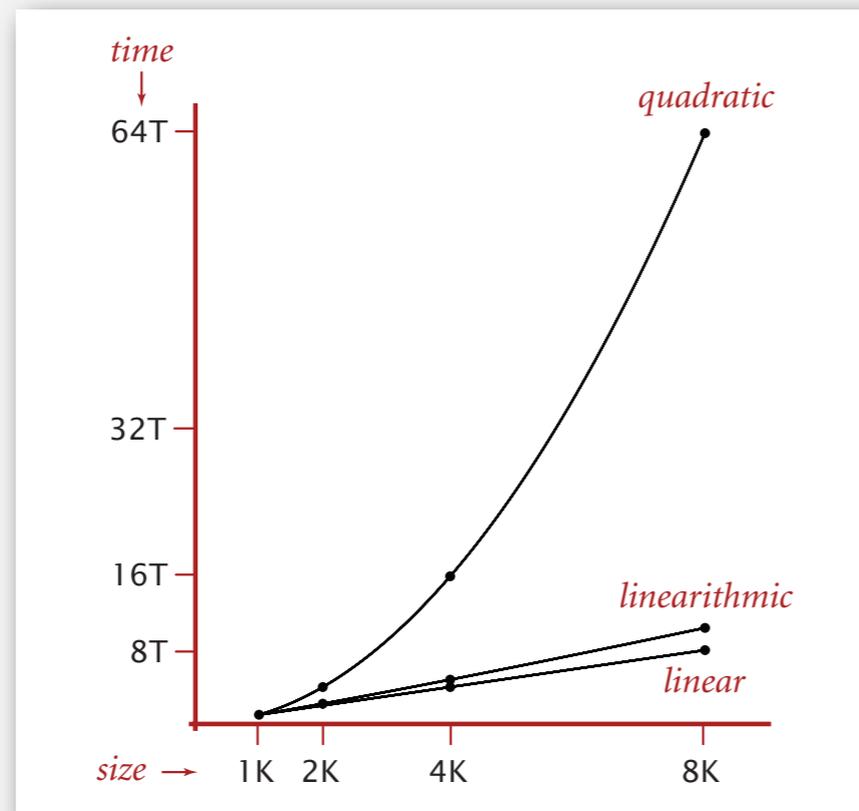
Some algorithmic successes

Discrete Fourier transform.

- Break down waveform of N samples into periodic components.
- Applications: DVD, JPEG, MRI, astrophysics,
- Brute force: N^2 steps.
- FFT algorithm: $N \log N$ steps, **enables new technology.**



Friedrich Gauss
1805



- sFFT: Sparse Fast Fourier Transform algorithm (Hassanieh et al., 2012)
 - A faster Fourier Transform: $k \log N$ steps (with k sparse coefficients)

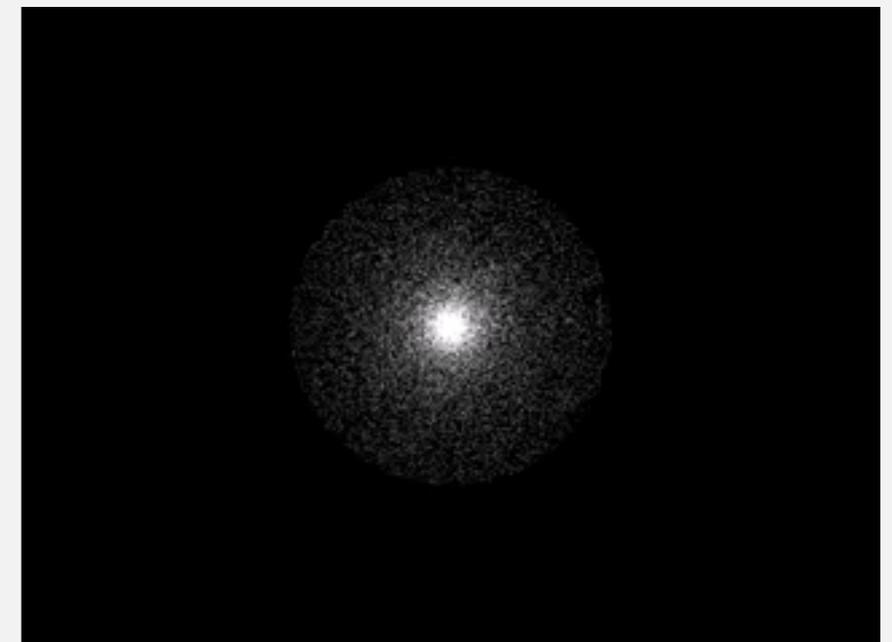
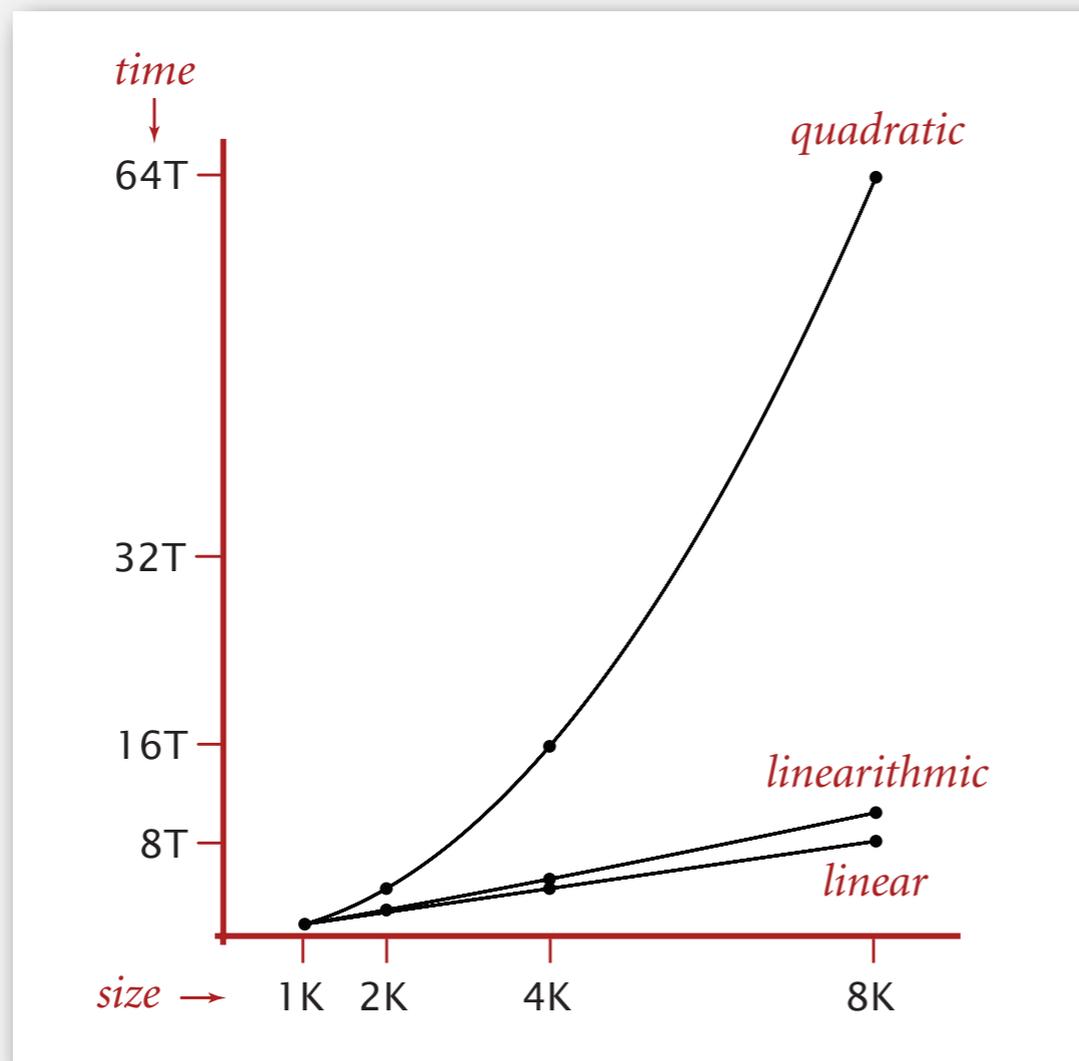
Some algorithmic successes

N-body simulation.

- Simulate gravitational interactions among N bodies.
- Brute force: N^2 steps.
- Barnes-Hut algorithm: $N \log N$ steps, **enables new research.**



Andrew Appel
PU '81



The challenge

Q. Will my program be able to solve a large practical input?

Why is my program so slow ?

Why does it run out of memory ?



Key insight. [Knuth 1970s] Use **scientific method** to understand performance.

Scientific method applied to analysis of algorithms

A framework for predicting performance and comparing algorithms.

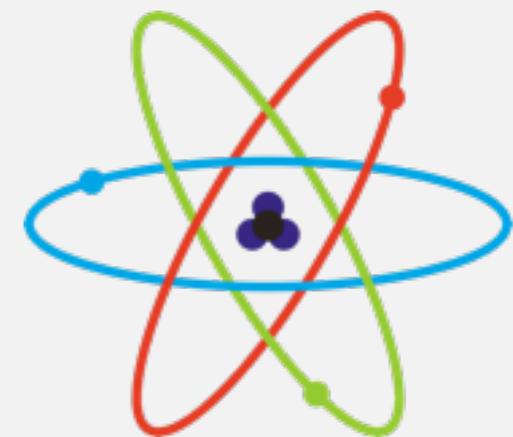
Scientific method.

- **Observe** some feature of the natural world.
- **Hypothesize** a model that is consistent with the observations.
- **Predict** events using the hypothesis.
- **Verify** the predictions by making further observations.
- **Validate** by repeating until the hypothesis and observations agree.

Principles.

Experiments must be **reproducible**.

Hypotheses must be **falsifiable**.



Feature of the natural world = computer itself.

ANALYSIS OF ALGORITHMS

- ▶ **Observations**
- ▶ Mathematical models
- ▶ Order-of-growth classifications
- ▶ Dependencies on inputs
- ▶ Memory

Example: 3-sum

3-sum. Given N distinct integers, how many triples sum to exactly zero?

```
% more 8ints.txt
8
30 -40 -20 -10 40 0 10 5

% java ThreeSum 8ints.txt
4
```

	a[i]	a[j]	a[k]	sum
1	30	-40	10	0
2	30	-20	-10	0
3	-40	40	0	0
4	-10	0	10	0

Context. Deeply related to problems in computational geometry.

3-sum: brute-force algorithm

```
public class ThreeSum
{
    public static int count(int[] a)
    {
        int N = a.length;
        int count = 0;
        for (int i = 0; i < N; i++)
            for (int j = i+1; j < N; j++)
                for (int k = j+1; k < N; k++)
                    if (a[i] + a[j] + a[k] == 0)
                        count++;
        return count;
    }

    public static void main(String[] args)
    {
        int[] a = In.readInts(args[0]);
        StdOut.println(count(a));
    }
}
```

← check each triple
← for simplicity, ignore integer overflow

Measuring the running time

Q. How to time a program?

A. Automatic.

```
public class Stopwatch    (part of stdlib.jar)
{
    Stopwatch()            create a new stopwatch
    double elapsedTime()   time since creation (in seconds)
}
```

```
public static void main(String[] args)
{
    int[] a = In.readInts(args[0]);
    Stopwatch stopwatch = new Stopwatch();
    StdOut.println(ThreeSum.count(a));
    double time = stopwatch.elapsedTime();
}
```

Measuring the running time

Q. How to time a program?

A. Automatic.

```
public class Stopwatch (part of stdlib.jar)
    Stopwatch() create a new stopwatch
    double elapsedTime() time since creation (in seconds)
```

```
public class Stopwatch
{
    private final long start = System.currentTimeMillis();

    public double elapsedTime()
    {
        long now = System.currentTimeMillis();
        return (now - start) / 1000.0;
    }
}
implementation (part of stdlib.jar)
```

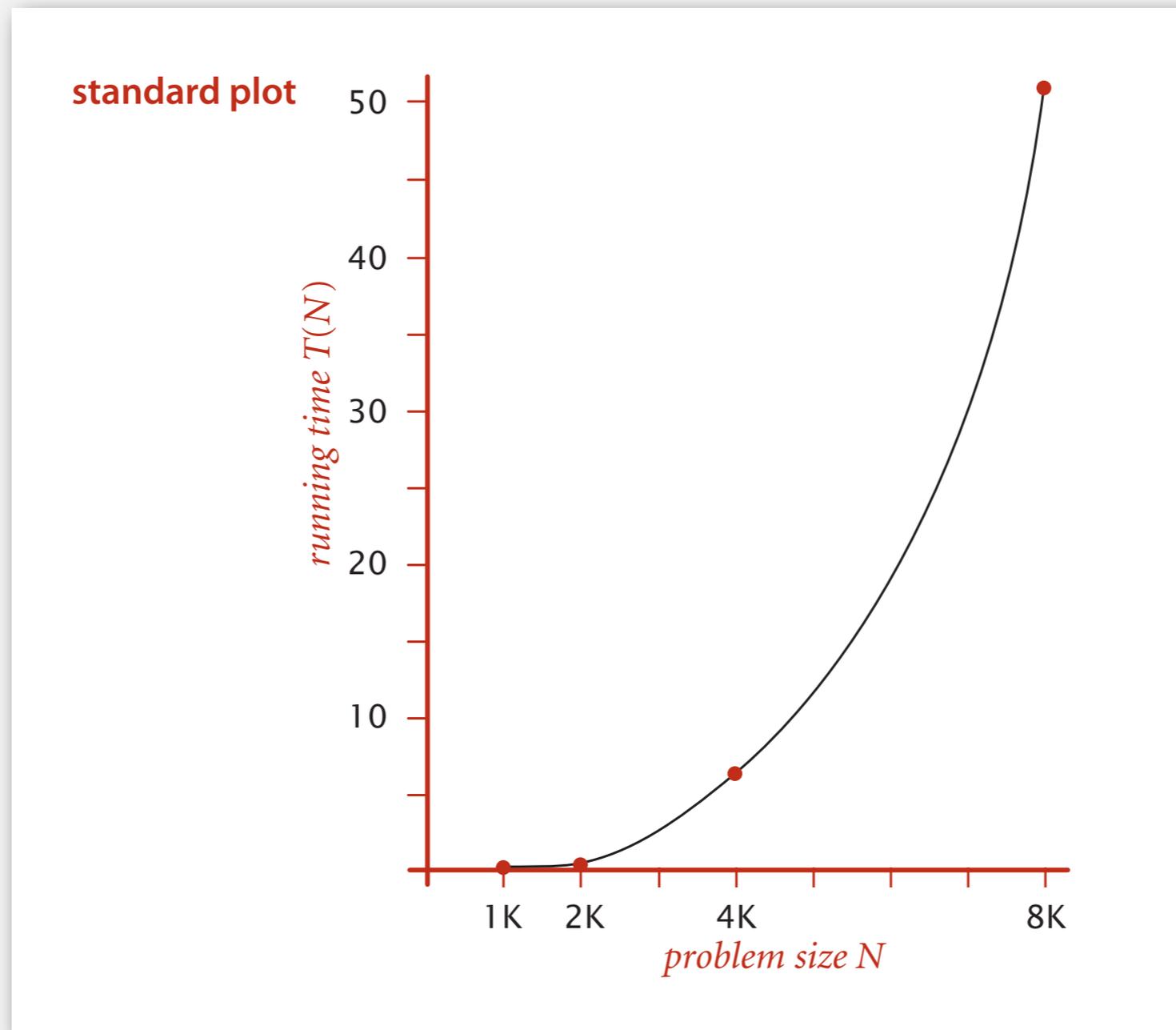
Empirical analysis

Run the program for various input sizes and measure running time.

N	time (seconds) †
250	0.0
500	0.0
1,000	0.1
2,000	0.8
4,000	6.4
8,000	51.1
16,000	?

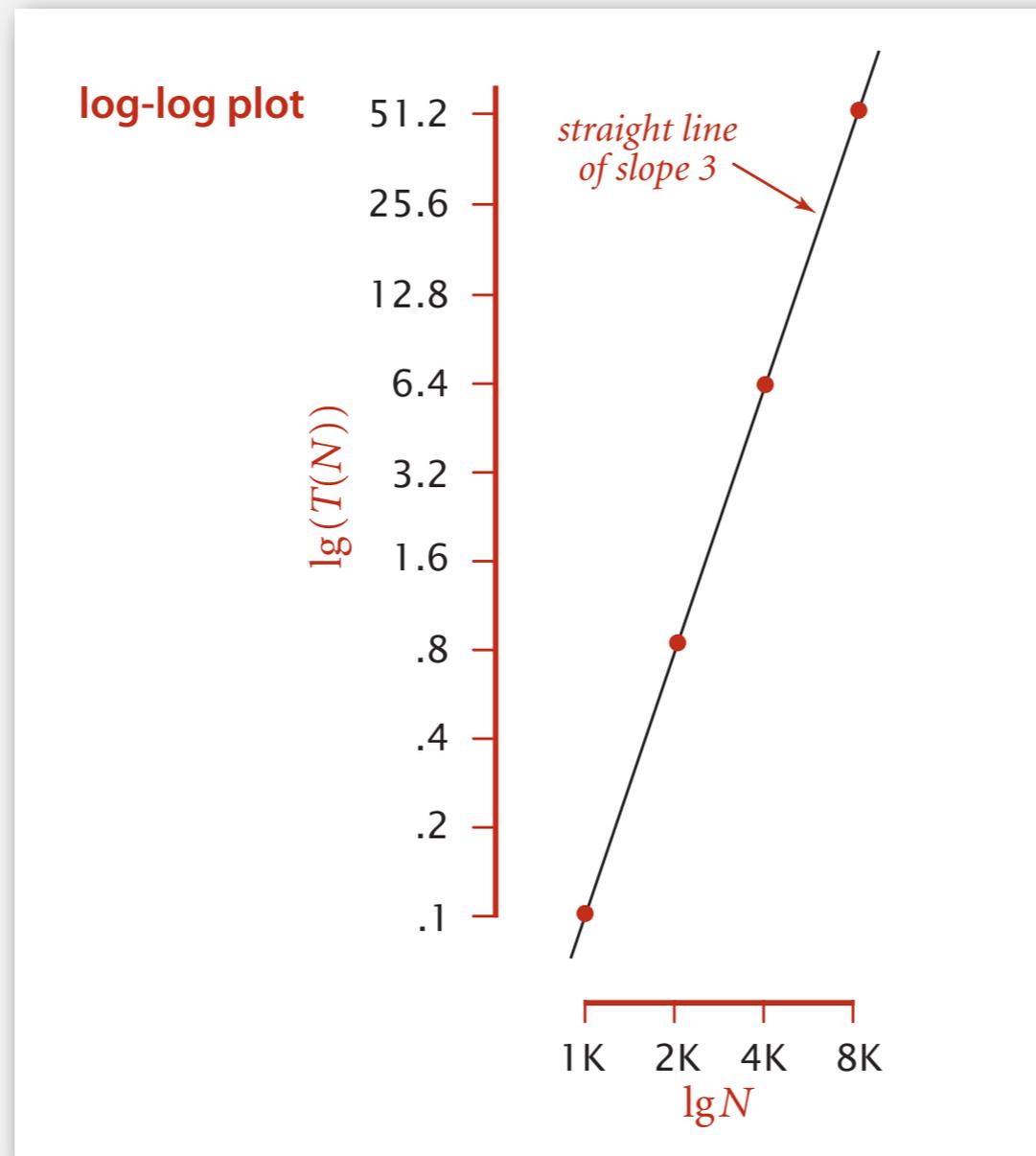
Data analysis

Standard plot. Plot running time $T(N)$ vs. input size N .



Data analysis

Log-log plot. Plot running time $T(N)$ vs. input size N using **log-log scale**.



$$\lg(T(N)) = b \lg N + c$$

$$b = 2.999$$

$$c = -33.2103$$

$$T(N) = a N^b, \text{ where } a = 2^c$$

Regression. Fit straight line through data points: $a N^b$.

Hypothesis. The running time is about $1.006 \times 10^{-10} \times N^{2.999}$ seconds.

power law

slope

Prediction and validation

Hypothesis. The running time is about $1.006 \times 10^{-10} \times N^{2.999}$ seconds.

"order of growth" of running time is about N^3 [stay tuned]

Predictions.

- 51.0 seconds for $N = 8,000$.
- 408.1 seconds for $N = 16,000$.

Observations.

N	time (seconds) †
8,000	51.1
8,000	51.0
8,000	51.1
16,000	410.8

validates hypothesis!

Doubling hypothesis

Doubling hypothesis. Quick way to estimate b in a power-law relationship.

Run program, **doubling** the size of the input.

N	time (seconds) †	ratio	lg ratio
250	0.0		-
500	0.0	4.8	2.3
1,000	0.1	6.9	2.8
2,000	0.8	7.7	2.9
4,000	6.4	8.0	3.0
8,000	51.1	8.0	3.0

↑
seems to converge to a constant $b \approx 3$

Hypothesis. Running time is about $a N^b$ with $b = \lg \text{ratio}$.

Caveat. Cannot identify logarithmic factors with doubling hypothesis.

Doubling hypothesis

Doubling hypothesis. Quick way to estimate b in a power-law hypothesis.

Q. How to estimate a (assuming we know b) ?

A. Run the program (for a sufficient large value of N) and solve for a .

N	time (seconds) †
8,000	51.1
8,000	51.0
8,000	51.1

$$51.1 = a \times 8000^3$$
$$\Rightarrow a = 0.998 \times 10^{-10}$$

Hypothesis. Running time is about $0.998 \times 10^{-10} \times N^3$ seconds.



almost identical hypothesis
to one obtained via linear regression

Experimental algorithmics

System independent effects.

- Algorithm.
 - Input data.
- } determines exponent b
in power law

System dependent effects.

- Hardware: CPU, memory, cache, ...
- Software: compiler, interpreter, garbage collector, ...
- System: operating system, network, other applications, ...

} determines constant a
in power law

Bad news. Difficult to get precise measurements.

Good news. Much easier and cheaper than other sciences.

↖ e.g., can run huge number of experiments

In practice, constant factors matter too!

Q. How long does this program take as a function of N ?

```
String s = StdIn.readString();
int N = s.length();
...
for (int i = 0; i < N; i++)
    for (int j = 0; j < N; j++)
        distance[i][j] = ...
...
```

N	time
1,000	0.11
2,000	0.35
4,000	1.6
8,000	6.5

Jenny $\sim c_1 N^2$ seconds

N	time
250	0.5
500	1.1
1,000	1.9
2,000	3.9

Kenny $\sim c_2 N$ seconds

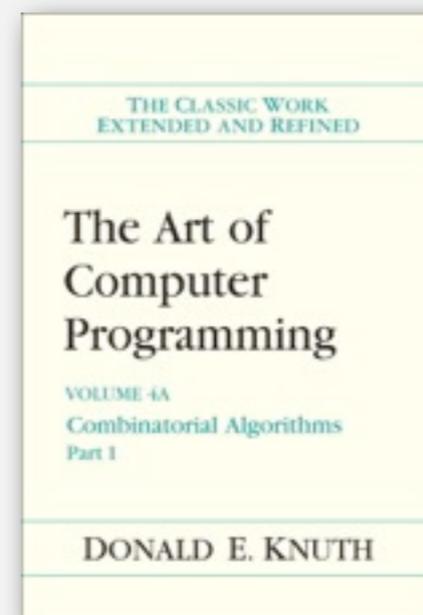
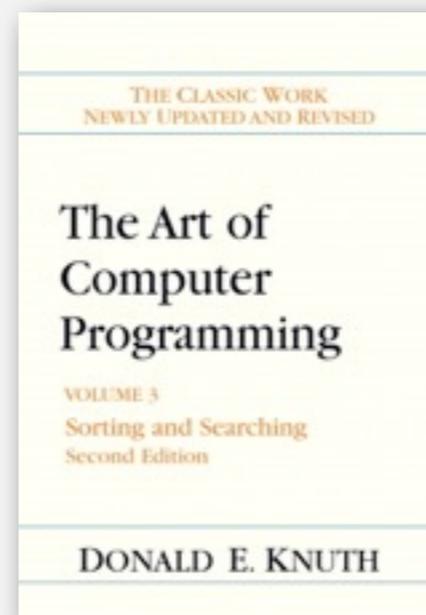
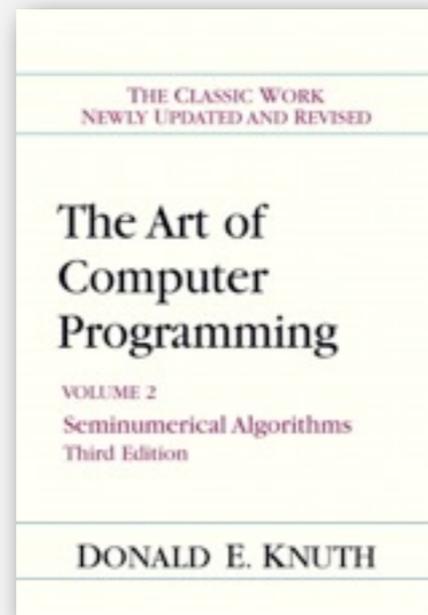
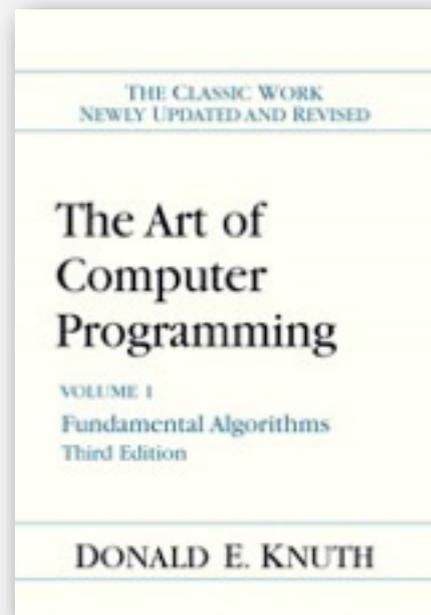
ANALYSIS OF ALGORITHMS

- ▶ Observations
- ▶ **Mathematical models**
- ▶ Order-of-growth classifications
- ▶ Dependencies on inputs
- ▶ Memory

Mathematical models for running time

Total running time: sum of cost \times frequency for all operations.

- Need to analyze program to determine set of operations.
- Cost depends on machine, compiler.
- Frequency depends on algorithm, input data.



Donald Knuth
1974 Turing Award

In principle, accurate mathematical models are available.

Cost of basic operations

operation	example	nanoseconds †
integer add	<code>a + b</code>	2.1
integer multiply	<code>a * b</code>	2.4
integer divide	<code>a / b</code>	5.4
floating-point add	<code>a + b</code>	4.6
floating-point multiply	<code>a * b</code>	4.2
floating-point divide	<code>a / b</code>	13.5
sine	<code>Math.sin(theta)</code>	91.3
arctangent	<code>Math.atan2(y, x)</code>	129.0
...

† Running OS X on Macbook Pro 2.2GHz with 2GB RAM

Cost of basic operations

operation	example	nanoseconds †
variable declaration	<code>int a</code>	C ₁
assignment statement	<code>a = b</code>	C ₂
integer compare	<code>a < b</code>	C ₃
array element access	<code>a[i]</code>	C ₄
array length	<code>a.length</code>	C ₅
1D array allocation	<code>new int[N]</code>	C ₆ N
2D array allocation	<code>new int[N][N]</code>	C ₇ N ²
string length	<code>s.length()</code>	C ₈
substring extraction	<code>s.substring(N/2, N)</code>	C ₉
string concatenation	<code>s + t</code>	C ₁₀ N

Novice mistake. Abusive string concatenation.

Example: I-sum

Q. How many instructions as a function of input size N ?

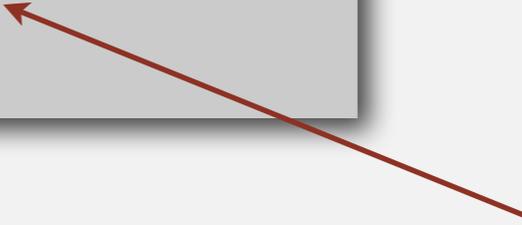
```
int count = 0;
for (int i = 0; i < N; i++)
    if (a[i] == 0)
        count++;
```

operation	frequency
variable declaration	2
assignment statement	2
less than compare	$N + 1$
equal to compare	N
array access	N
increment	N to $2N$

Example: 2-sum

Q. How many instructions as a function of input size N ?

```
int count = 0;
for (int i = 0; i < N; i++)
    for (int j = i+1; j < N; j++)
        if (a[i] + a[j] == 0)
            count++;
```


$$0 + 1 + 2 + \dots + (N - 1) = \frac{1}{2} N (N - 1)$$
$$= \binom{N}{2}$$

operation	frequency
variable declaration	$N + 2$
assignment statement	$N + 2$
less than compare	$\frac{1}{2} (N + 1) (N + 2)$
equal to compare	$\frac{1}{2} N (N - 1)$
array access	$N (N - 1)$
increment	$\frac{1}{2} N (N - 1)$ to $N (N - 1)$

tedious to count exactly

Simplifying the calculations

*“ It is convenient to have a **measure of the amount of work involved in a computing process**, even though it be a very **crude one**. We may count up the number of times that various elementary operations are applied in the whole process and then given them various weights. We might, for instance, count the number of additions, subtractions, multiplications, divisions, recording of numbers, and extractions of figures from tables. In the case of computing with matrices most of the work consists of multiplications and writing down numbers, and we shall therefore only attempt to count the number of **multiplications and recordings**. ” — Alan Turing*

ROUNDING-OFF ERRORS IN MATRIX PROCESSES

By A. M. TURING

(National Physical Laboratory, Teddington, Middlesex)

[Received 4 November 1947]

SUMMARY

A number of methods of solving sets of linear equations and inverting matrices are discussed. The theory of the rounding-off errors involved is investigated for some of the methods. In all cases examined, including the well-known ‘Gauss elimination process’, it is found that the errors are normally quite moderate: no exponential build-up need occur.



Simplification I: cost model

Cost model. Use some basic operation as a proxy for running time.

```
int count = 0;
for (int i = 0; i < N; i++)
    for (int j = i+1; j < N; j++)
        if (a[i] + a[j] == 0)
            count++;
```

$$0 + 1 + 2 + \dots + (N - 1) = \frac{1}{2} N (N - 1) \\ = \binom{N}{2}$$

operation	frequency
variable declaration	$N + 2$
assignment statement	$N + 2$
less than compare	$\frac{1}{2} (N + 1) (N + 2)$
equal to compare	$\frac{1}{2} N (N - 1)$
array access	$N (N - 1)$
increment	$\frac{1}{2} N (N - 1)$ to $N (N - 1)$

cost model = array accesses

Simplification 2: tilde notation

- Estimate running time (or memory) as a function of input size N .
- Ignore lower order terms.
 - when N is large, terms are negligible
 - when N is small, we don't care

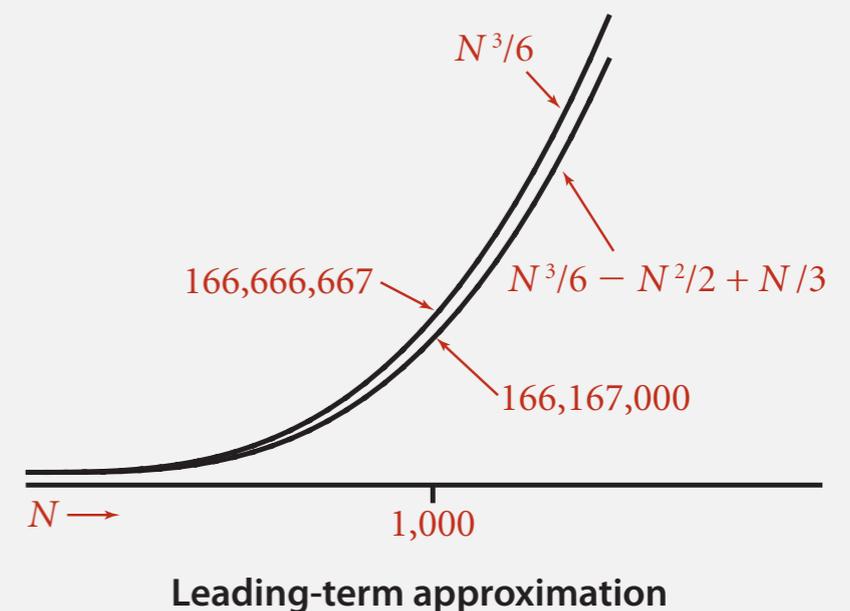
Ex 1. $\frac{1}{6} N^3 + 20 N + 16 \sim \frac{1}{6} N^3$

Ex 2. $\frac{1}{6} N^3 + 100 N^{4/3} + 56 \sim \frac{1}{6} N^3$

Ex 3. $\frac{1}{6} N^3 - \frac{1}{2} N^2 + \frac{1}{3} N \sim \frac{1}{6} N^3$

discard lower-order terms

(e.g., $N = 1000$: 500 thousand vs. 166 million)



Technical definition. $f(N) \sim g(N)$ means $\lim_{N \rightarrow \infty} \frac{f(N)}{g(N)} = 1$

Simplification 2: tilde notation

- Estimate running time (or memory) as a function of input size N .
- Ignore lower order terms.
 - when N is large, terms are negligible
 - when N is small, we don't care

operation	frequency	tilde notation
variable declaration	$N + 2$	$\sim N$
assignment statement	$N + 2$	$\sim N$
less than compare	$\frac{1}{2} (N + 1) (N + 2)$	$\sim \frac{1}{2} N^2$
equal to compare	$\frac{1}{2} N (N - 1)$	$\sim \frac{1}{2} N^2$
array access	$N (N - 1)$	$\sim N^2$
increment	$\frac{1}{2} N (N - 1)$ to $N (N - 1)$	$\sim \frac{1}{2} N^2$ to $\sim N^2$

Example: 2-sum

Q. Approximately how many array accesses as a function of input size N ?

```
int count = 0;
for (int i = 0; i < N; i++)
  for (int j = i+1; j < N; j++)
    if (a[i] + a[j] == 0)
      count++;
```

"inner loop"

A. $\sim N^2$ array accesses.

$$\begin{aligned} 0 + 1 + 2 + \dots + (N - 1) &= \frac{1}{2} N (N - 1) \\ &= \binom{N}{2} \end{aligned}$$

Bottom line. Use cost model and tilde notation to simplify frequency counts.

Example: 3-sum

Q. Approximately how many array accesses as a function of input size N ?

```
int count = 0;
for (int i = 0; i < N; i++)
  for (int j = i+1; j < N; j++)
    for (int k = j+1; k < N; k++)
      if (a[i] + a[j] + a[k] == 0)
        count++;
```

"inner loop"

A. $\sim \frac{1}{2} N^3$ array accesses.

$$\binom{N}{3} = \frac{N(N-1)(N-2)}{3!}$$
$$\sim \frac{1}{6} N^3$$

Bottom line. Use cost model and tilde notation to simplify frequency counts.

Estimating a discrete sum

Q. How to estimate a discrete sum?

A1. Take discrete mathematics course.

A2. Replace the sum with an integral, and use calculus!

Ex 1. $1 + 2 + \dots + N.$
$$\sum_{i=1}^N i \sim \int_{x=1}^N x dx \sim \frac{1}{2} N^2$$

Ex 2. $1 + 1/2 + 1/3 + \dots + 1/N.$
$$\sum_{i=1}^N \frac{1}{i} \sim \int_{x=1}^N \frac{1}{x} dx = \ln N$$

Ex 3. 3-sum triple loop.
$$\sum_{i=1}^N \sum_{j=i}^N \sum_{k=j}^N 1 \sim \int_{x=1}^N \int_{y=x}^N \int_{z=y}^N dz dy dx \sim \frac{1}{6} N^3$$

Mathematical models for running time

In principle, accurate mathematical models are available.

In practice,

- Formulas can be complicated.
- Advanced mathematics might be required.
- Exact models best left for experts.



costs (depend on machine, compiler)

$$T_N = c_1 A + c_2 B + c_3 C + c_4 D + c_5 E$$

A = array access

B = integer add

C = integer compare

D = increment

E = variable assignment

frequencies
(depend on algorithm, input)

Bottom line. We use **approximate** models in this course: $T(N) \sim c N^3$.

ANALYSIS OF ALGORITHMS

- ▶ Observations
- ▶ Mathematical models
- ▶ **Order-of-growth classifications**
- ▶ Dependencies on inputs
- ▶ Memory

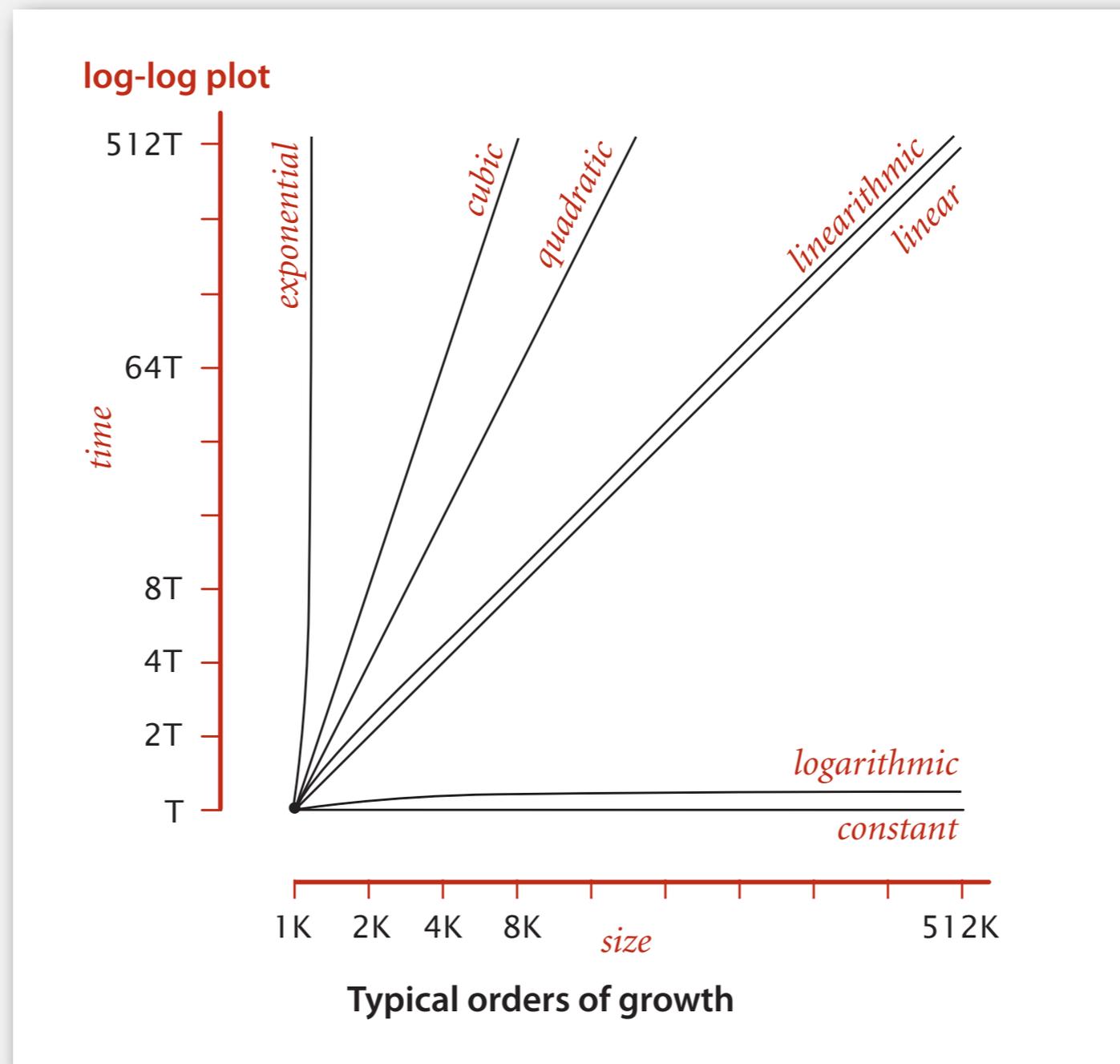
Common order-of-growth classifications

Good news. the small set of functions

1, $\log N$, N , $N \log N$, N^2 , N^3 , and 2^N

order of growth discards leading coefficient

suffices to describe order-of-growth of typical algorithms.



Common order-of-growth classifications

order of growth	name	typical code framework	description	example	$T(2N) / T(N)$
1	constant	<pre>a = b + c;</pre>	statement	add two numbers	1
log N	logarithmic	<pre>while (N > 1) { N = N / 2; ... }</pre>	divide in half	binary search	~ 1
N	linear	<pre>for (int i = 0; i < N; i++) { ... }</pre>	loop	find the maximum	2
N log N	linearithmic	[see mergesort lecture]	divide and conquer	mergesort	~ 2
N^2	quadratic	<pre>for (int i = 0; i < N; i++) for (int j = 0; j < N; j++) { ... }</pre>	double loop	check all pairs	4
N^3	cubic	<pre>for (int i = 0; i < N; i++) for (int j = 0; j < N; j++) for (int k = 0; k < N; k++) { ... }</pre>	triple loop	check all triples	8
2^N	exponential	[see combinatorial search lecture]	exhaustive search	check all subsets	$T(N)$

Practical implications of order-of-growth

growth rate	problem size solvable in minutes			
	1970s	1980s	1990s	2000s
1	any	any	any	any
log N	any	any	any	any
N	millions	tens of millions	hundreds of millions	billions
N log N	hundreds of thousands	millions	millions	hundreds of millions
N ²	hundreds	thousand	thousands	tens of thousands
N ³	hundred	hundreds	thousand	thousands
2 ^N	20	20s	20s	30

Practical implications of order-of-growth

growth rate	problem size solvable in minutes				time to process millions of inputs			
	1970s	1980s	1990s	2000s	1970s	1980s	1990s	2000s
1	any	any	any	any	instant	instant	instant	instant
log N	any	any	any	any	instant	instant	instant	instant
N	millions	tens of millions	hundreds of millions	billions	minutes	seconds	second	instant
N log N	hundreds of thousands	millions	millions	hundreds of millions	hour	minutes	tens of seconds	seconds
N ²	hundreds	thousand	thousands	tens of thousands	decades	years	months	weeks
N ³	hundred	hundreds	thousand	thousands	never	never	never	millennia

Practical implications of order-of-growth

growth rate	name	description	effect on a program that runs for a few seconds	
			time for 100x more data	size for 100x faster computer
1	constant	independent of input size	-	-
log N	logarithmic	nearly independent of input size	-	-
N	linear	optimal for N inputs	a few minutes	100x
N log N	linearithmic	nearly optimal for N inputs	a few minutes	100x
N ²	quadratic	not practical for large problems	several hours	10x
N ³	cubic	not practical for medium problems	several weeks	4-5x
2 ^N	exponential	useful only for tiny problems	forever	1x

Binary search

Goal. Given a sorted array and a key, find index of the key in the array?

Binary search. Compare key against middle entry.

- Too small, go left.
- Too big, go right.
- Equal, found.

6	13	14	25	33	43	51	53	64	72	84	93	95	96	97
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
↑							↑							↑
lo							mid							hi

Binary search demo

Goal. Given a sorted array and a key, find index of the key in the array?

Successful search. Binary search for 33.

6	13	14	25	33	43	51	53	64	72	84	93	95	96	97
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
↑							↑							↑
lo							mid							hi

Binary search demo

Goal. Given a sorted array and a key, find index of the key in the array?

Successful search. Binary search for 33.

6	13	14	25	33	43	51	53	64	72	84	93	95	96	97
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
↑			↑			↑								
lo			mid			hi								

Binary search demo

Goal. Given a sorted array and a key, find index of the key in the array?

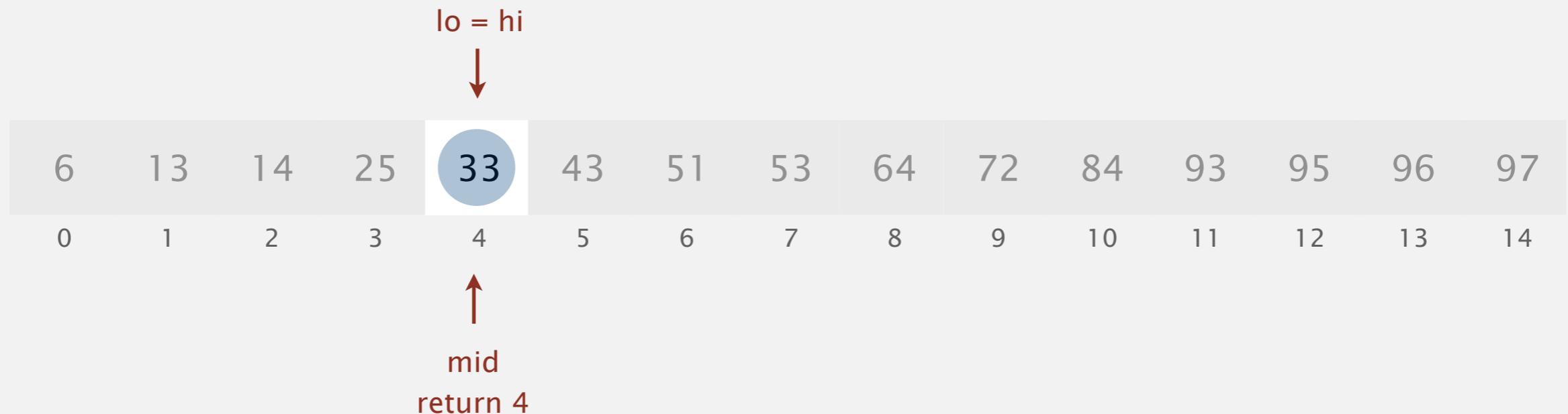
Successful search. Binary search for 33.

6	13	14	25	33	43	51	53	64	72	84	93	95	96	97
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
				↑	↑	↑								
				lo	mid	hi								

Binary search demo

Goal. Given a sorted array and a key, find index of the key in the array?

Successful search. Binary search for 33.



Binary search demo

Goal. Given a sorted array and a key, find index of the key in the array?

Unsuccessful search. Binary search for 34.

6	13	14	25	33	43	51	53	64	72	84	93	95	96	97
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
↑							↑							↑
lo							mid							hi

Binary search demo

Goal. Given a sorted array and a key, find index of the key in the array?

Unsuccessful search. Binary search for 34.

6	13	14	25	33	43	51	53	64	72	84	93	95	96	97
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
↑			↑			↑								
lo			mid			hi								

Binary search demo

Goal. Given a sorted array and a key, find index of the key in the array?

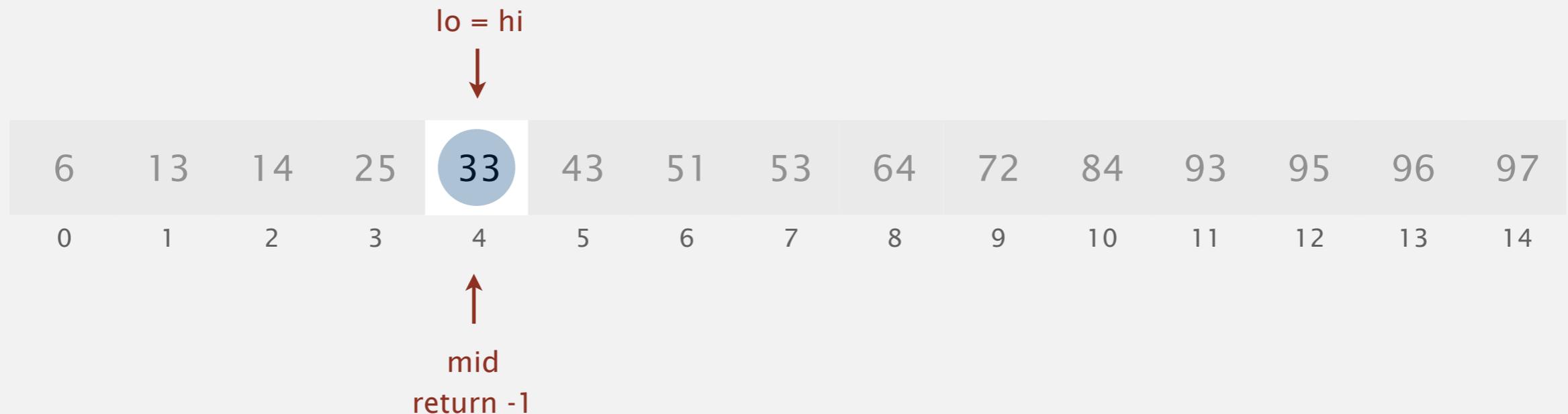
Unsuccessful search. Binary search for 34.

6	13	14	25	33	43	51	53	64	72	84	93	95	96	97
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
				↑	↑	↑								
				lo	mid	hi								

Binary search demo

Goal. Given a sorted array and a key, find index of the key in the array?

Unsuccessful search. Binary search for 34.



Binary search: Java implementation

Trivial to implement?

- First binary search published in 1946; first bug-free one published in 1962.
- Bug in Java's `Arrays.binarySearch()` discovered in 2006.

```
public static int binarySearch(int[] a, int key)
{
    int lo = 0, hi = a.length-1;
    while (lo <= hi)
    {
        int mid = lo + (hi - lo) / 2;
        if      (key < a[mid]) hi = mid - 1;
        else if (key > a[mid]) lo = mid + 1;
        else return mid;
    }
    return -1;
}
```

← one "3-way compare"

Invariant. If `key` appears in the array `a[]`, then $a[lo] \leq key \leq a[hi]$.

Binary search: mathematical analysis

Proposition. Binary search uses at most $1 + \lg N$ compares to search in a sorted array of size N .

Def. $T(N) \equiv \#$ compares to binary search in a sorted subarray of size at most N .

Binary search recurrence. $T(N) \leq T(N/2) + 1$ for $N > 1$, with $T(1) = 1$.

↑ left or right half ↑ possible to implement with one 2-way compare (instead of 3-way)

Pf sketch.

$$\begin{aligned} T(N) &\leq T(N/2) + 1 \\ &\leq T(N/4) + 1 + 1 \\ &\leq T(N/8) + 1 + 1 + 1 \\ &\dots \\ &\leq T(N/N) + 1 + 1 + \dots + 1 \\ &= 1 + \lg N \end{aligned}$$

given

apply recurrence to first term

apply recurrence to first term

stop applying, $T(1) = 1$

Binary search: mathematical analysis

Proposition. Binary search uses at most $1 + \lg N$ compares to search in a sorted array of size N .

Def. $T(N) \equiv \#$ compares to binary search in a sorted subarray of size at most N .

Binary search recurrence. $T(N) \leq T(\lfloor N/2 \rfloor) + 1$ for $N > 1$, with $T(0) = 0$.

For simplicity, we prove when $N = 2^n - 1$ for some n , so $\lfloor N/2 \rfloor = 2^{n-1} - 1$.

$$\begin{aligned} T(2^n - 1) &\leq T(2^{n-1} - 1) + 1 \\ &\leq T(2^{n-2} - 1) + 1 + 1 \\ &\leq T(2^{n-3} - 1) + 1 + 1 + 1 \\ &\dots \\ &\leq T(2^0 - 1) + 1 + 1 + \dots + 1 \\ &= n \end{aligned}$$

given

apply recurrence to first term

apply recurrence to first term

stop applying, $T(0) = 1$

An $N^2 \log N$ algorithm for 3-sum

Algorithm.

- **Sort** the N (distinct) numbers.
- For each pair of numbers $a[i]$ and $a[j]$, **binary search** for $-(a[i] + a[j])$.

Analysis. Order of growth is $N^2 \log N$.

- Step 1: N^2 with insertion sort.
- Step 2: $N^2 \log N$ with binary search.

input

30 -40 -20 -10 40 0 10 5

sort

-40 -20 -10 0 5 10 30 40

binary search

(-40, -20)	60
(-40, -10)	50
(-40, 0)	40
(-40, 5)	35
(-40, 10)	30
⋮	⋮
(-40, 40)	0
⋮	⋮
(-10, 0)	10
⋮	⋮
(-20, 10)	10
⋮	⋮
(10, 30)	-40
(10, 40)	-50
(30, 40)	-70

only count if
 $a[i] < a[j] < a[k]$
to avoid
double counting

Comparing programs

Hypothesis. The $N^2 \log N$ three-sum algorithm is significantly faster in practice than the brute-force N^3 algorithm.

N	time (seconds)
1,000	0.1
2,000	0.8
4,000	6.4
8,000	51.1

ThreeSum.java

N	time (seconds)
1,000	0.14
2,000	0.18
4,000	0.34
8,000	0.96
16,000	3.67
32,000	14.88
64,000	59.16

ThreeSumDeluxe.java

Guiding principle. Typically, better order of growth \Rightarrow faster in practice.

ANALYSIS OF ALGORITHMS

- ▶ Observations
- ▶ Mathematical models
- ▶ Order-of-growth classifications
- ▶ **Dependencies on inputs**
- ▶ Memory

Types of analyses

Best case. Lower bound on cost.

- Determined by “easiest” input.
- Provides a goal for all inputs.

Worst case. Upper bound on cost.

- Determined by “most difficult” input.
- Provides a guarantee for all inputs.

Average case. Expected cost for random input.

- Need a model for “random” input.
- Provides a way to predict performance.

Ex 1. Array accesses for brute-force 3 sum.

Best: $\sim \frac{1}{2} N^3$

Average: $\sim \frac{1}{2} N^3$

Worst: $\sim \frac{1}{2} N^3$

Ex 2. Compares for binary search.

Best: ~ 1

Average: $\sim \lg N$

Worst: $\sim \lg N$

Types of analyses

Best case. Lower bound on cost.

Worst case. Upper bound on cost.

Average case. “Expected” cost.

Actual data might not match input model?

- Need to understand input to effectively process it.
- Approach 1: design for the worst case.
- Approach 2: randomize, depend on probabilistic guarantee.

Theory of Algorithms

Goals.

- Establish “difficulty” of a problem.
- Develop “optimal” algorithms.

Approach.

- Suppress details in analysis: analyze “to within a constant factor”.
- Eliminate variability in input model by focusing on the worst case.

Optimal algorithm.

- Performance guarantee (to within a constant factor) for any input.
- No algorithm can provide a better performance guarantee.

Commonly-used notations

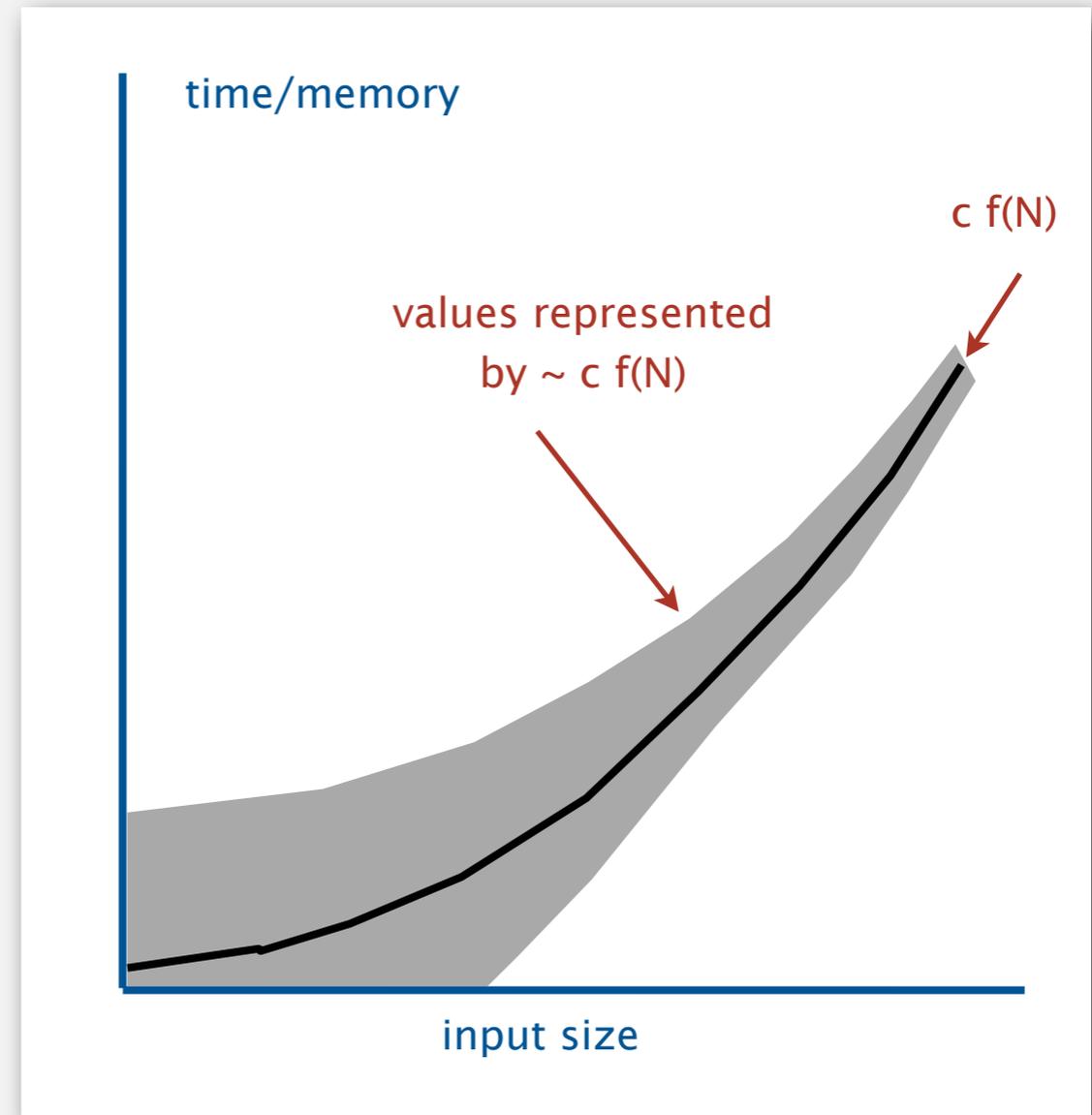
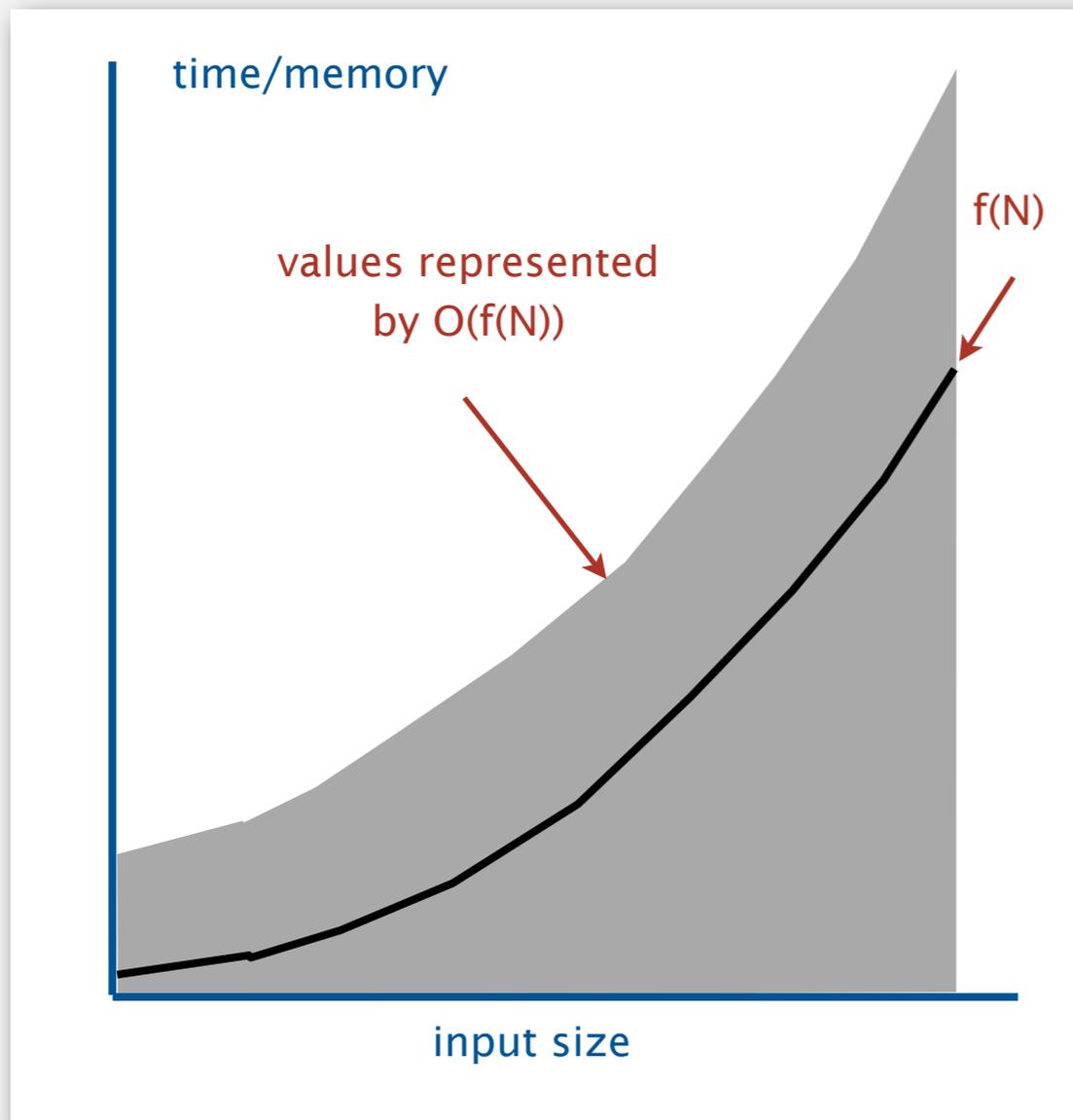
notation	provides	example	shorthand for	used to
Tilde	leading term	$\sim 10 N^2$	$10 N^2$ $10 N^2 + 22 N \log N$ $10 N^2 + 2 N + 37$	provide approximate model
Big Theta	asymptotic growth rate	$\Theta(N^2)$	$\frac{1}{2} N^2$ $10 N^2$ $5 N^2 + 22 N \log N + 3N$	classify algorithms
Big Oh	$\Theta(N^2)$ and smaller	$O(N^2)$	$10 N^2$ $100 N$ $22 N \log N + 3 N$	develop upper bounds
Big Omega	$\Theta(N^2)$ and larger	$\Omega(N^2)$	$\frac{1}{2} N^2$ N^5 $N^3 + 22 N \log N + 3 N$	develop lower bounds

Common mistake. Interpreting big-Oh as an approximate model.

Tilde notation vs. big-Oh notation

We use tilde notation whenever possible.

- Big-Oh notation suppresses leading constant.
- Big-Oh notation only provides upper bound (not lower bound).



Theory of algorithms: example I

Goals.

- Establish “difficulty” of a problem and develop “optimal” algorithms.
- Ex. I-SUM = “Is there a 0 in the array? ”

Upper bound. A specific algorithm.

- Ex. Brute-force algorithm for I-SUM: Look at every array entry.
- Running time of the optimal algorithm for I-SUM is $O(N)$.

Lower bound. Proof that no algorithm can do better.

- Ex. Have to examine all N entries (any unexamined one might be 0).
- Running time of the optimal algorithm for I-SUM is $\Omega(N)$.

Optimal algorithm.

- Lower bound equals upper bound (to within a constant factor).
- Ex. Brute-force algorithm for I-SUM is optimal: its running time is $\Theta(N)$.

Theory of algorithms: example 2

Goals.

- Establish “difficulty” of a problem and develop “optimal” algorithms.
- Ex. 3-SUM

Upper bound. A specific algorithm.

- Ex. Brute-force algorithm for 3-SUM
- Running time of the optimal algorithm for 3-SUM is $O(N^3)$.

Theory of algorithms: example 2

Goals.

- Establish “difficulty” of a problem and develop “optimal” algorithms.
- Ex. 3-SUM

Upper bound. A specific algorithm.

- Ex. **Improved** algorithm for 3-SUM
- Running time of the optimal algorithm for 3-SUM is $O(N^2 \log N)$.

Lower bound. Proof that no algorithm can do better.

- Ex. Have to examine all N entries to solve 3-SUM.
- Running time of the optimal algorithm for solving 3-SUM is $\Omega(N)$.

Open problems.

- Optimal algorithm for 3-SUM?
- Subquadratic algorithm for 3-SUM?
- Quadratic lower bound for 3-SUM?

Algorithm design approach

Start.

- Develop an algorithm.
- Prove a lower bound.

Gap?

- Lower the upper bound (discover a new algorithm).
- Raise the lower bound (more difficult).

Golden Age of Algorithm Design.

- 1970s-.
- Steadily decreasing upper bounds for many important problems.
- Many known optimal algorithms.

Caveats.

- Overly pessimistic to focus on worst case?
- Need better than “to within a constant factor” to predict performance.

ANALYSIS OF ALGORITHMS

- ▶ Observations
- ▶ Mathematical models
- ▶ Order-of-growth classifications
- ▶ Dependencies on inputs
- ▶ **Memory**

Basics

Bit. 0 or 1.

Byte. 8 bits.

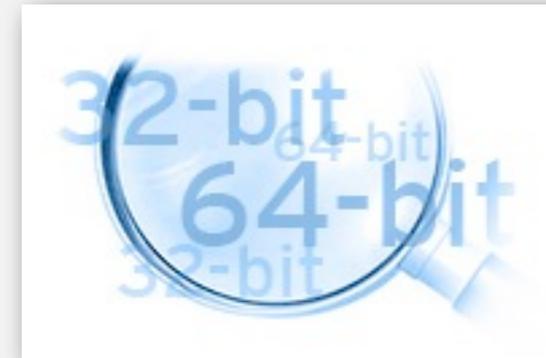
Megabyte (MB). 1 million or 2^{20} bytes.

Gigabyte (GB). 1 billion or 2^{30} bytes.

NIST



most computer scientists



Old machine. We used to assume a 32-bit machine with 4 byte pointers.

Modern machine. We now assume a 64-bit machine with 8 byte pointers.

- Can address more memory.
- Pointers use more space.



some JVMs "compress" ordinary object pointers to 4 bytes to avoid this cost

Typical memory usage for primitive types and

Primitive types.

type	bytes
<code>boolean</code>	1
<code>byte</code>	1
<code>char</code>	2
<code>int</code>	4
<code>float</code>	4
<code>long</code>	8
<code>double</code>	8

for primitive types

Array overhead. 24 bytes.

type	bytes
<code>char[]</code>	$2N + 24$
<code>int[]</code>	$4N + 24$
<code>double[]</code>	$8N + 24$

for one-dimensional arrays

type	bytes
<code>char[][]</code>	$\sim 2MN$
<code>int[][]</code>	$\sim 4MN$
<code>double[][]</code>	$\sim 8MN$

for two-dimensional arrays

Typical memory usage for objects in Java

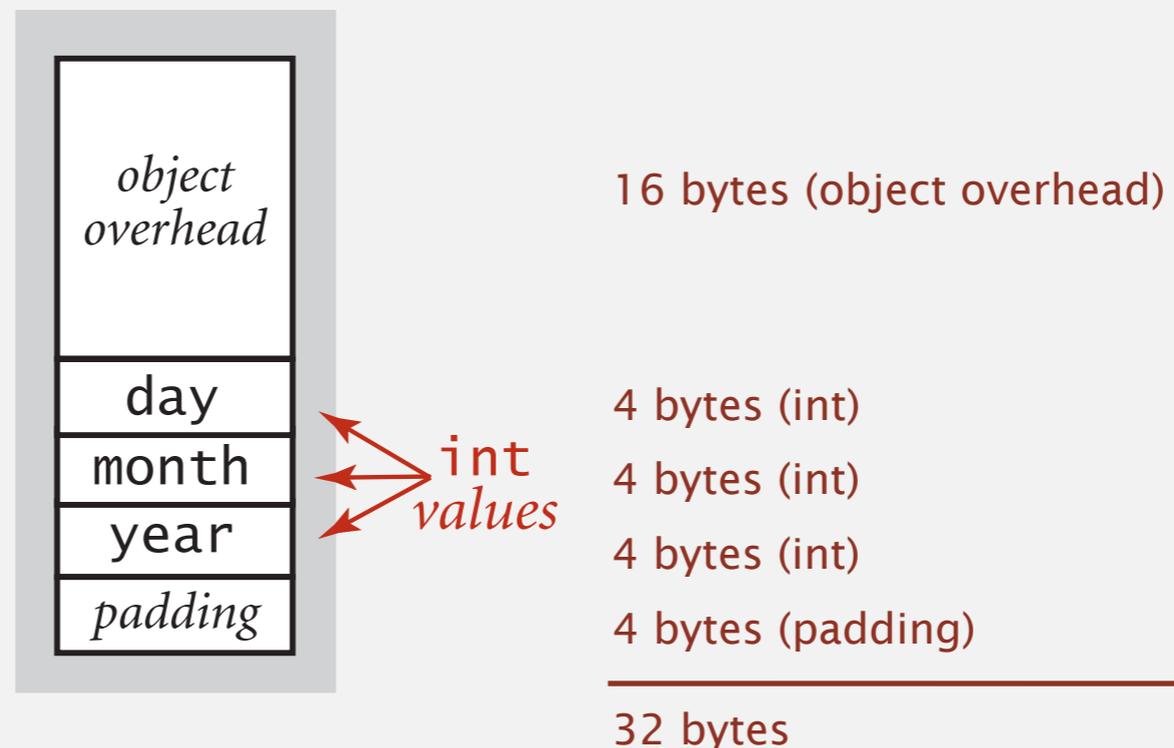
Object overhead. 16 bytes.

Reference. 8 bytes.

Padding. Each object uses a multiple of 8 bytes.

Ex 1. A `Date` object uses 32 bytes of memory.

```
public class Date
{
    private int day;
    private int month;
    private int year;
    ...
}
```



Typical memory usage for objects in Java

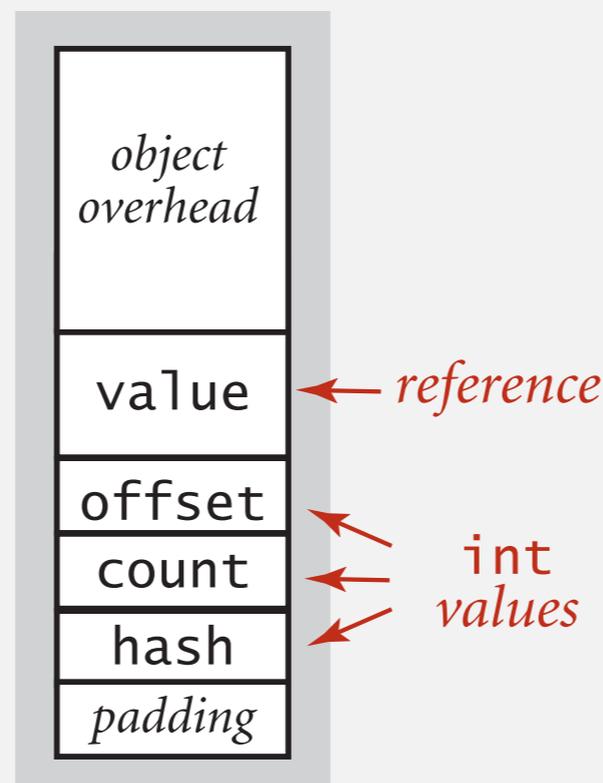
Object overhead. 16 bytes.

Reference. 8 bytes.

Padding. Each object uses a multiple of 8 bytes.

Ex 2. A virgin `String` of length N uses $\sim 2N$ bytes of memory.

```
public class String
{
    private char[] value;
    private int offset;
    private int count;
    private int hash;
    ...
}
```



16 bytes (object overhead)

8 bytes (reference to array)
2N + 24 bytes (char[] array)

4 bytes (int)

4 bytes (int)

4 bytes (int)

4 bytes (padding)

2N + 64 bytes

Typical memory usage summary

Total memory usage for a data type value:

- Primitive type: 4 bytes for `int`, 8 bytes for `double`, ...
- Object reference: 8 bytes.
- Array: 24 bytes + memory for each array entry.
- Object: 16 bytes + memory for each instance variable + 8 if inner class.

padding: round up
to multiple of 8



extra pointer to
enclosing class



Shallow memory usage: Don't count referenced objects.

Deep memory usage: If array entry or instance variable is a reference, add memory (recursively) for referenced object.

Memory profiler

Classmexer library. Measure memory usage of a Java object by querying JVM.

<http://www.javamex.com/classmexer>

```
import com.javamex.classmexer.MemoryUtil;

public class Memory {
    public static void main(String[] args) {
        Date date = new Date(12, 31, 1999);
        StdOut.println(MemoryUtil.memoryUsageOf(date));
        String s = "Hello, World";
        StdOut.println(MemoryUtil.memoryUsageOf(s));
        StdOut.println(MemoryUtil.deepMemoryUsageOf(s));
    }
}
```

← shallow

← deep

```
% javac -cp ../classmexer.jar Memory.java
% java -cp ../classmexer.jar -javaagent:../classmexer.jar Memory
32
40 ← don't count char[]
88 ← 2N + 64
```

use -XX:-UseCompressedOops
on OS X to match our model

Turning the crank: summary

Empirical analysis.

- Execute program to perform experiments.
- Assume power law and formulate a hypothesis for running time.
- Model enables us to **make predictions**.

Mathematical analysis.

- Analyze algorithm to count frequency of operations.
- Use tilde notation to simplify analysis.
- Model enables us to **explain behavior**.

Scientific method.

- Mathematical model is independent of a particular system; applies to machines not yet built.
- Empirical analysis is necessary to validate mathematical models and to make predictions.

