Algorithms: Truth, Beauty & Engineering

• Truth
  – History
  – Ethics

• Beauty
  – Elegance
  – Communication

• Engineering
  – Tricks of the trade
  – Eyes open to the world

History & Ethics

• Is China an Enron? (Part 2)
  By THOMAS L. FRIEDMAN, NY Times

If China forces out Google and suppresses the nation’s flow of knowledge, then it will be time to short the Chinese Communist Party.

Who is this guy??

Edward Snowden
NY Times Editorial

Edward Snowden, Whistle Bowler
January 2, 2014

“When someone reveals that government officials have routinely and deliberately broken the law, that person should not face life in prison at the hands of the same government. That’s why Rick Ledgett, who leads the N.S.A.’s task force on the Snowden leaks, recently told CBS News that he would consider amnesty if Mr. Snowden would stop any additional leaks. And it’s why President Obama should tell his aides to begin finding a way to end Mr. Snowden’s vilification and give him an incentive to return home.”
Elegance

- **Gordon Bell**: The cheapest, fastest and most reliable components are those that aren’t there.

- **Antoine de Saint Exupéry**: A designer knows he has arrived at perfection not when there is no longer anything to add, but when there is no longer anything to take away.

- **Albert Einstein**: Everything should be made as simple as possible, but no simpler.
January 13, 2018

Earlier Today

⚠️ EMERGENCY ALERTS

Emergency Alert
BALLISTIC MISSILE THREAT INBOUND TO HAWAII. SEEK IMMEDIATE SHELTER. THIS IS NOT A DRILL.

January 13, 2018

⚠️ EMERGENCY ALERTS

Emergency Alert
There is no missile threat or danger to the State of Hawaii. Repeat. False Alarm.
January 14-16, 2018

A Problem

Count the number of occurrences of all characters in a file.
A C++ Program

// count # occurrences of all characters in a file
// written: 8/5/94, Owen Astrachan, modified 5/1/99
void Print(const tvector<int> & counts, int total);
void Count(istream & input, tvector<int> & counts, int & total);

int main()
{
        int totalAlph = 0;
        string filename = PromptString("enter name of input file: ");
        ifstream input(filename.c_str());
        if (input.fail())
               cout << "could not open file " << filename << endl;
        exit(1);
        tvector<int> charCounts(CHAR_MAX+1,0); // all initialized to 0
        Count(input,charCounts,totalAlph);
        Print(charCounts,totalAlph);
        return 0;
}

void Count(istream & input, tvector<int> & counts, int & total)
// precondition: input open for reading
//               counts[k] == 0, 0 <= k < CHAR_MAX
// postcondition: counts[k] = # occurrences of character k
//                total = # alphabetic characters
{
        char ch;
        while (input.get(ch))               // read a character
               if (isalpha(ch))                // is alphabetic (a-z)?
               {   total++;
                   ch = tolower(ch);               // convert to lower case
                   counts[ch]++;                   // count all characters
               }
}

void Print(const tvector<int> & counts, int total)
// precondition: total = total of all entries in
//                counts['a']..counts['z']
// postcondition: all counts from 'a' to 'z' printed
{
        const int MIDALPH = 13;
        cout.setf(ios::fixed);    // print 1 decimal place
        cout.precision(1);
        char k;
        for(k = 'a'; k <= 'm'; k++)
               {   cout << k << setw(7) << counts[k] << " ";
                   cout << setw(4) << 100 * double(counts[k])/total
                               << "%
                     
               cout << char(k+MIDALPH) << setw(7)
                               << counts[k+MIDALPH] << " ";
                   cout << setw(4)
                               << 100 * double(counts[k+MIDALPH])/total
                               << "%" << endl;
               }
}

A Longer Program

The code is 3 1/2 pages long

Contains these Functions:

main, CountLetters, CountLettersInString,
RecordLetter, DisplayLetterCounts,
LetterIndex, ClearIntegerArray
Kernighan & Pike’s Version

unsigned long count[UCHAR_MAX+1];

/* freq main: display byte frequency counts */
int main(void)
{
    int c;

    while ((c = getchar()) != EOF)
    { count[c]++;

        for (c = 0; c <= UCHAR_MAX; c++)
            if (count[c] != 0)
                printf("\%.2x  %c  %lu\n", c, isprint(c) ? c : '-', count[c]);

    }
    return 0;
}
Communication

• How to talk about algorithms & computing?
• How to write about it?
• How to do presentations?
• How to exchange ideas?

Strunk & White’s Rule 17

Vigorous writing is concise. A sentence should contain no unnecessary words, a paragraph no unnecessary sentences, for the same reason that a drawing should have no unnecessary lines and a machine no unnecessary parts. This requires not that the writer make all sentences short, or avoid all detail and treat subjects only in outline, but that every word tell.
Algorithms: Truth, Beauty & Engineering

- **Truth**
  - History
  - Ethics

- **Beauty**
  - Elegance
  - Communication

- **Engineering**
  - Tricks of the trade
  - Eyes open to the world

A Quiz

- A TV Commercial
  - “U.S. college students eat 60 million slices of pizza per month.”
  - Is this reasonable?

- How much does a one-hour lecture cost?

- A program sorts 1 million integers in one second. How long to sort 2 million? 10 Million?

- How long will an exhaustive search take to solve a TSP of size 10? 20? 30?

- How much do CS110 texts cost?
Eyes open to the world

- **Cracking iPhones**
  - How do iPhones work?
    - Mathematics, science, technology
  - Back of the Envelope
    - Capacity of an iPhone (how many songs)?
  - History of music and protection
  - Ethics

iPhone 8 Plus
Algorithms

A Bird’s Eye View

Information Processing

Input  Process  Output
Problem Solving
(Investigation of Processes)

• Given a dictionary of english words, what are all the anagram classes? (e.g. earthling, haltering, lathering)
• Given the details of a tropical depression, can you predict if it will become a hurricane? What path will it take?
• Can you play the game of chess in a way that guarantees a win or a draw?
• How does the mind work?
• What is the most optimal way to get from here to there?
• What is the square root of 42?
• If the Fed. Raises the short term interest rates? What impact will it have on international currencies?
• What is the meaning of life?

Investigation: Asking Questions

Given: A problem/process.

• Does it have a model?
• Is the model solvable?
• Is it computable?
• What is the best algorithm for it?
• Write a computer program that implements the algorithm.
• Is the program equivalent to the model?
• Does the model lend any new insights into the problem/process?
For Lab This Week

• Make sure you know how to write and run a simple program in
  – C/C++ (gcc, g++)
  – Java (javac/java, or Eclipse), and
  – Python (idle)

• Make sure you can do it on a computer in the CS Lab.

Given: A Problem

Does it have a model?

Is the model solvable?

Is it computable?

What is the best algorithm?

Write a program that implements the algorithm that models the problem

Is the program correct?

Do the solutions offered by the program work in the real world?

Does the model/program offer any new insights into the problem?
Information Processing

Program that implements the model that solves the problem.

Input → Process → Output
As of June 2016, **Sunway TaihuLight** (神威·太湖之光) in China’s National Supercomputing Center in Wuxi, in Jiangsu province. It operates at ~94 Peta Flops/Second.
Computers & Clocks

- 1 flop = 1 floating-point operation (+, -, *, /) per second.
- 1 megaflop = 1 MF = 10^6 flops
- 1 gigaflop = 1 GF = 10^9 flops
- 1 teraflop = 1 TF = 10^12 flops
- 1 petaflop = 1 PF = 10^15 flops

- The fastest PC can give a peak performance of 1GF/second
- World’s fastest computer (Sunway TaiHu Light) operates at approximately ~94PF/sec.
  - Examining a single board of a chess game is roughly equivalent to 1000 FLOPS.
  - To make one move the computer will have to examine approximately 10^65 board situations.
  - This will take the Sunway TaiHu Light approximately 10^44 years to make one move!

Where the supercomputers are…
Computability:
Problems that can be solved by algorithms (Turing Machines)

Does it have a model?
  Is the model solvable?
    Is it computable?

What is the best algorithm?
  Write a program that implements the algorithm that models the problem
    Is the program correct?
      Do the solutions offered by the program work in the real world?
        Does the model/program offer any new insights into the problem?

Complexity Theory:
Computational Resources required (time & space)

Does it have a model?
  Is the model solvable?
    Is it computable?

What is the best algorithm?
  Write a program that implements the algorithm that models the problem
    Is the program correct?
      Do the solutions offered by the program work in the real world?
        Does the model/program offer any new insights into the problem?
What is an *algorithm*?

A set of instructions arranged in a specific order is a *procedure*.
Similar to a *recipe, process, method, technique, procedure, routine, rigmarole*, except the word “algorithm” connotes just a little something different. An algorithm is a finite, definite, effective procedure, with some output.


Algorithm (origins)

Abu ‘Abd Allah Muhammad ibn Musa al-Khwarizmi, ~780-850 A.D.
(Father of Abdullah, Mohammad, son of Moses, native of Khwarizm)
Was a member of Dar Al Hikmah (House of Wisdom) in Baghdad.
Kitab al jabr wa’l-muqabala (Rules of restoring and equating)
Algorithm (properties)

Finite
There must be an end to it within a reasonable time

Definite
Precisely definable in clearly understood terms, no "pinch of salt" type vagaries, or possible ambiguities

Effective
It must be possible to actually carry out the steps

Procedure
The sequence of specific steps

Output
Unless there is something coming out of the computation, the result will be unknown!

Problem Size

- Time complexity of a problem is the number of steps that it takes to solve an instance of the problem as a function of the size of the input. i.e. if the input is of size, \( n \), it will take \( f(n) \) steps to solve it.

![Problem size = 4](image)

Program that implements the model that solves the problem in \( f(n) \) steps.

Input  Process  Output
Complexity Classes
(# steps it takes to solve a problem)

- Constant time: $O(1)$
- Logarithmic time: $O(\log n)$
- Quadratic time: $O(n^2)$
- Polynomial time (P): $O(n^k)$
- Non-deterministic Polynomial time: $O(n^k)$ on some inputs
- Exponential time: $O(2^{p(n)})$
- Exponential time (in general): $O(n^k)$

Algorithms: Performance
Algorithm Performance vs. Program Performance

Actual performance depends on the computer running the program that implements the algorithm!

How long to sort 10 million numbers?

**Computer A**
- Speed: $10^{10}$ instructions/sec
- Running $O(n^2)$ sort
- Requires $2n^2$ instructions

How long will it take?

**Computer B**
- Speed: $10^7$ instructions/sec
- Running $O(n \log n)$ sort
- Requires $50n \log n$ instructions

How long will it take?
How long to sort 10 million numbers?

Computer A
Speed: $10^{10}$ instructions/sec
Running $O(n^2)$ sort
Requires $2n^2$ instructions

$$\frac{2 \times (10^7)^2}{10^{10}} \approx 20,000s$$

~5.5 hours

Computer B
Speed: $10^7$ instructions/sec
Running $O(n \log n)$ sort
Requires $50n \log n$ instructions

How long will it take?

How long to sort 10 million numbers?

Computer A
Speed: $10^{10}$ instructions/sec
Running $O(n^2)$ sort
Requires $2n^2$ instructions

$$\frac{2 \times (10^7)^2}{10^{10}} \approx 20,000s$$

~5.5 hours

Computer B
Speed: $10^7$ instructions/sec
Running $O(n \log n)$ sort
Requires $50n \log n$ instructions

$$\frac{50 \times 10^7 \times \log 10^7}{10^7} \approx 1163s$$

under 20 minutes!
How long to sort 10 million numbers?

**Computer A**

Speed: $10^{10}$ instructions/sec  
Running $O(n^2)$ sort  
Requires $2n^2$ instructions 

$$\frac{2 \times (10^7)^2}{10^{10}} \approx 20,000s$$

If running $50 \ n \ \log \ n$ program: < 2s!!

**Computer B**

Speed: $10^7$ instructions/sec  
Running $O(n \ \log \ n)$ sort  
Requires $50 \ n \ \log \ n$ instructions 

$$\frac{50 \times 10^7 \ \times \ \log 10^7}{10^7} \approx 1163s$$

under 20 minutes!

P = NP?
xkcd??

• More on xkcd.com