Lecture 2

Characteristics of Good Data Structures & Algorithms

Teera Siriteerakul

1

Recap: Data Structures & Algorithms

- Data Structures: Structures for keeping data
- Algorithms: Steps for solving problem
- Data Structures and Algorithms works together to solve problems.

2

Good Characteristics

- Good characteristics of Data Structures & Algorithms
  - Correctness
  - Efficiency
    - Time
    - Space

3

Correctness

- Proof of correctness
  - Empirical analysis
  - Formal reasoning, a.k.a. Mathematical proof

4

Empirical Analysis Example

Consider the following program

```
1  public class FindMax {
2     public static void main(String args[]) {
3        int a[] = {15, 2, 17, 4, 9, 12, 16};
4        int max = -1;
5        for(int i=0; i<a.length; i++) {
6           if(a[i]>max) {
7              max = a[i];
8           }
9        }
10       System.out.print("Max is "+max);
11    }
12 }
```

```
>java FindMax
Max is 17
```


5

Empirical Analysis Example

What happen when we change inputs?

```
1  public class FindMax {
2     public static void main(String args[]) {
3        int a[] = {-7, -12, -8, -5, -21, -6, -11};
4        int max = -1;
5        for(int i=0; i<a.length; i++) {
6           if(a[i]>max) {
7              max = a[i];
8           }
9        }
10       System.out.print("Max is "+max);
11    }
12 }
```

```
>java FindMax
Max is -1
```


6

Empirical Analysis

- Good for confirming theory
  - Beware of confirmation bias
- Good enough if we don't have theory?
  - Your "test cases" should be complete enough.
  - Should be careful with data outside test cases

7

Mathematical proof

- Sometimes call formal proof.
- Example of proving techniques
  - Direct proof
  - Proof by contradiction
  - Proof by contraposition
  - Proof by Mathematical Induction
- Good for 100% correctness.. In theory
- Details are outside this class.

8

Efficiency (in time and space)

- At the early age, we did care about space
- After memory becomes cheap, we tends to care for speed more
- In the age of Big Data, space becomes important again, in a whole new level.
- We typically use the same tool for evaluate efficiency in speed and space.

9

Measuring Efficiency

- Benchmark
  - Similar to empirical test
  - Good for measuring end products.
- Asymptotic analysis
  - Similar to Mathematical proof.
  - Concentrate on how time or space are increase as the size of input increase.
  - Compare the increasing trend to a familiar function.

10

Summary

- Characteristics of good Data Structures and Algorithms are **correctness** and **efficiency**
- To proof the correctness, we can test it (empirical analysis), or we can formally (mathematically) proof it.
  - Empirical analysis need good test cases.
  - Formal proof can guarantee the correctness.
- Measuring efficiency can be done by benchmark and asymptotic analysis.
  - Benchmark good for measuring end products
  - Asymptotic analysis measure growth rate of time or space comparing to the growth rate input.

11