CSCI 136
Data Structures &
Advanced Programming

Jeannie Albrecht

Lecture 8

Feb 25, 2014

Administrative Details

- Lab 3 is today
  - Lots of thinking…little typing
  - Problems can be done in any order!
  - Recursion can be frustrating…be patient!
- Using late days: Cannot work on last week's lab during this week's lab
- Hoping to return graded Lab 1 in lab today
  - General comments about Lab 1

Last Time

- Finished implementing vectors
- Started talking about the time-space tradeoff…

Growing Vectors

- Two ways to grow when adding n new elements to Vector:
  - Increase by 1 (or some other constant factor)
    - Requires ~$n^2/2$ operations (or copies)
  - Double
    - Requires ~n operations
  - Which is better?
  - Is there a tradeoff?

4

Vectors

- These questions relate to the time and space tradeoff
  - We could just as easily avoid all copy operations by making a huge Vector/array initially…
  - …but this wastes space and is inefficient

5

Today's Outline

- Wrap up Vectors
- Learn about Big-O analysis
- Review and discuss recursion
  - You'll get another chance to review recursion in lab this week…

Shrinking the Array

- When should we shrink the array in Vector implementation?
  - When 1/2 full?
  - When 1/4 full?
- We shrink when 1/4 full…
- Can get bad performance if array size changes too frequently

7

Vector Constructors

```
private Vector Object elementData[]; // the data
protected int elementCount; // number of elements in vector


public Vector() {
   this(10);
}

public Vector(int initialCapacity) {
   elementData = new Object[initialCapacity];
   elementCount = 0;

}
```

8

Vector Constructors

```
protected Vector Object elementData[]; // the data
protected int elementCount; // number of elements in vector
protected int capacityIncrement; // the rate of growth for vector

public Vector() {
   this(10);
}

public Vector(int initialCapacity) {
   elementData = new Object[initialCapacity];
   elementCount = 0;
   capacityIncrement = 0;
}

// pre: initialCapacity >= 0, capacityIncr >= 0
// post: constructs an empty vector with initialCapacity capacity
// that extends capacity by capacityIncr, or doubles if 0
public Vector(int initialCapacity, int capacityIncr) {
   elementData = new Object[initialCapacity];
   elementCount = 0;
   capacityIncrement = capacityIncr;
}
```

9

Observations about Vectors

- How long does it take to add an element?
  - Varies – sometimes takes a lot longer if we have to grow array before adding element
- How long does it take to insert/remove an element in the middle of the Vector?
  - Might take a long time if we have to move several other elements
- Key insight: The running time depends on the size of the Vector!

10

Running Time Analysis

- We want *general tools* for understanding how running time and memory usage changes as the amount of data increases
- Example:
  - If I double my Vector's size, how much longer will it take to:
    - Find an element?
    - Insert an element at the front?
    - Remove an element from the middle?
    - Etc.

11

Measuring Computational Cost

- How can we measure the cost of a computation?
  - Absolute clock time
    - Problems?
      - Different machines have different clocks
      - Lots of other stuff happening (network, OS, etc)
      - Not consistent. Need lots of tests to predict future behavior

12

Measuring Computational Cost

- How can we measure the cost of a computation?
  - Count how many "expensive" operations were performed (i.e., array copies in Vector)
  - Count number of times "x" happens
    - For a specific event or action "x"
    - i.e., How many times a certain variable changes
  - Problems?
    - 64 vs 65? 100 vs 105? Does it really matter??

13

Measuring Computational Costs

- Rather than keeping exact counts, we want to know the *order of magnitude* of occurrences
  - 60 vs 600 vs 6000, not 65 vs 68
- We want to make comparisons without looking at details and without running tests
- Avoid using specific numbers or values
- Look for overall trends

14

Looking for Trends

- Rule of thumb: ignore constants (most of the time…)
- Examples:
  - Treat n and n/2 as same order of magnitude
  - $n^2/1000$, $2n^2$, and $1000n^2$ are "pretty much" just n^2 (behave in same way)
  - $a_0n^k + a_1n^{k-1} + a_2n^{k-2} + \cdots a_k$ is roughly n^k
- The key is to find the most significant or dominant term

15

Asymptotic Bounds (Big-O Analysis)

- A function f(n) is O(g(n)) if and only if there exists positive constants c and n_0 such that

  $|f(n)| \leq c * g(n)$ for all $n \geq n_0$
- "g" is bigger than "f" **for large n**
- Example:
  - $f(n) = n^2/2$ is $O(n^2)$
  - $f(n) = 1000n^3$ is $O(n^3)$
  - $f(n) = n/2$ is $O(n)$

16

Determining Upper Bound

- We usually want the smallest upper bound to estimate running time
- Example:
  - $f(n) = 3n^2$
  - f(n) is $O(n^2)$
  - f(n) is $O(n^3)$
  - f(n) is $O(2^n)$
- Best estimate of running time is $O(n^2)$
- We only care about c and n_0 in practice, but focus on size of g when designing structures

17

Vector Operations

- For Object o, int i, and n elements:
  - set(i, o)
  - add(o)
  - add(i, o)
  - remove(i)
  - add(o) executed n times
  - add(i, o) executed n times

18

Vector Operations

- For Object o, int i, and n elements:
  - set(i, o) – O(1)
  - add(o) – O(1)
  - add(i, o) – O(n)
  - remove(i) – O(n)
  - add(o) executed n times – O(n)
  - add(i, o) executed n times – O(n^2)

19

Common Functions

For n = number of elements:

- O(1): constant time and space
- O(log n): divide and conquer algorithms, binary search
- O(n): linear dependence, simple list lookup
- O(n log n): divide and conquer sorting algorithms
- O(n^2): matrix addition, selection sort
- O(n^3): matrix multiplication
- O(n^k): cell phone switching algorithms
- O(2^n): color graph with 3 colors, satisfiability
- O(n!): traveling salesman problem

20

Input-dependent Running Times

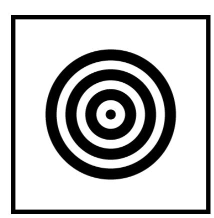
- Algorithms may have different running times for different input values
- Best case
  - Sort already sorted array in O(n)
  - Find item in first place that we look O(1)
- Worst case
  - Don't find item in list O(n)
  - Reverse order sort O(n^2)
- Average case
  - Linear search O(n)
  - Sort random array O(n log n)

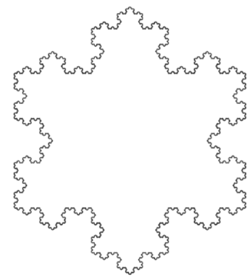
21

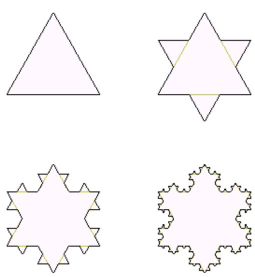
Moving on…

Recursion

- General problem solving strategy
  - Break problem into smaller pieces
  - Sub-problems may look a lot like original - may in fact by smaller versions of same problem
- Examples







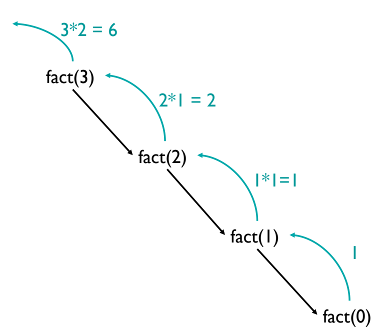
Recursion

- Many algorithms are recursive
  - Can be easier to understand (and prove correctness/state efficiency of) than iterative versions
- Today we will review recursion and then talk about techniques for reasoning about recursive algorithms

Factorial

- n! = n • (n-1) • (n-2) • … • 1
- How can we implement this?
  - We could use a while loop…
- But we could also write it recursively
  - n! = n • (n-1)!

Factorial



Factorial

- In recursion, we always use the same basic approach
- What's our base case?
  - n=0; fact(0) = 1
- What's our recursive case?
  - n>0; fact(n) = n • fact(n-1)

fact.java

```
public class fact{

    public static int fact(int n) {
       if (n==0) {
          return 1;
       }
       else {
          return n*fact(n-1);
       }
    }

    public static void main(String args[]) {
       System.out.println(fact(Integer.valueOf(args[0]).intValue()));
    }

}
```

Warm Up Problems

- Digit Sum
  - public static int digitSum(int n)
  - Base case?
  - Recursive case?
- Subset Sum
  - public static boolean canMakeSum(int set[], int target)
  - Helper:
    - public static boolean canMakeSumHelper(int set[], int target, int index)
  - Base case?
  - Recursive case?