Analysis of Algorithms
CS 3343 Lecture Ten

Prof. William Winsborough
October 23, 2008

Business

- Turn in Homework 5
  - Problem 7.1 (p.159)
  - Recall, I will also accept it Tuesday, 10/28.
- Recall Reading: 7.1, 7.2, 5.1, 5.2, 7.3, 7.4
- Homework 6: 5.2-3, 5.2-4, 7.2-1, 7.2-4, 7.3-1, 7.3-2, 7.4-1
  - Also: Justify each inequality in the derivation of A.10 on p.1066
  - And: explain the relationship between A.10 and the Norman Conquest (just kidding)

Quicksort Algorithm

- Partition(A, p, r)
  1. x ← A[r]
  2. i ← p – 1
  3. for j ← p to r – 1
  4. do if A[j] ≤ x
  5. then i ← i + 1
  6. exchange A[i] ↔ A[j]
  7. exchange A[i + 1] ↔ A[r]
  8. return i + 1

Randomized Quicksort

- What does randomizing achieve?
- Refer to algorithm on p.154

Worst-case Analysis of Quicksort

- Refer to the proof on p.155 that the worst-case running time is $\Theta(n^2)$

Expected-case Running time of Quicksort: Intuition Reprise

- We will assume the values in the input are distinct from one another
  - Why?
- We saw last time that if the splits created by partition are proportional to a constant (like 9 to 1), then the running time remains $\Theta(n \lg n)$
- The book also shows that if you introduce a couple really unbalanced splits between proportional splits, running time still remains $\Theta(n \lg n)$

Expected-case Formal Analysis

- There are at most n calls to Partition
  - This is because the pivot is always removed before making recursive calls to quicksort
- If X is the number of comparisons performed in line 4 throughout the entire execution of quicksort, then the running time is O(n + X)
- So our goal is to compute the expected value of X

Expected-case Formal Analysis, Cont.

- We analyze the total comparisons made during the *entire* computation, not just in one call to partition
- Need to understand when two elements of the array are compared and when they are not
- Let $z_1, z_2, \ldots, z_n$ enumerate the elements of the array from smallest to biggest
  - Let $Z_{i,j} = \{z_i, z_{i+1}, \ldots, z_j\}$
- When does the algorithm compare z_i and z_j?
- Note that each pair of elements is compared at most once
  - Elements are compared only to the pivot, and once a value is used as a pivot, it is never again compared to any element

Expected-case Formal Analysis, Cont.

- Let $X_{i,j} = I\{z_i \text{ is compared to } z_j\}$
  - Indicator random variable as discussed on Tuesday (sec. 5.2)
- $X = \Sigma_{1\le i\le n-1} \Sigma_{i+1\le j\le n} X_{i,j}$
- $E[X] = E[\Sigma_{1\le i\le n-1} \Sigma_{i+1\le j\le n} X_{i,j}]$
  $= \Sigma_{1\le i\le n-1} \Sigma_{i+1\le j\le n} E[X_{i,j}]$
  $= \Sigma_{1\le i\le n-1} \Sigma_{i+1\le j\le n} \Pr\{z_i \text{ is compared to } z_j\}$
- Remains to compute $\Pr\{z_i \text{ is compared to } z_j\}$

Expected-case Formal Analysis, Cont.

- When are two elements not compared?
  - Suppose values 1..10 (in any order) are to be partitioned
  - If 7 is selected as the pivot, the numbers 1..6 will never be compared with the numbers 8..10
  - In general, if $z_i < x < z_j$, and x is chosen as a pivot, z_i and z_j will never be compared later
- On the other hand, if z_i is chosen as a pivot, it will be compared to each element of $Z_{i+1,j}$
  - Similarly if z_j is chosen
  - In the example, 7 and 9 are compared, because 7 is the first item from $Z_{7,9}$ to be chosen as a pivot
  - On the other hand, 2 and 9 are not compared because the first item from $Z_{2,9}$ chosen as a pivot is 7

Expected-case Formal Analysis, Cont.

- In general, z_i and z_j will be compared just in case one of them is the first pivot chosen from $Z_{i,j}$
- Recall:
  - The probability of any item being chosen as a pivot is the same as the probability of any other element being chosen
  - There are j-i+1 elements in $Z_{i,j}$
  - So the probability of z_i being chosen before any other element in $Z_{i,j}$ is 1/(j-i+1)
  - Likewise for z_j

Expected-case Formal Analysis, Cont.

- So,
  $\Pr\{z_i \text{ is compared to } z_j\}$
  $= \Pr\{z_i \text{ or } z_j \text{ is the first pivot chosen from } Z_{i,j}\}$
  $= \Pr\{z_i \text{ is the first pivot chosen from } Z_{i,j}\}$
  $+ \Pr\{z_j \text{ is the first pivot chosen from } Z_{i,j}\}$
  (because the two events are mutually exclusive)
  $= 1/(j-i+1) + 1/(j-i+1)$
  $= 2/(j-i+1)$
- Now refer to equation 7.4 on p.158