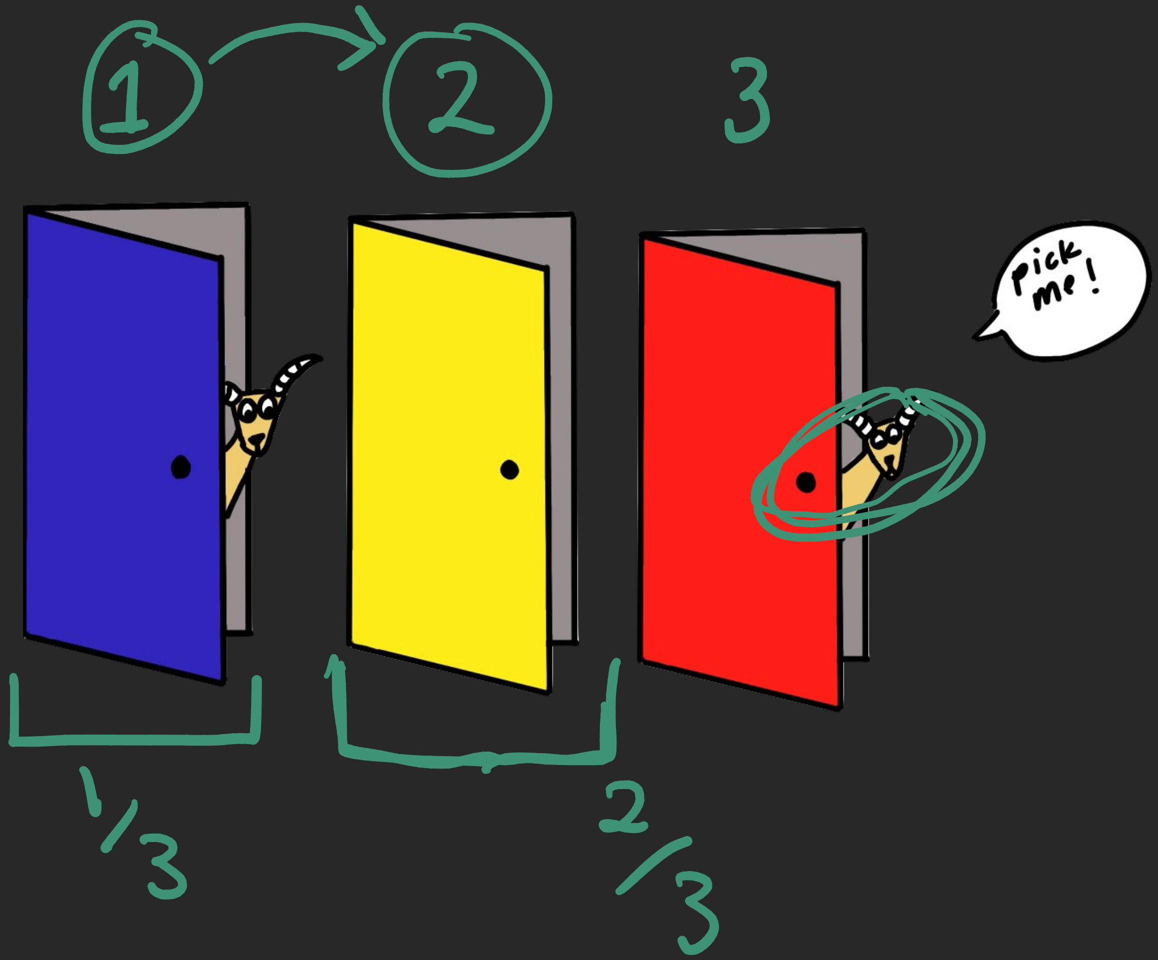
CSCI 3104
Lecture 5: Randomized Quicksort

Caleb Escobedo

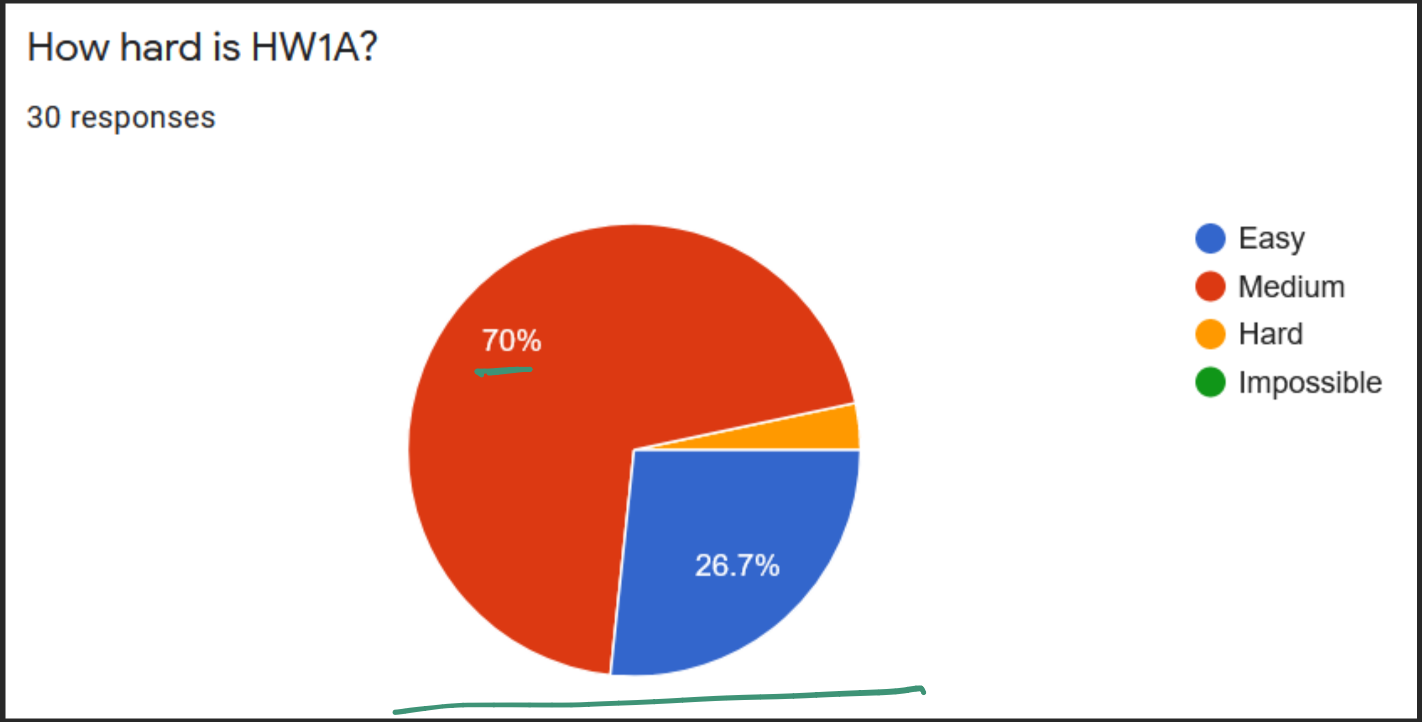
Caleb.Escobedo@colorado.edu

Lecture Outline

- Poll review/Administrative
- Review last class
  - Book resource location
  - Master Theorem
- Random Quicksort
- Introduction to Probability
- Analysis of Randomized Quicksort
- Alternative Quicksort
- Example Problems for HW1



Poll Review!

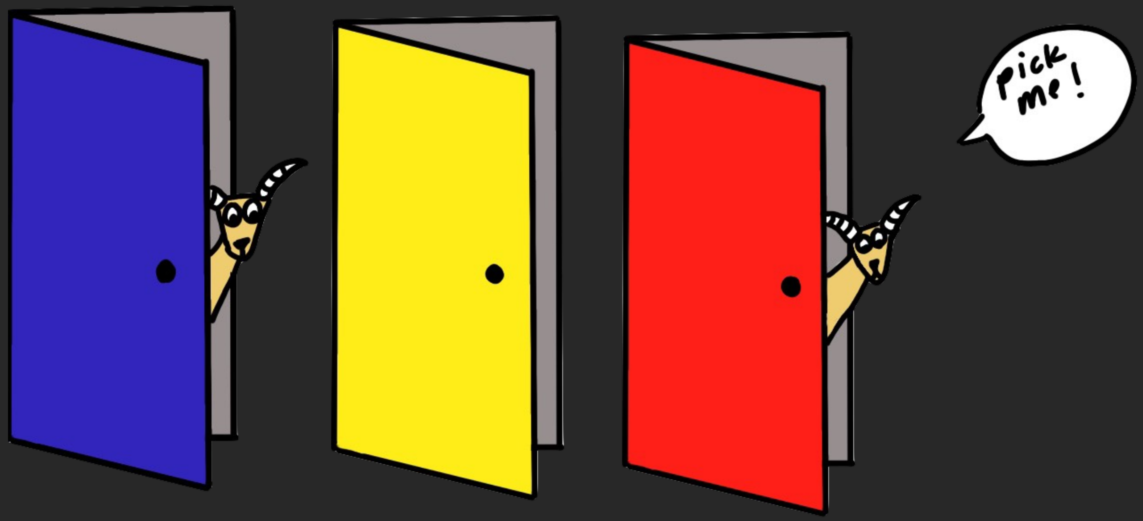

How hard is HW1A?
30 responses
70%
26.7%
Easy
Medium
Hard
Impossible

Administrative

- Video recordings error – will now be uploading mp4
- Examples today
  - 1 for Master Theorem
  - 2 ~~3~~ for HW1A
- HW2A released by tomorrow morning

Lecture Outline

- ~~Poll review/Administrative~~
- Review last class
  - Book resource location
  - Master Theorem
- Random Quicksort
- Introduction to Probability
- Analysis of Randomized Quicksort
- Alternative Quicksort
- Example Problems for HW1



Last Class: Quicksort

- Quicksort – Ch. 7
  - Worst-case analysis $= O(n^2)$
    - Book: Ch. 7.2 page 175 and Ch. 7.4.1 page 180
    - In class: Unrolling
  - Best/Expected-case analysis: $O(n \log n)$
    - Book: Ch. 7.2 page 175
    - In class: Master Theorem

Last Class: Master Theorem

Master theorem. Let $a \geq 1$ and $b > 1$ be constants, let $f(n)$ be a function and let $T(n)$ be defined on the non-negative integers by the recurrence,

$$T(n) = a\,T(n/b) + f(n) \ ,$$

where we interpret n/b to mean either $\lfloor n/b \rfloor$ or $\lceil n/b \rceil$. Then $T(n)$ can be bounded asymptotically as follows:

1. If $f(n) = O(n^{\log_b a - \epsilon})$ for some constant $\epsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$
2. If $f(n) = \Theta(n^{\log_b a})$, then $T(n) = \Theta(n^{\log_b a} \log n)$
3. If $f(n) = \Omega(n^{\log_b a + \epsilon})$ for some constant $\epsilon > 0$, and if $a\,f(n/b) \leq c\,f(n)$ for some constant $c < 1$ and all sufficiently large n, then $T(n) = \Theta(f(n))$. □

- Example of how students will use it!

$$f(n) = \Theta(n^2)$$

$$T(n) = 24T\left(\frac{n}{5}\right) + \Theta(n^2)$$

$$f(n) \ ? \ n^{\log_b a}$$

$$\boxed{T(n) = \Theta(n^2)}$$

$$\log_5(25) = 2$$

$$\log_5(24) < 2$$

$$n^2 < n^{\log_5(24)}$$

$$\epsilon = 1$$

$$24\left(\frac{n}{5}\right)^2 < c\,n^2$$

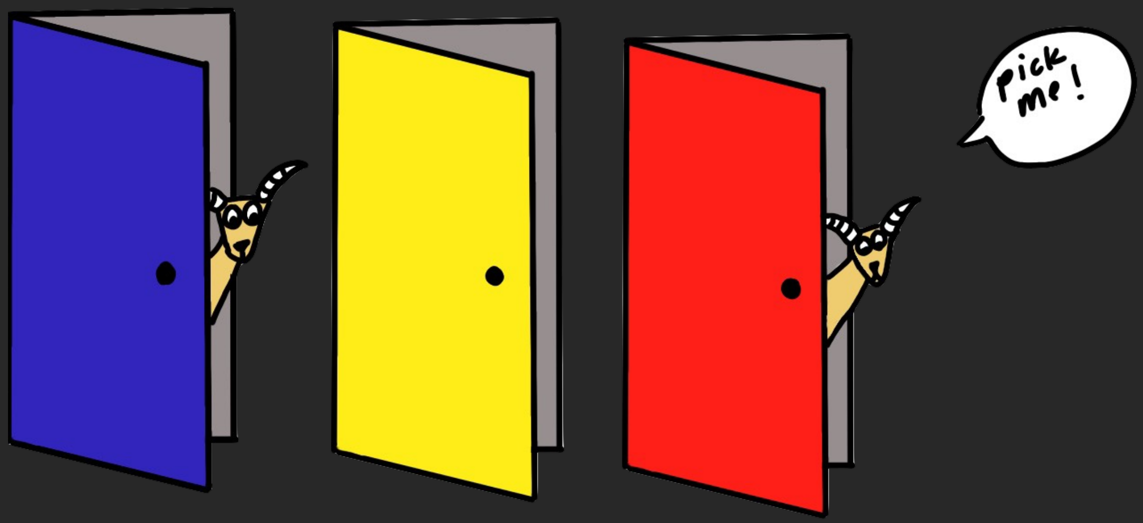
$$\frac{24}{25} n^2 < c n^2$$

$$\frac{24}{25} < c < 1$$

$$c = \frac{24.1}{25}$$

Lecture Outline

- ~~Poll review/Administrative~~
- ~~Review last class~~
  - ~~Book resource location~~
  - ~~Master Theorem~~
- Random Quicksort
- Introduction to Probability
- Analysis of Randomized Quicksort
- Alternative Quicksort
- Example Problems for HW1



Randomized Quicksort

- Goal: Show that we can use randomness to achieve the average case consistently
- How often do we get the worst case?

max $O(n^2)$

- How often do we get the best case?

x_1 .... x_n

```
Partition(A,p,r) {
   x = A[r]
   i = p-1
   for (j=p; j<=r-1;j++) {
      if A[j]<=x {
         i++
         exchange(A[i],A[j])
   }}
   exchange(A[i+1],A[r])
   return i+1
}
```

Randomized Quicksort

- Question: How many ways can n numbers be ordered?

[1, 2, 3] $n!$

$3 \cdot 2 \cdot 1$ $3!$

```
Partition(A,p,r) {
   x = A[r]
   i = p-1
   for (j=p; j<=r-1;j++) {
      if A[j]<=x {
         i++
         exchange(A[i],A[j])
   }}
   exchange(A[i+1],A[r])
   return i+1
}
```

Randomized Quicksort

$O(n^2)$

- Can we disconnect the behavior of our algorithm from the input form? yes

p=lower
r=upper

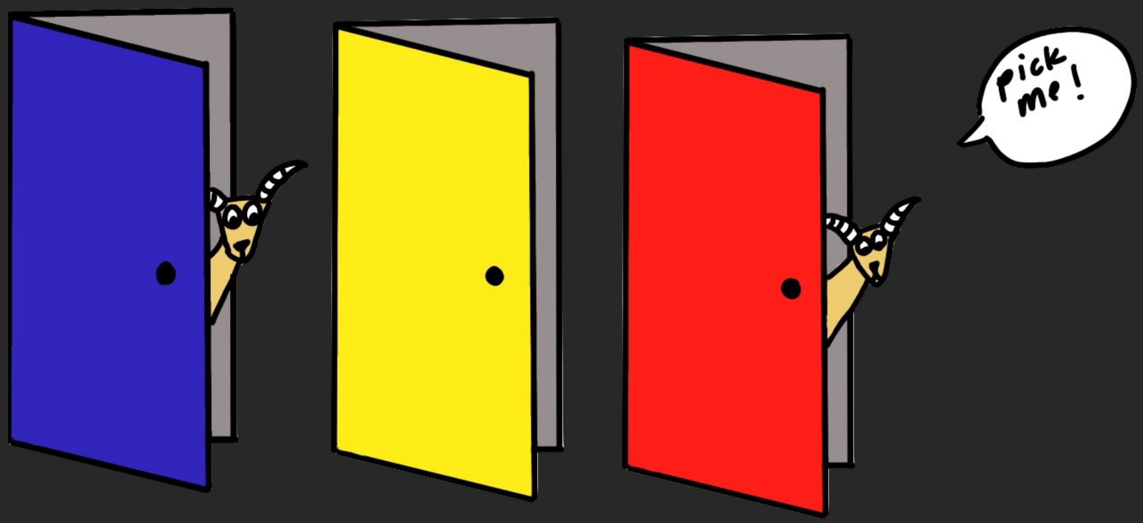
```
Randomized-Quicksort(A,p,r) {
    if (p<r){
        q = RPartition(A,p,r)
        Randomized-Quicksort(A,p,q-1)
        Randomized-Quicksort(A,q+1,r)
}}.
```

new

```
RPartition(A,p,r) {
  i = random-int(p,r)     // NEW: choose uniformly random integer on [p..r]
  swap(A[i],A[r])         // NEW: swap corresponding element with last element
  x = A[r]                // pivot is now a uniformly random element
  i = p-1                 // code from deterministic Partition
  for (j=p; j<=r-1;j++) {
     if A[j]<=x {             // same as before
        i++                   //
        swap(A[i],A[j])       //
  }}
  swap(A[i+1],A[r])       //
  return i+1              //
},
```

Lecture Outline

- ~~Poll review/Administrative~~
- ~~Review last class~~
  - ~~Book resource location~~
  - ~~Master Theorem~~
- ~~Random Quicksort~~
- Introduction to Probability
- Analysis of Randomized Quicksort
- Alternative Quicksort
- Example Problems for HW1



Introduction to Probability

- X is a random variable that can take several values, each with some probability.

100%

$$E(X) = \sum_x x \Pr(X = x) \qquad \text{given that } \sum_x \Pr(X = x) = 1 \text{ and } x \text{ is discrete}$$

$$E[X] = 1 \cdot \left(\tfrac{1}{6}\right) + 2\left(\tfrac{1}{6}\right) + 3\left(\tfrac{1}{6}\right) + 4\left(\tfrac{1}{6}\right) + 5\left(\tfrac{1}{6}\right) + 6\left(\tfrac{1}{6}\right)$$

$$\frac{1}{6}\sum_{i=1}^{6} i = \frac{1}{6}\,\frac{6(6+1)}{2} = \frac{21}{6} = \boxed{3.5} \qquad \sum_{i=1}^{n} = \frac{n(n+1)}{2}$$

Introduction to Probability

- The outcome of our die are independent, formally written as:

$$\Pr(\underbrace{X = x}_{1}, \underbrace{Y = y}_{2}) = \Pr(X = x)\Pr(Y = y)$$

- What is the probability of getting "snake eyes"?

$$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

Introduction to Probability

- What is the probability that *X* or *Y* *is 1?*

$$\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

- What is the probability of getting a Yahtzee? (5 die land on 6)

0.00077

Introduction to Probability

- Indicator random variable: used to indicate the amount of time some event has occurred

did we roll a 6?

$$I(A) = \begin{cases} 1 & \text{if the event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}.$$

Introduction to Probability

- Linearity of Expectations

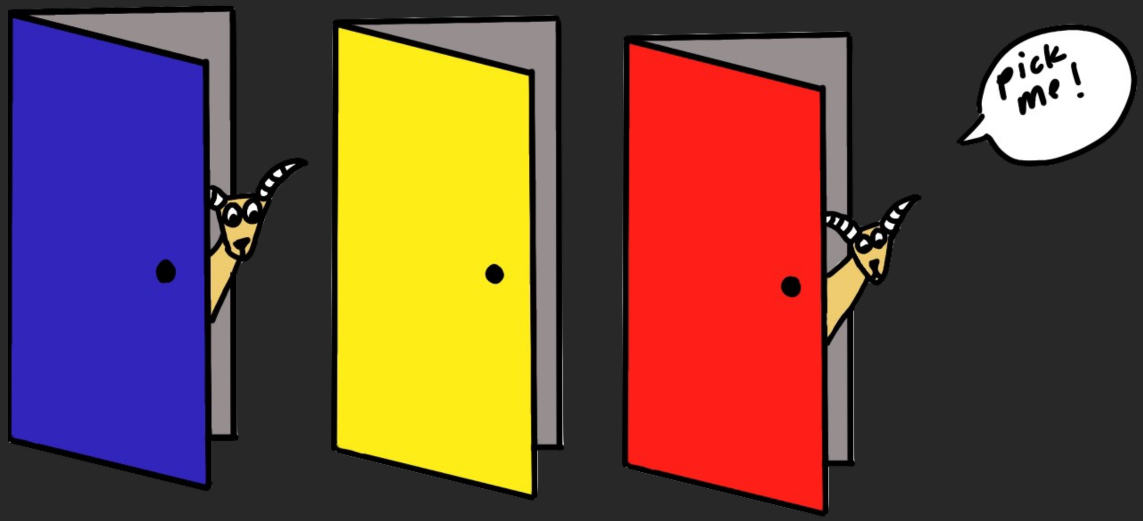
$$\overset{1-6}{E(X} + \overset{1-6}{Y}) = E(X) + E(Y)$$

$3.5 + 3.5 = 7$

$$E\left(\sum_i X_i\right) = \sum_i E(X_i)$$

Lecture Outline

- ~~Poll review/Administrative~~
- ~~Review last class~~
  - ~~Book resource location~~
  - ~~Master Theorem~~
- ~~Random Quicksort~~
- ~~Introduction to Probability~~
- Analysis of Randomized Quicksort
- Alternative Quicksort
- Example Problems for HW1



Randomized Quicksort Analysis

- Input: $A = [z_1, z_2, \ldots, z_n]$

1) z_i and z_j are in the same sub problem

Small ↓ large

- Contains the set: → $Z_{ij} = \{z_i, z_{i+1}, \ldots, z_{j-1}, z_j\}$

z_i is the smallest element

z_j is the largest element

2) z_i or z_j is the pivot.

Randomized Quicksort Analysis

$O(n^2)$

runtime

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$$

$$X_{ij} = I\{z_i \text{ is compared to } z_j\}$$

Randomized Quicksort Analysis

random runtime

- Apply linearity of expectations to

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$$

$$\begin{aligned} E(X) &= E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right] \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}] \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Pr(z_i \text{ is compared to } z_j) \end{aligned}$$

Randomized Quicksort Analysis

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Pr(z_i \text{ is compared to } z_j)$$

$j - i + 1 = \text{length}$

$$\Pr(z_i \text{ is compared to } z_j) = \Pr(\text{either } z_i \text{ or } z_j \text{ are the first pivots chosen in } Z_{ij})$$

$$= \Pr(z_i \text{ is first pivot chosen in } Z_{ij}) + \Pr(z_j \text{ is first pivot chosen in } Z_{ij})$$

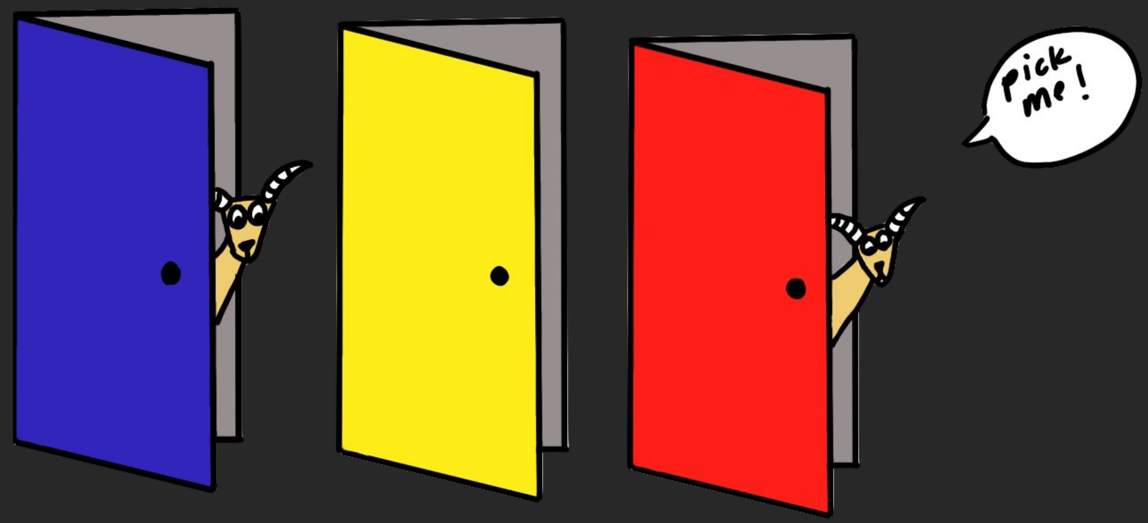
$$\frac{1}{j-i+1} + \frac{1}{j-i+1} = \frac{2}{j-i+1}$$

Randomized Quicksort Analysis

$$
\begin{aligned}
E(X) &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Pr(z_i \text{ is compared to } z_j) \;. \\
&= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} \\
&= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} && \text{change of variables } k = j - i \\
&< \sum_{i=1}^{n-1} \sum_{k=1}^{n} \frac{2}{k} && \text{lower bound} \\
&= \sum_{i=1}^{n-1} O(\log n) && \text{harmonic series bound} \\
&= O(n \log n) \;.
\end{aligned}
$$

Lecture Outline

- ~~Poll review/Administrative~~
- ~~Review last class~~
  - ~~Book resource location~~
  - ~~Master Theorem~~
- ~~Random Quicksort~~
- ~~Introduction to Probability~~
- ~~Analysis of Randomized Quicksort~~
- Alternative Quicksort ➜
- Example Problems for HW1



Median-of-k algorithm

$k = 3$

$A = [8, 2, 5, 4, 3, 6, 7, 1]$

$[8, 4, 1]$

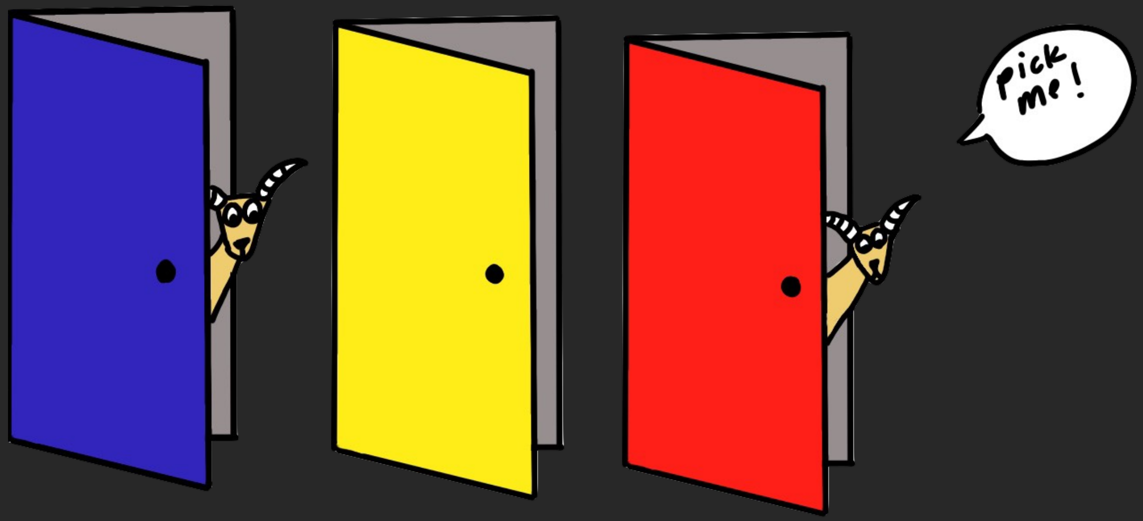
pivot

Closing Thoughts: Random Quicksort

- Introduction: Will not test on this method
- Randomness can be added to many algorithms
- Stick around if you want to review some HW1 problems
- Please will out the poll! I really like hearing from you all

Lecture Outline

- ~~Poll review/Administrative~~
- ~~Review last class~~
  - ~~Book resource location~~
  - ~~Master Theorem~~
- ~~Random Quicksort~~
- ~~Introduction to Probability~~
- ~~Analysis of Randomized Quicksort~~
- ~~Alternative Quicksort~~
- Example Problems for HW1



HW1: Example 1

$f(n) = n^2 + n^2 + n$

$g(n) = n^2$

$\Theta(g(n)) = \{f(n) :$ there exist positive constants c_1, c_2, and n_0 such that $0 \le c_1 g(n) \le f(n) \le c_2 g(n)$ for all $n \ge n_0\}$.[1]

3.1-1

Let $f(n)$ and $g(n)$ be asymptotically nonnegative functions. Using the basic definition of Θ-notation, prove that $\max(f(n), g(n)) = \Theta(f(n) + g(n))$.

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \frac{2n^2+n}{n^2} = \lim_{n\to\infty} \frac{4n+1}{2n} = \lim_{n\to\infty} \frac{4}{2} = 2$$

$$f(n) = \Theta(g(n))$$

HW1: Example 2

$\Theta(g(n)) = \{f(n) :$ there exist positive constants c_1, c_2, and n_0 such that $0 \le c_1 g(n) \le f(n) \le c_2 g(n)$ for all $n \ge n_0\}$.[1]

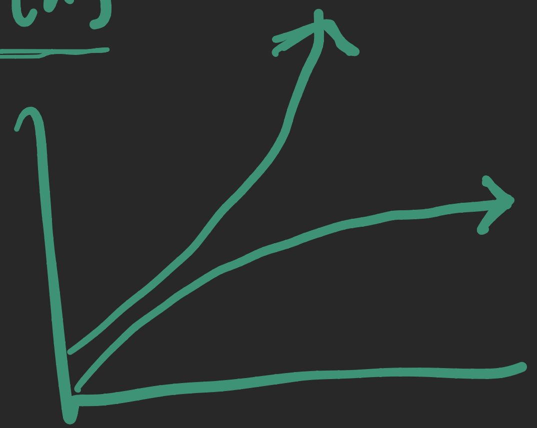
3.1-2

Show that for any real constants a and b, where $b > 0$,

$$(n + a)^b = \Theta(n^b) .$$

$g(n) = n^{10}$

$f(n) = \underline{n^{10} + \log(n)}$



HW1: Example 2

$\Theta(g(n)) = \{f(n) :$ there exist positive constants c_1, c_2, and n_0 such that $0 \le c_1 g(n) \le f(n) \le c_2 g(n)$ for all $n \ge n_0\}$.[1]

3.1-2

Show that for any real constants a and b, where $b > 0$,

$$(n + a)^b = \Theta(n^b) .$$