Random Binary Search
(Slides for ESTR2102)

Yufei Tao

Department of Computer Science and Engineering
Chinese University of Hong Kong
The Dictionary Search Problem

Input

An array $A$ of $n$ integers, sorted in ascending order. And a search value $q$.

Output

Determine whether $q \in S$. 
$i = \text{RANDOM}(1, n)$
If $A[i] = q$, then done.
If $A[i] < q$, recurse on $A[1 : i - 1]$
Otherwise, recurse on $A[i + 1 : n]$


Remark 2: For our discussion, we will refer to $A[i]$ as a pivot.
Random Binary Search

We will prove that the algorithm finishes in $O(\log n)$ time in expectation.

We will focus on only the scenario where $q \notin S$.

Suppose that $A = (e_1, e_2, \ldots, e_n)$. For each $i \in [1, n]$, define random variable $X_i$:

- 1 if $e_i$ is compared to $q$ in the algorithm (i.e., $e_i$ is one of the pivots picked).
- 0, otherwise.
Random Binary Search

The expected running of the algorithm is

\[ O \left( \sum_{i=1}^{n} E[X_i] \right). \]

We will prove

\[ \sum_{i=1}^{n} E[X_i] = O(\log n). \]
Focus on a particular $i \in [1, n]$. Without loss of generality, assume that $e_i < q$. Suppose that $e_{i+1}, e_{i+2}, \ldots, e_{i+t}$ are less than $q$, for some $t \geq 0$.

**Lemma:** $Pr[X_i = 1] = 1/(t + 1)$.

**Proof:** Define $Y$ be the first pivot falling in $[e_i, e_{i+t}]$. Note that $Y$ definitely exists (think: why?).

$X_i = 1$ if and only if $Y = e_i$. $Y$ can be any of $e_i, e_{i+1}, \ldots, e_{i+t}$ with the same probability. We thus complete the proof. **QED**
Random Binary Search

It thus follows from the previous lemma that

$$\sum_{i=1}^{n} E[X_i] \leq 2 \sum_{i=1}^{n} \frac{1}{i} = O(\log n).$$

Think: why?

**Remark:** $1 + 1/2 + 1/3 + ... + 1/n$ is the harmonic series. The value is between $\ln(n + 1)$ and $1 + \ln n$. 