randomization
What is your best strategy to survive & become King??
CHECK PROCEDURE:

RANDOMLY PICK 50 MATCHES AND LIGHT THEM

IF ONE FAILS, REJECT THE BOX.

IF ALL SUCCEED, ACCEPT IT.
PR THAT TEST FAILS

THREE CASES TO CONSIDER:

Suppose the uncle tampers with

< 50 matches ⇒ You camp + live, or you don't camp + live.

≥ 50 matches ⇒ Kill the uncle!! b/c you catch him

50 matches ⇒ one case of failure =

you test the 50 good matches &

camp w/ 50 bad ones.
What is the pr of failure??

100 matches, 50 good ones. You pick those so good ones!!

\[
\frac{50}{100} \left( \frac{49}{99} \right) \left( \frac{48}{98} \right) \ldots \left( \frac{1}{50} \right)
\]

\[
\frac{1}{\binom{100}{50}} \quad \approx \quad \frac{1}{\sqrt{\frac{2\pi n}{n}} n^{n+1/2} e^{-n}}
\]

via Stirling identity

\[
\frac{1}{\binom{100}{50}} < 2^{-98}
\]
PR OF DEATH:

9.91165302141833906737674969688360149
5412210270643283767892785256889073029
997327393587632943101698342E−30
PR OF DEATH:

9.91165302141833906737674969688360149
5412210270643283767892785256889073029
997327393587632943101698342E-30

PR OF ROYAL FLUSH:

1.53908E-6

10^{-5}
<table>
<thead>
<tr>
<th>Age in 1990</th>
<th>Total U.S.</th>
<th>White Male</th>
<th>White Female</th>
<th>Black Male</th>
<th>Black Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.102%</td>
<td>0.128%</td>
<td>0.045%</td>
<td>0.307%</td>
<td>0.074%</td>
</tr>
</tbody>
</table>

$10^{-3}$
FINGERPRINTING

Alice

Bob

Disk

$10^8$

Disk

$10^8$
FINGERPRINTING

Alice

Pick prime $p$

$A \mod p$

Bob

$B \mod p \equiv x$

$128$ bits

$64$ bits
FINGERPRINTING

Alice

Bob

PICK PRIME P

SEND p, A mod p

if A=B, then this protocol certainly succeeds!!

if A=B, then

this protocol certainly succeeds!!
FINGERPRINTING

Alice

PICK PRIME $p$

SEND $p$, $A \mod p$

Bob

COMPARE WITH $B \mod p$

$A = B$

$A = 26$

$B = 39$

IF $A \neq B$ THEN there is some small chance that the protocol makes an error and convinces Alice Bob that their disks are equal.

$p = 13$. 
NUMBER OF PRIMES

00  EUCLID.

how many primes are there that are ≤ 2^64
NUMBER OF PRIMES
THERE ARE CERTAINLY INFINITELY MANY

$\pi(x) : \# \text{ OF PRIMES } < x$

$\pi(x) > \frac{x}{\log x}$

$x \sim 2^{6^\gamma}$

$\pi(2^{128}) > \frac{2^{128}}{128} \sim 2^{121}$

$\pi(2^{6^g}) > \frac{2^{6^g}}{6^g} \sim 2^{58}$

$\pi(x) : \# \text{ of primes } < x$
LEMMA: # OF PRIME DIVISORS OF \( x \) < \( \log(x) \)

2 is the smallest prime.

If \( x \) has \( t \) prime divisors then

\[ x \geq 2^t. \]
PR OF FALSE MATCH:

Suppose that \( A \neq B \), but the protocol concludes "match".

\[ A \mod p = B \mod p \]

\[ \Rightarrow (A-B) = 0 \mod p \Rightarrow p \text{ divides } (A-B) \]

\[ \Rightarrow \text{how many prime divisors can } (A-B) \text{ have?} \]

\[ \log(A-B) \sim \log \left( 2^{2^{2^{32}}} \right) \sim 2^{2^{32}} \]

\[ \Rightarrow \text{how many 64-bit primes are there?} \quad 2^{58} \]

\[ \text{Pr failure is } \leq \frac{2^{2^{32}}}{2^{58}} \sim 2^{-18} \]
EXAMPLE PARAMS

RANDOMLY PICK 64BIT PRIME P

SEND $p$, $A \mod p$

COMPARE WITH $B \mod p$
A squabble between a group fighting spam and a Dutch company that hosts Web sites said to be sending spam has escalated into one of the largest computer attacks on the Internet, causing widespread congestion and jamming crucial infrastructure around the world. Millions of ordinary Internet users have experienced delays in services like Netflix or could not reach a particular Web site for a short time.

However, for the Internet engineers who run the global network the problem is more worrisome. The attacks are becoming increasingly powerful and computer security...
STRING MATCHING

PATTERN

CORPUS
for (int i = 0, j=0; i < n-m; i++) {
    while (j < m && t[i+j] == p[j]) { j++; }
    if (j == m) return i;
}
return -1;

Running time \(O(n \cdot m)\)
SIMPLE ALGORITHM

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAB

BRUTE FORCE WORST CASE:
KMP algorithm

abcdabcdabcdefhabcdabhi
KMP algorithm

ABCDABCDABCDABCDABabcdefh

ABCDABHII
SLIDING RULE

Given that $P[1...Q]$ matches $T[m+1...m+Q]$, but a mismatch occurs at $Q+1$, then:

Text
SLIDING RULE

Given that $P[1...Q]$ matches $T[m+1...m+q]$, but a mismatch occurs at $q+1$, then:

- Find the longest prefix of $P[1...Q]$ that is also a suffix of $P[1...Q]$.
- Slide so that $P[1...p]$ matches $T[i-p+1,...]$. 

$\Theta(n+m)$
NEW IDEA FOR STRING MATCH
STRING MATCHING

PICK RANDOM T-BIT PRIME

COMPUTE PATTERN MOD PRIME

FOR i=1...n

COMPUTE NEXT CORPUS MOD PRIME

COMPARE, OUTPUT MATCH IF SAME

(Alice operation)

Bob's operations over sliding window of size $m$
PICK AN 80-BIT PRIME P

What is the probability of a mismatch at the first position?

If the pattern is of size \( m \) bits and the pattern has \( \leq m \) prime divisors,

\[
\begin{align*}
\Pr \left[ \text{mismatch @ position 0} \right] &< \frac{m}{2^{121}} \quad \# \text{ of 128-bit primes.} \\
\Pr \left[ \text{another mismatch} \right] &< \frac{m}{2^{121}} \quad \ldots
\end{align*}
\]

There are \( n \) positions to match, so

\[
\Pr \left[ \text{At least one fails} \right] < \frac{n \cdot m}{2^{121}}
\]
PR OF ANY MISMATCH:
STRING MATCHING

PICK RANDOM T-BIT PRIME

COMPUTE PATTERN MOD PRIME

FOR I=1...N

   COMPUTE NEXT CORPUS MOD PRIME
   COMPARE, OUTPUT MATCH IF SAME
STRING MATCHING EXAMPLE

PATTERN

Text

3141592653589793123127398

\[ p = 13 \]

\[ 26535 \mod 13 = 2 \]
Given that $31415 \mod 13 = 7$,
How can I compute $14159 \mod 13$?

Hint: $10000 \mod 13 = 3$

$10 \cdot \left[ \frac{31415}{13} - 3 \cdot \frac{10000}{13} \right] + 9 = 14159 \mod 13$

$\left[ \frac{7 - 3 \cdot 3}{13} \right] \cdot 10 + 9 = 15 = 2 \mod 13$

$\Theta(1)$ single mod $p$ operations.
<table>
<thead>
<tr>
<th>Pattern</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>26535</td>
<td>3141592653589793123127398</td>
</tr>
</tbody>
</table>

14159 =
STRING MATCHING

PICK RANDOM T-BIT PRIME

COMPUTE PATTERN MOD PRIME

FOR I=1...N

    COMPUTE NEXT CORPUS MOD PRIME
    COMPARE, OUTPUT MATCH IF SAME

\[ \Theta(n + m) \mod p \text{ operations} \]
GOAL:

DEVISE A RELIABLE METHOD FOR NODES TO SEND MESSAGE TO THE SERVER WITH AS LITTLE COORDINATION AS POSSIBLE.
FIRST EXAMPLE
SIMPLE ALGORITHM

AT TIME T, FLIP A COIN THAT IS HEADS WITH PR $\frac{1}{n}$

IF HEADS, THEN BROADCAST. IF SUCCESS, THEN STOP.
ELSE WAIT AND TRY AGAIN.
REPEAT $cn \log n$ TIMES
ANALYZE THE SIMPLE ALGORITHM
\[ S_{i,t} = \]

\[ \Pr[S_{i,t} = 1] = \]
\[
\Pr[S_{i,t} = 1] = \frac{1}{n} \left(1 - \frac{1}{n}\right)^{n-1}
\]
FACT: IF \[ f(n) = \left(1 - \frac{1}{n}\right)^n \] THEN
FACT: IF $f(n) = \left(1 - \frac{1}{n}\right)^n$ THEN
$S_{i,t} = \text{NODE } i \text{ SUCCEEDS IN SENDING AT TIME } t$

$\frac{1}{en} \leq \Pr[S_{i,t} = 1] \leq \frac{1}{2n}$
\[ F_{i,t} = \text{FAILURE} \]
\[ F_{i,t} = \begin{cases} \text{NODE } i \text{ fails to send at times } 1, 2, \ldots, t \\ \end{cases} \]

\[ \Pr[F_{i,t}] = \bigwedge_{j=1}^{t} \Pr[S_{i,j}] \]
\[ F_{i,t} = \ \text{NODE } i \text{ FAILS TO SEND AT TIMES } 1,2,\ldots,t \]

\[
\Pr[F_{i,t}] = \bigwedge_{j=1}^{t} \Pr[S_{i,j}] = \prod_{j=1}^{t} \Pr[S_{i,j}]
\]
\[ \Pr[F_{i,t}] = \bigwedge_{j=1}^{t} \Pr[S_{i,j}] = \prod_{j=1}^{t} \Pr[S_{i,j}] \]
\[ F_{i,t} = \text{NODE } i \text{ FAILS TO SEND AT TIMES } 1,2,...,t \]

\[
\Pr[F_{i,t}] = \bigwedge_{j=1}^{t} \Pr[S_{i,j}] = \prod_{j=1}^{t} \Pr[S_{i,j}]
\]

\[ t = O(n \ln n) \]

\[ \Pr[F_{i,t}] = n^{-c} \]
\[ F_t = \]

\[ \Pr[F_t] = \]
\[ F_t = \text{SOME NODE } i \text{ FAILS TO SEND AT TIMES } 1,2,\ldots,t \]

\[ \Pr[F_t] = \bigcap_{i=1}^{n} \Pr[F_{i,t}] \]
\[ F_t = \text{SOME NODE } i \text{ Fails to send at times } 1, 2, \ldots, t \]

\[ \Pr[F_t] = \bigvee_{i=1}^{n} \Pr[F_{i,t}] \leq \sum_{i=1}^{n} \Pr[F_{i,t}] \leq \sum_{i=1}^{n} n^{-c} \]
SUMMARY

AT TIME T, FLIP A COIN THAT IS HEADS WITH PROBABILITY $\frac{1}{n}$

IF HEADS, THEN BROADCAST. IF SUCCESS, THEN STOP.

ELSE WAIT AND TRY AGAIN.
REPEAT $O(n \ln n)$ TIMES

WITH PROBABILITY

EVERY NODE SUCCEEDS IN SENDING MESSAGE.
TOOLS WE USED

ANALYSIS OF

\[
\left(1 - \frac{1}{n}\right)^n
\]

PROBABILITY THAT MANY INDEPENDENT EVENTS ALL OCCUR:

PROBABILITY THAT ONE OUT OF N EVENTS OCCURS:
SECOND EXAMPLE:

MEDIAN
SELECT \((i, A[1, \ldots, n])\)

PICK FIRST ELEMENT

PARTITION LIST ABOUT THIS ONE

IF PIVOT IS POSITION \(i\), RETURN PIVOT

ELSE IF PIVOT IS IN POSITION \(> i\)  SELECT \((i, A[1, \ldots, p - 1])\)

ELSE  SELECT \(((i - p - 1), A[p + 1, \ldots, n])\)
PROBLEM: WHAT IF WE ALWAYS PICK BAD PARTITIONS?
\( \text{PARTITION } (A[1, \ldots, n]) \)

\( B[1, \ldots, \lfloor n/5 \rfloor] \)

\( \text{SELECT } (\lfloor n/5 \rfloor/2, B[1, \ldots, \lfloor n/5 \rfloor]) \)
A NICE PROPERTY OF OUR PARTITION

\[
3 \left( \left\lfloor \frac{1}{2} \left\lfloor \frac{n}{5} \right\rfloor \right\rfloor - 2 \right) \\
\geq \frac{3n}{10} - 6
\]

This implies there are at most \( \frac{7n}{10} + 6 \) numbers larger than \( n \).
**SELECT** \((i, A[1, \ldots, n])\)

**PICK FIRST ELEMENT**

**PIVOT = PARTITION** \((A[1, \ldots, n])\)

**if PIVOT IS POSITION** \(i\) **, RETURN PIVOT**

**else if PIVOT IS IN POSITION** \(> i\) **SELECT** \((i, A[1, \ldots, p - 1])\)

**else SELECT** \(((i - p - 1), A[p + 1, \ldots, n])\)

\[
S(n) = S(\left\lceil n/5 \right\rceil) + O(n) + S(7n/10 + 6)
\]

\[\Theta(n)\]
**RANDOMIZEDSELECT**

(i, A[1, . . . , n])

**PICK RANDOM PARTITION ELEMENT**

**PARTITION LIST ABOUT THIS ONE**

**IF PIVOT IS POSITION i, RETURN PIVOT**

**ELSE IF PIVOT IS IN POSITION > i **

**SELECT (i, A[1, . . . , p − 1])**

**ELSE**

**SELECT ((i − p − 1), A[p + 1, . . . , n])**
RUNNING TIME ANALYSIS

RECURSIVE CALLS
PHASES
PHASES

ALGORITHM IS IN PHASE J IF

SIZE OF INPUT LIST IS <

\[ \left( \frac{3}{4} \right)^j \cdot n \]
RandomizedSelect $(i, A[1, \ldots, n])$

PICK RANDOM PARTITION ELEMENT
PARTITION LIST ABOUT THIS ONE
....
$X_j \equiv \text{NUMBER OF STEPS IN PHASE } J$

$E[X_j] =$
$X_j \equiv \text{NUMBER OF STEPS IN PHASE J}$

$E[X_j] = \sum_{j=0}^{\infty} j \cdot \Pr[X_j = j]$

$\Pr[X_j = 1] = $

$\Pr[X_j = 2] = $

$\Pr[X_j = j] = $
LINEARITY OF EXPECTATION

∀X, Y, \quad E[X + Y] = E[X] + E[Y]
EXPECTED RUNNING TIME

\[ \mathbb{E}[X] = \]