Tutorial II
Outline

- Sampling using fewer random bits
- Solution of Assignment 1
- Hint of Assignment 2
Sampling using fewer random bits

- Let $L$ be a language and $A$ be a randomized algorithm for deciding whether an input string $x$ belongs to $L$ or not.
- Given any $x$, suppose that $A$ can pick a random number $r$ from the range $Z_n = \{ 0, 1, \ldots, n-1 \}$ where $n$ is a prime, with the following property:
  - If $x \in L$, $A(x,r) = 1$ for at least half the possible choices of $r$.
  - If $x \notin L$, $A(x,r) = 0$ for all possible choices of $r$. 
Sampling using fewer random bits (cont)

• We want to increase the correct probability
  ➔ repeat the algorithm multiple times

• Pick \( t > 1 \) values, \( r_1, r_2, \ldots, r_t \in \mathbb{Z}_n \)

• Compute \( A(x, r_j) \) for \( j = 1, \ldots, t \)

• If for any \( j, A(x, r_j) = 1 \), we declare \( x \in L \)

• The error probability of this algo is at most \( 2^{-t} \)

• Uses \( t \log n \) random bits
Sampling using fewer random bits (cont)

• In fact, we can use fewer random bits, and still increase the probability

• Choose \( a, b \) randomly from \( \mathbb{Z}_n \)
• Let \( r_j = aj + b \mod n, j = 1, \ldots, t \)
• Compute \( A(x, r_j) \) for \( j = 1, \ldots, t \)
• If for any \( j, A(x, r_j) = 1 \), we declare \( x \in L \)
• Uses \( 2 \log n \) random bits

What is the error probability?
Sampling using fewer random bits (cont)

• Claim: \( r_j \)'s are pairwise independent (why?)

• Proof: For any \( j \) and \( k \),

1. \( \Pr(r_j = c) = \Pr(a_j + b = c) = \frac{n}{n^2} = \frac{1}{n} \)
   when \( j \) is fixed, there are exactly \( n \) choices of \((a,b)\)
such that \( r_j = c \)

   Similarly, \( \Pr(r_k = d) = \frac{1}{n} \)

2. \( \Pr((r_j = c) \cap (r_k = d)) = \frac{1}{n^2} \)
   when \( j \) and \( k \) are fixed, there is exactly 1 choice of \((a,b)\)
such that \( r_j = c \) and \( r_k = d \)

So, \( \Pr((r_j = c) \cap (r_k = d)) = \Pr(r_j = c) \Pr(r_k = d) \)

\( \Rightarrow \) \( r_j \)'s are pairwise independent
Sampling using fewer random bits (cont)

• Let $Y = \sum_{j=1}^{t} A(x, r_j)$
• Let $Z$ be the value of $Y$ when given $x \in L$
• $E[Z] \geq t/2$, $\text{Var}[Z] = t/4$, $\sigma[Z] = \sqrt{t/2}$
• $\text{Pr}(\text{error}) = \text{Pr}(Z=0)$
\[ \leq \text{Pr}( | Z - E[Z] | \geq t/2 ) \]
\[ = \text{Pr}( | Z - E[Z] | \geq \sqrt{t} (\sigma[Z]) ) \]
\[ \leq 1/t \]

where the last inequality follows from Chebyshev
Assignment 1 – problem 1

• The proof of principle of inclusion-exclusion

\[
Pr \left( \bigcup_{i=1}^{n} E_i \right) = \sum_{i=1}^{n} Pr(E_i) - \sum_{i<j} Pr(E_i \cap E_j) + \sum_{i<j<k} Pr(E_i \cap E_j \cap E_k) \\
- \cdots + (-1)^{\ell+1} \sum_{i_1 < i_2 < \cdots < i_{\ell}} Pr \left( \bigcap_{r=1}^{\ell} E_{i_r} \right) \\
+ \cdots + (-1)^{n} Pr \left( \bigcap_{i=1}^{n} E_i \right).
\]

• Hint: by induction
Assignment 1 – problem 1 (cont)

Another simple proof:
• \( x \) in \( \bigcup_{i=1}^{n} E_i \), \( x \) is exactly \( k \) of sets \( E_i \), the number of times \( x \) contributes \( \Pr(x) \) to the RHS is equal to:

\[
\Pr \left( \bigcup_{i=1}^{n} E_i \right) = \sum_{i=1}^{n} \Pr(E_i) - \sum_{i<j} \Pr(E_i \cap E_j) + \sum_{i<j<k} \Pr(E_i \cap E_j \cap E_k) - \cdots + (-1)^{\ell+1} \sum_{i_1<i_2<\cdots<i_\ell} \Pr \left( \bigcap_{r=1}^{\ell} E_{i_r} \right) + \cdots + (-1)^n \Pr \left( \bigcap_{i=1}^{n} E_i \right) .
\]
Assignment 1 – problem 1 (cont)

C(k,1) - C(k,2) + C(k,3) - ... - (-1)^kC(k,k) = ?

Firstly, 0 = (1-1)^k
Also, (1-1)^k = 1 - C(k,1) + C(k,2) + ... + (-1)^kC(k,k)
So, C(k,1) - C(k,2) + C(k,3) - ... - (-1)^kC(k,k) = 1

⇒ x contributes Pr(x) exactly once on both sides of the equation
Assignment 1 – problem 2

• the values of $F$ are stored in a lookup table, $1/5$ of the lookup table entries are changed

• $F((x + y) \mod n) = (F(x) + F(y)) \mod m$

• Give input $z$, $F(z)$?

• Hint: If $F(z)$ is changed, you never get correct answer. You can use the above formula.
(a)

- Randomly choose a number $x$, and get $y$ such that $z = ((x+y) \mod n)$
- Return $(F(x)+F(y)) \mod m$ as $F(z)$
- The probability that $F(z)$ is correct is at least

$$1 - \Pr((F(x) \text{ is changed}) \cup (F(y) \text{ is changed})) \geq 1 - \Pr(F(x) \text{ is changed}) - \Pr(F(y) \text{ is changed})$$

$$= 1 - \frac{1}{5} - \frac{1}{5} = \frac{3}{5}$$
Assignment 1 – problem 2 (cont)

(b)
• Repeat three times, and choose the repeated values. If all the values are different, pick one randomly
• $\Pr(\text{three times are the same and correct}) \geq (3/5)^3 = 27/125$
• $\Pr(\text{exactly two times are the same and correct}) \geq 3 \times (3/5)^2(2/5) = 54/125 \ldots \text{(why?)}$
• $\Pr(F(z) \text{ is correct}) \geq 81/125 \geq 3/5$
• Describe a randomized algorithm for finding an r-cut with minimum number of edges.

• Hint: r-cut is a general case of 2-cut.
Assignment 1 – problem 3 (cont)

- 2-cut: reduce the number of vertexes until the graph consists of 2 remaining vertices.
- r-cut: reduce the number of vertexes until the graph consists of r remaining vertices.
- 2-cut: contracted n-2 times
- r-cut: contracted n-r times
Assignment 1 – problem 3 (cont)

• Pr(the algorithm is correct) is a bit tricky to analyze.

• Please see the solution of HW 1 when it is posted.
Assignment 1 – problem 4

• The expected number of fixed points \((\pi(x)=x)\) in permutation \(\pi\)
  
  \(- X_i = 1 \text{ if } \pi(i)=i\)
  
  \(- X_i = 0 \text{ otherwise}\)
  
  \(- E[X_i] = 1 \times \Pr(\pi(i)=i) = 1 \times ((n-1)!/n!) = 1/n\)
  
  \(- \text{The expected number of fixed points in } \pi\)
  
  \[= E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = n \times (1/n) = 1\]
Assignment 1 – problem 5

• Interview problem:
  – First interview m candidates but reject them all
  – From the (m+1)th candidate, hire the first candidate who is better than all of the previous candidates you have interviewed

• Hint: $E_i$ be the event that the $i^{th}$ candidate is the best and we hire him
Assignment 1 – problem 5 (cont)

• Let $E_i$ be the event that the $i$th candidate is the best and we hire him
  – $A_i$ : the event that the $i$th candidate is the best
    $\Rightarrow$ $\Pr(A_i) = \frac{1}{n}$
  – $B_i$ : the event that we hire him
    $\Rightarrow$ $\Pr(B_i) = 0$ (if $i \leq m$)
    $\Pr(B_i) = \frac{m}{i-1}$ (otherwise)
  – the best of the first $i-1$ people is between 1 to $m$
  – $A_i$ and $B_i$ are independent $\Rightarrow$ $\Pr(E_i) = \Pr(A_i) \times \Pr(B_i)$
  – $E_i$ are disjoint $\Rightarrow$ $\Pr(E) = \sum_{i=1}^{n} \Pr(E_i) = \frac{m}{n} \sum_{j=m+1}^{n} \frac{1}{j-1}$. 
Assignment 1 – problem 6

• Prove that $E[X^k] \geq E[X]^k$ for any positive even integer $k$.

Solution 1: (Directly from Jensen’s Inequality)

Fact: If $f$ is a convex function, $E[f(x)] \geq f(E[x])$

Note: $f(x) = X^k$ is convex, since for pos. even $k$

$f''(x) = k(k-1)X^{k-2} \geq 0$
Assignment 1 – problem 6 (cont)

Solution 2: (By induction)

- Base case: true for $k = 2$
- Inductive case:

Claim: $E[(X^k - E[X]^k)(X^2 - E[X]^2)] \geq 0$ ...(why?)

Also, $E[(X^k - E[X]^k)(X^2 - E[X]^2)]$

$= E[ X^{2+k} - X^k E[X]^2 - X^2 E[X]^k + E[X]^{2+k} ]$

$= E[X^{2+k}] - E[X^k E[X]^2] - E[X^2 E[X]^k] + E[E[X]^{2+k}]$

$= E[X^{2+k}] - E[X^k] E[X]^2 - E[X^2] E[X]^k + E[X]^{2+k}$

$\leq E[X^{2+k}] - E[X]^k E[X]^2 - E[X]^2 E[X]^k + E[X]^{2+k}$

$= E[X^{2+k}] - E[X]^{2+k} \Rightarrow$ proof completes
Assignment 2 – problem 1

• The variance in the number of fixed points \((\pi(x)=x)\) in permutation \(\pi\).

• Hint: You cannot use linearity to find the variance, but you can calculate it directly.
Assignment 2 – problem2

• Generalize the median-finding algorithm to find the $k^{th}$ largest item in a set of $n$ items. Prove that your resulting algorithm is correct, and bound its running time.

• Hint: 1. How to get the $d$ and $u$ in $R$

2. You must be careful, when you bound the probability that the algorithm outputs FAIL.
Assignment 2 – problem 3

• Proof the weak law of large numbers

\[ \lim_{n \to \infty} \Pr \left( \left| \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right| > \varepsilon \right) = 0. \]

• Hint: Chebyshev's inequality
Assignment 2 – problem 4

• Consider a collection $X_1, \ldots, X_n$ of $n$ independent integers chosen uniformly at random from the set \{0, 1, 2\}. Let $X = \sum_{i=1}^{n} X_i$ and $0 < \delta < 1$.

• Derive a Chernoff bound for $\Pr(X \geq (1 + \delta)n)$ and $\Pr(X \leq (1 - \delta)n)$.

• Hint: define new random variable $Y$ which is related to $X$ (the idea is the same as Corollary 4.9)
Assignment 2 – problem 5

• Weighted sum of Poisson trials. Let $a_1, a_2, \ldots, a_n$ be real numbers in $[0,1]$. $W = \sum_{i=1}^{n} a_i X_i$, show the bound

$$\Pr(W \geq (1 + \delta)\nu) < \left( \frac{e^\delta}{(1 + \delta)(1+\delta)} \right)^\nu.$$ 

• Hint: to prove $e^{\tau a_i} - 1 \leq a_i (e^\tau - 1)$