Recap :

Analyzing randomized processes :

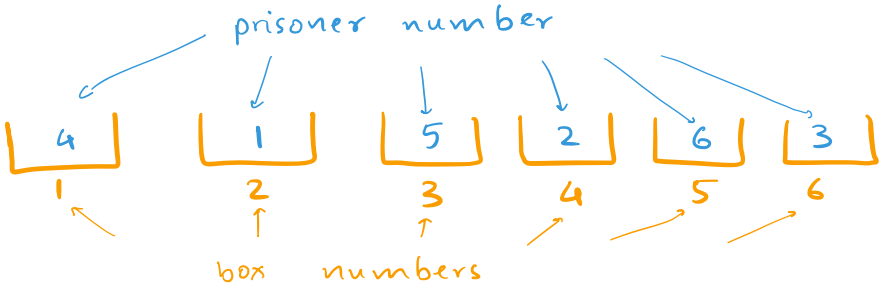
2n prisoners puzzle

There are $2n$ prisoners in a jail, numbered 1 to $2n$.
Jailer has $2n$ boxes, numbered $1, 2, \ldots, 2n$. Jailer picks a unif. rand. perm. π, and for each $i \in \{1, 2, \ldots, 2n\}$, puts name of $\pi(i)^{th}$ prisoner in box i.

Example : Suppose $n = 3$,

$\pi(1) = 4$ $\pi(2) = 1$ $\pi(3) = 5$ $\pi(4) = 2$

$\pi(5) = 6$ $\pi(6) = 3$



Prisoners can decide some strategy. After the strategy is finalized, the prisoners are not allowed to communicate.
They must go in, one by one, and are allowed to open at most n boxes.

All $2n$ prisoners are released if EVERYONE finds their name in one of the n boxes that they opened.

Naive Strategy: every prisoner opens n uniformly random boxes.

$$\Pr[\text{all prisoners are released}] = \frac{1}{2^{2n}}.$$

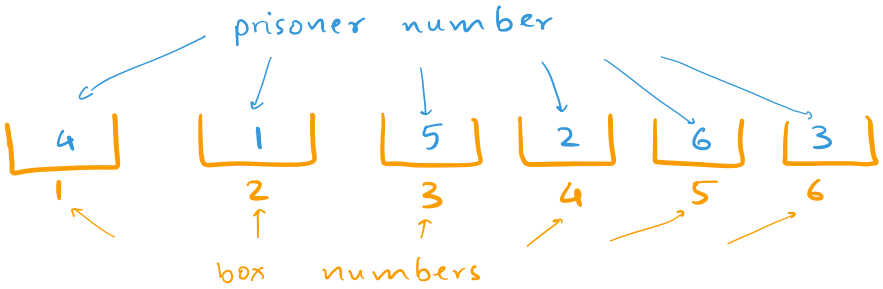
Thm: There exists a deterministic strategy s.t. all prisoners released w.p. ≥ 0.3.

Strategy for i^{th} prisoner:

Let $p_0 = i$.

For $j = 1$ to n, do:

1. open box numbered p_{j-1}. Let z denote the number in box p_{j-1}.
2. If $z = i$, then exit. Else set $p_j = z$.



Example :

- Prisoner 1 opens boxes numbered 1, 4 and 2

- Prisoner 2 opens boxes numbered 2, 1 and 4.
- Prisoner 3 opens boxes numbered 3, 5 and 6
- Prisoner 4 opens boxes numbered 4, 2 and 1.
- Prisoner 5 opens boxes numbered 5, 6 and 3.
- Prisoner 6 opens boxes numbered 6, 3, and 5.

All of them find their numbers in at most n box openings.

This puzzle is essentially a problem on random permutations. Random functions / permutations are often used in the design and analysis of algorithms and data structures. The prisoners puzzle itself appeared in a paper by Anna Gál and Peter Miltersen in 2003, on the complexity of succinct data structures.

Def^n [Cycle Decomposition of a permutation]

Given permutation π, we can partition $\{1, 2, \ldots, n\}$ into disjoint subsets $P_1 \ldots P_t$ s.t. for every $i \in [t]$, for every $x \in P_i$,

$$\{x, \pi(x), \pi(\pi(x)), \ldots\} = P_i .$$

Examples:

$\pi(1) = 4$, $\pi(2) = 6$ $\pi(3) = 2$ $\pi(4) = 1$

$\pi(5) = 5$ $\pi(6) = 7$ $\pi(7) = 3$.

$P_1 = \{4, 1\}$, $P_2 = \{6, 7, 3, 2\}$, $P_3 = \{5\}$.

The partitioning is unique, and can be achieved as follows: start with 1, compute $\pi(1)$, $\pi(\pi(1))$, ... until the elements start repeating. These form the first partition. Remove these from $\{1, 2, \ldots, n\}$.

Then take one of the remaining numbers, say z, compute $\pi(z)$, $\pi(\pi(z))$, ..., and these form the second partition.

The success probability of the prisoners relies on the following two claims.

Claim 1: Suppose the jailer picks permutation π. π has a unique cycle decomposition. If all partitions in this decomposition have size $\leq n$, then all prisoners find their number within n steps (and therefore all are released).

Claim 2: For a unif. rand. permutation π,

$$\Pr\left[\begin{array}{l}\text{the cycle decomp. of } \pi \text{ has} \\ \text{a partition of size } > n\end{array}\right] = \frac{1}{n+1} + \cdots + \frac{1}{2n}$$

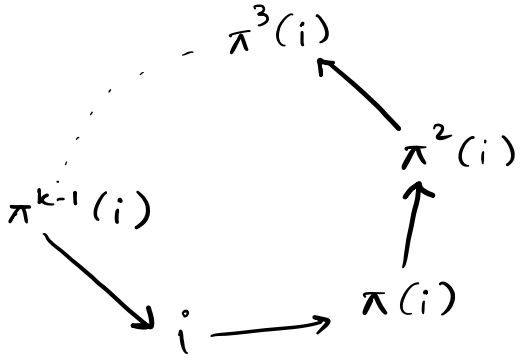
$$\sim \ln 2$$

Therefore, all prisoners are released w.p. $\sim 1 - \ln 2$.

Proof of Claim 1:

Consider any permutation π of $\{1, 2, \ldots, 2n\}$ s.t. all cycles of π have length at most n.

Consider any $i \in \{1, 2, \ldots, 2n\}$. The number i is in some cycle of length $k \leq n$.



$\pi(i)$	$\pi^2(i)$	$\pi^3(i)$	. . .	i
i	$\pi(i)$	$\pi^2(i)$		$\pi^{k-1}(i)$

i^{th} prisoner finds his/her name in the k^{th} box.

∎

Claim 2: For a unif. rand. permutation π,

$$\Pr\left[\begin{array}{l}\text{the cycle decomp. of } \pi \text{ has} \\ \text{a partition of size } > n\end{array}\right] = \frac{1}{n+1} + \cdots + \frac{1}{2n}$$

$$\sim \ln 2$$

Proof: For any $\ell > n$, let

$$\mathcal{E}_\ell = \text{the cycle decomp. of } \pi \text{ has a partition of size } \ell$$

Note that there can be at most one partition of size greater than n. Therefore for $\ell \neq \ell'$, $\ell > n$, $\ell' > n$,

$$\mathcal{E}_\ell \cap \mathcal{E}_{\ell'} = \phi.$$

$$\Pr\left[\begin{array}{l}\text{the cycle decomp. of } \pi \text{ has} \\ \text{a partition of size } > n\end{array}\right]$$

$$= \sum_{\ell = n+1}^{2n} \Pr[\mathcal{E}_\ell]$$

Number of permutations with a cycle of length $2n$:

Equivalent problem: number of ways of arranging $2n$ people at a round table

$$(2n-1)!$$

Number of permutations on $\{1, 2, \ldots, 2n\}$ with a cycle of length $\ell > n$.

- There can be only one cycle of length greater than n.
- $\binom{2n}{\ell}$ ways of picking the ℓ numbers

  $(\ell-1)!$ ways of arranging them in a cycle.

  $(2n-\ell)!$ ways of permuting the remaining

$\therefore$ Number of permutations with ℓ-cycle, $\ell > n$ is $\binom{2n}{\ell} (\ell-1)! (2n-\ell)!$

$$= \frac{(2n)!}{\ell}$$

$$\Pr\left[\text{unif. rand. perm. has cycle of length } > n\right]$$

$$= \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}$$

■

More Properties of Uniformly Random Permutations:

a) σ : unif. rand. perm. over $\{1, 2, \ldots, n\}$.
What is the probability that, for all i, $\sigma(i) \neq i$?

Easier to compute the complement:

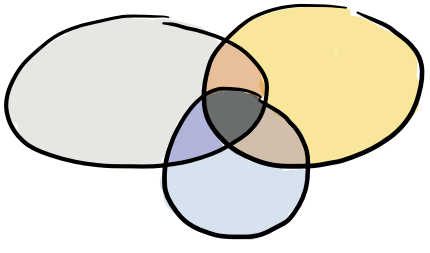
$$\Pr\left[\exists\, i \in \{1, 2, \ldots, n\} \text{ s.t. } \sigma(i) = i\right]$$

S_n : set of all permutations over $\{1, 2, \ldots, n\}$

$A_i \subseteq S_n$: set of all perm. s.t. $\sigma(i) = i$

Want: $|A_1 \cup A_2 \cup \ldots \cup A_n|$

Idea: use Principle of Inclusion-Exclusion



$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cup B \cup C|.$$

$$A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}$$

$$= \left\{ \sigma \quad \text{s.t.} \quad \sigma(i_1) = i_1, \ \sigma(i_2) = i_2, \dots, \ \sigma(i_k) = i_k \right\}$$

$$= (n-k)!$$

k of the values are fixed, the rest are arbitrary (subject to σ being perm.)

Using Principle of Inc.-Exc.,

$$\left| A_1 \cup A_2 \cup \dots \cup A_n \right|$$

$$= \sum_{i=1}^{n} |A_i| - \sum_{i<j} |A_i \cap A_j|$$

$$+ \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \dots$$

$$+ (-1)^{n-1} |A_1 \cap A_2 \cap \dots \cap A_n|$$

$$= n \cdot (n-1)! - \frac{n(n-1)}{2} \cdot (n-2)!$$

$$+ \frac{n(n-1)(n-2)}{3!} - \frac{n(n-1)(n-2)(n-3)}{4!} (n-4)!$$

$$+ \quad \cdots$$

$$= \quad n! \left[1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + \frac{(-1)^{n-1}}{n!} \right]$$

$$\approx \quad n! \left(1 - \frac{1}{e} \right)$$

$$\Pr\left[\exists\, i \in \{1, 2, \ldots, n\} \text{ s.t. } \sigma(i) = i \right] \approx 1 - \frac{1}{e}.$$

$$\Pr\left[\forall\, i, \ \sigma(i) \neq i \right] \approx \quad 1/e.$$

∎

b) σ : unif. rand. perm. over $\{1, 2, \ldots, n\}$.
What is the expected number of elements i s.t. $\sigma(i) = i$?

It is not too difficult to compute the probability that there are exactly k elements i s.t. $\sigma(i) = i$. However, computing the expected value this way is a bit complicated.

Let $X_i = \begin{cases} 1 & \text{if } \sigma(i) = i \\ 0 & \text{otherwise} \end{cases}$

$$X = X_1 + X_2 + \cdots + X_n$$

$$E[X] = E[X_1] + \cdots + E[X_n]$$

$$E[X_i] = Pr\left[\sigma(i) = i\right] = 1/n.$$

$$\therefore E[X] = n \cdot \left(\frac{1}{n}\right) = 1.$$

(c) σ : unif. rand. perm. over $\{1, 2, \ldots, n\}$.
What is the expected length of cycle containing 1 ?

Claim: For every $k \in \{1, 2, \ldots, n\}$,

$$\Pr\left[\text{cycle containing } 1 \text{ has length } k\right] = \frac{1}{n}$$

Proof: Count the number of permutations σ s.t. $\sigma^k(1) = 1$, but $\sigma^j(1) \neq 1$ for all $j < k$.

$\sigma(1)$ can take $n-1$ values (everything except 1)
$\sigma(\sigma(1))$ can take $n-2$ values (everything except 1, $\sigma(1)$)
$\vdots$
$\sigma^{k-1}(1)$ can take $n-k+1$ values
$\sigma^k(1) = 1$

Once these k values are fixed, the remaining $n-k$ can be permuted among themselves in $(n-k)!$ ways.

Hence, the number of permutations where 1 is in cycle of length k is $(n-1)\cdots(n-k+1)(n-k)! = (n-1)!$

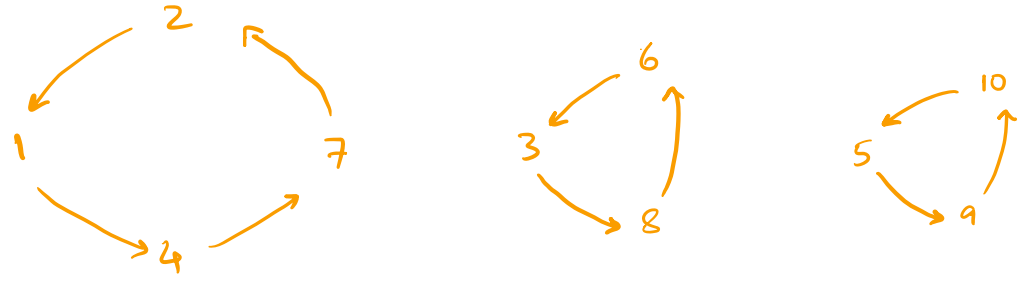
$\therefore \Pr[\text{cycle containing } 1 \text{ has length } k] = 1/n.$ ∎

(d) σ : unif. rand. perm. over $\{1, 2, \ldots, n\}$.
What is the expected number of cycles?

X : number of cycles

Can we decompose X as sum of simpler random variables?

Example:



Number of cycles = 3

Can we decompose $X = X_1 + X_2 + \ldots + X_n$?

$X_i = \frac{1}{k}$ if i is in a k-cycle.

$$X = X_1 + X_2 + \ldots + X_n.$$

$$E[X] = E[X_1] + \ldots + E[X_n]$$

For any $i \in \{1, 2, \ldots, n\}$, $k \in \{1, 2, \ldots, n\}$,

$$\Pr[\, i \text{ is in cycle of length } k \,] = 1/n.$$

$$E[X_i] = \frac{1}{n}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)$$

$$E[X] = \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) \sim \ln n.$$

∎

The following elegant alternate solution was proposed by Prakhar Gupta and Vijay.

Let E_n denote the expected number of cycles in a random perm. over n elements.

The number 1 can be in a cycle of length 1, 2, ... or n with equal prob.

This gives the following recursion :

$$E_n = \frac{1}{n}\left(1 + E_{n-1}\right) + \frac{1}{n}\left(1 + E_{n-2}\right) + \ldots + \frac{1}{n}(1)$$

the formal proof of this uses conditional expectation.

$$= 1 + \frac{1}{n}\sum_{i=1}^{n-1} E_i$$

$$n E_n = n + \sum_{i=1}^{n-1} E_i \quad \ldots (i)$$

$$(n+1) E_{n+1} = (n+1) + \sum_{i=1}^{n} E_i \quad \ldots (ii)$$

$(ii) - (i) \Rightarrow \quad E_{n+1} = \frac{1}{n+1} + E_n$

END OF RANDOM PERMUTATIONS

Analyzing Simple Randomized Processes

1. Tossing a p-biased coin. $\Pr[H] = p$.
   Number of coin tosses needed to get the first H.

$$\Omega = \{ H, TH, TTH, TTTH, \ldots \}$$

countably infinite.

X: Number of coin tosses needed to get the first H.

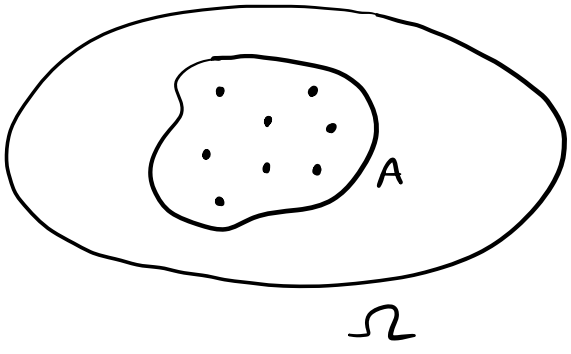
$$Pr[X = i] = (1-p)^{i-1} p$$

$$E[X] = \sum_{i=1}^{\infty} i (1-p)^{i-1} p$$

⋮

Summation isn't very complicated. However, there's a cleaner analysis using conditional expectation.

Conditional Expectation



$p : \Omega \to \mathbb{R}^{\geq 0}$, $\sum_{w \in \Omega} p(w) = 1$.

$X : \Omega \to \mathbb{R}$ - random var.

$A \subseteq \Omega$ - event

$$E[X \mid A] = \sum_{w \in A} X(w) \cdot \frac{p(w)}{Pr[A]}.$$

Partitioning Ω: Let $\Omega_1, \ldots, \Omega_t$ be partitioning of Ω.

$$E[X] = E[X \mid \Omega_1] \cdot Pr[\Omega_1] + \ldots + E[X \mid \Omega_t] \cdot Pr[\Omega_t]$$

Back to our coin tossing problem:

$$\Omega = \{ H, TH, TTH, TTTH, \dots \}$$

$$= \{H\} \cup \{ TH, TTH, TTTH, \dots \}$$

$$\Omega_1 \qquad\qquad \Omega_2$$

Obs 1: $Pr[\Omega_1] = p$, $Pr[\Omega_2] = 1-p$

Obs 2: $E[X \mid \Omega_1] = 1$

Obs 3: $E[X \mid \Omega_2] = 1 + E[X]$

Putting these together,

$$E[X] = p \cdot E[X \mid \Omega_1] + (1-p)\, E[X \mid \Omega_2]$$

$$= p \cdot 1 + (1-p)(1 + E[X])$$

$$\Rightarrow \quad E[X] = 1/p.$$

2. The Coupon Collector Problem

There are n distinct coupons. Each time you sample, you will get one of these coupons uniformly at random.

r.v. X: number of samples needed to collect all n coupons.

$$E[X] = ?$$

Sample space Ω is countably infinite.
For every $\omega \in \Omega$, we can compute $\Pr[\omega]$.

For every $k \geq n$, we can also compute $\Pr[X = k]$. cumbersome expression.

Is there a decomposition of X into r.v.s that are easier to analyze?

e.g. $n = 5$

sample : $\underbrace{4}_{X_1}\ \underbrace{2}_{X_2}\ \underbrace{4\ 1}_{X_3}\ \underbrace{2\ 3}_{X_4}\ \underbrace{1\ 4\ 2\ 5}_{X_5}$

X_i : number of samples needed for the i^{th} new coupon

$$X = X_1 + X_2 + \cdots + X_n$$

$$E[X] = E[X_1] + \cdots + E[X_n]$$

$E[X_i]$:

Recall the coin tossing experiment.
You toss a biased coin s.t. $Pr[H] = p$.
Y : number of coin tosses for the first H.
$E[Y] = 1/p$.

$$E[X_i] = \frac{n}{i-1}$$

$$\therefore E[X] = n\left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n-1}\right)$$

$$\sim n \cdot \ln n$$

■

The coupon collector problem is often used in the analysis of randomized algorithms. We will (hopefully) see one application in one of the upcoming lectures.

Summary :

1. The prisoners puzzle can be solved using the following property of permutations:

Given a uniformly random permutation σ on $\{1, 2, \ldots, 2n\}$

$$\Pr\left[\text{every cycle in } \sigma \text{ has length} \leq n\right] = 1 - \left(\frac{1}{n+1} + \cdots + \frac{1}{2n}\right)$$

$$\approx 1 - \ln 2$$

Some other properties of uniform dist." over S_n (the set of all permutations on $\{1, 2, \ldots, n\}$)

(a) $\Pr\left[\text{no cycle of length } 1\right] \approx 1 - \frac{1}{e}$

probability that $\forall i,\ \sigma(i) \neq i$

(b) $E\left[\text{number of elements s.t. } \sigma(i) = i\right] = 1$

(c) $E\left[\text{length of cycle containing } 1\right] = \frac{n}{2}$.

nothing special about 1 here
Holds for any fixed $x \in \{1, 2, \ldots, n\}$

(d) $E[\text{number of cycles}] \approx \ln n$.

2. Biased coin, $\Pr[H] = p$. Suppose you keep tossing this coin until you get H.

X: number of tosses to get H.

$E[X] = 1/p$.

3. Coupon Collector Problem: n distinct coupons. In each draw, you receive a unif. rand. coupon.

X: number of draws to collect all n coupons.

$$E[X] = n\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-1}\right)$$

$$\sim n \ln n.$$