# Randomized Algorithms

## The Secretary Problem

Joseph Chuang-Chieh Lin

Dept. CSE,
National Taiwan Ocean University

# Review

- Geometry random variable.
- Coupon Collector’s Problem

# Geometric Distribution

- Imagine: *flip a coin until it lands on a head.*
  - What's the distribution of the number of flips?

- Definition. A geometric random variable $X$ with parameter $p$ is

$$\Pr[X = n] = (1-p)^{n-1}p.$$

for $n = 1, 2, \ldots$

# The mean of a geometric r.v. $X(p)$

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

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{j=1}^{\infty} j \Pr[X=j] \\
&= \sum_{j=1}^{\infty}\sum_{i=1}^{j} \Pr[X=j] \\
&= \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} \Pr[X=j] \\
&= \sum_{i=1}^{\infty} \Pr[X \geq i].
\end{aligned}$$

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

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

# The mean of a geometric r.v. $X(p)$

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

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{j=1}^{\infty} j \Pr[X=j] \\
&= \sum_{j=1}^{\infty}\sum_{i=1}^{j} \Pr[X=j] \\
&= \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} \Pr[X=j] \\
&= \sum_{i=1}^{\infty} \Pr[X \ge i].
\end{aligned}$$

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

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

$$\begin{aligned}
&\Pr[X=1] \\
&\Pr[X=2] + \Pr[X=2] \\
&\Pr[X=3] + \Pr[X=3] + \Pr[X=3] \\
&\Pr[X=4] + \Pr[X=4] + \Pr[X=4] + \Pr[X=4] \\
&\Pr[X=5] + \Pr[X=5] + \Pr[X=5] + \Pr[X=5] + \Pr[X=5] \\
&\vdots
\end{aligned}$$

# Other methods (1)

$$\mathbf{E}[X] = \sum_{i\geq 1} i \cdot (1-p)^{i-1} \cdot p = p(1 + 2 \cdot (1-p) + 3 \cdot (1-p)^2 + \ldots)$$

$$(1-p)\mathbf{E}[X] = \sum_{i\geq 1} i \cdot (1-p)^{i} \cdot p = p((1-p) + 2 \cdot (1-p)^2 + \ldots)$$

$$\therefore p \cdot \mathbf{E}[X] = p(1 + (1-p) + (1-p)^2 + \ldots) = p \cdot \frac{1}{1-(1-p)} = 1$$

# Other methods (2)

$$
\begin{aligned}
\mathbf{E}[X] &= \sum_{k\geq 1} k \cdot (1-p)^{k-1} \cdot p \\
&= p \cdot \sum_{k\geq 1} kq^{k-1} \qquad (\text{let } \ q = (1-p)) \\
&= p \cdot \sum_{k\geq 1} \frac{d}{dq}(q^k) \\
&= p \cdot \frac{d}{dq}\left(\frac{q}{1-q}\right) \\
&= p \cdot \frac{1}{(1-q)^2} \\
&= \frac{1}{p}.
\end{aligned}
$$

# Memoryless

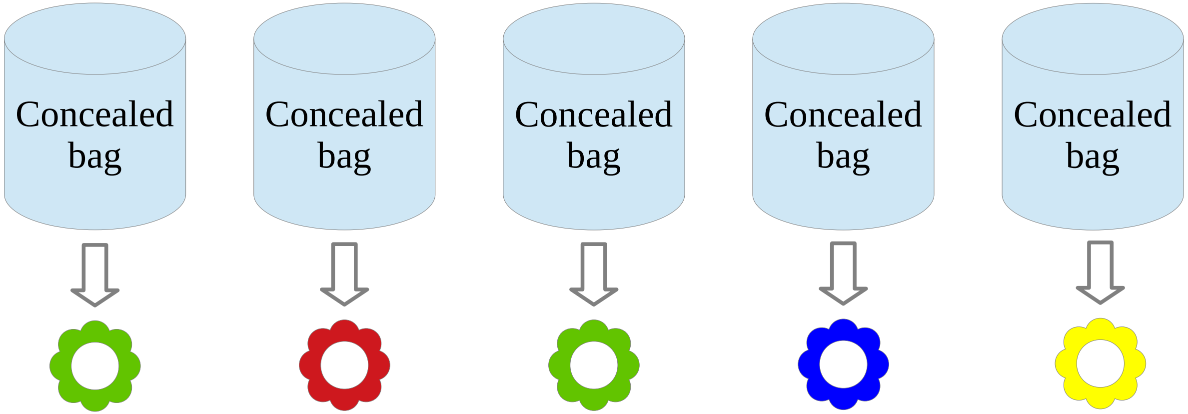
- Let $X$ be a geometric random variable $X$ with parameter $p > 0$.
- For any $n, k > 0$, $\Pr[X = n + k \mid X > k] = \Pr[X = n]$.

# Coupon Collector’s Problem


▲一名網友在臉書社團「爆廢公社公開版」表示，多年前為了收集連鎖便利商店 7-11 的一款贈品，花了不少金額，還拿到許多重複的款式，貼文引發 3 千多名網友共鳴。（圖／翻攝自爆廢公社公開版）

# Coupon Collector's Problem

- Have you already got all of them (totally *n* types)?
- Have you ever thought about how much you should pay for them?




- Each bag is chosen independently and uniformly at random from the *n* possibilities.

# Coupon Collector's Problem

- Let $X$ be the number of bags bought until every type of coupon is obtained.
- Let $X_i$ be the number of bags bought while you had already got exactly $i-1$ different coupons.
  - Geometric random variables.
  - What about $X = \sum_{i=1}^{n} X_i$?

# Coupon Collector's Problem

- When exactly $i-1$ coupons have been collected, the probability of obtaining a new one is

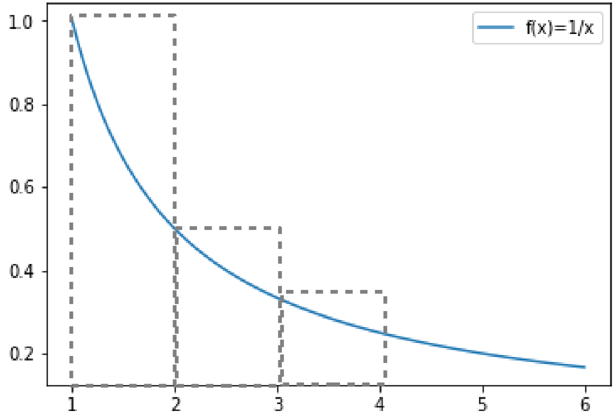
$$p_i = 1 - \frac{i-1}{n}$$

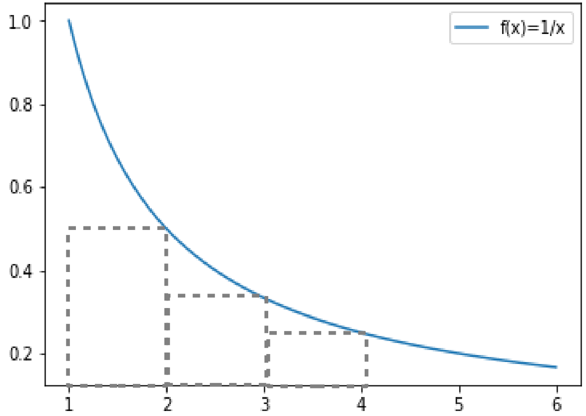
- $X_i$ is a geometric random variable, so

$$\mathbf{E}[X_i] = \frac{1}{p_i} = \frac{n}{n-i+1}.$$

# Coupon Collector's Problem (contd.)

- $$\begin{aligned} \mathbf{E}[X] &= \mathbf{E}\left[\sum_{i=1}^{n} X_i\right] \\ &= \sum_{i=1}^{n} \mathbf{E}[X_i] \\ &= \sum_{i=1}^{n} \frac{n}{n-i+1} \\ &= n \cdot \boxed{\sum_{i=1}^{n} \frac{1}{i}}. \end{aligned}$$




$$\sum_{k=1}^{n} \frac{1}{k} \geq \int_{x=1}^{n} \frac{1}{x} dx = \ln n$$

$$\sum_{k=2}^{n} \frac{1}{k} \leq \int_{x=1}^{n} \frac{1}{x} dx = \ln n$$

$$H(n) = \sum_{i=1}^{n} \frac{1}{i} = \ln n + \Theta(1).$$

# Coupon Collector's Problem (contd.)

- Expected $n \ln n + \Theta(n)$ bags to buy for collecting all the coupons (stickers)!

# The Secretary Problem

- Or, we call it the 'Girlfriend/Boy friend Choosing Problem'.
- Consider the problem of hiring an office secretary.
  - We interview candidates, coming one by one, on a rolling basis.
  - Let's say the *i*th candidate has a value $v_i \in \mathbf{R}$ which stands for how much we like her.
  - At some time point, we would like to hire the best candidate we have seen so far.
  - Suppose we can fire the old one and hire a new better candidate.
  - Assume that we only want to interview at most *n* candidates.

# The Secretary Problem (contd.)

- The whole hiring process will be just like:

```
Randomly shuffle the n candidates.
Set TheOne ← 0
for i ← 1 to n do:
    interview candidate i
    if v_i > v_TheOne then:
        TheOne ← i
        Hire candidate i
```

- Isn't it very simple?

# The cost

- No pain, no gain. We cannot reap without sowing.

- Let $c_{\mathrm{I}}$ be the cost associated with interviewing a candidate.
- Let $c_{\mathrm{H}}$ be the cost associated with hiring a candidate.

- So, if totally we have ever hired $m$ people ($m-1$ was fired though…), what is the total cost of the algorithm?

# The cost

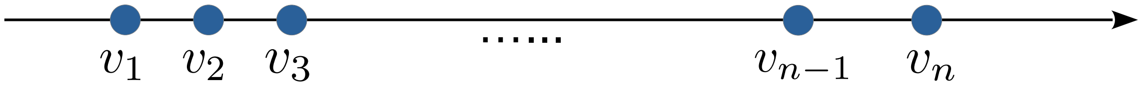
- No pain, no gain. We cannot reap without sowing.

- Let $c_I$ be the cost associated with interviewing a candidate.
- Let $c_H$ be the cost associated with hiring a candidate.

- So, if totally we have ever hired $m$ people ($m-1$ was fired though…), what is the total cost of the algorithm? $O(c_I n + c_H m)$

# The cost

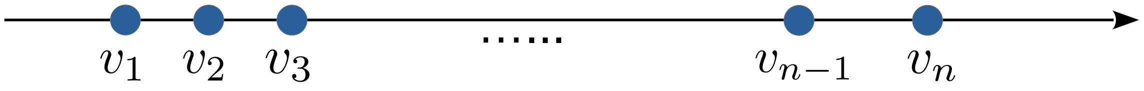
- In the worst scenario,




- The cost in the worst case:

# The cost

- In the worst scenario,




- The cost in the worst case:

$$O((c_I + c_H)n)$$

- What's about the EXPECTED cost?

# The expected cost analysis

- Let $X_i$ be an indicator random variable such that

$$\begin{cases} X_i = 1 & \text{if candidate } i \text{ is hired} \\ X_i = 0 & \text{otherwise} \end{cases}$$

- $X = \sum_{i=1}^{n} X_i$ : the number of times we hire a new candidate.

- $\Pr[X_i] = ?$

# The expected cost analysis

- Let $X_i$ be an indicator random variable such that

$$\begin{cases} X_i = 1 & \text{if candidate } i \text{ is hired} \\ X_i = 0 & \text{otherwise} \end{cases}$$

- $X = \sum_{i=1}^{n} X_i$ : the number of times we hire a new candidate.

- $\Pr[X_i] = \Pr[\text{candidate } i \text{ is better than previous } i-1 \text{ candidates}] = \frac{1}{i}.$

# The expected cost analysis

- Let $X_i$ be an indicator random variable such that

$$\begin{cases} X_i = 1 & \text{if candidate } i \text{ is hired} \\ X_i = 0 & \text{otherwise} \end{cases}$$

- $X = \sum_{i=1}^{n} X_i$ : the number of times we hire a new candidate.

- $\Pr[X_i] = \Pr[\text{candidate } i \text{ is better than previous } i-1 \text{ candidates}] = \frac{1}{i}.$
  - Imagine we have randomly chosen *i* numbers, what's the probability that the *i*th number is the biggest?
  - It's NOT a conditional probability.

# The expected cost analysis

- Therefore,

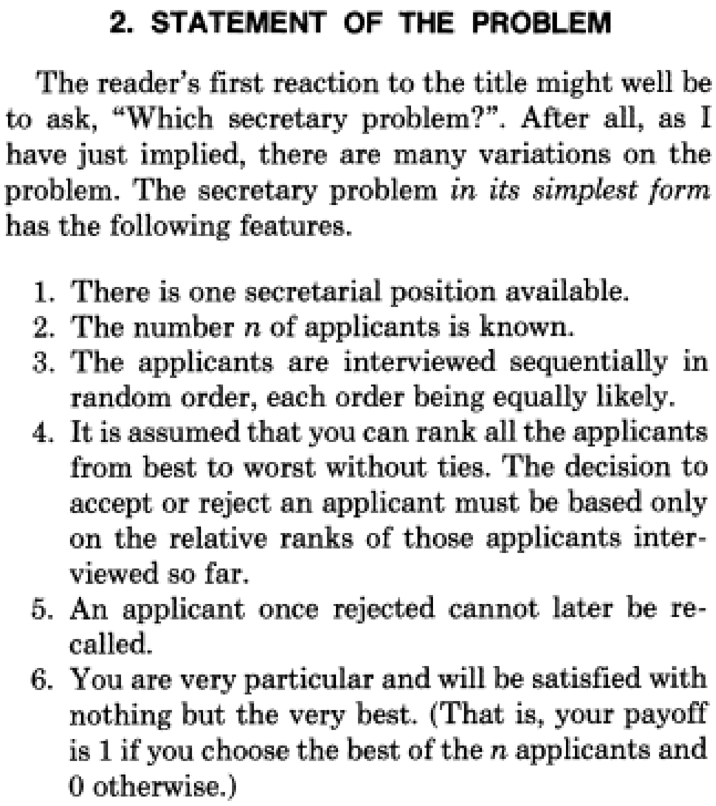
$$\begin{aligned} \mathbf{E}[X] = \sum_{i=1}^{n} \mathbf{E}[X_i] &= \sum_{i=1}^{n} \Pr[X_i = 1] \\ &= \sum_{i=1}^{n} \frac{1}{i} \\ &= H_n \\ &= \ln n + \Theta(1). \end{aligned}$$

- The expected cost is

$$O(c_H \ln n + c_I n).$$

# The classic version

- Reference:
  - Thomas S. Ferguson: Who solved the Secretary Problem? *Statistical Science*, Vol. 4 (1989), pp. 282−289.



**2. STATEMENT OF THE PROBLEM**

The reader's first reaction to the title might well be to ask, "Which secretary problem?". After all, as I have just implied, there are many variations on the problem. The secretary problem *in its simplest form* has the following features.

1. There is one secretarial position available.
2. The number $n$ of applicants is known.
3. The applicants are interviewed sequentially in random order, each order being equally likely.
4. It is assumed that you can rank all the applicants from best to worst without ties. The decision to accept or reject an applicant must be based only on the relative ranks of those applicants interviewed so far.
5. An applicant once rejected cannot later be recalled.
6. You are very particular and will be satisfied with nothing but the very best. (That is, your payoff is 1 if you choose the best of the $n$ applicants and 0 otherwise.)

# A simple solution

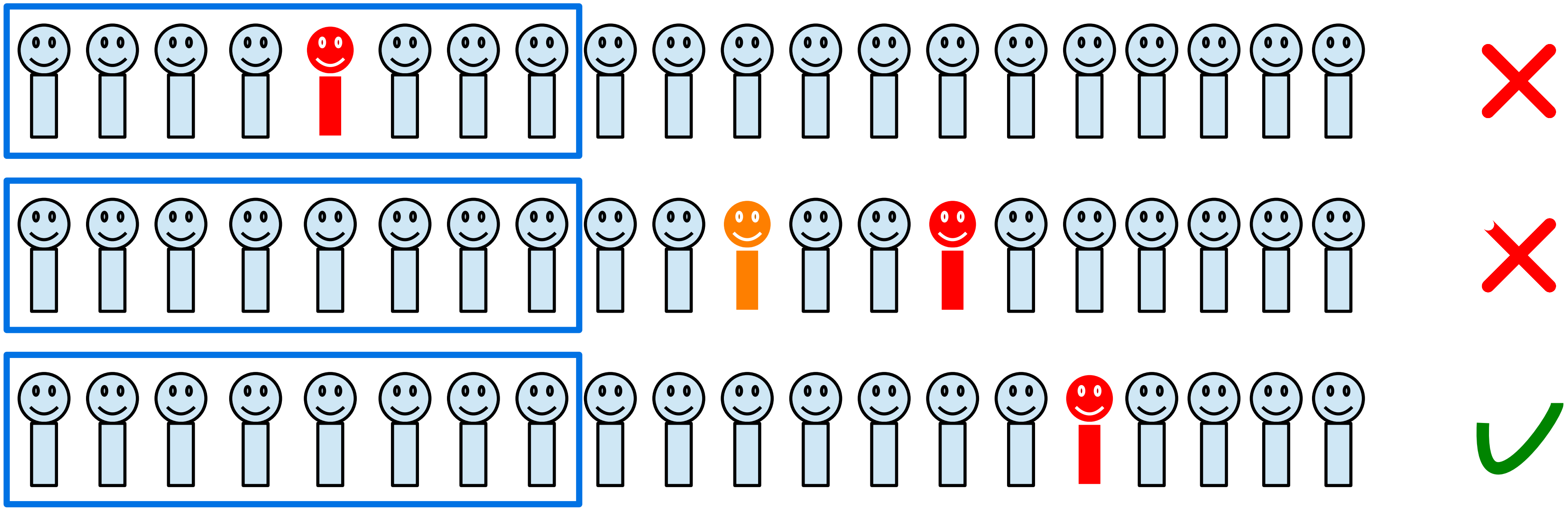
- Reject the first $r-1$ applicants.
- Choose the next applicant who is the best in the relative ranking of the observed applicants.

- The famous 37% rule.

$$\frac{r}{n} \approx 37\% \quad \text{as } n \text{ is large}$$


Refer to
https://www.books.com.tw/products/F014054315

# Illustration



# Analysis

- $\phi_n(r)$: the probability that the best applicant is selected.
- For $r = 1$, $\phi_n(r) = 1/n$.
- For $r > 1$,

$$\begin{aligned}
\phi_n(r) &= \Pr\left[\bigcup_{j=r}^{n}\{\text{applicant } j \text{ is the best and is selected}\}\right] \\
&= \sum_{j=r}^{n}\frac{1}{n}\cdot\frac{r-1}{j-1} = \frac{r-1}{n}\sum_{j=r}^{n}\frac{1}{j-1} \\
&= \frac{r-1}{n}\sum_{j=r}^{n}\left(\frac{n}{j-1}\right)\cdot\left(\frac{1}{n}\right) \\
&\approx x\int_{x}^{1}\left(\frac{1}{t}\right)dt \\
&= -x\ln x.
\end{aligned}$$

We want $x$,

$$x = \lim_{n\to\infty}\frac{r}{n}$$

# Analysis

- $\phi_n(r)$: the probability that the best applicant is selected.
- For $r = 1$, $\phi_n(r) = 1/n$.
- For $r > 1$,

$$\begin{aligned}
\phi_n(r) &= \Pr\left[\bigcup_{j=r}^{n}\{\boxed{\text{applicant } j \text{ is the best}} \text{ and } \boxed{\text{is selected}}\}\right] \\
&= \sum_{j=r}^{n} \boxed{\frac{1}{n}} \cdot \boxed{\frac{r-1}{j-1}} = \frac{r-1}{n}\sum_{j=r}^{n}\frac{1}{j-1} \\
&= \frac{r-1}{n}\sum_{j=r}^{n}\left(\frac{n}{j-1}\right)\cdot\left(\frac{1}{n}\right) \\
&\approx x\int_{x}^{1}\left(\frac{1}{t}\right)dt \\
&= -x\ln x.
\end{aligned}$$

Each of the $j-1$ candidates is the best of them with prob. $1/(j-1)$, but we count only the $r-1$ of them.

We want $x$,

$$x = \lim_{n\to\infty}\frac{r}{n}$$

# Analysis (contd.)

- Find the value of $x$ which maximizes $\phi_n(r)$.
  - Solve $d(\phi_n(r))/dx = 0$.

# Analysis (contd.)

- Find the value of $x$ which maximizes $\phi_n(r)$.
  - Solve $d(\phi_n(r))/dx = 0$.

$$\frac{d}{dx}\phi_n(r) = -x \cdot \frac{1}{x} + (\ln x) \cdot (-1) = -1 - \ln x.$$

# Analysis (contd.)

- Find the value of $x$ which maximizes $\phi_n(r)$.
  - Solve $d(\phi_n(r))/dx = 0$. $$\frac{d}{dx}\phi_n(r) = -x \cdot \frac{1}{x} + (\ln x) \cdot (-1) = -1 - \ln x.$$

- We can find optimal $x = 1/e = 0.367879\ldots \approx 37\%$.

# Analysis (contd.)

- Find the value of $x$ which maximizes $\phi_n(r)$.
  - Solve $d(\phi_n(r))/dx = 0$. $\frac{d}{dx}\phi_n(r) = -x \cdot \frac{1}{x} + (\ln x) \cdot (-1) = -1 - \ln x.$
- We can find optimal $x = 1/e = 0.367879\ldots \approx 37\%$.
- What about your strategy of finding your wife/husband?
  - This lecture assumes that you can never go back to your ex's...