

Randomized Algorithms

Discrete Random Variables and Expectation

Joseph Chuang-Chieh Lin

Dept. CSE,
National Taiwan Ocean University

Review

- Why randomized algorithms?
- Types of randomized algorithms.

Randomized algorithms



Why?

- Randomized algorithms are
 - often much ***simpler*** than the best known deterministic ones.
 - often much ***more efficient*** (faster or using less space) than the best known deterministic ones.

Two types of randomized algorithms

- The **accuracy** is guaranteed.
 - Las Vegas algorithms.
- The **running time** is guaranteed.
 - Monte Carlo algorithms.

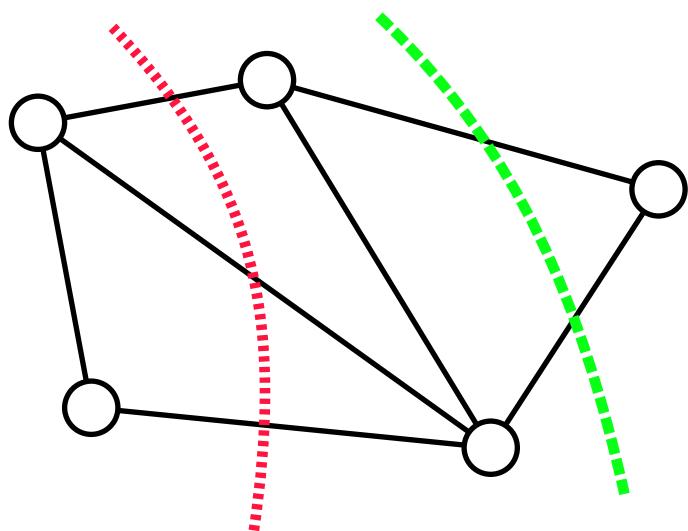


Mike Boening@mbphotography



Pasquale Iovino@pask_07

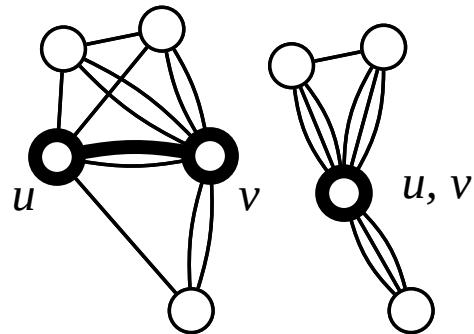
Min-Cut



$$|V| = n, |E| = m$$

- A graph $G = (V, E)$ and its two “cuts”.
 - **Cut:** a partition of the vertices in V into two non-empty, disjoint sets S and T such that
 - $S \cup T = V$
- The **cutset** of a cut:
 - $\{uv \in E \mid u \in S, v \in T\}$.
- The size of the cut:
 - the cardinality of its cutset.

Edge contraction



$$e = (u, v)$$

$$G \rightarrow G/e$$

By Thore Husfeldt - Python, networkx, neato and TikZ, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=21117500>

Karger's edge-contraction algorithm (1993)

Procedure contract ($G = (V, E)$):

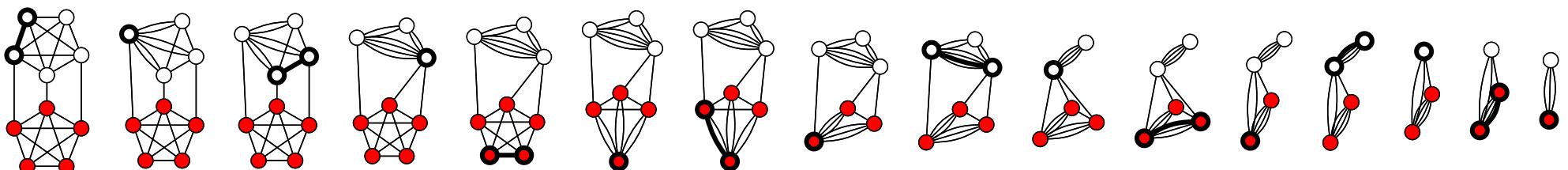
while $|V| > 2$:

 choose $e \in E$ uniformly at random

$G \leftarrow G/e$

return the only cut in G

Time complexity: $O(m)$ or $O(n^2)$.



By Thore Husfeldt - Created in python using the networkx library for graph manipulation, neato for layout, and TikZ for drawing., CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=21103489>

Analysis

- C : a specific cut of G .
- k : the size of the cut C .
- The minimum degree of G must be $\geq k$. (WHY?)
 - So, $|E| \geq nk/2$.
- The probability that the algorithm picks an edge from C to contract is

$$\frac{k}{|E|} \leq \frac{k}{nk/2} = \frac{2}{n}.$$

Analysis (contd.)

- Let p_n be the probability that the algorithm on an n -vertex graph avoids C .
- Then,

$$p_n \geq \left(1 - \frac{2}{n}\right) \cdot p_{n-1}$$

- The recurrence can be expanded as

$$p_n \geq \prod_{i=0}^{n-3} \left(1 - \frac{2}{n-i}\right) = \frac{n-2}{n} \cdot \frac{n-3}{n-1} \cdots \frac{2}{4} \cdot \frac{1}{3} = \frac{1}{\binom{n}{2}}.$$

Analysis (contd.)

- Repeat the contract algorithm for $T = \binom{n}{2} \ln n$ times, and then choose the minimum of them.
- The probability of NOT finding a min-cut is $\left[1 - \left(\frac{n}{2}\right)^{-1}\right]^T \leq \frac{1}{e^{\ln n}} = \frac{1}{n}$.
- Let's take a look at the exponential function e^x .

Facts on e^x

$$e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots$$

$$\therefore e = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

$$\left(1 + \frac{1}{n}\right)^n = \binom{n}{0} 1 + \binom{n}{1} \frac{1}{n} + \binom{n}{2} \frac{1}{n^2} + \binom{n}{3} \frac{1}{n^3} + \dots + \binom{n}{n} \frac{1}{n^n}$$

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$$

Probability basics

A *random variable* X on a sample space Ω is a real-valued function. That is,

$$X : \Omega \mapsto \mathbf{R}$$

A *discrete random variable* is a random variable that takes on only a finite or countably infinite number of values.

Kolmogorov axioms

- 1. For any event $A \subset S$, $P(A) \geq 0$.
- 2. $\Pr[S] = 1$.
- 3. If A_1, A_2, \dots are mutually exclusive events, then
$$\Pr[A_1 \cup A_2 \cup \dots] = \Pr[A_1] + \Pr[A_2] + \dots$$

Useful theorems

- $\Pr[\emptyset] = 0$ for any experiment.
- For any event $A \subseteq S$, $\Pr[A] = 1 - \Pr[\bar{A}]$.
- If $A \subseteq S, B \subseteq S$ are any two events, then
$$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B].$$
- If $A \subset B$, then $\Pr[A] \leq \Pr[B]$.

Conditional probability

- In an experiment with sample space S , let B be any event such that $\Pr[B] > 0$.
- Then the conditional probability of A occurring, given that B has occurred, is

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

for any $A \subset S$.

Example

- Suppose you are going to buy milk in a supermarket.
 - There are a total of 40 boxes for you to choose from.
 - 10 of them are corrupted (not visible on the outside).
- Then, you are asked to buy two boxes of milk.
What is the probability that both boxes are good?

Example (contd.)

- A : the event that the first box you choose is good.
 B : the event that the second box you choose is good.

Then

$$\Pr[A] = \frac{30}{40}$$

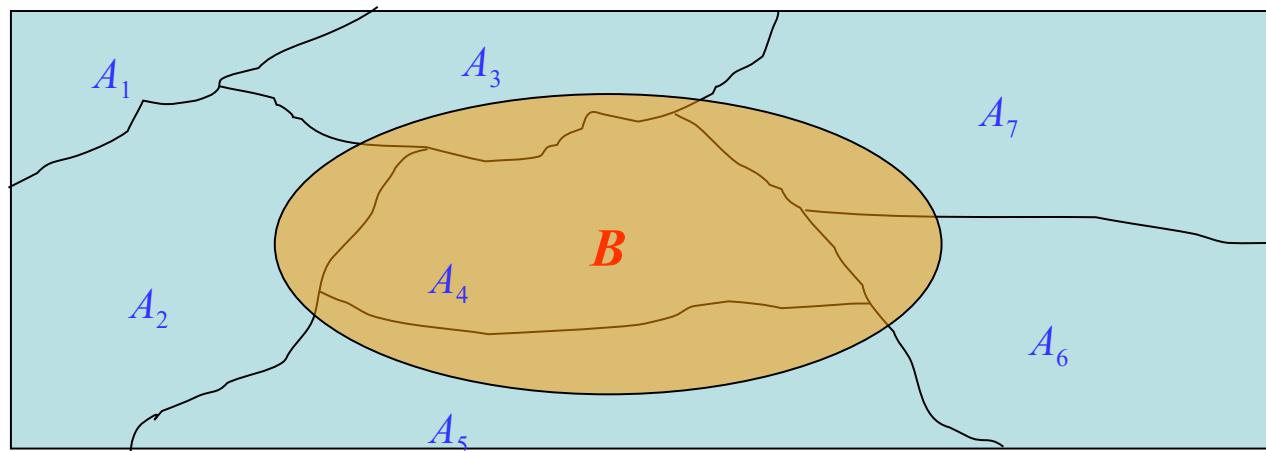
$$\Pr[B \mid A] = \frac{29}{39}$$

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B \mid A] = \frac{30}{40} \cdot \frac{29}{39} = \frac{87}{156}.$$

Theorem of total probability

- If A_1, A_2, \dots, A_n is a partition of S , and B is any event, then

$$\Pr[B] = \sum_{i=1}^n \Pr[B|A_i] \Pr[A_i].$$



Bayes' theorem

- From the theorem of total probability, and granted that A_1, A_2, \dots, A_n is a partition of S , we have

$$\begin{aligned}\Pr[A_i \mid B] &= \frac{\Pr[A_i \cap B]}{\Pr[B]} \\ &= \frac{\Pr[A_i] \cdot \Pr[B \mid A_i]}{\sum_{i=1}^n \Pr[A_i] \cdot \Pr[B \mid A_i]}\end{aligned}$$

- This result is known as *Bayes' theorem*.

Example

- Assuming a jury selected to participate in a criminal trial.
- Whether the defendant is guilty or not, there is a 95% chance of making the correct verdict.
- It is also assumed that the local police law enforcement is very strict, such that 99% of the people being tried are actually guilty.
- **If a jury is known to sentence a defendant not guilty, what is the probability that the defendant is really not guilty?**

Example (contd.)

- A_1 : the defendant is guilty
- $A_2 = \bar{A}_1$: the defendant is not guilty.
- Let B be the event that the defendant is sentenced to unguilty.
- We want to know $\Pr[A_2 | B]$.

Example (contd.)

$$\begin{aligned}\Pr[A_2|B] &= \frac{\Pr[A_2]\Pr[B|A_2]}{\Pr[A_1]\Pr[B|A_1]+\Pr[A_2]\Pr[B|A_2]} \\ &= \frac{(0.01)(0.95)}{(0.99)(0.05)+(0.01)(0.95)} \\ &= 0.161\end{aligned}$$

- Before the sentence, this defendant is supposed to be unguilty with probability 1%.
- After the sentence of unguilty, the probability is increased to be **16.1%**.

Independent events

If $A \subset S$ and $B \subset S$ are any two events with nonzero probabilities,
 A and B are called independent if and only if $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$

That is, $\Pr[A] = \Pr[A | B]$ and $\Pr[B] = \Pr[B | A]$.

Independent trials

- An experiment is said to consist of n ***independent*** trials if and only if
 - $S = T_1 \times T_2 \times \cdots \times T_n$.
 - For every $(x_1, x_2, \dots, x_n) \in S$,
 $\Pr[\{(x_1, x_2, \dots, x_n)\}] = \Pr[\{x_1\}] \cdot \Pr[\{x_2\}] \cdots \Pr[\{x_n\}]$,
where $\Pr[\{x_i\}]$ is the probability of $x_i \in T_i$ occurring on trial i .

Expectation

- The expectation of a discrete random variable X , denoted by $\mathbf{E}[X]$, is

$$\mathbf{E}[X] = \sum_i i \cdot \Pr[X = i]$$

- Example: Let X denote the sum of of dices:

$$\mathbf{E}[X] = \frac{1}{36} \cdot 2 + \frac{2}{36} \cdot 3 + \frac{3}{36} \cdot 4 + \dots + \frac{1}{36} \cdot 12 = 7.$$



Adam Fejes@adamfjs

Linearity of Expectation

- For any finite collection of discrete random variables X_1, X_2, \dots, X_n with finite expectations,

$$\mathbf{E} \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \mathbf{E}[X_i].$$

- For any constant c and discrete random variable X ,

$$\mathbf{E}[cX] = c \cdot \mathbf{E}[X].$$

- Why is it useful?

Example

- Consider the dice-throwing example again.

- X_1 : the outcome of die 1
- X_2 : the outcome of die 2

$$\mathbf{E}[X_1] = \mathbf{E}[X_2] = \frac{1}{6} \cdot \sum_{j=1}^6 j = \frac{7}{2}$$

$$\mathbf{E}[X] = \mathbf{E}[X_1 + X_2] = 7.$$



Adam Fejes@adamfjs

Bernoulli random variable

- Suppose we run an experiment that succeeds with probability p and fails with probability $1-p$.

$$Y = \begin{cases} 1 & \text{if the experiment succeeds,} \\ 0 & \text{otherwise.} \end{cases}$$

- Y : Bernoulli random variable.
 - or *indicator random variable*.



Binomial random variable

- A binomial random variable X with parameters n and p , denoted by $B(n, p)$, is defined as

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

for $j = 0, 1, 2, \dots, n$.

- Exercise: Show that $\sum_{j=0}^n \Pr[X = j] = 1$.

Binomial random variable (contd.)

- $$\begin{aligned}\mathbf{E}[X] &= \sum_{j=0}^n j \binom{n}{j} p^j (1-p)^{n-j} \\ &= \sum_{j=0}^n j \frac{n!}{j!(n-j)!} p^j (1-p)^{n-j} \\ &= \sum_{j=1}^n \frac{n!}{(j-1)!(n-j)!} p^j (1-p)^{n-j} \\ &= np \sum_{j=1}^n \frac{(n-1)!}{(j-1)!((n-1)-(j-1))!} p^{j-1} (1-p)^{(n-1)-(j-1)} \\ &= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!((n-1)-k)!} p^k (1-p)^{(n-1)-k} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} \\ &= np.\end{aligned}$$

Binomial random variable (contd.)

$$\begin{aligned}\bullet \quad \mathbf{E}[X] &= \sum_{j=0}^n j \binom{n}{j} p^j (1-p)^{n-j} \\&= \sum_{j=0}^n j \frac{n!}{j!(n-j)!} p^j (1-p)^{n-j} \\&= \sum_{j=1}^n \frac{n!}{(j-1)!(n-j)!} p^j (1-p)^{n-j} \\&= np \sum_{j=1}^n \frac{(n-1)!}{(j-1)!((n-1)-(j-1))!} p^{j-1} (1-p)^{(n-1)-(j-1)} \\&= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!((n-1)-k)!} p^k (1-p)^{(n-1)-k} \\&= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} \\&= np.\end{aligned}$$

$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}.$



Let's make it simpler!

- Denote a set of n Bernoulli random variables X_1, X_2, \dots, X_n .
 - $X_i = 1$ if the i th trial is successful and 0 otherwise.

$$\mathbf{E}[X] = \mathbf{E} \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \mathbf{E}[X_i] = np.$$