

随机算法 Spring 2025

Lecture 1 - 2025 / 2 / 17

Checking Matrix Multiplication

输入：三个 $n \times n$ 矩阵 A, B, C 。

输出：是否 $AB = C$ 。

随机选定向量 $r = (r_1, r_2, \dots, r_n)$ ，其中 r_i 独立同分布于 $U(S)$ ， $2 \leq |S| < |\mathbb{N}|$

如果 $(AB)r \neq Cr$ 则输出 No，否则输出 Yes。

确定算法 $O(n^3)$ ，或者最优秀的是 $O(n^{2.376})$ 。

算法时间复杂度 $O(n^2)$ 。

Claim: 如果 $AB \neq C$ ，则 $\Pr[(AB)r = Cr] \leq \frac{1}{|S|}$ 。

设 $D = AB - C \neq 0$ ，则不失一般性设 $d_{11} \neq 0$ 。

如果 $Dr = 0$ ，则 $(Dr)_1 = \sum_{i=1}^n d_{1i}r_i = 0$

于是 $r_1 = -\frac{1}{d_{11}}(d_{12}r_2 + \dots + d_{1n}r_n)$

于是对于 r_2, \dots, r_n 的每种选择， r_1 只有唯一的可能性有可能使 $Dr = 0$ ，于是 $\Pr[Dr = 0] \leq \frac{1}{|S|}$ 。

Checking Associativity

输入：在一个大小为 n 的集合 X 上定义二元运算 \circ 。

输出：是否满足结合律 $\forall i, j, k \in X, i \circ (j \circ k) = (i \circ j) \circ k$ 。

确定性算法 $O(n^3)$ 。

不妨规定 $X = \{1, 2, \dots, n\}$ 。

首先可以构造一种 \circ 使得不满足条件的三元组是常数组。

事实上，构造 $1 \circ 2 = 1$ ，其余运算结果全部为 3，则只有 $(1 \circ 2) \circ 2 \neq 1 \circ (2 \circ 2)$ 。

记 $\mathcal{X} = 2^X$ ，对于 $R \in \mathcal{X}$ ，可以用 $R = r_1 r_2 \cdots r_n$ 表示，其中 $r_i \in \mathbb{F}_2$ 表示 i 有没有在 R 中出现。

从而 R 可以写成 $\sum_{i=1}^n r_i \cdot i$ 。

我们在 \mathcal{X} 上定义一种 $+$ 运算，并扩展 \circ 运算

$$R + S = \sum_{i=1}^n (r_i + s_i) \cdot i$$

$$R \circ S = \sum_{i=1}^n \sum_{j=1}^n (r_i s_j) \cdot (i \circ j)$$

我们将算法规定为：

均匀随机选择 $R, S, T \in \mathcal{X}$ ，如果 $(R \circ S) \circ T \neq R \circ (S \circ T)$ 输出 No，否则输出 Yes。

可以看出 \circ 在 X 上是结合的，等价于 \circ 在 \mathcal{X} 上是结合的。

⇒ 可以通过展开得到，⇐ 是因为单元素集 $\in \mathcal{X}$ 。

Claim: 如果 \circ 不结合，那么 $\Pr[(R \circ S) \circ T = R \circ (S \circ T)] \leq \frac{7}{8}$ 。

假设存在 i^*, j^*, k^* 不结合。

任取一组 R_0, S_0, T_0 使得 $i^* \notin R_0, j^* \notin S_0, k^* \notin T_0$ 。

令 $R_1 = R_0 \cup \{i^*\}, S_1 = S_0 \cup \{j^*\}, T_1 = T_0 \cup \{k^*\}$ 。

则设 $f(\alpha, \beta, \gamma) = (\alpha \circ \beta) \circ \gamma + \alpha \circ (\beta \circ \gamma)$ 。

不结合即 $f(\{i^*\}, \{j^*\}, \{k^*\}) \neq \emptyset$ 。

根据容斥原理

$$f(\{i^*\}, \{j^*\}, \{k^*\}) = \sum_{r,s,t \in \{0,1\}} f(R_r, S_s, T_t) \neq \emptyset$$

从而 $\exists r, s, t \in \{0, 1\}$ 使得 $f(R_r, S_s, T_t) \neq \emptyset$ 。

由于这样的 (R_0, S_0, T_0) 以及衍生出的 8 个集合构成了 \mathcal{X}^3 的一个划分，所以一定有 $\frac{1}{8}$ 的 \mathcal{X} 的三元组是不满足结合律的。

Testing Polynomial Identities

给定某个域下 2 个 n 元多项式 P, Q ，判定是否 $P \equiv Q$ 。

作差后问题等价于判定 $P \equiv 0$ 是否成立。

我们在有限集 $|S|$ 上均匀随机采样 r_1, \dots, r_n , 并带入 P 计算。

Claim: 如果 $P \neq 0$, 则 $\Pr[P(r_1, \dots, r_n) = 0] \leq \frac{d}{|S|}$, 其中 $d = \deg P$ 。

对于 n 归纳。 $n = 1$ 时显然至多 d 个根, 结论成立。

设 k 是 P 关于 x_1 的最大度数。

$$P(x_1, \dots, x_n) = M(x_2, \dots, x_n)x_1^k + N(x_1, \dots, x_n)$$

其中 $\deg M \leq d - k$, N 中 x_1 的度数 $< k$ 。

设 \mathcal{E} 表示 $M(r_2, \dots, r_n) = 0$ 。

1. 如果 \mathcal{E} 发生, 则对 M 由归纳, $\Pr[\mathcal{E}] \leq \frac{d - k}{|S|}$ 。
2. 如果 \mathcal{E} 不发生, 则当固定 r_2, \dots, r_n 时, P 是关于 x_1 的 k 次多项式, 从而能使 $P = 0$ 的 x_1 不超过 k 个, 于是 $\Pr[P(r_1, \dots, r_n) = 0 \mid \neg\mathcal{E}] \leq \frac{k}{|S|}$ 。

根据 union bound 立刻得证。

Lecture 2 - 2025 / 2 / 20

Bipartite Matching

给定一个二分图 $G = (V_1, V_2, E)$, 且 $|V_1| = |V_2| = n$, 求 G 是否包含一个完美匹配?

Definition (Tutte matrix): 二分图 G 的 Tutte 矩阵定义为 $n \times n$ 矩阵 $A_G = [a_{ij}]$, 其中如果 $(i, j) \in G$ 那么 $a_{ij} = x_{ij}$ 为一个变量, 否则 $a_{ij} = 0$ 。

Claim: G 包含完美匹配当且仅当 $|A_G| \neq 0$ 。

由行列式定义

$$|A_G| = \sum_{\sigma} \text{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)}$$

G 包含完美匹配, 也就是存在排列 σ 使得 $\forall 1 \leq i \leq n, a_{i\sigma(i)} \neq 0$ 。换言之 $\prod_{i=1}^n a_{i\sigma(i)} \neq 0$ 。

根据 Tutte 矩阵的定义, 每个 σ 对应的乘积包含的变量均不相同, 因此只要有一项非 0, 就有 $|A_G| \neq 0$ 。反之亦然。

小知识: $n \times n$ 矩阵的行列式可以通过并行算法, 在 $O(n^{3.5})$ 个处理器上用 $O(\log^2 n)$ 的时间计算。

那么利用 Lecture 1 判定多项式是否为 0 的方法即可。

Finding a Perfect Matching in Parallel

Lemma (Isolation Lemma): 设 $S_1, S_2, \dots, S_k \subseteq S$, 给 S 中的每个元素均匀随机赋值 $\{1, 2, \dots, l\}$, 则

$$\Pr[\exists \text{unique } S_i \text{ of minimal sum of weights}] \geq \left(\frac{l-1}{l}\right)^{|S|} \geq 1 - \frac{|S|}{l}$$

我们可以不妨设集合是没有包含关系的。

我们记所有赋值方法 $w = \{w_x \mid x \in S\}$ 构成的集合为 \mathcal{W} , 如果 $\forall x, w_x > 1$, 那么这样的赋值方法构成的集合为 \mathcal{W}^+ 。

易知 $|\mathcal{W}| = l^{|S|}$ 。

接下来我们构造一个从 \mathcal{W}^+ 到“最小集合唯一的赋值方式”的单射。

对于 $w \in \mathcal{W}^+$, 我们任取一个此时的最小集合 S_* , 构造 w' 为

$$w'_x = \begin{cases} w_x - 1 & (x \in S_*) \\ w_x & (x \notin S_*) \end{cases}$$

此时 w' 是一个有唯一最小集合 (S_*) 的赋值方式。

而且对于 w' , 可以通过取出唯一最小集合 +1 返回得到 w , 因此该映射为单射。从而“最小集合唯一的赋值方式”不少于 $|\mathcal{W}^+| = (l-1)^{|S|}$ 种。

由于 $\frac{|\mathcal{W}^+|}{|\mathcal{W}|} = \frac{(l-1)^{|S|}}{l^{|S|}}$, 立刻得证。

于是我们给每条边 e 随机赋值 $w_e \in \{1, 2, \dots, l\}$, 根据 Isolation Lemma 有很大把握认为最小权完美匹配是唯一的。假设确实唯一。

从而我们令 $x_{ij} = 2^{w_{(i,j)}}$, 称带入值之后的为矩阵 B , 则当 $|A_G| \neq 0$ 即完美匹配存在时:

$$\text{lowbit}(|B|) = 2^{\text{minimal weights perfect match}}$$

求出一个完美匹配的并行算法:

首先计算 $2^w = \text{lowbit}(|B|)$ 。

然后并行的对于每条边 (i, j) , 如果 $2^{w_{(i,j)}} \times \text{lowbit}(|B_{ij}|) = 2^w$, 那么输出 (i, j) 。这里 B_{ij} 表示余子式。

Fingerprinting

给定 n -bit 数 a 和 b , 判断是否相等。

假设这两个数可以快速取模，那么我们在不超过 T 的素数中，随机一个素数 p 。

由于 $|a - b|$ 的素因子个数不超过 $\log_2 |a - b| \leq n$ 个，因此 $a \equiv b \pmod{p}$ 的概率不超过 $\frac{n}{\pi(T)}$ 。

Theorem (Prime Number Theorem): 用 $\pi(x)$ 表示 $\leq x$ 的素数的个数，

$$\forall x \geq 17, \quad \frac{x}{\ln x} \leq \pi(x) \leq 1.26 \frac{x}{\ln(x)}$$

我们随机生成一个不超过 T 的素数，发生错误的概率不超过 $\frac{n \ln T}{T}$ 。

因此取 $T = cn \ln n$ ，则有错误概率 $\leq \frac{1}{c} + o(1)$ 。

更紧的，有结论： n -bit 数的素因子数量不超过 $\pi(n)$ ，因此取 $T = cn$ 就能达到效果。

Fingerprinting 算法直接应用：Pattern matching。

Lecture 3 - 2025 / 2 / 24

Primality Testing

费马素数测试：随机选择 $a \in \{1, 2, \dots, n-1\}$ ，如果 $\gcd(a, n) \neq 1$ 直接输出 n 不是素数，否则如果 $a^{p-1} \equiv 1 \pmod{n}$ ，则输出 Yes，否则为 No。

Definition (Carmichael number): 对于所有 $1 \leq a < n$ ，都有 $a^{p-1} \equiv 1 \pmod{n}$ ，则 n 为 Carmichael 数。

Theorem: 如果 n 是合数且不是 Carmichael 数，则 $\Pr[\text{Error in Fermat test}] \leq \frac{1}{2}$ 。

后文称 $G = \{a \mid (a, n) = 1\} = \mathbb{Z}_n^*$ 。

令 $H = \{a \in G \mid a^{n-1} \equiv 1 \pmod{n}\}$ ，显然有 $H \subseteq G$ ，从而根据拉格朗日定理，

$$\Pr[\text{Error in Fermat test}] \leq \frac{|H|}{|G|} \leq \frac{1}{2}。$$

现在考虑 Carmichael 数，首先我们处理掉 $n = p^k$ 的情况。

Claim: 可以在 $O(\log^2 n)$ 的时间内，判断一个数是不是 p^k 。

首先 $k < O(\log n)$ ，所以每次二分 p 即可。

Lemma: 对于素数 p ，一定不存在 $x \not\equiv \pm 1 \pmod{p}$ ， $x^2 \equiv 1 \pmod{p}$ 。

$$(x-1)(x+1) \equiv 0 \pmod{p}$$

我们试图通过寻找非平凡 1 的平方根的方式来判定素数。

记 $n - 1 = 2^w O$ 。随机选择 $a \in G$ 。

- 首先根据 Carmichael, $a^{2^w O} \equiv 1 \pmod{n}$ 。
- 计算 $a^{2^{w-1} O} \pmod{n}$, 如果是 -1 , 输出 Yes, 如果是 1, 继续; 否则输出 No。
- 计算 $a^{2^{w-2} O} \pmod{n}$, 如果是 -1 , 输出 Yes, 如果是 1, 继续; 否则输出 No。
-
- 如果 $a^O \equiv 1 \pmod{n}$ 依然成立, 输出 Yes。

显然素数一定能通过这个测试。对于合数, 如果 a 能够成功淘汰它, 则称 a 为一个 witness。

Claim: 对于存在两个不同素因子 p_1, p_2 的合数 n , $\Pr[a \text{ is a non-witness}] \leq \frac{1}{2}$ 。

第一步构造一个包含所有 non-witness 的 G 的子群。

记 $s^* \in \{O, 2O, \dots, 2^w O\}$ 为最大的满足, $\exists x \in G, x^{s^*} \equiv -1 \pmod{n}$ 的数。

s^* 一定是良定义的, 因为 $(-1)^O \equiv -1 \pmod{n}$ 。

构造 $H = \{a \in G \mid a^{s^*} \equiv \pm 1 \pmod{n}\} \leq G$ 。易见所有 non-witness 都包含于 H 。下面说明

$H \leq G$, 即可由拉格朗日定理得到 $\Pr[a \text{ is a non-witness}] \leq \frac{|H|}{|G|} = \frac{1}{2}$ 。

考虑中国剩余定理, 取出一个 $(x^*)^{s^*} \equiv -1 \pmod{n}$, 我们构造满足如下方程的 $a \in G$ 。

$$\begin{cases} a \equiv x^* \pmod{p_1^{k_1}} \\ a \equiv 1 \pmod{p_2^{k_2}} \end{cases}$$

由于 $a \notin H, a \in G$, 从而 H 是真子群, 原命题得证。

Probabilistic Method

Theorem (Ramsey): 对于 $n \leq 2^{k/2}$ 个点的图, 存在二染色方案, 使得任意 k 完全子图都不是同色的。

Theorem (Max Cut): 对于图 $G = (V, E)$, 存在一个割的大小 $\geq \frac{|E|}{2}$ 。

Independent Set

Claim: 对于图 $G = (V, E)$, 存在独立集大小 $\geq \sum_v \frac{1}{\deg(v) + 1}$ 。

随机对点赋实数值, 如果一个点是自己和邻居的最小值, 就将其选入独立集。

可以看出不会选到相邻的点。 v 被选入的概率是 $\frac{1}{\deg(v) + 1}$, 从而期望即右式。

Crossing Number

Definition (crossing number): 把 $G = (V, E)$ 嵌入平面，交叉数 $c(G)$ 为最少的边的交点数量。

Theorem (Euler's formula): 对于平面图， $|V| + |R| = |E| + 2$ 。同时 $|R| \geq \frac{2|E|}{3}$ 从而 $|E| \leq 3|V| - 6$ 。

Claim: $c(G) \geq |E| - 3|V| + 6$

容易验证，最佳的嵌入方式满足：

- 边不自交
- 两条边至多一个交点
- 有公共点的边不交

于是，对于原图每一组相交的 $(a, b), (c, d)$ ，构造新的点 v ，断开原来的边并将 $(a, v), (b, v), (c, v), (d, v)$ 连边。

新图为平面图， $|E'| = |E| + 2c(G)$ ， $|V'| = |V| + c(G)$ ，从而

$$|E| + 2c(G) \leq 3|V| + 3c(G) - 6 \Rightarrow c(G) \geq |E| - 3|V| + 6$$

用概率方法加强这个结论。我们以 p 的概率保留一个点， $1 - p$ 的概率把点删去。

从而每条边有 p^2 的概率保留下来，每个原来的交点有 p^4 的概率被保留下来。

从而

$$p^4 c(G) \geq \mathbb{E}[c(G)] \geq \mathbb{E}[|E| - 3|V| + 6] = p^2 |E| - 3p |V| + 6$$

$$c(G) \geq \frac{p^2 |E| - 3p |V| + 6}{p^4} \geq \frac{p |E| - 3 |V|}{p^3}$$

Claim: 对任何 $|E| \geq 4|V|$ 的图 G ，有 $c(G) \geq \frac{|E|^3}{64|V|^2}$ 。

取 $p = \frac{4|V|}{|E|}$ 即可。

Lecture 4 - 2025 / 2 / 27

Unbalancing Lights

对于 $n \times n$ 的灯泡矩阵，每行、每列各有一个开关，作用是翻转完整的一行、一列。

现在对于一个初始状态，试图通过操作开关最大化亮灯数。

Claim: 对于每一种初始状态, 存在操作方式使亮灯数量当 $n \rightarrow \infty$ 时渐进

$$\frac{n^2}{2} + \sqrt{\frac{1}{2\pi}} n^{3/2}$$

首先均匀随机操作每一列的开关。用 $X_{ij} = \pm 1$ 表示 (i, j) 位置的灯是否亮。

对于第 i 行, 用 $Z_i = \sum_j X_{ij}$, 由于 X_{i1}, \dots, X_{in} 在 $\{1, -1\}$ 中均匀随机, 因此由随机游走结论:

$$\mathbb{E}[|Z_i|] \sim \sqrt{\frac{2}{\pi}} n$$

对于每一行的开关, 如果操作后亮灯数量增多就操作它。从而根据期望的线性性:

$$\mathbb{E}[\#\text{on} - \#\text{off}] \sim \sqrt{\frac{2}{\pi}} n^{3/2}$$

从而 $\mathbb{E}[\#\text{on}] \sim \frac{n^2}{2} + \sqrt{\frac{1}{2\pi}} n^{3/2}$ 。

Large Girth and Chromatic Number

Definition (girth): 一个图 G 的周长为其中最小环的长度。

Definition (chromatic number): 一个图 G 的染色数为同色不相邻染色, 最少需要的颜色数。

Theorem: $\forall k, l$, 存在一张图的周长 $\geq l$, 染色数 $\geq k$ 。

取随机图 $G \sim \mathcal{G}_{n,p}$, 这里 $p = n^{-1+1/l}$ 。

用 X 表示 G 的 $< l$ 的环数量, Y 表示最大独立集的大小。

首先

$$\mathbb{E}[X] = \sum_{i=3}^{l-1} \frac{n^i}{2^i} p^i \leq \sum_{i=3}^{l-1} \frac{(np)^i}{2^i} = \sum_{i=3}^{l-1} \frac{n^{i/l}}{2^i} = O(n^{1-1/l}) = o(n)$$

从而 $\Pr[X \geq \frac{n}{2}] = o(1)$ 。

另一方面, 任取 y ,

$$\begin{aligned} \Pr[Y \geq y] &\leq \binom{n}{y} (1-p)^{\binom{y}{2}} \\ &\leq n^y \cdot e^{-p \binom{y}{2}} \leq (e^{\ln n - py/4})^y \end{aligned}$$

取 $y = \frac{8 \ln n}{p} = 8 \ln n \cdot n^{1-1/l} = o(n)$, 就有 $\Pr[Y \geq y] \leq e^{-\ln n \cdot y} = o(1)$ 。

因此，根据 union bound，当 n 足够大， G 有 $\geq \frac{1}{2}$ 的概率满足：

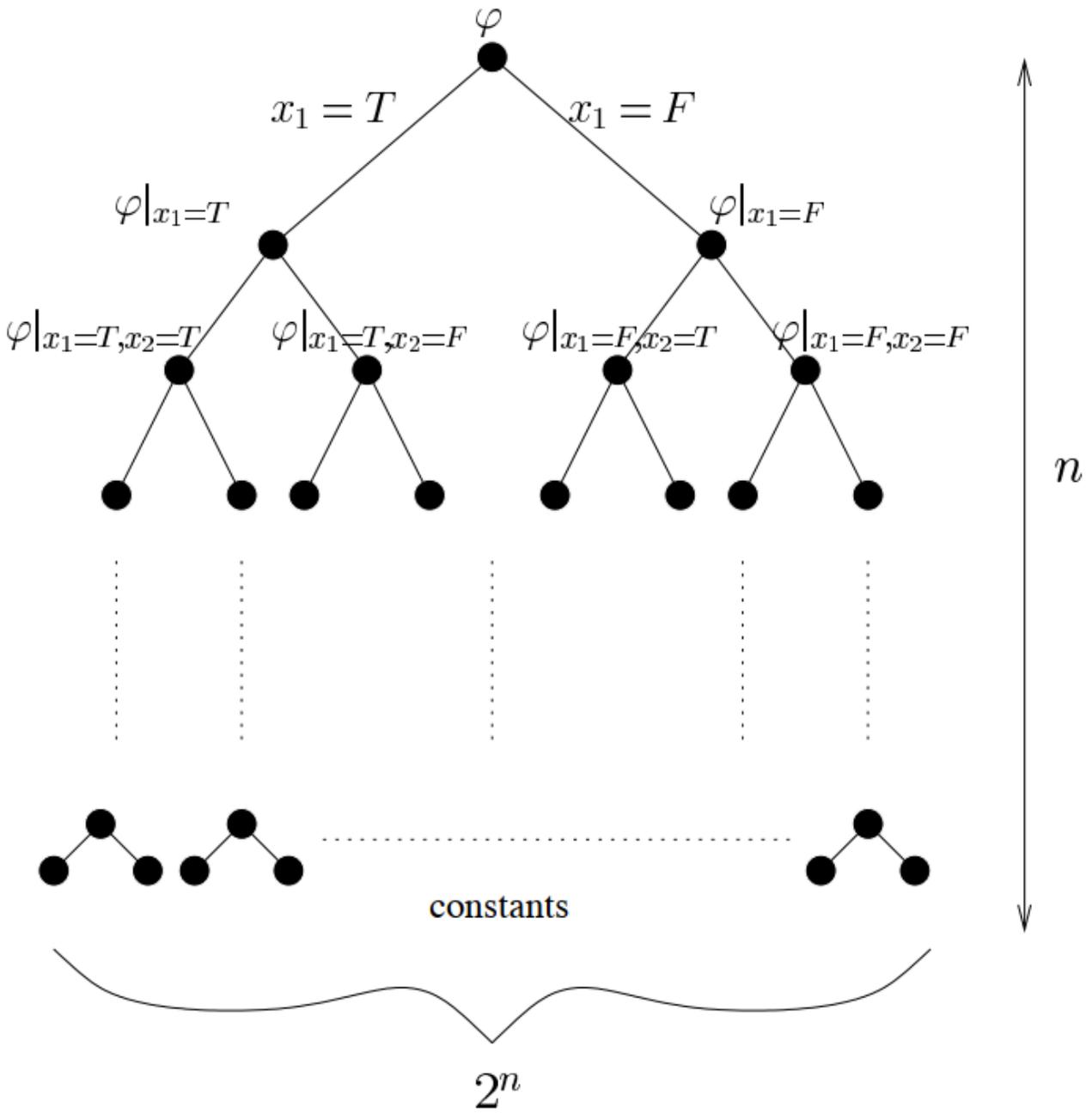
- $< l$ 的环的数量不超过 $\frac{n}{2}$
- 最大独立集的大小不超过 $y = o(n)$

从每个环中删去一个点，剩下的图 G' 周长 $\geq l$ ，染色数 $\geq \frac{n}{y} = \omega(1)$ ，从而 n 充分大一定可以满足染色数 $\geq k$ 。

MAX3SAT

记 $\varphi = \{(x_1 \vee \neg x_2 \vee x_3), \dots\}$ ，其中的每一项称为一个 clause。

Claim: 对于任一个 φ ，存在一种赋值方法使至少 $\frac{7}{8}|\varphi|$ 的 clause 被满足。并且可以高效找出。



存在性只需要随机赋值即可证明。

依次考虑每一个 x_i ，由于

$$\frac{7}{8}|\varphi| = \mathbb{E}[\varphi] = \Pr[x_1 = T] \cdot \mathbb{E}[\varphi|x_1 = T] + \Pr[x_1 = F] \cdot \mathbb{E}[\varphi|x_1 = F]$$

从而一定能有一种条件期望 $\geq \frac{7}{8}|\varphi|$ ，递归下去寻找即可。

这种方法叫做 **Method of conditional probabilities**。

4-Cliques / Triangles

Definition (threshold): 称 $p(n)$ 是性质 Q 的 threshold, 当且仅当:

$$\begin{aligned} p \gg p(n) &\implies \Pr[G \in \mathcal{G}_{n,p} \text{ has } Q] \rightarrow 1 \text{ as } n \rightarrow \infty \\ p \ll p(n) &\implies \Pr[G \in \mathcal{G}_{n,p} \text{ has } Q] \rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

对于图 $G \sim \mathcal{G}_{n,p}$, 设 X 为其中的 4-Clique 的个数, $X_C = 0/1$ 代表 C 是不是 4-Clique。

$$\mathbb{E}[X] = \binom{n}{4} p^6 = \Theta(n^4 p^6)$$

Theorem: $p(n) = n^{-2/3}$ 是包含 4-Clique 的 threshold。

首先 $p \ll p(n)$ 时, 由于 $\mathbb{E}[X] \rightarrow 0$, 因此 $\Pr[X \geq 1] \leq \mathbb{E}[X] \rightarrow 0$ 。

当 $p \gg p(n)$ 时, $\Pr[X = 0] \leq \Pr[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]] \leq \frac{\text{Var}[X]}{\mathbb{E}[X]^2}$ 。

由于

$$\begin{aligned} \text{Var}[X] &= \sum_C \text{Var}[X_C] + \sum_{C,D} \text{Cov}[X_C, X_D] \\ &\leq \Theta(n^4 p^6) + \binom{n}{6} \binom{6}{2} p^{11} + \binom{n}{5} \binom{5}{3} p^9 \\ &= \Theta(n^4 p^6) + \Theta(n^6 p^{11}) + \Theta(n^5 p^9) \end{aligned}$$

从而 $\frac{\text{Var}[X]}{\mathbb{E}[X]^2} = \Theta(n^{-4} p^{-6}) + \Theta(n^{-2} p^{-1}) + \Theta(n^{-3} p^{-3}) \rightarrow 0$ 。

该方法不适用于密集程度“不均匀”的图。

Lecture 5 - 2025 / 3 / 3

Monotone Circuits for the Majority Function

Definition (Boolean circuit): $f : \{0, 1\}^n \rightarrow \{0, 1\}$, 通过门进行计算, 每个门即 $\{0, 1\}^2 \rightarrow \{0, 1\}$ 的函数 (共 16 种门)。

Claim: 几乎所有 n 个输入的 Boolean function 需要 $\Omega(2^n/n)$ 个门 (包括输入门)。

首先 n 个输入的 Boolean function 有 2^{2^n} 种。

考虑 S 个门能够表达多少种 Boolean function。首先每个门可以选择 S^2 种输入, 以及自身有 16 种计算方法, 故函数数量不超过 $(16S^2)^S$ 。

将 S 用 $\frac{2^n}{16n}$ 带入, 由于

$$\begin{aligned} S \ln(16S^2) &= \frac{2^n}{16n} \ln \left(16 \cdot \frac{4^n}{16^2 n^2} \right) = \frac{2^n}{16n} (-\ln 16 + n \ln 4 - 2 \ln n) \\ &= 2^n \frac{\ln 2}{8} + \dots \end{aligned}$$

另一方面 $\ln 2^{2^n} = 2^n \ln 2$, 因此 $S < \frac{2^n}{16n}$ 时, $\lim_{n \rightarrow \infty} \frac{(16S^2)^S}{2^{2^n}} = 0$ 。

Definition (monotone circuits): 一个电路是单调的, 当且仅当它的所有门都是单调函数, 即:

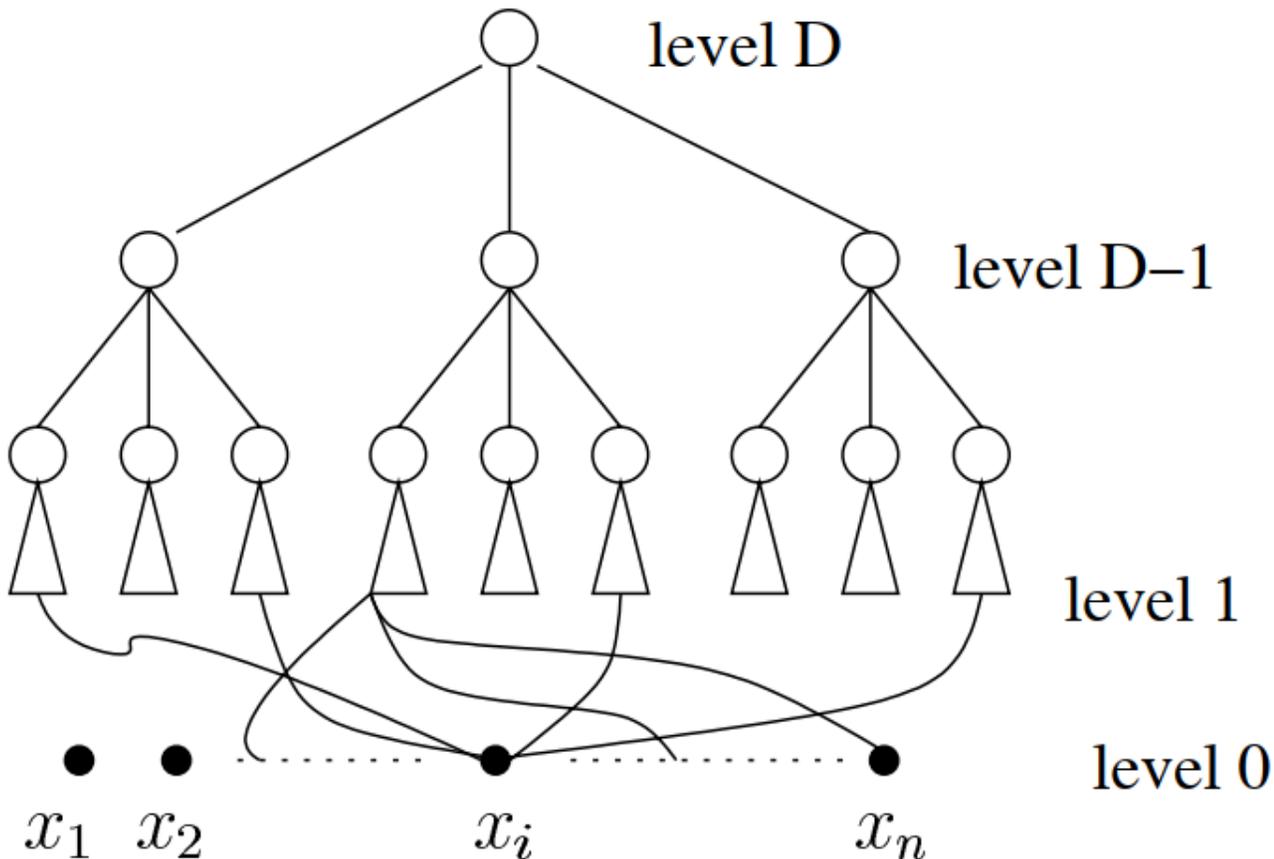
$$f(x_1, \dots, x_n) = 1, \forall i, y_i \geq x_i \Rightarrow f(y_1, \dots, y_n) = 1$$

现在考虑众数函数 $\text{Maj}_n(x_1, \dots, x_n)$, 试图找到一个单调电路来实现它。

一个最优的实现 Maj_3 的电路为 (因为只用到了单调的 \wedge, \vee , 故这个电路也是单调的):

$$(x_1 \wedge (x_2 \vee x_3)) \vee (x_2 \wedge x_3)$$

Theorem: 存在一个单调电路计算 Maj_n , n 为奇数, 门的数量是 $\text{poly}(n)$, 深度是 $O(\log n)$ 。



考虑一个随机电路 C , 包含 $D = O(\log n)$ 层的 Maj_3 , 底层每个 Maj_3 随机从 x_1, \dots, x_n 中选择 3 个输入。

不妨设众数为 1，那么底层每个门输入 1 的概率至少为 $p_0 = \frac{n+1}{2n} = \frac{1}{2} + \frac{1}{2n}$ 。

如果一个 Maj_3 的每个输入有 p 的概率为 1，那么其输出为 1 的概率为

$$f(p) = p^3 + 3p^2(1-p) = 3p^2 - 2p^3$$

考虑迭代过程 $p_1 = f(p_0), p_2 = f(p_1), \dots$ ，目标为证明在 $O(\log n)$ 次迭代后， $p \geq 1 - 2^{-(n+1)}$ ，从而根据 union bound， $\Pr[\exists \mathbf{x}, C(\mathbf{x}) \neq \text{Maj}_n(\mathbf{x})] \leq 2^n \cdot 2^{-(n+1)} = \frac{1}{2}$ ，根据概率方法立刻得证。

1. 第一阶段， $\frac{1}{2} + \frac{1}{2n} \leq p_t \leq \frac{3}{4}$ ，由于步长增大，计算得

$$\left(p_{t+1} - \frac{1}{2}\right) \geq \frac{11}{8} \left(p_t - \frac{1}{2}\right)$$

故在 $O(\log n)$ 步内， p_t 可以达到 $\frac{3}{4}$ 。

2. 第二阶段： $p_t \geq \frac{3}{4}$ ，设第一次达到这个要求为 p_{t_0} ，则：

$$(1 - p_{t+1}) \leq 3(1 - p_t)^2 \leq 3(1 - p_{t_0})^{2^{t+1-t_0}} \leq \frac{3}{4^{2^{t+1-t_0}}}$$

故在 $O(\log n)$ 步内， p_t 可以达到 $1 - \frac{1}{2^{n+1}}$ 。

从而总共只需 $D = O(\log n)$ 次迭代即可。

Lecture 6 - 2025 / 3 / 6

Probability Amplification Using Pairwise Independence

Claim: 随机变量 $a, b \sim U(\mathbb{Z}_q)$ ， q 是质数，则

$$\{ax + b \mid x \in \mathbb{Z}_q\}$$

是一组两两独立的随机变量，且同分布于 $U(\mathbb{Z}_q)$ 。

首先 $\forall x, c \in \mathbb{Z}_q, \Pr[ax + b = c] = \frac{1}{q}$ ，故 $ax + b \sim U(\mathbb{Z}_q)$ 。

考虑 $\forall x, y, c_1, c_2 \in \mathbb{Z}_q, x \neq y$ ，则 $\Pr[ax + b = c_1, ay + b = c_2] = \frac{1}{q^2} = \Pr[ax + b = c_1] \Pr[ay + b = c_2]$ 。（因为关于 a, b 的方程有唯一解）从而两两独立。

假设现在已有一个随机算法 A ，依赖 m 个随机 bits，用来判断 $x \in L \subseteq \{0, 1\}^n$ 是否成立。而且满足：

$$\begin{aligned} x \in L &\Rightarrow \Pr[A \text{ output Yes}] \geq \frac{1}{2} \\ x \notin L &\Rightarrow \Pr[A \text{ output Yes}] = 0 \end{aligned}$$

现在试图将这个算法泛化到任何正确性。如果独立重复 t 次，可以做到 $\Pr[\mathcal{E}] \leq 2^{-t}$ ，从而如果需要达到 $\frac{1}{r}$ 的正确率，则需要生成 $m \log r$ 个随机 bits。

Theorem: 对于 $r \leq 2^m$ ，可以只生成 $2m$ 随机 bits，在 $O(rm)$ 的时间复杂度内达到 $\Pr[\mathcal{E}] \leq 2^{-t}$ 的效果。

考虑生成 r 组两两独立的长度为 m 的随机 bits。形式化的说，每组随机 bits 可以看作从 $U(\{0, 1\}^m) \cong U(\mathbb{Z}_{2^m})$ 采样的随机变量，这 r 个随机变量两两独立。一个不太完美的做法可以利用上面 **Claim** 的算法，取质数 $2^m \leq q \leq 2^{m+1}$ ，通过 rejection sampling 可以通过生成期望 $O(m)$ 个随机 bits 得到 $U(\{0, 1\}^m)$ 中 r 组两两独立的比特串。

然后运行算法 A r 次。用 $X_i = 0/1$ 代表第 i 次 A 的输出，输出 Yes 时 $X_i = 1$ 。定义 $X = \sum_{i=1}^r X_i$ 。

当 $x \in L$ 时，发生错误的概率为

$$\Pr[\mathcal{E}] = \Pr[X = 0] \leq \Pr[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]] \leq \frac{\text{Var}[X]}{\mathbb{E}[X]^2}$$

其中由于两两独立， $\text{Var}[X] = \sum_{i=1}^r \text{Var}[X_i] \leq \frac{r}{4}$ ， $\mathbb{E}[X] \geq \frac{r}{2}$ ，从而 $\Pr[\mathcal{E}] \leq \frac{1}{r}$ 。

Derandomization Using k -wise Independence

考虑给一张完全图 K_n 的边二染色，要求没有同色 k -clique，这里 $n = 2^{k/2}$ ，根据之前的概率方法，染色方案是存在的。

如果要求出一种方案，一种暴力的策略是枚举 $2^{\binom{n}{2}}$ 种染色方案。

回顾概率证法，设 X 为同色 k -clique 数量，

$$\mathbb{E}[X] = \binom{n}{k} \frac{2}{2^{\binom{k}{2}}} < 1$$

这里其实并不要求所有边的染色全部独立。事实上，只要每 $\binom{k}{2}$ 条边的染色是相互独立的即可。

考虑一族 $\binom{k}{2}$ -wise 独立的染色方案，其中每条边的颜色边际分布是均匀的。根据上述概率证法， $\mathbb{E}[X]$ 不变，从而这族染色方案中一定存在一个合法方案。

推广 **Claim** 到 $ax^2 + bx + c$ ，不难看出，要生成服从 $U(\mathbb{Z}_q)$ 的 $\binom{k}{2}$ -wise 独立的随机变量，只需要采样 $\binom{k}{2}$ 个服从 $U(\mathbb{Z}_q)$ 的变量。这里需要 $q \geq \binom{n}{2}$ ，以保证能够生成足够数量的随机变量。

从而我们枚举 $q^{\binom{k}{2}}$ 种采样的可能性，然后通过固定的解码策略得到唯一对应的 $\binom{k}{2}$ -wise 独立的边染色方案。在这族方案上， $\mathbb{E}[X] < 1$ ，从而其中必有可行解。

由于 $q^{\binom{k}{2}} \simeq n^{O(k^2)}$ ，相较于暴力做法 $2^{O(n)}$ ，我们将复杂度降到了多项式级别。

Universal Hashing

Definition (2-universal): 一个 $U \rightarrow T$ 的函数集 \mathcal{H} 是 2-universal 的当且仅当 $\forall x, y \in U, x \neq y$ ，有

$$\Pr_{h \in \mathcal{H}} [h(x) = h(y)] \leq \frac{1}{|T|}$$

例如， $h_{a,b}(x) = (ax + b) \bmod q \bmod |T|$ ，其中 $a, b \sim U(\mathbb{Z}_q), q > |U|$ 。

$$\begin{aligned} \Pr[h_{a,b}(x) = h_{a,b}(y)] &\leq \sum_{c_1 = c_2 \pmod{|T|}} \Pr[h_{a,b}(x) = c_1] \Pr[h_{a,b}(y) = c_2] \\ &= \frac{q^2}{|T|} \cdot \frac{1}{q} \cdot \frac{1}{q} = \frac{1}{|T|} \quad \forall x \neq y \end{aligned}$$

Lecture 7 - 2025 / 3 / 10

Double Hashing

Claim: 对于一组 2-universal hashing 把 $S \subseteq U$ 的元素投影到 T ，且 $|T| = |S|^2$ ，则存在碰撞的概率 $\leq \frac{1}{2}$ 。

$$\mathbb{E}[\text{collision}] \leq \binom{|S|}{2} \frac{1}{|T|} \leq \frac{1}{2}$$

当然哈希表大小为 $O(|S|^2)$ 还是过大，希望能压缩到 $O(|S|)$ 。

Claim: 对于一组 2-universal hashing 把 $S \subseteq U$ 的元素投影到 T ，且 $|T| = |S|$ 。设有 b_i 个元素 $h(x) = i$ ，则 $\Pr \left[\sum_{i=1}^{|S|} b_i^2 \geq 4|S| \right] \leq \frac{1}{2}$ 。

首先注意到：

$$\#\text{collision} = \sum_{i=1}^{|S|} \binom{b_i}{2} = \frac{1}{2} \left(\sum_{i=1}^{|S|} b_i^2 - |S| \right)$$

另一方面 $\mathbb{E}[\#\text{collision}] = \binom{|S|}{2} \frac{1}{|T|} \leq \frac{|S|}{2}$ ，从而 $\mathbb{E} \left[\sum_{i=1}^{|S|} b_i^2 \right] \leq 2|S|$ 。

从而由 Markov 不等式立刻得证。

从而可以通过第一次 hash 将值域映射到 $|S|$ ，对于有 b_i 个冲突的组，再进行一次 hash 将值域映射到 b_i^2 。从而我们可以在期望 $O(S)$ 次抽取哈希函数，构造一个值域为 $O(S)$ 的无冲突 hash。

Buffon's Needle

平面上一组两两距离为 1 的平行线，现在随机投掷（中心点均匀随机、角度均匀随机）一根长度为 1 的针，那么针与线相交的概率是多少？

$$\frac{2}{\pi} \int_{\theta=0}^{\pi} \int_{d=0}^{1/2 \sin \theta} 1 \, dd \, d\theta = \frac{1}{\pi} \int_{\theta=0}^{\pi} \sin \theta \, d\theta = \frac{2}{\pi}$$

Median Trick

Theorem (Unbiased Estimator Theorem): 对于两两独立的 X_1, \dots, X_t ，期望为 μ ，方差为 σ^2 ， $X = \frac{1}{t} \sum_{i=1}^t X_i$ ，则当 $t \geq \frac{1}{\delta} \cdot \frac{\sigma^2}{\epsilon^2 \mu^2}$ 时，

$$\Pr[|X - \mu| \geq \epsilon \mu] \leq \frac{\text{Var}[X]}{\epsilon^2 \mu^2} = \frac{\sigma^2}{t \epsilon^2 \mu^2} \leq \delta$$

现在所以，达到 δ 的错误率需要通过 $O(\frac{1}{\delta})$ 次采样。现在考虑增加一部分随机性，能否通过 $O(\log \frac{1}{\delta})$ 的样本实现同样的错误率。

Lemma: 对于一枚 $\Pr[\text{Head}] \geq \frac{3}{4}$ 的硬币，在 $2s + 1$ 次相互独立投掷中， $\Pr[\#\text{Head} \leq s] \leq (\frac{3}{4})^s$ 。

$$\begin{aligned} \Pr[\#\text{Head} \leq s] &\leq \sum_{i=0}^s \binom{2s+1}{s} \left(\frac{3}{4}\right)^i \left(\frac{1}{4}\right)^{2s+1-i} \\ &\leq \left(\sum_{i=0}^s \binom{2s+1}{i}\right) \left(\frac{3}{4}\right)^s \left(\frac{1}{4}\right)^{s+1} \\ &\leq \left(\frac{3}{4}\right)^s \times \frac{2^{2s+1}}{4^{s+1}} \leq \left(\frac{3}{4}\right)^s \end{aligned}$$

从而我们组间完全独立、组内两两独立的生成 $2 \log_{3/4} \frac{1}{\delta} + 1$ 组、每组 $\frac{4\sigma^2}{\epsilon^2 \mu^2}$ 个样本。对于每组求平均值、再对所有组求中位数。从而可以在 $O(\log \frac{1}{\delta})$ 次采样实现 δ 的错误率。

Lecture 8 - 2025 / 3 / 13

DNF Counting

Definition (Disjunctive Normal Form): 称形如 $(x_1 \wedge x_2 \wedge \dots) \vee (\bar{x}_3 \wedge \dots) \vee \dots$ 为 DNF。

类似的, CNF 就是常见的 SAT 问题, 有 $\#\text{SAT}(\varphi) = 2^n - \#\text{DNF}(\neg\varphi)$ 。

我们试图设计一个算法在多项式时间内估算 DNF 的解的比例的 FPRAS。

Definition (fully polynomial randomized approximation scheme): 针对 $f: \Sigma^* \rightarrow \mathbb{N}$ 的 FPRAS 是一个算法, 读入 (x, ε) , 在关于 $|x|, \varepsilon^{-1}$ 多项式时间内输出随机变量 Z 满足:

$$\Pr[(1 - \varepsilon)f(x) \leq Z \leq (1 + \varepsilon)f(x)] \geq \frac{3}{4}$$

给定 DNF $\varphi_1 \vee \dots \vee \varphi_n$, 共涉及 x_1, \dots, x_m 。设第 i 个 term 的解集为 S_i , 显然 $S_i = 2^{m-|\varphi_i|}$, 目标即为求 $|\bigcup S_i|$ 。具体而言, 构造集合

$$U = \{(a, i) \mid a \in S_i\}$$

从而 $|U| = \sum_{i=1}^n |S_i|$ 。我们可以在 U 中均匀随机采样 (a, i) 。我们称一个样本 (a, i) 是 *special* 的, 当且仅当 $\forall j < i, a \notin S_j$ 。换言之, a 最早出现在 S_i 中。

从而

$$\mathbb{E} \left[\frac{\#\text{special}}{\#\text{total}} \right] = \frac{|\bigcup S_i|}{|U|}$$

由于 $\mu \geq \frac{1}{n}$, 从而由 Unbiased Estimator Theorem, 可以有效得到 $\frac{3}{4}$ 正确率的 ε 误差估计。实际上根据 Chernoff bound, 只需要 $O(n/\varepsilon^2)$ 次独立采样即可。

Network Reliability (1)

对于一张图 G , 有 n 个点 m 条边, 每条边有 p 的概率割断, 记 p_{fail} 为 G 不连通的概率, 即“网络鲁棒性”。

Theorem: 存在关于 n, ε^{-1} 多项式时间的 FPRAS 估测 p_{fail} 。(对于每条边隔断概率不同的情况, 依然存在)

设 c 为最小割的长度。

如果 $p^c \geq \frac{1}{n^4}$, 则直接使用 Monte Carlo 方法, 根据 Unbiased Estimator Theorem, 由于 $\mu \geq p^c$, 可以在 $O(\frac{1}{\mu^2 \varepsilon^2}) = O(n^8 \varepsilon^{-2})$ 次采样中得到估计。后文假设该性质不成立, 即 $p^c = n^{-(4+\delta)}$ 。

用 α -最小割表示大小不超过 αc , 且仅将 G 分为两部分的割。

考虑如下算法 RMinCut: 均匀随机抽取图中一条边 (u, v) , 将两个点缩点 (保留重边), 直到只剩下 2 个点, 返回它们之间所有的边。

Theorem: 设 $C \subset E$ 是任一个最小割, 则 $\Pr[\text{RMinCut returns } C] \geq \binom{n}{2}^{-1}$ 。

由于最小割大小为 c , 所以任何点的度数都 $\geq c$, 也就是 $|E(G)| \geq \frac{cn}{2}$ 。

从而第 1 轮选中 C 中边的概率 $\leq \frac{c}{cn/2} = \frac{2}{n}$

容易看出等价于一直没有选择 C 中的边, 而且缩点并不会导致新图最小割变小, 从而

$$\begin{aligned} \Pr[C \text{ survive all rounds}] &\geq \left(1 - \frac{2}{n}\right) \left(1 - \frac{2}{n-1}\right) \cdots \left(1 - \frac{2}{3}\right) \\ &= \frac{2}{n(n-1)} \end{aligned}$$

Corollary: 任意图 G 的最小割的数量不超过 $\binom{n}{2}$, 因为 RMinCut 输出任何一个最小割是互斥事件。

Claim: 只有至多 $n^{2\alpha}$ 个 α -最小割, 这些割可以在关于 n, ϵ^{-1} 的多项式时间内列举出。

类似上面的证明方法, 对于任意一个 α -最小割 C , 有

$$\begin{aligned} \Pr[C \text{ survive until } 2\alpha \text{ vertices remain}] &\geq \left(1 - \frac{2\alpha}{n}\right) \cdots \left(1 - \frac{2\alpha}{2\alpha+1}\right) \\ &= \binom{n}{2\alpha}^{-1} \end{aligned}$$

对于剩下 2α 个点的图, 任意输出一个割, 则

$$\Pr[C \text{ survive}] \geq \binom{n}{2\alpha}^{-1} \frac{1}{2^{2\alpha-1}} \geq \frac{1}{n^{2\alpha}}$$

最后, 根据 coupon-collector, 可以在期望 $O(n^{2\alpha} \log n^{2\alpha})$ 次实验内, 列举出所有的 α -最小割。

从而, 我们对于 $\text{poly}(n)$ 个 α -最小割, 可以用加权的 DNF Counting 的方式, 估算至少一个发生的概率。具体而言, 割掉 x 条边的方案权值为 $p^x(1-p)^{m-x}$, 从而一个 term 的总权值为 $p^{|\varphi_i|}$, 可以构造

$$\mathbb{E} \left[\frac{\sum_{\text{special } (a,i)} p^{|a|} (1-p)^{m-|a|}}{\sum_{(a,i)} p^{|a|} (1-p)^{m-|a|}} \right] = \frac{\sum_{\text{cut } a} p^{|a|} (1-p)^{m-|a|}}{\sum_{i=1}^n p^{|\varphi_i|}}$$

的 $(1 \pm \epsilon)$ 估计, 而右边分子正是我们想求的概率。

Lecture 9 - 2025 / 3 / 17

Network Reliability (2)

接下来，对于 $\geq \alpha c$ 个点的割，我们只需要通过说明

$$\Pr[\text{some cut of size } \geq \alpha c \text{ fails}] \leq \varepsilon p_{\text{fail}}$$

即可。

将 $\geq \alpha c$ 个点的割从小到大排序 $c_1 \leq c_2 \leq \dots$ 。假设至少有 $n^{2\alpha}$ 个割（如若不然，直接得到总概率不超过 $n^{2\alpha} p^{c\alpha}$ ），那么前面这部分 fail 的概率不超过

$$n^{2\alpha} p^{c\alpha} \leq n^{2\alpha} n^{-(4+\delta)\alpha} = n^{-(2+\delta)\alpha} \quad (1)$$

对于任意 $\beta > 0$ ，我们知道 $\leq \beta c$ 的割不超过 $n^{2\beta}$ 个，从而 $c_{n^{2\beta}} \geq \beta c$ ，换言之

$$c_k \geq \frac{c}{2} \log_n k \quad \Rightarrow \quad p^{c_k} \leq p^{\frac{c}{2} \log_n k} = k^{-2+\frac{\delta}{2}}$$

所以

$$\begin{aligned} \Pr[\exists i > n^{2\alpha}, c_i \text{ fails}] &\leq \sum_{i > n^{2\alpha}} k^{-2+\frac{\delta}{2}} \leq \int_{n^{2\alpha}}^{\infty} x^{-2+\frac{\delta}{2}} dx \\ &= \frac{n^{-2\alpha(1+\frac{\delta}{2})}}{1+\frac{\delta}{2}} \leq n^{-(2+\delta)\alpha} \end{aligned} \quad (2)$$

结合 (1)(2)， $\Pr[\text{some cut of size } \geq \alpha c \text{ fails}] \leq 2n^{-(2+\delta)\alpha}$ 。

取 $\alpha = 2 + \frac{1}{2} \log_n \left(\frac{2}{\varepsilon}\right)$ ，立刻得到

$$2n^{-(2+\delta)\alpha} \leq 2n^{-(2+\delta)(2+\frac{1}{2} \log_n(\frac{2}{\varepsilon}))} \leq \varepsilon n^{-(4+\delta)} \leq \varepsilon p_{\text{fail}}$$

综上，直接忽略这些大割，算法可以在 $O(n^{2\alpha} \log n^{2\alpha}) = O(n^4 \varepsilon^{-1} (\log n + \log \varepsilon^{-1}))$ 次调用 RMinCut 内，得到关于 p_{fail} 的 $(1 \pm \varepsilon)^2$ 估计。

Chernoff Bounds

Theorem: 让 X_1, \dots, X_n 为独立 $[0, 1]$ 变量 $\mathbb{E}[X_i] = p_i$ ， $X = \sum_{i=1}^n X_i$ ， $\mu = \mathbb{E}[X] = \sum_{i=1}^n p_i$ ， $p = \frac{\mu}{n}$

- $\Pr[X \geq \mu + \lambda] \leq \exp(-nH_p(p + \frac{\lambda}{n}))$ ，对于 $0 < \lambda < n - \mu$
- $\Pr[X \leq \mu - \lambda] \leq \exp(-nH_{1-p}(1 - p + \frac{\lambda}{n}))$ ，对于 $0 < \lambda < \mu$

其中 $H_p(x) = x \ln \frac{x}{p} + (1-x) \ln \frac{1-x}{1-p}$ 为 KL 散度。

通过矩生成函数证明。

Corollary:

$$\Pr[X \leq \mu - \lambda] \\ \Pr[X \geq \mu + \lambda] \leq \exp\left(-\frac{2\lambda^2}{n}\right)$$

对指数部分求导比较即可。

Corollary:

- 对 $0 < \beta < 1$, $\Pr[X \leq (1 - \beta)\mu] \leq \exp\left(-\frac{\beta^2 \mu}{2}\right)$
- 对 $\beta > 0$, $\Pr[X \geq (1 + \beta)\mu] \leq \begin{cases} \exp\left(-\frac{\beta^2 \mu}{2 + \beta}\right) & \beta > 0 \\ \exp\left(-\frac{\beta^2 \mu}{3}\right) & 0 < \beta \leq 1 \end{cases}$

Corollary: 对于 X_i 在 $[a_i, b_i]$ 中取值时,

$$\Pr[X \leq \mu - \lambda] \\ \Pr[X \geq \mu + \lambda] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

Lecture 10 - 2024 / 3 / 20

Randomized Routing

考虑 n 维超立方体, 网络的顶点为 $\{0, 1\}^n$, 共 $N = 2^n$ 个, 每条边双向, 令 π 是任意排列, 目标是从每个 i 发送一个数据包到对应的 $\pi(i)$, 但是同一个条边每个时间只能有 1 个数据包通过。

现在要设计一种路径规划算法, 最小化最大传输时间。这里要求 i 的路径只取决于 i 和 $\pi(i)$ 我们称之为 *oblivious*, 这是具备现实意义的。

Theorem: 对于任何确定性 *oblivious* 的路径规划算法, 存在一种排列需要 $\Omega(\sqrt{N/n}) = \Omega(\sqrt{2^n/n})$ 。

Theorem: 存在一种 *oblivious* 随机路径规划算法, w.h.p 在 $O(n)$ 步停止。

该算法的思路是“随机中转”, 即对于每一个 i , 等概率采样一个 $\delta(i)$, 算法分为两个阶段

1. 从 $i \rightarrow \delta(i)$
2. 从 $\delta(i) \rightarrow \pi(i)$

在两个阶段中，都采用 bit-fixing 方式，例如 $x \rightarrow y$ ，就是从左到右逐位比较，如果 $x_i \neq y_i$ ，那么当前就从对应边前进。

不失一般性，我们只分析第 1 阶段的长度。

用 $D(i)$ 表示 i 在路径中等待的时间长短，那么总时长一定不超过 $n + \max_i D(i)$ ，我们接下来将证明

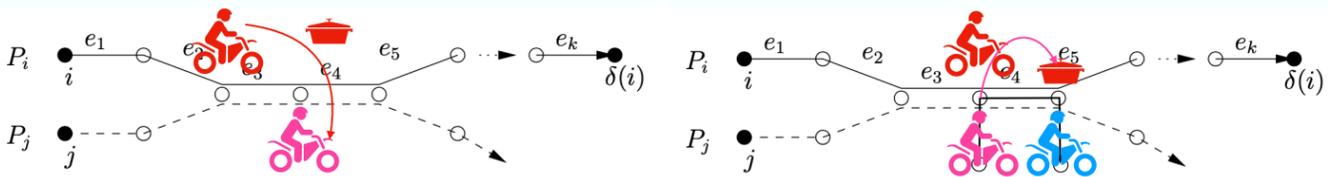
$$\forall i, \Pr[D(i) > cn] \leq e^{-2n}$$

根据 union bound, $\Pr[\exists i, D(i) > cn] \leq 2^n e^{-2n} < 2^{-n}$

用 P_i 表示 $i \rightarrow \delta(i)$ 的路径上所经过的点，用 $S_i = \{j \neq i \mid P_j \cap P_i \neq \emptyset\}$ ，也即路径相交的个数。直观上我们得到以下结论：

Claim: $D(i) \leq |S(i)|$

证明的思路是“锅不能停”。



在等价意义下，至多只会被每个人阻碍一次，因此 $D(i) \leq |S(i)|$ 。

Lemma: $\forall i, \Pr[D(i) > cn] \leq e^{-2n}$

定义 $H_{ij} = \begin{cases} 1 & P_i \cap P_j \neq \emptyset \\ 0 & P_i \cap P_j = \emptyset \end{cases}$ ，从而 $D(i) \leq \sum_{j \neq i} H_{ij} = |S(i)|$ 。

对于图中每一条边，期望经过它的路径条数为 $\frac{Nn/2}{Nn} = \frac{1}{2}$ ，从而 $\mathbb{E}[|S(i)|] \leq \frac{n}{2}$ ，即路径长度乘每条边的期望路径个数。根据 Chernoff bound，有

$$\Pr[D(i) \geq (1 + \beta)\mu] \leq \exp\left(-\frac{\beta^2}{2 + \beta}\mu\right)$$

容易看出，当我们想分析一个固定的 tail 时， $\mu = \frac{n}{2}$ 一定是最坏的，取 $\beta = 6$ ，有 $\Pr[D(i) \geq \frac{7}{2}n] \leq \exp\left(-\frac{9}{4}n\right) \leq \exp(-2n)$ 。

Hamilton Cycles (1)

对于 $G = (V, E) \in \mathcal{G}(n, p)$ ，其中 $p \geq \frac{72 \ln n}{n-1}$ ，则存在一个多项式时间的随机算法 w.h.p. 找到一个 Hamiltonian 圈。

算法的思想是利用 coupon collector，算法陈述如下

```

start with  $P = \{v_1\}$  where  $v_1 = s$  is an arbitrary vertex
repeat for at most  $4(n-1)\ln(n-1)$  steps
  if  $|P| = n$  and  $\{v_1, v_n\} \in E$  then output  $P$ 
  else choose vertex  $y$  with  $\{v_k, y\} \in E$  where  $v_k$  is the current path endpoint
    if  $y \notin P$  then extend( $P, y$ )
    else rotate( $P, y$ )
  if cycle not found then output "fail"
  
```

这里有 *extent* 和 *rotate* 两个函数，分别表示枚举到新端点 v 时，若其之前还没出现 / 已经出现，现在应该如何操作



于是重点在于 *choose* 这一步，我们需要保证的是从“观察者视角”，所有的 V 中的点具有均等概率被选作下一个点。从而我们可以利用 coupon collector 的结论，期望在前 $2n \ln n$ 步收集到所有点形成 Hamiltonian 路，后 $2n \ln n$ 步枚举到与 v_1 相邻的点，找到 Hamiltonian 圈。

严谨的来说，我们要保证

$$\Pr_G[v \in V \text{ is next endpoint} \mid \text{Path history}] \text{ are the same}$$

为了达到这一点，我们需要将边退化为有向边，以确保双向选择的独立性，对于 $(u, v) \in E$ 时，定义 $N(x)$ 为 G' 中 x 的邻居结点

$$\begin{aligned} \{y \in N(x) \wedge x \in N(y)\} & \text{ w.p. } \frac{p}{4} \\ \{y \notin N(x) \wedge x \in N(y)\} & \text{ w.p. } \frac{1}{2} - \frac{p}{4} \\ \{y \in N(x) \wedge x \notin N(y)\} & \text{ w.p. } \frac{1}{2} - \frac{p}{4} \\ \{y \notin N(x) \wedge x \notin N(y)\} & \text{ w.p. } \frac{p}{4} \end{aligned}$$

我们将在实现 *choose* 时从 $N(x)$ 中挑选邻居。定义 $OLD(x) = \{y \mid \text{choose picked } y \text{ when } x \text{ was endpoint}\}$ ，选点策略如下：

- 以 $\frac{|OLD(x)|}{n-1}$ 的概率，在 $OLD(x)$ 中等概率随机挑选一个点。
- 以剩下的概率，在 $N(x) \setminus OLD(x)$ 中等概率随机挑选一个点。

算法的正确性将在下一讲证明。

Lecture 11 - 2025 / 3 / 24

Hamilton Cycles (2)

Claim: 在 G' 中, $\{y \in N(x)\}$ 是互相独立的事件, 且每个以 $\frac{p}{2}$ 的概率发生。

$$\text{首先 } \Pr[y \in N(x)] = p \cdot \left(\frac{p}{4} + \left(\frac{1}{2} - \frac{p}{4}\right)\right) = \frac{p}{2}.$$

其次由于下式, 立刻得到独立性:

$$\Pr[y \in N(x) \wedge x \in N(y)] = p \cdot \frac{p}{4} = \frac{p^2}{4} = \Pr[y \in N(x)] \Pr[x \in N(y)]$$

Claim: 取 $p \geq 72 \frac{\ln n}{n-1}$, $G \in \mathcal{G}(n, p)$, w.h.p. *choose* 的实现方法保证了每个点以均等的概率 $\frac{1}{n-1}$ 作为新的端点。

我们假设始终有 $N(x) \setminus OLD(x) \neq \emptyset$ (接下来会证明), 那么

1. 对于 $y \in OLD(x)$, 有 $\Pr[\text{choose picks } y] = \frac{|OLD(x)|}{n-1} \cdot \frac{1}{|OLD(x)|} = \frac{1}{n-1}$

2. 对于 $y \notin OLD(x)$, 类似可知也为 $\frac{1}{n-1}$ 。

值得注意的是, 这里 G 也为随机性来源之一, 我们是从观察者视角计算概率, 也即我们只能根据 *choose history* 对 G 进行假设。

Claim: 在 $4n \ln n$ 步内, w.h.p $\forall x, N(x) \setminus OLD(x) \neq \emptyset$ 。

我们对于一个 fixed x , 说明 $\Pr[N(x) \setminus OLD(x) = \emptyset] = O\left(\frac{1}{n^2}\right)$ 即可通过 union bound 证明原结论。

首先 $\Pr[|N(x)| \leq 24 \ln n] \leq \frac{1}{n^2}$, 这是因为 $|N(x)| \sim B(n-1, \frac{p}{2})$, 所以 $\mathbb{E}[|N(x)|] = 36 \ln n$ 。我们根据 Chernoff bound 即可得证。

接下来 $\Pr[|OLD(x)| \geq 24 \ln n] \leq \frac{1}{n^2}$, 这是因为, x 作端点的次数 $\sim B(4n \ln n, \frac{1}{n-1})$, 而 $|OLD(x)|$ 显然不会超过这个次数, 故由 Chernoff bound 再次得证。

Balls and Bins (1)

考虑将 m 个球独立均匀放进 n 个桶里, 设第 i 个桶里 X_i 个球, 那么

$$\Pr[X_1 = k_1, \dots, X_n = k_n] = \frac{1}{n^m} \frac{m!}{k_1! \dots k_n!}$$

另一方面, 假设 Y_1, \dots, Y_n 是一列独立服从 $\pi(\lambda)$ 的变量,

$$\Pr[Y_1 = k_1, \dots, Y_n = k_n] = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{k_i}}{k_i!}$$

$$\Pr\left[\sum_{i=1}^n Y_i = m\right] = \frac{e^{-\lambda n} (\lambda n)^m}{m!}$$

从而我们有 $\Pr[X_1 = k_1, \dots, X_n = k_n] = \Pr[Y_1 = k_1, \dots, Y_n = k_n \mid \sum_{i=1}^n Y_i = m]$ 。

Theorem: 将 n 个球独立均匀放进 n 个桶里, 最大负载量 w.h.p 是 $O\left(\frac{\ln n}{\ln \ln n}\right)$ 。

记 \mathcal{E}_1 表示某个桶的球个数 $> (1 + \varepsilon) \frac{\ln n}{\ln \ln n}$ 我们需要证明 $\Pr[\mathcal{E}_1] = 1/\text{poly}(n)$ 。

由于 $X_1 \sim B(n, \frac{1}{n})$, 有

$$\begin{aligned} \Pr\left[X_1 > (1 + \varepsilon) \frac{\ln n}{\ln \ln n}\right] &\leq \left(\frac{e \ln \ln n}{(1 + \varepsilon) \ln n}\right)^{(1 + \varepsilon) \ln n / \ln \ln n} \\ &= \exp\left(\left(1 + \varepsilon\right) \frac{\ln n}{\ln \ln n} \cdot (1 + \ln \ln \ln n - \ln(1 + \varepsilon) - \ln \ln n)\right) \\ &= \exp(-\Theta((1 + \varepsilon) \ln n)) = n^{-\Theta(1 + \varepsilon)} \end{aligned}$$

从而根据 union bound 得证。

Lecture 12 - 2025 / 3 / 27

Balls and Bins (2)

Lemma: 设 \mathcal{E} 是关于 bin loads 的事件, 且 $\Pr[\mathcal{E}]$ 关于 m 递增是单调上升 / 单调下降的, 则 $\Pr_X[\mathcal{E}] \leq 4 \Pr_Y[\mathcal{E}]$, 其中 X 为 Balls and Bins 模型, Y 为 n 个独立的 $\pi(m/n)$ 。

不妨设 $\Pr[\mathcal{E}]$ 单调上升, 则

$$\begin{aligned} \Pr_Y[\mathcal{E}] &= \sum_{k=0}^{\infty} \Pr_Y\left[\mathcal{E} \mid \sum_{i=1}^n Y_i = k\right] \Pr\left[\sum_{i=1}^n Y_i = k\right] \\ &\geq \sum_{k=m}^{\infty} \Pr_Y\left[\mathcal{E} \mid \sum_{i=1}^n Y_i = m\right] \Pr\left[\sum_{i=1}^n Y_i = k\right] \\ &\geq \Pr_Y\left[\mathcal{E} \mid \sum_{i=1}^n Y_i = m\right] \Pr\left[\sum_{i=1}^n Y_i \geq m\right] \\ &\geq \Pr_X[\mathcal{E}] \cdot \frac{1}{4} \end{aligned}$$

最后一步用到对于 $\lambda \in \mathbb{N}$, 对于 $X \sim \pi(\lambda)$, 有 $\Pr[X \geq \lambda] \geq 1/4$ 。

Corollary: $\Pr[\forall i, X_i \leq c] \leq 4 \Pr[\forall i, Y_i \leq c]$

Theorem: 将 n 个球独立均匀放进 n 个桶里, 最大负载量 w.h.p 是 $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ 。

记 \mathcal{E}_2 表示所有 $Y_i \leq (1 - \varepsilon) \frac{\ln n}{\ln \ln n}$ 我们需要证明 $\Pr[\mathcal{E}_2] = 1/\text{poly}(n)$ 。

由于 $Y_1 \sim \pi(1)$, 所以 $\Pr[Y_1 \geq k] = \sum_{j=k}^{\infty} \frac{e^{-1}}{j!} \leq \frac{1}{k!}$ 。这是因为 $e = 1 + 1/2 + 1/3! + \dots$ 。当然, 更直接的有 $\Pr[Y_1 \geq k] \geq \frac{1}{ek!}$ 。

$$\begin{aligned} \Pr[\mathcal{E}_2] &= (1 - \Pr[Y_1 \geq k])^n \\ &\leq \left(1 - \frac{1}{ek!}\right)^n \\ &\leq \exp\left(-\frac{n}{ek!}\right) \\ &\leq \exp(-\exp(\Theta(\varepsilon \ln n))) \\ &= \exp(-n^{\Theta(\varepsilon)}) \end{aligned}$$

于是以指数速度趋于 0。

综上所述, 最大负载量 w.h.p 是 $\Theta\left(\frac{\ln n}{\ln \ln n}\right)$ 。

Stochastic Dominance

Definition (SD w.r.t. random variables): 对于两个在 $[a, b]$ 上的随机变量 X, Y , 如果 $\forall c \in [a, b], \Pr[Y \geq c] \geq \Pr[X \geq c]$, 则称 Y stochastic dominates X , 记作 $X \preceq Y$ 。

Definiton (SD w.r.t. functions): 对于两个在 $[a, b]$ 上的函数 f, g , 如果 $\forall c \in [a, b]$

$$\int_{x \geq c} f(x) dx \leq \int_{y \geq c} g(y) dy$$

则称 f stochastic dominates g , 记作 $f \preceq g$ 。

Lemma: $X_1 \preceq Y_1, X_2 \preceq Y_2$, 且 X_1, X_2 独立, Y_1, Y_2 独立, 则 $X_1 + X_2 \preceq Y_1 + Y_2$ 。

对于任何 c , 我们只需证明 $Y_1 + X_2 \preceq Y_1 + Y_2$, 则根据对称性得证。

$$\begin{aligned}
\Pr[Y_1 + Y_2 \geq c] &= \sum_{y_1} \Pr[Y_1 = y_1] \Pr[Y_2 \geq c - y_1] \\
&\geq \sum_{y_1} \Pr[Y_1 = y_1] \Pr[X_2 \geq c - y_1] \\
&= \Pr[Y_1 + X_2 \geq c]
\end{aligned}$$

Corollary: 如果函数列 $\{g_j\}_{j=1}^m$ 和 $\{f_j\}_{j=1}^m$ 满足 $f_j(\cdot; x_1, \dots, x_{i-1}) \preceq g_j(\cdot)$, 则

$$\int_{\sum x_j \geq c} f_1(x_1) \cdots f_m(x_m; x_1, \dots, x_{m-1}) dx \leq \int_{\sum x_j \geq c} g_1(x_1) \cdots g_m(x_m) dx$$

归纳法, 先固定 x_1, \dots, x_{m-1} , 将 $f_m(\cdot; x_1, \dots, x_{m-1})$ 替换为 $g(\cdot)$, 然后重复上述过程。

Power of 2 Choices (1)

将 m 个球独立放入 n 个桶中, 每个球随机选择两个桶, 放入负载较小的那个桶。

Theorem: $m = n$ 时, 最大负载量 w.h.p 不超过 $\frac{\ln \ln n}{\ln 2} + \Theta(1)$ 。

证明的大体思路是, 设 B_i 为负载量 $\geq i$ 的桶的个数。我们试图找到一系列 bound β_i , 使得 w.h.p $B_i \leq \beta_i$, 则对于任何一个特定的球, 其落在负载 $\geq i$ 的桶的概率 $\leq \left(\frac{\beta_i}{n}\right)^2$ 。从而 $B_{i+1} \leq \mathcal{B}(n, (\beta_i/n)^2)$, 均值为 β_i^2/n , 可以根据 Chernoff bound 取 $\beta_{i+1} = c\beta_i^2/n$, 于是有 $\frac{\beta_{i+1}}{n} = c\left(\frac{\beta_i}{n}\right)^2$, 即 β_i/n 平方速度下降, 当 $i \approx \frac{\ln \ln n}{\ln 2}$ 时有 $\beta_i < 1$, 这便是最大负载量。

Lecture 13 - 2025 / 3 / 31

Power of 2 Choices (2)

Theorem: 将 n 个球独立放入 n 个桶中, 每个球随机选择两个桶, 放入负载较小的那个桶, 最大负载量 w.h.p 不超过 $\frac{\ln \ln n}{\ln 2} + \Theta(1)$ 。

分两个阶段对该定理进行证明。不妨设 $\beta_6 = \frac{n}{2e}$, 则 $B_6 \leq \beta_6$ 是 trivial 的, 因为 ≥ 6 的桶的个数不超过 $\frac{n}{6} < \frac{n}{2e}$ 。对于 $i > 6$, 定义 $\beta_{i+1} = \frac{e\beta_i^2}{n}$ 。

Claim: 对于任意 $i > 6$, $\beta_i^2 \geq 2n \ln n$ 时, 有 $\Pr[B_i > \beta_i] \leq \frac{i}{n^2}$ 。

归纳法。显然有 $\Pr[B_{i+1} > \beta_{i+1}] \leq \Pr[B_{i+1} > \beta_{i+1}, B_i \leq \beta_i] + \Pr[B_i > \beta_i]$ 。后项根据归纳假设 $\leq \frac{i}{n^2}$ 。

接下来试图说明前一项不超过 $\Pr[\mathcal{B}(n, (\beta_i/n)^2) > \beta_{i+1}]$ ，从而根据 Chernoff bound $\Pr[X \geq e\mu] \leq e^{-\mu}$ 得知不超过 $\exp(-\beta_i^2/n) \leq \frac{1}{n^2}$ ，于是即可证毕。

定义 $B_i^{(j)}$ 代表当我们放置第 j 个球之前，负载量 $\geq i$ 的桶的个数。用 X_j 作为第 j 个球的高度是否 $\geq i+1$ 的 indicator。则易见 $B_{i+1} \leq \sum X_j$ 。

$$\Pr[B_{i+1} > \beta_{i+1}, B_i \leq \beta_i] \leq \Pr\left[\sum_{j=1}^n X_j > \beta_{i+1}, B_i \leq \beta_i\right]$$

从而将右侧写作 $\sum_{\sum x_j > \beta_{i+1}} \Pr[X_1 = x_1, \dots, X_m = x_m, B_i \leq \beta_i]$ ，对于其中每一项

$$\begin{aligned} & \Pr[X_1 = x_1, \dots, X_m = x_m, B_i \leq \beta_i] \\ &= \Pr[X_1 = x_1, \dots, X_m = x_m, B_i^{(1)} \leq \beta_i, \dots, B_i^{(m)} \leq \beta_i] \\ &\leq \Pr[X_1 = x_1, B_i^{(1)} \leq \beta_i] \cdots \Pr[X_m = x_m, B_i^{(m)} \leq \beta_i \mid X_j = x_j, B_i^{(j)} \leq \beta_i] \\ &= f_1(x_1) \cdots f_m(x_m; x_1, \dots, x_{m-1}) \end{aligned}$$

其中 $f_j(x_j; x_1, \dots, x_{j-1}) = \Pr[X_j = x_j, B_i^{(j)} \leq \beta_i \mid X_1 = x_1, B_i^{(1)} \leq \beta_i, \dots, X_{j-1} = x_{j-1}, B_i^{(j-1)} \leq \beta_i]$ 。下面说明 $f_j(1, x_1, \dots, x_{j-1}) \leq (\beta_i/n)^2$ 。

实际上这是因为，我们可以通过全概率公式枚举 $B_i^{(j)}$ 的值，而 condition on $B_i^{(j)}$ 的值后， $\Pr[X_j = x_j]$ 是完全与 x_1, \dots, x_{j-1} 的情况无关的。也就是说，

$$\begin{aligned} f_j(1, x_1, \dots, x_{j-1}) &= \sum_{b_i^{(j)} \leq \beta_i} \Pr[X_j = 1 \mid B_i^{(j)} = b_i^{(j)}] \Pr[B_i^{(j)} = b_i^{(j)} \mid \dots] \\ &\leq \left(\frac{\beta_i}{n}\right)^2 \sum_{b_i^{(j)} \leq \beta_i} \Pr[B_i^{(j)} = b_i^{(j)} \mid \dots] \\ &= \left(\frac{\beta_i}{n}\right)^2 \end{aligned}$$

至此，我们证明了 $f_j(x_j \mid x_1, \dots, x_{j-1}) \leq \mathcal{B}(1, (\beta_i/n)^2)$ 。根据 stochastic dominance 的 Lemma 证毕。

设 i^* 是第一个满足 $\beta_i^2 < 2n \ln n$ 的 i ，则容易看出 $i^* = \frac{\ln \ln n}{\ln 2} + O(1)$ ，我们手动分析最后两个阶段。

Claim: $\Pr[B_{i^*+1} \geq 6 \ln n] = O(1/n)$

根据 $\Pr[B_{i^*+1} \geq 6 \ln n] \leq \Pr[B_{i^*+1} \geq 6 \ln n, B_{i^*} \leq \sqrt{2n \ln n}] + \Pr[B_{i^*} > \sqrt{2n \ln n}]$, 后面一项由前面的归纳法 $\leq 1/n$, 而前面一项使用前面类似的 Stochastic dominance 的技术 $\leq \Pr[\mathcal{B}(n, 2 \ln n/n) \geq 6 \ln n] \leq 1/n^2$ (Chernoff bound)。

Claim: $\Pr[B_{i^*+2} \geq 1] \leq O(\log^2 n/n)$ 。

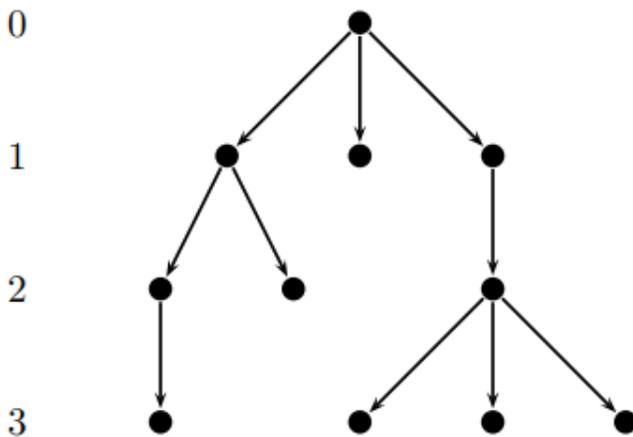
根据 $\Pr[B_{i^*+2} \geq 1] \leq \Pr[B_{i^*+2} \geq 1, B_{i^*+1} \leq 6 \ln n] + \Pr[B_{i^*+1} > 6 \ln n]$, 后面一项由前面的 Claim $\leq O(1/n)$, 而前面一项使用 Stochastic dominance 的技术 $\leq \Pr[\mathcal{B}(n, (6 \ln n/n)^2) \geq 1] \leq (6 \ln n)^2/n$ (union bound)。

事实上, 最大负载量的下界也是 w.h.p $\Omega(\ln \ln n)$ 的。如果每次选择 d 个桶, 则最大负载量 w.h.p 是 $\frac{\ln \ln n}{\ln d} + O(1)$ 。

Galton-Watson Branching Process

设 X 是一个非负整数 r.v., X 定义的分支过程从时间 0 的一个单点开始, 每次分支生成 $x \sim X$ 个儿子, 并对每个儿子分别独立进行下去。

Time



用 Z_i 代表时间 i 的结点数量, 则 $Z_0 = 1$, 将灭绝的概率定义为

$$\Pr[\text{extinction}] = \lim_{n \rightarrow \infty} \Pr[Z_n = 0]$$

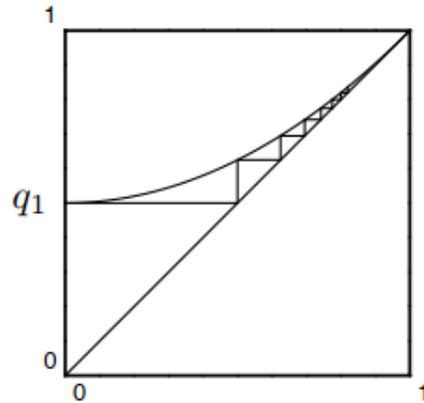
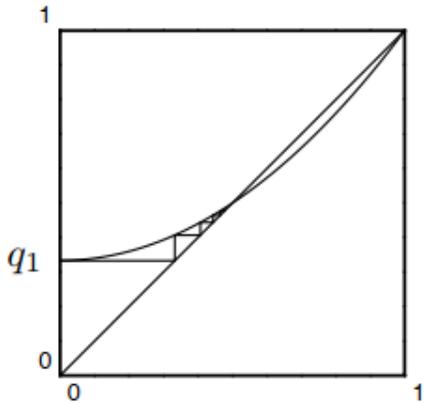
Theorem: 对于一个 X 定义的分支过程, $\Pr[X = 1] < 1, \Pr[X = 0] > 0$, 有

- 如果 $\mathbb{E}[X] \leq 1$ 则 $\lim_{n \rightarrow \infty} \Pr[Z_n = 0] = 1$
- 如果 $\mathbb{E}[X] > 1$ 则 $\lim_{n \rightarrow \infty} \Pr[Z_n = 0] = p^* < 1$, 其中 p^* 是 $(0, 1)$ 之间的方程 $f(x) = x$ 的唯一解,

$$f(x) = \sum_{i \geq 0} \Pr[X = i]x^i$$

设 q_n 为时间 n 灭绝的概率，即 $q_n = \Pr[Z_n = 0]$ ，其中 $q_0 = 0$ ，我们可以针对第 1 步分裂情况进行讨论，从而列出递推方程 $q_n = f(q_{n-1})$ 。

根据实际含义容易看出 $0 < q_1 \leq q_2 \leq q_3 \leq \dots \leq 1$ ，也就是 (q_n) 单调递增且有界，故必然收敛到 $q^* \leq 1$ 。



注意到 $f(x)$ 是在 $[0, 1]$ 内的严格递增函数，且严格凸的函数，我们针对 $y = f(x)$ 和 $y = x$ 的关系进行讨论。注意 $\mathbb{E}[X] = f'(1)$ 。

- 对于第二种情况，因为 $\mathbb{E}[X] = f'(1) > 1$ ，所以 $y = f(x)$ 和 $y = x$ 第一次相交于 $a < 1$ ，根据左图 $q^* = a < 1$ 。
- 对于第一种情况，因为 $\mathbb{E}[X] = f'(1) \leq 1$ ，所以 $y = f(x)$ 和 $y = x$ 第一次相交于 1，根据右图 $q^* = 1$ 。

Lecture 14 - 2025 / 4 / 7

Giant Component (1)

Theorem: 对于 $G \in \mathcal{G}_{n,p}$ ，其中 $p = \frac{c}{n}$ ， $c < 1$ 是一个常数，则 a.a.s. G 的最大的连通分支大小是 $O(\log n)$ 的。

对于一个结点 v ，通过 BFS 找出 v 所在的连通块大小的过程，可以看作从 v 开始的一个 branching process。

- 即根节点为 v ，为 v 采样 $\mathcal{B}(n-1, p)$ 个邻居（儿子）结点，假设这里是 2 个 v_1, v_2 。
- 为 v_1 采样 $\mathcal{B}(n-3, p)$ 个邻居（儿子结点），即忽略掉 v, v_1, v_2 的影响，假设是 3 个。
- 为 v_2 采样 $\mathcal{B}(n-6, p)$ 个邻居（儿子结点），即忽略掉所有上述已知连通的点的影响……

从而“ v 在一个大小为 k 的连通块”即“branching process 可以展出 k 个结点”的概率。注意这里每次针对一个结点展开，而不是针对一层展开。

我们将上述每一步展开放缩为 $\mathcal{B}(n, p)$ ，这给出了一个上界。进而上述概率不低于“ k 次采样 $\mathcal{B}(n, p)$ 之和不低于 $k - 1$ 的概率”。

$$1 + \mathcal{B}(n, p) + \cdots + \mathcal{B}(n, p) \geq k$$

我们基于这一点给出一个 upper bound。

设 $X_i \sim \mathcal{B}(n, c/n)$ i.i.d $i = 1, 2, \dots, k$ ，则

$$\Pr \left[\sum_{i=1}^k X_i \geq (k-1) \right] = \Pr \left[\sum_{i=1}^k X_i \geq ck + (1-c)k - 1 \right]$$

注意 $\mu = ck, \beta = \frac{(1-c)k-1}{ck} = \Theta(1)$ ，根据 Chernoff bound

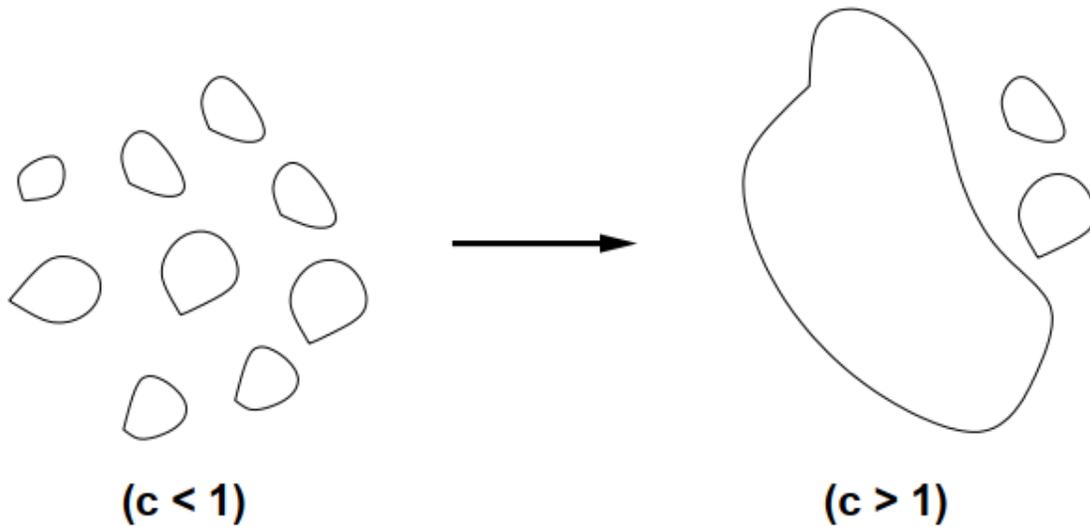
$$\begin{aligned} \Pr \left[\sum_{i=1}^k X_i \geq (k-1) \right] &\leq \exp \left(-\frac{((1-c)k-1)^2}{c^2 k^2 (2 + ((1-c)k-1)/ck)} ck \right) \\ &= \exp \left(-\frac{((1-c)k-1)^2}{((c+1)k-1)} \right) \\ &= \exp \left(-\frac{(1-c)^2}{c+1} k + O(1) \right) \end{aligned}$$

从而取 $k = 2 \cdot \frac{(1+c)}{(1-c)^2} \ln n$ ，则有上述概率 $\leq O(n^{-2})$ 。对所有 n 个初始的 v union bound，得到原命题 w.p. $1 - O(n^{-1})$ 成立。

Lecture 15 - 2025 / 4 / 10

Giant Component (2)

Theorem: 对于 $G \in \mathcal{G}_{n,p}$ ，其中 $p = \frac{c}{n}$ ， $c > 1$ 是一个常数，则 a.a.s. G 存在唯一一个最大的连通分支大小是 $\beta n(1 + o(1))$ ，其中 β 是 $(0, 1)$ 之间 $\beta + e^{-\beta c} = 1$ 的唯一解。其余的连通块大小都是 $O(\log n)$ 级别。



Claim: 对于所有结点 v , a.a.s 以下两者之一成立:

1. 从 v 开始的 branching process 在 k^- 步内停止。
2. $\forall k$ s.t. $k^- \leq k \leq k^+$, 从 v 开始的 branching process 在 k 步后, 至少有 $(c-1)k/2$ 个已探索但是没有饱和的结点。

对于后者, 实际上只需要证明从 v 开始总共至少探索到了

$$\frac{(c-1)k}{2} + k = \frac{(c+1)k}{2}$$

个点。我们定义一个点 v 是 k -bad 的, 如果从 v 开始的 branching process 在 k 步后停止或者探索到了少于 $(c+1)k/2$ 个点。

因此, 当 v 是 k -bad 时, 从 v 开始的 branching process 被每次展开服从 $\mathcal{B}\left(n - \frac{(c+1)k^+}{2}, \frac{c}{n}\right)$ 的过程支配 (因为总共涉及到的点数不超过 $(c+1)k^+/2$, 因此 $\mathcal{B}(n-?, p)$ 的 $?$ 处不高于这个值)。

进而从 v 开始的 branching process 在 k 步内展开的点数, 不低于 k 次采样 $\mathcal{B}\left(n - \frac{(c+1)k^+}{2}, \frac{c}{n}\right)$ 展开的点数。

「上面这一步并没有理解, 如果 branching process 提前终止了, 为什么还能 dominate 固定次数采样的求和? 」

从而 a.a.s 从任何一个点 v 开始的 branching process 要么在 $k^- = O(\log n)$ 轮终止, 要么持续至少 $k^+ = n^{2/3}$ 轮。记前面的一类点是 *small* 的, 后面的一类是 *large* 的。

Lemma: a.a.s. 存在唯一的一个连通块, 包含了所有 large 点。

考虑两个 large 的 $u \neq v$ 。分别从 u, v 独立进行 branching process, 则在 k^+ 轮后, 两者已探索未饱和的点分别记作 $U(u), U(v)$, 则这两个集合大小都 $\geq \frac{c-1}{2}k^+$ 。

如果前 k^+ 步已经遇到公共点了, 则 u, v 已连通。否则我们证明 w.h.p. $U(u), U(v)$ 之间有边。

$$\begin{aligned} \Pr[\nexists \text{edge between } U(u), U(v)] &\leq (1-p)^{\left(\frac{c-1}{2}k^+\right)^2} \\ &\leq \exp\left(-p\left(\frac{c-1}{2}k^+\right)^2\right) \\ &\leq \exp\left(-\frac{c(c-1)^2}{4}n^{1/3}\right) \\ &= o(n^{-2}) \end{aligned}$$

从而对所有 u, v 进行 union bound 立刻得到总概率是 $o(1)$ 。

至此已经证明了最大连通块的唯一性, 以及所有小连通块都是 $O(\log n)$, 只剩下判断最大连通块的大小了。我们通过对 small 点计数来证明此。

Lemma: a.a.s. small 点的个数是 $(1+o(1))(1-\beta)n$ 。

根据 small 点的定义, 可以知道 $\Pr[v \text{ is small}]$:

- (\geq) 服从 $\mathcal{B}(n, c/n)$ 的 branching process 在 k^- 步内终止的概率。
这是因为利用 $\mathcal{B}(n - \dots, p) \leq \mathcal{B}(n, p)$, 展出的点变多, 终止概率变低。
- (\leq) 服从 $\mathcal{B}(n - k^-, c/n)$ 的 branching process 在 k^- 步内终止的概率。
这是因为 small 的点总共展出了 $\leq k^-$ 个点, 所以 $\mathcal{B}(n - \dots, p) \geq \mathcal{B}(n - k^-, p)$, 展出的点变少, 终止概率增大。

更进一步, 用 $d(n, p)$ 表示服从 $\mathcal{B}(n, p)$ 的 branching process 终止的概率:

- (\geq) 根据 claim, 我们知道 w.h.p. 如果不在 k^- 步终止, 则最终不会终止, 故下界为 $d(n, c/n) + o(1)$ 。
- (\leq) 不限制终止步数, 终止概率自然增大, 故上界为 $d(n - k^-, c/n)$ 。

当 $n \rightarrow +\infty$ 时, 根据泊松分布的结论, $d(n, c/n) \rightarrow 1 - \beta$, 其中 β 是 $(0, 1)$ 之间 $\beta + e^{-\beta c} = 1$ 的解。同时因为 $k^- \ll n$, 所以 $d(n - k^-, c/n) \rightarrow 1 - \beta$ 。根据 sandwiching 定理, 可以知道

$$\Pr[v \text{ is small}] \rightarrow 1 - \beta =: \alpha$$

用 $Z = \sum_v Z_v$ 代表 small 点的个数, 我们通过 Chebyshev 给 Z 一个 concentration bound。则 $\mathbb{E}[Z_v] \rightarrow \alpha, \mathbb{E}[Z] = (1+o(1))\alpha n$ 。

$$\begin{aligned}\mathbb{E}[Z^2] &= \mathbb{E}[Z] + \sum_{u \neq v} \mathbb{E}[Z_u Z_v] \\ &= \mathbb{E}[Z] + \sum_v \Pr[v \text{ is small}] \sum_{u \neq v} \Pr[u \text{ is small} \mid v \text{ is small}]\end{aligned}$$

对于最后一个 \sum ，可以拆分为 u 和 v 在同一连通块、 u 和 v 在不同连通块的两类分别计数。

- 和 v 在同一连通块的 u 不超过 k^- 个
- 和 v 在不同连通块的任何一个 u 满足

$$\begin{aligned}& \Pr[u \text{ is small} \mid v \text{ is small}] \\ &= \Pr[u \text{ is small in } \mathcal{G}(n - |\text{Comp}(v)|, p)] \\ &\leq \Pr[u \text{ is small in } \mathcal{G}(n - k^-, p)] \\ &\leq d(n - k^-, c/n) \sim d(n, c/n) \rightarrow \alpha\end{aligned}$$

从而 $\mathbb{E}[Z^2] \leq \mathbb{E}[Z] + n(\alpha + o(1))(k^- + n(\alpha + o(1))) \sim \mathbb{E}[Z] + n^2\alpha^2(1 + o(1)) = \mathbb{E}[Z]^2(1 + o(1))$ 。

从而根据 Chebyshev 不等式

$$\Pr[|Z - \mathbb{E}[Z]| > \gamma \mathbb{E}[Z]] \leq \frac{1}{\gamma^2} \left(\frac{\mathbb{E}[Z^2]}{\mathbb{E}[Z]^2} - 1 \right) = \frac{1}{\gamma^2} o(1)$$

只需取 $\gamma = o(1)$ 但下降足够缓慢，则上式昭示了 a.a.s. 最大连通分支大小是 $(1 + o(1))\beta n$ 。

综合以上两个 Lemma，原 Theorem 得证。

Lecture 16 - 2025 / 4 / 14

Johnson & Lindenstrauss Lemma

Theorem (JL Lemma). 对于任何 \mathbb{R}^d 上 n 个点的集合 X ，任何 $\varepsilon \in (0, 1)$ ，存在一个 $\mathbb{R}^d \rightarrow \mathbb{R}^k$ 的映射 φ ，其中

$$k = \left\lceil \frac{4 \ln n}{\varepsilon^2/2 - \varepsilon^3/3} \right\rceil \leq \left\lceil \frac{24 \ln n}{\varepsilon^2} \right\rceil$$

使得 $\forall u, v \in X$,

$$(1 - \varepsilon)\|u - v\|_2^2 \leq \|\varphi(u) - \varphi(v)\|_2^2 \leq (1 + \varepsilon)\|u - v\|_2^2$$

考虑随机选择一个坐标系，并保留 u 在其中的前 k 个坐标（的一个倍数）作为 $\varphi(u)$ 。为了分析这个过程，我们可以对称的看作，对于个固定的标准正交坐标系， u 在 \mathbb{S}^{d-1} 上均匀随机采样。

于是我们生成一个随机向量 $X = (X_1, \dots, X_d)$, 其中 $X_i \sim \mathcal{N}(0, 1)$, 可以将 u 表示为 $Z = \frac{1}{\|X\|_2}(X_1, \dots, X_d)$, 降维后的向量定义为 $Y = \varphi(X) = \sqrt{\frac{k}{d}} \cdot \frac{1}{\|X\|_2}(X_1, \dots, X_k)$ 。

需要分析 $L = \frac{X_1^2 + \dots + X_k^2}{X_1^2 + \dots + X_d^2}$ 的分布。根据对称性, 显然有 $\mathbb{E}[L] = k/d$, 于是 $\mathbb{E}[\|Y\|_2^2] = 1$ 。

根据 Chernoff bound 可以得到

- $\Pr[\|\varphi(u)\|_2^2 \geq (1 + \varepsilon)] \leq \exp(-\frac{k}{2}(\frac{\varepsilon^2}{2} - \frac{\varepsilon^3}{3}))$
- $\Pr[\|\varphi(u)\|_2^2 \leq (1 - \varepsilon)] \leq \exp(-\frac{k}{4}\varepsilon^2)$

证明过程主要利用了 $\ln(1 - \varepsilon) < (-\varepsilon - \frac{\varepsilon^2}{2})$ 和 $\ln(1 + \varepsilon) < (\varepsilon - \frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3})$ 。

于是, 当 k 满足条件时, $\Pr[|\|\varphi(u)\|_2^2 - 1| > \varepsilon] \leq 2 \exp(-2 \ln n) = 2/n^2$ 。从而根据 union bound, 对于所有 $\binom{n}{2}$ 个点 (u, v) , 都保距的概率 $\geq \frac{1}{n}$ 。根据 probabilistic method, 可以得到 JL 引理。

Embedding into ℓ_p metrics

Theorem. 设 (X, d) 是一个度量空间, $|X| = n$, 则 (X, d) 可以被嵌入一个 ℓ_1 空间, 保距比为 $O(\log n)$, 维度 $k = O(\log^2 n)$ 。

我们通过构造 $m = O(\log^2 n)$ 个随机的 $A_i \subseteq X$, 并定义

$$\varphi(x) = \frac{1}{m}(d(x, A_1), d(x, A_2), \dots, d(x, A_m))$$

其中 $d(x, A_i) = \min_{y \in A_i} d(x, y)$ 。我们从两个方向分别证明这个构造的合理性。

Claim. $\forall x, y \in X, \|\varphi(x) - \varphi(y)\|_1 \leq d(x, y)$

$$\begin{aligned} \|\varphi(x) - \varphi(y)\|_1 &= \frac{1}{m} \sum_{i=1}^m |d(x, A_i) - d(y, A_i)| \\ &\leq \frac{1}{m} \sum_{i=1}^m d(x, y) = d(x, y) \end{aligned}$$

上式中, 第二个不等式是因为, 不妨设 $d(x, A_i) \geq d(y, A_i)$, 设 $d(y, A_i) = d(y, z)$, 其中 $z \in A_i$, 则有 $d(x, A_i) - d(y, A_i) \leq d(x, z) - d(y, z) \leq d(x, y)$ 。

我们构造 $\{A_i\}$ 的方法是, 对于每个 $t \in \{1, 2, \dots, \log n\}$, 构造 $r \log n$ 个随机集合 $\{A_i^{(t)}\}_{i=1}^{r \log n}$, 其中每个 $x \in X$ 都独立均匀的以 2^{-t} 的概率包含在 $A_i^{(t)}$ 中。因此 $A_i^{(t)}$ 的期望大小为 $\frac{n}{2^t}$, 总共有 $r \log^2 n$ 个集合。

Claim. $\exists c, \forall x, y \in X, \|\varphi(x) - \varphi(y)\|_1 \geq \frac{1}{c \log n} d(x, y)$

为了证明这个 claim, 我们首先定义“球”:

$$\begin{aligned} B(x, \rho) &= \{z \in X \mid d(x, z) \leq \rho\} \\ B^\circ(x, \rho) &= \{z \in X \mid d(x, z) < \rho\} \end{aligned}$$

定义一系列半径 $0 = \rho_0 < \rho_1 < \dots$, 其中 ρ_t 定义为

$$\rho_t = \min\{\rho \mid B(x, \rho), B(y, \rho) \text{ both contain } \geq 2^t \text{ points of } X\}$$

持续定义这样的 ρ_t , 直到某一项 $\rho_{t^*} \geq \frac{1}{4}d(x, y)$ 时, 修改定义这一项为 $\rho_{t^*} = \frac{1}{4}d(x, y)$, 定义结束。可以看出 $B(x, \rho_t), B(y, \rho_t)$ 永远是不交的。

我们称 $A_i^{(t)}$ 是 good 的当且仅当 (两者之一):

- ρ_t 对于 $B(x, \rho_t)$ 是紧的, 而 $A_i^{(t)}$ 与 $B(y, \rho_{t-1})$ 相交但与 $B^\circ(x, \rho_t)$ 不交。
- ρ_t 对于 $B(y, \rho_t)$ 是紧的, 而 $A_i^{(t)}$ 与 $B(x, \rho_{t-1})$ 相交但与 $B^\circ(y, \rho_t)$ 不交。

注意, 一个 good 的集合将为 $\|\varphi(x) - \varphi(y)\|_1$ 贡献 $\frac{1}{m}(\rho_t - \rho_{t-1})$ 。

对于任何集合 $A_i^{(t)}$, 它 good 的概率有

$$\begin{aligned} \Pr[A_i^{(t)} \text{ is good for } x, y] &= \Pr[A_i^{(t)} \cap B^\circ(x, \rho_t) = \emptyset \wedge A_i^{(t)} \cap B(y, \rho_{t-1}) \neq \emptyset] \\ &\geq \Pr[A_i^{(t)} \cap B^\circ(x, \rho_t) = \emptyset] \cdot \Pr[A_i^{(t)} \cap B(y, \rho_{t-1}) \neq \emptyset] \\ &\geq (1 - 2^{-t})^{2^t} \cdot (1 - (1 - 2^{-t})^{2^{t-1}}) \\ &\geq \frac{1}{4} \cdot \left(1 - \frac{1}{\sqrt{e}}\right) \end{aligned}$$

第一个不等号是因为两个事件是正相关的, 最后一个不等号是因为前者单调递增, 后者单调递减。

因此 $A_i^{(t)}$ 以常数概率是 good 的, 对于每个固定的 t , $\mathbb{E}[\#\text{good sets}] \geq \frac{r \log n}{12} = \mu$, 根据 Chernoff bound, $\Pr[\#\text{good sets} \leq \mu/2] \leq \exp(-\mu/8) = \exp(-r \log n/96) \leq n^{-3}$, 这里取 $r = 288$ 。从而根据 union bound, 对于所有的 x, y, t 都成立的概率 $\geq 1 - \log n/n_0$ 。

因此, 当上述事件发生时,

$$\begin{aligned}
\|\varphi(x) - \varphi(y)\|_1 &= \frac{1}{m} \sum_{t=1}^{\log n} \sum_{i=1}^{r \log n} |d(x, A_i^{(t)}) - d(y, A_i^{(t)})| \\
&\geq \frac{1}{m} \frac{r \log n}{24} \sum_{t=1}^{\log n} (\rho_t - \rho_{t-1}) \\
&= \frac{1}{m} \frac{r \log n}{24} (\rho_{t^*} - \rho_0) \\
&= \frac{1}{96 \log n} d(x, y)
\end{aligned}$$

Lecture 17 - 2025 / 4 / 17

Martingale

Definition (filter): $\emptyset = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}_n$ 是一个概率空间上的递增 σ -代数。

例如 $\mathcal{F}_n = Z_1, \dots, Z_n$, 其中 Z_i 是随机变量。

Definition (martingale): (X_i) 是关于 (\mathcal{F}_i) 的鞅, 如果满足

$$\mathbb{E}[X_i | \mathcal{F}_{i-1}] = X_{i-1}$$

Azuma Inequality

Lemma: 对于 r.v. X , 若 $|X| \leq 1, \mathbb{E}[X] = 0$, 则 $\mathbb{E}[e^{tX}] \leq e^{t^2/2}$ 。

根据凸性和 Taylor 展开, $\mathbb{E}[e^{tX}] \leq \frac{1}{2}(e^t + e^{-t}) \leq e^{t^2/2}$ 。

Theorem: 设 (X_i) 是关于 (\mathcal{F}_i) 的鞅, $Y_i = X_i - X_{i-1}$ 是“差异”序列, 如果 $c_i > 0$ 使得 $|Y_i| \leq c_i$, 则

$$\begin{aligned}
\Pr[X_n \geq X_0 + \lambda] \\
\Pr[X_n \leq X_0 - \lambda] &\leq \exp\left(-\frac{\lambda^2}{2 \sum_{i=1}^n c_i^2}\right)
\end{aligned}$$

当 $n = 1$ 时, $|X_1 - X_0| \leq c_1$, 则

$$\begin{aligned}
\Pr[X_1 \geq X_0 + \lambda] &= \min_t \Pr[e^{t(X_1 - X_0)} \geq e^{t\lambda}] \\
&\leq \min_t \frac{\mathbb{E}[e^{t(X_1 - X_0)}]}{e^{t\lambda}} \\
&\leq \min_t \exp\left(\frac{c_1^2 t^2}{2} - t\lambda\right) = \exp\left(-\frac{\lambda^2}{2c_1^2}\right)
\end{aligned}$$

接下来归纳,

$$\begin{aligned}
\Pr[X_n \geq X_0 + \lambda] &\leq \min_t \frac{\mathbb{E}[e^{t(X_n - X_{n-1})} \cdot e^{t(X_{n-1} - X_0)}]}{e^{t\lambda}} \\
&= \min_t \frac{\mathbb{E}_{\mathcal{F}_{n-1}}[\mathbb{E}[e^{t(X_n - X_{n-1})} \mid \mathcal{F}_{n-1}] \cdot e^{t(X_{n-1} - X_0)}]}{e^{t\lambda}} \\
&\leq \min_t \frac{e^{c_n^2 t^2 / 2} \cdot \mathbb{E}[e^{t(X_{n-1} - X_0)}]}{e^{t\lambda}} \\
&\leq \min_t \frac{e^{c_n^2 t^2 / 2} \cdot \exp(-\lambda^2 / 2 \sum_{i=1}^{n-1} c_i^2)}{e^{t\lambda}} \\
&\leq \exp\left(-\frac{\lambda^2}{2 \sum_{i=1}^n c_i^2}\right)
\end{aligned}$$

Doob Martingale

Claim: 设 $A, (Z_i)$ 是 r.v., 则 $X_i = \mathbb{E}[A \mid Z_1, \dots, Z_i]$ 是鞅, 称之为 A 的 Doob 鞅。

验证定义即可:

$$\begin{aligned}
\mathbb{E}[X_i \mid Z_1, \dots, Z_{i-1}] &= \mathbb{E}_{Z_i}[\mathbb{E}[X_i \mid Z_1, \dots, Z_i] \mid Z_1, \dots, Z_{i-1}] \\
&= \mathbb{E}_{Z_i}[\mathbb{E}[A \mid Z_1, \dots, Z_i] \mid Z_1, \dots, Z_{i-1}] \\
&= \mathbb{E}[A \mid Z_1, \dots, Z_{i-1}] = X_{i-1}
\end{aligned}$$

Definition: $f(Z_1, \dots, Z_n)$ 是 c -Lipschitz 函数, 当且仅当改变 f 的任何一个坐标值, f 的变化绝对值不超过 $\pm c$ 。

Lemma: 如果 f 是 c -Lipschitz 函数, 给定 Z_1, \dots, Z_{i-1} 的条件下, Z_i 与 Z_{i+1}, \dots, Z_n 相互独立, 则 f 关于 Z_i 的 Doob 鞅 (X_i) 满足 $|X_i - X_{i-1}| \leq c$ 。

我们根据定义对 $|X_i - X_{i-1}|$ 进行展开

$$\begin{aligned}
&= |\mathbb{E}_{Z_{i+1}, \dots, Z_n}[f \mid Z_1, \dots, Z_i] - \mathbb{E}_{Z_i, \dots, Z_n}[f \mid Z_1, \dots, Z_{i-1}]| \\
&= |\mathbb{E}_{Z_{i+1}, \dots, Z_n}[f \mid Z_1, \dots, Z_i] - \mathbb{E}_{Z_{i+1}, \dots, Z_n}[\mathbb{E}_{Z_i}[f \mid Z_1, \dots, Z_{i-1}, Z_{i+1}, \dots, Z_n] \mid Z_1, \dots, Z_{i-1}]| \\
&= |\mathbb{E}_{Z_{i+1}, \dots, Z_n}[f(Z_1, \dots, Z_i, \dots, Z_n) \mid Z_1, \dots, Z_{i-1}] \\
&\quad - \mathbb{E}_{Z_{i+1}, \dots, Z_n}[\mathbb{E}_{Z_i}[f(Z_1, \dots, Z_n) \mid Z_1, \dots, Z_{i-1}, Z_{i+1}, \dots, Z_n] \mid Z_1, \dots, Z_{i-1}]| \\
&= |\mathbb{E}_{Z_{i+1}, \dots, Z_n}[f(Z_1, \dots, Z_i, \dots, Z_n) \\
&\quad - \mathbb{E}_{Z_i}[f(Z_1, \dots, Z_i, \dots, Z_n) \mid Z_1, \dots, Z_{i-1}, Z_{i+1}, \dots, Z_n] \mid Z_1, \dots, Z_{i-1}]| \\
&= |\mathbb{E}_{Z_{i+1}, \dots, Z_n}[\mathbb{E}_{Z_i}[f(Z_1, \dots, Z_i, \dots, Z_n) - f(Z_1, \dots, Z_i, \dots, Z_n) \\
&\quad \mid Z_1, \dots, Z_{i-1}, Z_{i+1}, \dots, Z_n] \mid Z_1, \dots, Z_{i-1}]|
\end{aligned}$$

注意这里 Z_i 是已知量, 而 Z_i 是未知量, 可以看作两者是独立同分布的变量。从而每一项均 $\leq c$, 由此结论成立。

Applications: Balls and Bins

m 个球 n 个桶, Z_i 是 i 号球选择的桶, $X = f(Z_1, \dots, Z_m)$ 是空桶的个数。容易看出 f 是 1-Lipschitz 的, 从而

$$\Pr[|X - \mathbb{E}[X]| \geq \lambda] \leq 2 \exp\left(-\frac{\lambda^2}{2m}\right)$$

这是 Chernoff bound 所不能得到的结论。

Applications: Chromatic Number of $\mathcal{G}_{n,1/2}$

染色数 $\chi(G)$ 代表最少需要的颜色数量, 使得存在一组同色不相邻的方案。

对于随机图我们有两种常见的鞅。

Edge Exposure Martingale: 用 $Z_i = 0/1$ 表示第 i 条边是否在图中出现, 则 $A = f\left(Z_1, \dots, Z_{\binom{n}{2}}\right)$ 的

Doob 鞅是 edge exposure martingale。

Vertex Exposure Martingale: 用 $Z_i \in \{0, 1\}^{n-i}$ 代表是否 i 和 j (满足 $j > i$) 的边是存在的, 则 $A = f(Z_1, \dots, Z_n)$ 的 Doob 鞅是 vertex exposure martingale。

这里我们使用后者, 用 $X = f(Z_1, \dots, Z_n)$ 代表 $\chi(G)$, 则容易看出 f 是 1-Lipschitz 的。从而

$$\Pr[|X - \mathbb{E}[X]| \geq \lambda] \leq 2 \exp\left(-\frac{\lambda^2}{2n}\right)$$

注意我们不依赖任何关于 $\mathbb{E}[X]$ 的知识, 给出了一个 concentration bound。

Lecture 18 - 2025 / 4 / 21

Quick Sort

考虑随机版本的快速排序算法

```
def QuickSort(a : list[int])
    x = random element in a
    a1 = [ y in a | y < x ]
    a2 = [ y in a | y > x ]
    QuickSort(a1)
    QuickSort(a2)
```

定义 Q_n 为对于大小为 n 的集合 S 进行快速排序所需要的比较次数, $q_n = \mathbb{E}[Q_n]$, 经典地, 有:

$$q_n = (n-1) + \frac{1}{n} \sum_{j=1}^n (q_{j-1} + q_{n-j})$$

$$q_n = 2n \ln n - (4 - 2\gamma)n + 2 \ln n + O(1)$$

其中 γ 是欧拉常数，现在考虑给 Q_n 一个 concentration bound。一个构造鞅的想法是记递归树上前 k 层的分割结果为 \mathcal{F}_k ，取 Q_n 关于 (\mathcal{F}_i) 的 Doob 鞅。但是以第 1 层划分为例，划分在最边上和最中间造成的差异远超常数级别。因此 $\mathbb{E}[Q_n | \mathcal{F}_k]$ 并不满足 Azuma inequality 的使用条件。

回归 Azuma inequality 的证明过程，我们需要给予 $\mathbb{E}[e^{t(X_k - X_{k-1})} | \mathcal{F}_{k-1}]$ 一个上界。假设 \mathcal{F}_{k-1} 中记录了第 $k-1$ 层时，各段长度为 L_1, L_2, \dots, L_m ，则各个段之间相互独立。定义 $T_j := \mathbb{E}[Q_{L_j} | \mathcal{F}_k^{(j)}] - \mathbb{E}[Q_{L_j}]$ ，则显然

$$|T_j| = |(L_j - 1 + q_{L_j'} + q_{L_j''}) - q_{L_j}| \leq L_j$$

上式在“最不平均”的分割时贴近取等，从而

$$\begin{aligned} \mathbb{E}[e^{t(X_k - X_{k-1})} | \mathcal{F}_{k-1}] &= \mathbb{E}[e^{t \sum_{j=1}^m T_j}] \\ &= \prod_{j=1}^m \mathbb{E}[e^{t T_j}] \\ &\leq \prod_{j=1}^m \exp\left(\frac{1}{2} t^2 L_j^2\right) \\ &\leq \exp\left(\frac{1}{2} t^2 (\max_{j=1}^m L_j) n\right) \quad (*) \end{aligned}$$

第一个 \leq 使用了和证明 Azuma 相同的 Lemma，第二个 \leq 把每一项的一个 L_j 放缩成了 $\max L_j$ 。

Lemma: $\forall 0 < \alpha < 1$ ，当 $k > \ln \frac{1}{\alpha}$ ，对于第 k 层的 L_1, L_2, \dots, L_m ，有

$$\Pr[\max_{j=1}^m L_j \geq \alpha n] \leq \alpha \left(\frac{2e \ln \frac{1}{\alpha}}{k}\right)^k$$

看作如下过程：第 1 层随机采样 $U_1 \sim U[0, 1]$ ，将长度为 n 的区间划分为长度为 $U_1 n$ 和 $(1 - U_1)n$ 的两段，然后第二层采样 $U_2, U_3 \sim U[0, 1]$ ，分别表示左、右区间的划分点，然后第三层再采样 $U_4, U_5, U_6, U_7 \sim U[0, 1] \dots$

第 k 层划分结束产生 2^k 个区间， L_j 的长度可以视作 $n \cdot U_1 \cdot U_{2/3} \cdots U_{2^{k-1}/\dots/2^{k-1}}$ ，上式即

$$\begin{aligned} \Pr \left[\left(\max_{j=1}^{2^k} \prod_{i=1}^k n \cdot U_i \cdots \right) \geq \alpha n \right] &\leq 2^k \cdot \Pr \left[\prod_{i=1}^k U_i \geq \alpha \right] \\ &\leq 2^k \cdot \Pr \left[\sum_{i=1}^k \ln U_i \geq \ln \alpha \right] \end{aligned}$$

注意到 $-\ln U_i \sim \text{Exp}(1)$, 从而 $-\sum_{i=1}^k \ln U_i \sim \Gamma(k, 1)$,

$$\begin{aligned} \Pr \left[\sum_{i=1}^k \ln U_i \geq \ln \alpha \right] &\leq \Pr_{X \sim \Gamma(k, 1)} [-X \geq \ln \alpha] \\ &= \min_{t > 0} \Pr_{X \sim \Gamma(k, 1)} [(\alpha e^X)^t \leq 1] \\ &= \min_{t > 0} \mathbb{E}_{X \sim \Gamma(k, 1)} [(\alpha e^X)^t] \\ &= \min_{t > 0} \alpha^t (1 - t)^{-k} \end{aligned}$$

当 $1 - t = k / \ln \frac{1}{\alpha}$ 时, 上式为 $\alpha \left(\frac{e \ln \frac{1}{\alpha}}{k} \right)^k$, 结合 union bound 给出的 2^k 原命题得证。

接下来我们分 3 个阶段分析快速排序过程:

1. 对于前 k_1 层, 比较次数不超过 $k_1 n$
2. 对于 $k_1 + 1 \sim k_2$ 层, 高概率有 k_1 层的 $\max L_j \leq \alpha n$ (1), 从而 (*) $\leq \exp \left(\frac{1}{2} t^2 \alpha n^2 \right)$
3. 对于 k_2 层, 高概率有 $\max L_j < 2$ (2), 从而算法停止。

Theorem: $\forall \varepsilon > 0, \Pr[|Q_n - q_n| \geq \varepsilon q_n] \leq n^{-(2+o(1))\varepsilon \ln \ln n}$

根据上述 Lemma, 事件 (1), (2) 均发生的概率 \geq

$$1 - \alpha \left(\frac{2e \ln \frac{1}{\alpha}}{k_1} \right)^{k_1} - \frac{2}{n} \left(\frac{2e \ln \frac{n}{2}}{k_2} \right)^{k_2}$$

假设这两个事件发生, 对于 $k_1 + 1 \sim k_2$ 层, 根据 (*), 类比于 Azuma inequality 得到

$$\Pr[|Q_n - q_n| \geq k_1 n + \lambda] \leq 2 \exp \left(-\frac{\lambda^2}{2(k_2 - k_1)\alpha n^2} \right)$$

只需取得 $k_1 n + \lambda \leq \varepsilon q_n$, 并让上述 3 个概率之和为 $n^{-(2+o(1))\varepsilon \ln \ln n}$ 时, 原命题即证毕。

接下来为琐碎的调参工作, 首先希望 k_2 尽量小, 取 $k_2 = (\ln n)(\ln \ln n)$, 则

$$\frac{2}{n} \left(\frac{2e \ln \frac{n}{2}}{k_2} \right)^{k_2} \sim \exp((\ln n)(\ln \ln n)(-\ln \ln \ln n))$$

接下来为了 k_1 尽量大, 但必须有 $k_1 \leq n^{-1}(\varepsilon q_n - \lambda) \sim 2\varepsilon \ln n - \frac{\lambda}{n}$, 这里希望 $2\varepsilon \ln n$ 是主导项, 需要

$\lambda = o(\varepsilon n \ln n)$ 。从而可以令 $k_1 = 2\varepsilon \ln n - \frac{2\lambda}{n}$

$$2 \exp \left(-\frac{\lambda^2}{2(k_2 - k_1)\alpha n^2} \right) \sim \exp \left(-\frac{\lambda^2}{(\ln n)(\ln \ln n)n^2 \alpha} \right) \quad (\text{A})$$

同时有（注意 $\alpha < 1$ ）

$$\alpha \left(\frac{2e \ln \frac{1}{\alpha}}{k_1} \right)^{k_1} \sim \exp \left(2\varepsilon \ln n \ln \ln \frac{1}{\alpha} \right) \quad (\text{B})$$

(A) 式希望 λ 尽可能大一些，故取 $\lambda = \frac{\varepsilon n \ln n}{\ln \ln n}$ ，(A) 式变为

$$\exp \left(\frac{\varepsilon^2 \ln n \ln \ln n}{\alpha} \right)$$

通过权衡两式，取 $\alpha = \frac{\varepsilon^2}{\ln \ln n}$ ，则 (A) 式为 $\exp(-\ln n (\ln \ln n)^2)$ ，(B) 式为

$$\exp(-2\varepsilon \ln n \ln \ln n + O(\ln \ln \ln n))$$

Corollary: $\forall \varepsilon > 0$, $\Pr[|Q_n - q_n| \geq \varepsilon q_n] = n^{-(2+o(1))\varepsilon \ln \ln n}$

Optional Stopping Theorem

Definition (Stopping time): (\mathcal{F}_i) 是一组 filter，一个 r.v. $T \in \{0, 1, \dots\} \cup \{\infty\}$ 是一个 (\mathcal{F}_i) 的停时如果事件 $T = i$ 是 \mathcal{F}_i -可测的。

Theorem (Optimal stopping theorem): (X_i) 是一个鞅， T 是一个关于 (\mathcal{F}_i) 的停时，则当下面条件成立时：

1. $\Pr[T < \infty] = 1$
2. $\mathbb{E}[|X_T|] < \infty$
3. $\mathbb{E}[X_i \cdot 1\{T > i\}] \rightarrow 0$ 当 $i \rightarrow \infty$ 时。

或者更强一些，满足：

1. $\mathbb{E}[T] < \infty$
2. $\mathbb{E}[|X_i - X_{i-1}| \mid \mathcal{F}_i] \leq c$ 对任意 i

则此时有 $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ 。

Gambler's Ruin

考虑从 0 处开始随机游走， $1/2$ 概率 $+1$ ， $1/2$ 概率 -1 。第一次到达 $-a$ 或 b 的时候停止。

定义 T 为上述停时，可以验证坐标位置 (X_i) 是一组鞅，并且满足停时定理的条件，则

$$\mathbb{E}[X_T] = p \cdot (-a) + (1 - p) \cdot b = 0$$

解出 $p = \frac{b}{a + b}$ ，即首先碰到 $-a$ 的概率。

接下来定义 $Y_i = X_i^2 - i$ 以分析 $\mathbb{E}[T]$ 。

Claim: (Y_i) 是一组关于 (X_i) 的鞅。

$$\mathbb{E}[Y_i | X_1, X_2, \dots, X_{i-1}] = \frac{1}{2} ((X_{i-1} - 1)^2 + (X_{i-1} + 1)^2) - i = X_{i-1}^2 - (i - 1) = Y_{i-1}$$

从而 $\mathbb{E}[Y_T] = \mathbb{E}[X_T^2] - \mathbb{E}[T] = \mathbb{E}[Y_0] = 0$, 即 $\mathbb{E}[T] = \mathbb{E}[X_T^2] = a^2 \frac{b}{a+b} + b^2 \frac{a}{a+b} = ab$ 。

Lecture 19 - 2025 / 4 / 24

Ballot

有两个竞选者 A, B, 分别收到 a, b 张票。假设选票按随机顺序计入, $a > b$, 则 A 的选票数量一直 $>$ B 的选票数量的概率是多少。

定义 S_k 为 k 轮后 A, B 的选票数量之差, 则 $S_n = a - b$ 。定义 $X_k = \frac{S_{n-k}}{n-k}$, 即倒过来看, $X_0 = \frac{a-b}{a+b}$ 。

Claim: (X_k) 是鞅。

在给定 X_{k-1} 的情况下, 此时 A, B 的选票数量 a', b' 满足 $X_{k-1} = \frac{a' - b'}{a' + b'}$ 。因此

$$\begin{aligned} \mathbb{E}[X_k | X_{k-1}] &= \frac{a'}{a' + b'} \cdot \frac{(a' - 1) - b'}{a' + b' - 1} + \frac{b'}{a' + b'} \cdot \frac{a' - (b' - 1)}{a' + b' - 1} \\ &= \frac{a'(a' - 1) - b'(b' - 1)}{(a' + b')(a' + b' - 1)} \\ &= \frac{(a' - b')(a' + b' - 1)}{(a' + b')(a' + b' - 1)} = X_{k-1} \end{aligned}$$

定义 $T = \min\{k | X_k = 0\}$ 或者 $n - 1$ 如果 k 不存在。

- 如果 A 一直领先, 则 $T = n - 1$, 故 $X_T = X_{n-1} = S_1 = 1$
- 如果存在平票的时刻, 则 $X_T = 0$ 。

从而第一种情况的概率, 即答案为 $\mathbb{E}[X_T] = \mathbb{E}[X_0] = \frac{a-b}{a+b}$ 。

Submartingale

Definition (sub/supmartingale): (X_i) 是关于 filter (\mathcal{F}_i) 的**下鞅**如果

$$\mathbb{E}[X_i | \mathcal{F}_{i-1}] \geq X_{i-1}$$

反之, 是**上鞅**如果

$$\mathbb{E}[X_i | \mathcal{F}_{i-1}] \leq X_{i-1}$$

在满足相应条件下，关于下鞅，有 $\mathbb{E}[X_T] \geq \mathbb{E}[X_0]$ ；对于上鞅，有 $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$ 。

基于此可以有一种 bound $\mathbb{E}[T]$ 的方式：

记 $D_i = X_i - X_{i-1}$ ，假设 (X_i) 是一个鞅，即 $\mathbb{E}[D_i | X_1, \dots, X_{i-1}] = 0$ ，并且有 $\mathbb{E}[D_i^2 | X_1, \dots, X_{i-1}] \geq \sigma^2$ 。那么设 $Y_i = X_i^2 - \sigma^2 \cdot i$ ，从而

$$\begin{aligned} \mathbb{E}[Y_i | X_1, \dots, X_{i-1}] &= \mathbb{E}[X_i^2 | X_1, \dots, X_{i-1}] - \sigma^2 \cdot i \\ &= \mathbb{E}[D_i^2 | X_1, \dots, X_{i-1}] + X_{i-1}^2 - \sigma^2 \cdot i \\ &\geq \sigma^2 + (Y_{i-1} + \sigma^2 \cdot (i-1)) - \sigma^2 \cdot i \\ &= Y_{i-1} \end{aligned}$$

这表明 (Y_i) 是一个下鞅，从而对于一个停时 T ，

$$\mathbb{E}[Y_T] \geq \mathbb{E}[Y_0] \quad \Rightarrow \quad \mathbb{E}[T] \leq \frac{\mathbb{E}[X_T^2] - \mathbb{E}[X_0^2]}{\sigma^2}$$

现在考虑一个上鞅 (X_i) ，定义在区间 $[0, n]$ 上， $X_0 = s$ ，满足：

$$\begin{aligned} \mathbb{E}[D_i | X_1, \dots, X_{i-1}] &\leq 0 \\ \mathbb{E}[D_i^2 | X_1, \dots, X_{i-1}] &\geq \sigma^2 \end{aligned}$$

Claim: 设 T 是第一次到达 0 的时刻， $\mathbb{E}[T] \leq \frac{2ns - s^2}{\sigma^2} \leq \frac{n^2}{\sigma^2}$

构造 $Y_i = X_i^2 - 2nX_i - \sigma^2 i$ ，可以验证 Y_i 是一个下鞅，从而

$$\mathbb{E}[Y_T] \geq \mathbb{E}[Y_0] \quad \Rightarrow \quad \mathbb{E}[T] \leq \frac{2ns - s^2}{\sigma^2} \leq \frac{n^2}{\sigma^2}$$

Random 2-SAT

对于一个有 n 个变量的 2-CNF ϕ ，任意选定一个起始赋值 a_0 。如果 ϕ 不满足，则任取一个没满足的 clause C_0 ，任选其中的一个 literal 并翻转之。

Claim: 如果 ϕ 是可满足的，则上述随机算法在期望 $O(n^2)$ 次找到一个合法赋值。

任取一个合法赋值 a^* ，用 X_i 代表 i 轮后的赋值 a_i 和 a^* 的 Hamming 距离，则当 a_i 仍是不满足的赋值时，

$$|X_i - X_{i-1}| = 1, \quad \Pr[X_i - X_{i-1} = -1] \geq \frac{1}{2}$$

后者是因为一个错误的 clause 当中所涉及的两个变量，不妨在 a^* 中的赋值是 00，则在 a_{i-1} 中的赋值只可能是 01, 10, 11。对于前两者 Hamming 距离期望不变，而对于最后一种情况 Hamming 距离一定 -1

。

因此设 $D_i = X_i - X_{i-1}$ ，则有

$$\begin{aligned}\mathbb{E}[D_i | X_1, \dots, X_{i-1}] &\leq 0 \\ \mathbb{E}[D_i^2 | X_1, \dots, X_{i-1}] &= 1\end{aligned}$$

从而根据前述结论，有 $\mathbb{E}[\text{steps to } a^*] \leq n^2$ 。

注：事实上在上述迭代过程中可能中途即出现 $a_i \neq a^*$ 已经满足了 ϕ 的情况，此时迭代会收敛，因为找不到“错误的 clause”，但这是有助于结论的，故不做考虑。

Lecture 20 - 2025 / 4 / 28

Percolation on d -Regular Graphs

Theorem: G 为 n 顶点的 d -正则图，其中 $3 \leq d \leq n - 1$ 。用 \mathcal{C}_1 代表 G 上的 p -渗滤的最大的连通分支，其中 $p = \frac{1}{d-1}$ ，则对任意 $A > 0$:

$$\Pr[|\mathcal{C}_1| \geq An^{2/3}] \leq \frac{\alpha}{A^{3/2}}$$

其中 α 是一个 universal 常数。

考虑选定一个点 v 开始分支过程，用 X_t 表示当前“前沿”点的数量，每次展开一个“前沿”点。初始 $X_0 = 1$ ，于是

$$X_t = X_{t-1} - 1 + \mathcal{B}\left(d-1, \frac{1}{d-1}\right)$$

可以看出 (X_t) 是鞅，我们关注的是 $X_T = 0$ 的时刻。

Lemma: 假设 (X_t) 是关于 (\mathcal{F}_t) 的鞅， $X_0 = 1, X_t \geq 0$ ，定义停时 $T = \min\{k, \min\{t \mid X_t = 0 \vee X_t \geq h\}\}$ ，那么如果满足

- (方差有下界) $\text{Var}[X_t | \mathcal{F}_{t-1}] \geq \sigma^2 > 0$ ，对于 $X_t > 0$
- (越界不太多) $\mathbb{E}[X_T^2 | X_T \geq h] \leq Dh^2$

那么就有 $\Pr[\forall t \leq k, X_t > 0] \leq \frac{1}{h} + \frac{Dh}{k\sigma^2}$ 。

首先所求即 $\Pr[X_T \neq 0] \leq \Pr[T \geq k] + \Pr[X_T \geq h]$ 。

容易根据 Markov 不等式得到 $\Pr[X_T \geq h] \leq \frac{\mathbb{E}[X_T]}{h} = \frac{1}{h}$ 。

考虑 $Y_t := X_t^2 - hX_t - \sigma^2 t$, 易见 (Y_t) 是下鞅, 从而 $1 - h = \mathbb{E}[Y_0^2] \leq \mathbb{E}[Y_T^2] \leq \mathbb{E}[X_T^2] - h\mathbb{E}[X_T] - \sigma^2\mathbb{E}[T]$ 。

注意到 $\mathbb{E}[X_T^2] - h\mathbb{E}[X_T]$ 在 $X_T < h$ 时是负的, 故 $\leq \Pr[X_T \geq h] \cdot (Dh^2 - h^2) \leq (D - 1)h$ 。于是立刻可以得到 $\mathbb{E}[T] \leq Dh/\sigma^2$ 。

再根据 Markov 不等式, 有 $\Pr[T \geq k] \leq \frac{Dh}{k\sigma^2}$ 。

我们考虑将上述引理应用到 (X_t) 上。易见方差 $\sigma^2 = \frac{d-2}{d-1} \geq \frac{1}{2}$, 于是只需关注 $X_T \geq h$ 时的情况, 我们针对最后一步展开。

$$\begin{aligned} \mathbb{E}[X_T^2 \mid X_T \geq h] &\leq \mathbb{E}_{Z \sim \mathcal{B}(d-1, 1/(d-1))}[(h + Z)^2] \\ &\leq h^2 + 2h + 2 \leq 2h^2 \quad (\forall h \geq 3) \end{aligned}$$

于是根据 Lemma, 对于任何 $h \geq 3$, 都有 $\Pr[\forall t \leq k, X_t > 0] \leq \frac{1}{h} + \frac{4h}{k}$, 取 $h = \frac{\sqrt{k}}{2}$ 得到最优概率 $\frac{2}{\sqrt{k}}$, 即设 $C(v)$ 表示从 v 开始分支过程的连通分支大小, 则有 $\Pr[C(v) \geq k] \leq \frac{2}{\sqrt{k}}$ 。下证 Theorem。

如果直接对所有 v 使用 union bound, 则将得到 $\Pr[\exists v, C(v) \geq k] \leq \frac{2n}{\sqrt{k}}$, 这显然对于 $k = O(n^{2/3})$ 是一个不好的界限。我们可以巧妙地将分母再乘一个 k 。

考虑用 N_k 代表位于 $\geq k$ 个点的连通分支的点数, 则 $\mathbb{E}[N_k] = n \Pr[C(v) \geq k] = \frac{2n}{\sqrt{k}}$ 。根据 Markov 不等式, 有 $\Pr[N_k \geq k] \leq \frac{2n}{k^{3/2}}$ 。取 $k = An^{2/3}$, 则有

$$\Pr[|\mathcal{C}_1| \geq An^{2/3}] \leq \frac{2}{A^{3/2}}$$

Lecture 21 - 2025 / 5 / 8

Lovász Local Lemma

Lemma: 设 A_1, \dots, A_n 是一系列“坏事件”, $\Pr[A_i] \leq p$, 并且每个 A_i 独立于除最多 d 个其他事件 A_j 之外的所有事件。如果 $ep(d+1) \leq 1$, 则

$$\Pr\left[\bigcap_{i=1}^n \overline{A_i}\right] > 0$$

Claim: 对于任意任何 $S \subseteq \{1, \dots, n\}$, 对任意 i , 有 $\Pr\left[A_i \mid \bigcap_{j \in S} \overline{A_j}\right] \leq \frac{1}{d+1}$ 。

对 $m := |S|$ 归纳, $m = 0$ 时 $\Pr[A_i] \leq p \leq \frac{1}{e(d+1)} < \frac{1}{d+1}$ 。

将 S 分为 $S_1 = S \cap D_i, S_2 = S \setminus S_1$, 其中 D_i 为和 A_i 有关的事件集合。

$$\Pr \left[A_i \mid \bigcap_{j \in S} \overline{A_j} \right] = \frac{\Pr \left[A_i \cap \bigcap_{j \in S_1} \overline{A_j} \mid \bigcap_{k \in S_2} \overline{A_k} \right]}{\Pr \left[\bigcap_{j \in S_1} \overline{A_j} \mid \bigcap_{k \in S_2} \overline{A_k} \right]}$$

分子 $\leq \Pr \left[A_i \mid \bigcap_{k \in S_2} \overline{A_k} \right] \leq \Pr[A_i]$ 。

对于分母, 不妨设 $S_1 = \{1, 2, \dots, |S_1|\}$ 。

$$\begin{aligned} \Pr \left[\bigcap_{j \in S_1} \overline{A_j} \mid \bigcap_{k \in S_2} \overline{A_k} \right] &= \prod_{j=1}^{|S_1|} \left(1 - \Pr \left[A_j \mid \bigcap_{j' < j} \overline{A_{j'}} \cap \bigcap_{k \in S_2} \overline{A_k} \right] \right) \\ &\geq \left(1 - \frac{1}{d+1} \right)^{|S_1|} \\ &\geq \left(1 - \frac{1}{d+1} \right)^d > \frac{1}{e} \end{aligned}$$

从而原式 $\leq \frac{p}{1/e} = ep \leq \frac{1}{d+1}$, 根据归纳法原命题得证。

根据 Claim, 我们有

$$\begin{aligned} \Pr \left[\bigcap_{i=1}^n \overline{A_i} \right] &= \prod_{i=1}^n \left(1 - \Pr \left[A_i \mid \bigcap_{j < i} \overline{A_j} \right] \right) \\ &\geq \left(1 - \frac{1}{d+1} \right)^n > 0 \end{aligned}$$

从而 LLL 得证。

Example: k -SAT

Claim: 任何 k -CNF φ , 如果每个变量都出现在至多 $\frac{2^{k-2}}{k}$ 个 clause 里, 则 φ 是可被满足的。

$A_i :=$ 第 i 个 clause 不满足, 则 $\Pr[A_i] = 2^{-k} = p$, 同时 $d = k \cdot \frac{2^{k-2}}{k} = 2^{k-2}$ 。容易验证此时 LLL 的条件满足。

Lecture 22 - 2025 / 5 / 12

Packet Routing

考虑给定一张无向图 G ，第 i 个数据包想从 $s_i \rightarrow t_i$ ，沿着固定的路径 P_i 。但是每条边每个时刻只能通过一个数据包。我们想要设计一个调度方案，使得传输完所有数据包的总时间最少。

定义 c_e 为经过 e 的路径数量， $c = \max\{c_e\}$ ， d 为所有路径 P_i 长度的最大值。显然答案必须 $\geq \max\{c, d\}$ 。

Theorem: 存在一种调度方案满足时间为 $O(c + d)$ 且只有常数大小的缓冲区。

Theorem': 存在一种调度方案满足时间为 $O((c + d)2^{O(\log^*(c+d))})$ 且只有 $O((\log d)2^{O(\log^*(c+d))})$ 大小的缓冲区。这里 \log^* 的意思是通过不断取 \ln 直到变成常数规模所需要的次数。

不失一般性设 $c = d$ 。考虑尝试安排数据包 i 在起点等待 Z_i 时间，然后直接不等待地沿着路径 P_i 完成传输。这里 Z_i 是独立均匀从 $\{1, 2, \dots, \alpha d\}$ 中抽取， $\alpha > 1$ 是待确定常数，显然这种做法的时间开销是 $(1 + \alpha)d$ ，正确性待证。

Claim: 将时间切分为 $\ln d$ 长度的帧，可以将问题分割为若干子问题，其中每个数据包想从这个帧内的起点到这个帧内的终点。以正概率每个子问题中的边的冲突次数（经过的路径数量）为 $\ln c$ 。

对于每条边 e 定义坏事件 A_e 代表在某个帧内经过 e 的路径数量超过 $\ln c$ 。

注意到 A_e 只和 $A_{e'}$ 相关，其中 e, e' 存在公共经过的数据包。由于只有至多 c 个数据包经过 e ，每个数据包经过的路径长度至多 d ，因此 A_e 至多依赖 $cd = d^2$ 个坏事件。

接下来分析 $\Pr[A_e]$ 。对于任何一个数据包，因为帧的长度是 $\ln d$ ，所以对于一个特定的帧，在其中任何数据包经过 e 的概率仅为 $\ln d / \alpha d$ 。于是该帧内经过 e 的总边数 $\sim \mathcal{B}(c, \ln d / \alpha d)$ ，因此 $\Pr[A_e] = (1 + \alpha)d \cdot \Pr[\mathcal{B}(c, \ln d / \alpha d) > \ln c]$ ，即对所有 $< (1 + \alpha)d$ 个帧 union bound。

根据 Chernoff bound,

$$\Pr[A_e] \leq (1 + \alpha)d \cdot \left(\frac{ce \ln d}{d\alpha \ln c} \right)^{\ln d} = (1 + \alpha)d^{2 - \ln \alpha}$$

因此只需要取 α 足够大，即可满足 $\Pr[A_e] < 1/e(d^2 + 1)$ 。

利用这一性质，可以将问题拆分为 $(1 + \alpha)d / \ln d$ 个子问题，参数分别为 $\ln c$ 和 $\ln d$ ，然后分别递归解决。通过不断取 \ln ，最终问题会变成常数规模，于是我们可以构造一个确定的调度方案。最终通过合并解决原问题。由于递归层数是 $O(\log^*(c + d))$ 的，每层总长度会伸长 $1 + \alpha$ 倍，因此总时间为 $d2^{O(\log^*(c+d))}$ ，同时不同帧之间不会影响，因此缓冲区大小为 $O((\log d)2^{O(\log^*(c+d))})$ 。

Asymmetric LLL

Lemma (General LLL): 设 A_1, \dots, A_n 是一系列坏事件, $D_i \subseteq \{A_1, \dots, A_n\}$ 是 A_i 相关的事件集合, 如果存在实数 $x_1, \dots, x_n \in [0, 1)$ 使得对所有的 i , 有 $\Pr[A_i] \leq x_i \prod_{j \in D_i} (1 - x_j)$, 则 $\Pr[\bigcap_{i=1}^n \overline{A_i}] \geq \prod_{i=1}^n (1 - x_i) > 0$ 。

通过带入 $x_i = 2 \Pr[A_i]$, 有

Corollary (Asymmetric LLL): 同上, 如果 $\sum_{j \in D_i} \Pr[A_j] \leq 1/4$, 则 $\Pr[\bigcap_{i=1}^n \overline{A_i}] \geq \prod_{i=1}^n (1 - 2 \Pr[A_i]) > 0$ 。

Frugal Graph Coloring

Definition: 称 G 的一个合法染色是 β -frugal 的, 如果对于任何 $v \in G$ 的邻居, 都没有一种颜色出现了多于 β 次。

Theorem: 如果 G 的最大度数 $\Delta \geq \beta^\beta$, 则 G 有一种用 $16\Delta^{1+1/\beta}$ 种颜色的 β -frugal 染色。

对于 $\beta = 1$, 有 $16\Delta^2$ 种颜色, 这是容易做到的。

对于 $\beta \geq 2$, 对 G 随机均匀 $Q := 16\Delta^{1+1/\beta}$ 染色。下面证明有正数概率是满足条件的即可。有两类坏事件:

1. A_{uv} : 相邻的两点 u, v 染成同一种颜色。
2. $B_{u_1, u_2, \dots, u_{\beta+1}}$: 某一个点的邻居 $u_1, u_2, \dots, u_{\beta+1}$ 染成了同一种颜色。

容易看出 $\Pr[A_{uv}] = 1/Q$, $\Pr[B_{u_1, \dots, u_{\beta+1}}] = 1/Q^\beta$ 。对于 A 类事件, 它与至多 2Δ 个 A 类事件、 $2\Delta \binom{\Delta}{\beta}$ 个 B 类事件相关; 对于 B 类事件, 它与至多 $(\beta + 1)\Delta$ 个 A 类事件、 $(\beta + 1)\Delta \binom{\Delta}{\beta}$ 个 B 类事件相关。可见 B 类事件的相关性更强, 我们对其验证 Asymmetric LLL 的使用条件:

$$\begin{aligned} & \left((\beta + 1)\Delta \cdot \frac{1}{Q} \right) + \left((\beta + 1)\Delta \binom{\Delta}{\beta} \cdot \frac{1}{Q^\beta} \right) \\ & \leq \frac{(\beta + 1)\Delta}{Q} + \frac{(\beta + 1)\Delta^{\beta+1}}{\beta! Q^\beta} \\ & \leq \frac{\beta + 1}{16\Delta^{1/\beta}} + \frac{\beta + 1}{\beta! 16^\beta} \leq \frac{\beta + 1}{16\beta} + \frac{\beta + 1}{\beta! 16^\beta} < 1/4 \end{aligned}$$

因此满足 Asymmetric LLL 的使用条件, 得证。

Lecture 23 - 2025 / 5 / 15

Markov Chains

Definition: 一个 Markov 链是一列随机变量 $(X_t)_{t=0}^\infty$, 满足

$$\Pr[X_t = y \mid X_{t-1} = x, X_{t-2}, \dots, X_0] = \Pr[X_t = y \mid X_{t-1} = x] = P(x, y)$$

其中 $P(x, y)$ 是一个转移概率， P 是行和为 1 的矩阵。

我们有 $p_x^{(t)} = p_x^{(0)} P^t$ ，其中 $p_x^{(0)}$ 是从 x 出发的 one-hot 初始分布。

Definition (irreducible): $\forall x, y, \exists t$ s.t. $p_x^{(t)}(y) > 0$

Definition (aperiodic): $\forall x, y, \gcd\{t \mid p_x^{(t)}(y) > 0\} = 1$

Stationary Distribution

Theorem (Fundamental Theorem): 如果 P 是不可约且非周期的，则存在唯一的平稳分布 π ，满足 $\pi P = \pi$ ，且 $p_x^{(t)}(y) \xrightarrow{t \rightarrow \infty} \pi(y) \quad \forall x, y$ 。这里 π 实际上是 P 特征值为 1 的唯一左特征向量。

Observation 1: 如果 P 是对称的，则 π 是均匀分布。

Observation 2: 如果 P 列和也为 1，则 π 是均匀分布。

Observation 3: 如果 P 关于某个分布 π 可反的，即 $\pi(x)P(x, y) = \pi(y)P(y, x)$ ，则 π 是平稳分布。

Metropolis Process

给定一个大集合 Ω 和权重 $w : \Omega \rightarrow \mathbb{R}^+$ ，希望设计一个稳态分布为 $\pi(x) = w(x)/Z$ 的 Markov 链，其中 $Z = \sum_{x \in \Omega} w(x)$ ，并且我们假定 Z 是不知道的，或者正是我们想求的。

大空间采样过程给定将 Ω 连接起来的无向图，以及位于 x 时抽取邻居的分布 $\kappa(x, y) > 0$ ，并且有 $\kappa(x, y) = \kappa(y, x)$ ，我们构造 Markov 链如下：

- 在 x 时，抽取一个邻居 y ，概率为 $\kappa(x, y)$ 。
- 以概率 $\min\{1, w(y)/w(x)\}$ 接受 y ，否则停留在 x 。

Claim: 由大空间采样构造出的 Markov 链的平稳分布为 $\pi(x) = w(x)/Z$ 。

不妨设 $w(x) \geq w(y)$ 。当 x, y 不是邻居时， $\pi(x)P(x, y) = \pi(y)P(y, x) = 0$ 。当 x, y 是邻居时，

$$\pi(x)P(x, y) = \frac{w(x)}{Z} \cdot \kappa(x, y) \frac{w(y)}{w(x)} = \frac{w(y)}{Z} \kappa(x, y) = \pi(y)P(y, x)$$

最后一个等号是因为 $\kappa(x, y) = \kappa(y, x)$ 。

事实上，如果不满足 $\kappa(x, y) = \kappa(y, x)$ ，我们只需将接受概率修改为 $\min\{1, (w(y)\kappa(y, x))/(w(x)\kappa(x, y))\}$ 。

Lecture 24 - 2025 / 5 / 19

Mixing Time

Definition (Variation Distance): 对于两个 Ω 上的分布 μ, ξ , 定义

$$\|\mu - \xi\| = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \xi(x)| = \max_{A \subseteq \Omega} |\mu(A) - \xi(A)|$$

Definition: 对于一个不可约无周期的 Markov 链, 定义时间 t 的距离为 $\Delta(t) = \max_{x \in \Omega} \|\pi - p_x^{(t)}\|$ 。

Definition (Mixing Time): 定义 τ_{mix} 为混合时间: $\tau_{\text{mix}} = \min\{t \mid \Delta(t) \leq 1/2e\}$ 。

Fact: $\Delta(\tau_{\text{mix}} \lceil \ln \epsilon^{-1} \rceil) \leq \epsilon$

通过 coupling 的方式可以证明 $\Delta(kt) \leq (2\Delta(t))^k$ 。

Definition (Strong Stationary Time): 停时 T 是一个强稳定时间, 如果停下来时可以保证收敛 $\Pr[X_t = y \mid T = t] = \pi(y)$ 。

Claim: $\Delta(t) \leq \Pr[T > t]$

虽然 $\Delta(t)$ 是一个固定的数, 但我们可以对它求期望

$$\begin{aligned} \mathbb{E}[\Delta(t)] &= \Pr[T > t] \cdot \mathbb{E}[\Delta(t) \mid T > t] + \Pr[T \leq t] \cdot \mathbb{E}[\Delta(t) \mid T \leq t] \\ &\leq \Pr[T > t] \cdot 1 + \Pr[T \leq t] \cdot 0 = \Pr[T > t] \end{aligned}$$

Example: Top-in-at-Random

考虑一种洗牌方式: 每次把最顶上的牌插入随机位置。

Claim: 这种洗牌方式的混合时间为 $O(n \log n)$ 。

用 T 表示原本最底下的牌被随机插入的时刻, 则 T 是一个强稳定时间。可见 $T = T_1 + T_2 + \dots + T_{n-1} + 1$, 其中 T_i 表示从位置 i 变动到 $i + 1$ 所需要的时间。每个 T_i 的分布是几何分布, 期望为 n/i , 故 $\mathbb{E}[T] = O(n \log n)$ 。根据 Markov 不等式, $\tau_{\text{mix}} \leq O(n \log n)$ 。

Example: Riffle Shuffle

考虑一种洗牌方式: 每次把牌按照 $\mathcal{B}(n, 1/2)$ 分成两堆, 然后随机均匀交叉。它的逆过程是, 随机将每张牌标记为 0/1, 然后将 0 的牌挪到上面, 1 的牌挪到下面。

Claim: 这种洗牌方式的混合时间 $\leq 2 \log_2 n + O(1)$ 。

将每轮的编号串联为一个二进制串, 用 T 表示每张牌被唯一标号确定的时间, 也即给每张牌随机抽样 $[0, 2^T)$ 内的编号, 能够做到不重复的时间。

根据生日悖论, n 个人从 cn^2 大小的集合抽取生日, 有生日冲突的概率渐进趋向 $1 - \exp(-1/2c)$ 。因此, 只需 $1 - \exp(-1/2c) \leq 1/2e$ 且 $2^t \geq cn^2$, 则有 $\tau_{\text{mix}} \leq 2 \log_2 n + O(1)$ 。

另一种看法是, 对于固定的两张牌 (x, y) , 无法被分开的概率为 2^{-t} , 根据 union bound, 只需要 $t = O(\log n)$ 即可使得 $n^2 2^{-t} \leq 1/2e$ 。

Coupling

Definition (Coupling): 设 $(X_t), (Y_t)$ 为一个 Markov 链的两个样本, 称它们是一个耦合, 如果

1. 边际上 X_t 和 Y_t 的分布相同, 即 $\Pr[X_t = y] = \Pr[Y_t = y]$;
2. $X_t = Y_t$ 时, $X_{t+1} = Y_{t+1}$ 。

Definition (Meeting Time): T_{xy} 是从 x, y 开始的两个 Markov 链的耦合的第一次相遇时间。即 $T_{xy} = \min\{t \mid X_t = Y_t, X_0 = x, Y_0 = y\}$ 。

Claim: $\Delta(t) \leq \max_{x,y} \Pr[T_{xy} \geq t]$

首先注意到, 对于任何两个 r.v. X, Y , 都有 $\Pr[X \neq Y] \geq \|P_X - P_Y\|$ 。

从而 $\Delta(t) = \max_x \|P_x^{(t)} - \pi\| \leq \max_{x,y} \|P_x^{(t)} - P_y^{(t)}\| \leq \max_{x,y} \Pr[X_t \neq Y_t \mid X_0 = x, Y_0 = y] \leq \max_{x,y} \Pr[T_{xy} \geq t]$ 。其中第一个不等号是因为 π 可以写作 $P_y^{(t)}$ 的线性组合 $\pi = \sum_y \pi(y) P_y^{(t)}$:

$$\pi(x) = (\pi P^t)(x) = \sum_y \pi(y) P^t(y, x) = \sum_y P_y^{(t)}(x) \pi(y)$$

Corollary: $\tau_{\text{mix}} \leq 2e \max_{x,y} \mathbb{E}[T_{xy}]$

根据 Markov 不等式, $\Pr[T_{xy} \geq t] \leq \mathbb{E}[T_{xy}]/t$, 因此 $\Delta(t) \leq \max_{x,y} \mathbb{E}[T_{xy}]/t$ 。当 $t = 2e \max_{x,y} \mathbb{E}[T_{xy}]$ 时, $\Delta(t) \leq 1/2e$ 。

Example: Random Transposition Shuffle

考虑一种洗牌方式: 每次随机选择两个位置交换。这个洗牌方式的等价描述是, 选择一个位置和一张牌 c , 将 c 交换到位置 i 。

Claim: 这种洗牌方式的混合时间为 $O(n^2)$ 。

用 Coupling 来分析, 用 D_t 表示 X_t, Y_t 不同的位置, 目标是分析多久之后 $D_t = 0$ 。

考虑一次选中 (i, c) ,

- 如果 c 已经匹配了, 则 D_t 不会改变
- 如果 c 没有匹配, 则 D_t 不会上深, 且如果 i 位置之前不匹配, 将会至少减少 1。

因此，如果当前 $D_t = d$ ，则 $\Pr[D_t \text{ decreases}] \geq (d/n)^2$ 。于是 $\mathbb{E}[T_{xy}] \leq \sum_{d=1}^n (n/d)^2 = O(n^2)$ 。

注：实际上为 $\Theta(n \log n)$ 。

Lecture 25 - 2025 / 5 / 22

Graph Colorings

给定一张无向图 $G = (V, E)$ ，最大度数为 Δ ， k 种颜色。目标是随机生成一个 k -着色，使得同色不相邻。

考虑如下过程：

1. 随机选择结点 v 和颜色 c
2. 如果 v 可以用 c 染色，即染

Theorem: 如果 $k \geq 4\Delta + 1$ 则这个 Markov 链的混合时间为 $O(n \log n)$ 。

定义一个 coupling: X_t 和 Y_t 每次选择同样的 v, c ，用 D_t 表示 X_t, Y_t 不同色的结点， $d_t = |D_t|$ ，目标则是计算 $d_t = 0$ 所需的时间。

- 好的操作：如果 $v \in D_t$ ，且 c 对 X_t, Y_t 都合法，则 $d_{t+1} = d_t - 1$ 。好的操作数量 $\geq d_t(k - 2\Delta)$ 。
- 坏的操作：如果 $v \in V \setminus D_t$ ，且 c 对 X_t, Y_t 中的一个合法、另一个不合法，则 $d_{t+1} = d_t + 1$ 。坏的操作数量 $\leq 2d_t\Delta$ 。这可以通过枚举 v 的异色邻居计数。

从而 $\mathbb{E}[d_{t+1} | d_t] \leq d_t + d_t \frac{4\Delta - k}{kn} \leq d_t(1 - 1/kn)$ 。进而 $\mathbb{E}[d_t | d_0] \leq d_0(1 - 1/kn)^t$ 。取 $t = Ckn \log n$ ，结合 $d_0 \leq n$ 有 $\mathbb{E}[d_t] \leq 1/2e$ 。

Theorem: 如果 $k \geq 3\Delta + 1$ 则这个 Markov 链的混合时间为 $O(n \log n)$ 。

我们通过设计一个更好的 coupling 来证明。具体而言， X_t 和 Y_t 每次选择同样的 v ，但 X_t 选择颜色 c 时：

- 如果 X_t, Y_t 中都可以用 c 染色，则 Y_t 也选择颜色 c 。
- 如果 X_t, Y_t 中都不可以用 c 染色，则 Y_t 也选择颜色 c 。
- 如果 X_t 可以用 c 染色， Y_t 不可以，则 Y_t 尽量选择一个可以染色的颜色。
- 如果 X_t 不可以用 c 染色， Y_t 可以，则 Y_t 尽量选择一个不可以染色的颜色。

上述定义的思路是“将 $N_X(v) \setminus N_Y(v)$ 和 $N_Y(v) \setminus N_X(v)$ ”尽量配对起来，其中 $N(v)$ 表示与 v 邻居的颜色集合。从而好的操作数量仍然为 $d_t(k - 2\Delta)$ ，而坏的操作数量 $\leq d_t\Delta$ ，缩小了一半。从而好坏操作的差 $\leq d_t(3\Delta - k)$ 。

Theorem: 如果 $k \geq 2\Delta + 1$ 则这个 Markov 链的混合时间为 $O(n \log n)$ 。

我们只需对上面的 coupling 进行更为精细的分析。事实上，好坏操作的差为

$$d_t k - \sum_{v \in D_t} |N_X(v) \cup N_Y(v)| - \sum_{v \in V \setminus D_t} \max\{|N_X(v) \setminus N_Y(v)|, |N_Y(v) \setminus N_X(v)|\}$$

采用贡献法，对于每个 $v \in D_t$ 及其邻居构成的有序二元组 (v, u) ，如果 $u \in D_t$ ，则这条边分别在第一个求和 v 时贡献两次，如果 $u \in V \setminus D_t$ ，则这条边在第一个求和 v 时贡献一次，在第二个求和 u 时贡献一次。从而总贡献量不超过 $2d_t \Delta$ ，即好坏操作的差 $\leq d_t(2\Delta - k)$ 。

Algorithmic LLL

Theorem: 对于任何 k -SAT 问题 ϕ ，如果每个变量至多在 $\frac{2^{k-d}}{k}$ 个子句出现，则该实例是可满足的，且赋值可以在多项式时间内构造得到。

解的存在性是 LLL 的经典应用，考虑如何构造。首先给 ϕ 随机赋值，然后每次取出一个尚未满足的子句 C ，将其中的每个变量重新随机赋值。直到所有子句都满足为止。

Solve(ϕ):

Pick a random assignment of ϕ

while there is an unsatisfiable clause C

Fix(C)

Fix(C):

Replace the variables of C with new random values

while there is clause D that shares a variable with C that is not satisfied

Fix(D)

下面从 Kolomogrov 复杂度的角度给出证明这个算法终止性证明。

考虑随机串是“不可压缩的”，那么进行 F 次修复就需要 Fk 个 bit。但是现在更换方式为记录最终赋值和修复历史 C_1, C_2, \dots, C_F ，可以看出通过这些信息足够恢复出所用到的所有随机 bit。因为修复一个 clause 前这个 clause 一定是完全不满足的，而最后一次被修复的信息又可通过最终赋值获得。

从而记录随机串只需要 $c + n + F(k - d)$ 个 bit，其中 c 是常数。最后一项是因为被 solve 调用的 fix 可以用 $m \log m$ bit 记录， m 是子句数目。而递归调用的 fix 涉及的 clause D 是与 C 有交的，因此只需要 $\log 2^{k-d} = k - d$ bit 记录。

综上 $c + n + F(k - d) \geq Fk$ 可以推出 F 是多项式级别的。结合随机串高概率 Kolomogrovly random 可知结论成立。