高等演算法 HW3

許博翔

May 24, 2024

Notation 1. Let n be a positive integer. $[n] := \{1, 2, \ldots, n\}$.

Problem 0.

Problem 1, 2, 3, 4: All by myself.

Problem 5, 6: Discuss with: B10401113 張有朋

Problem 1. Consider $(x_1 \vee x_2) \wedge (x_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_2) \wedge (\neg x_1 \vee \neg x_2)$.

Equivalent ILP:

$\max(z_1 + z_2 + z_3 + z_4)$.

$$\text{subject to :} \begin{cases} y_1 + y_2 \geq z_1 \\ y_1 + 1 - y_2 \geq z_2 \\ 1 - y_1 + y_2 \geq z_3 \\ 1 - y_1 + 1 - y_2 \geq z_4 \\ y_i, z_c \in \{0, 1\} \end{cases}.$$

One can see that in LP, we can set $y_1 = y_2 = \frac{1}{2}$, and get $z_1 = z_2 = z_3 = z_4 = 1$, which maximizes $z_1 + z_2 + z_3 + z_4 = 4$.

But since exactly one of the 4 clauses above must be false, $\max(z_1 + z_2 + z_3 + z_4) = 3$.

$\therefore$ the integrality gap is $\frac{3}{4}$ in this case.

Note that it can't be more than $\frac{3}{4}$ since in the following problem (or in class), we've find a solution ALG that satisfies $\frac{3}{4}OPT \leq \frac{3}{4}OPT(LP) \leq ALG \leq OPT$.

$\therefore$ MAX-SAT has integrality gap $\frac{3}{4}$.

Problem 2. Let the graph decribed in the problem be G.

Consider $H := G \setminus C_m$. That is, H is the single path from u to v with n edges.

Let h'_{uv} be the hitting time on graph H.

Let $R'_{eff}(u, v)$ be the equivalent resistance of u, v on H.

In class, we know that $h'_{uv} + h'_{vu} = 2n \times R'_{eff}(u, v) = 2n \times n$.

Since u, v are symmetric in H.

$\therefore h'_{uv} = h'_{vu}$.

$\therefore h'_{vu} = \frac{1}{2} \times 2n^2 = n^2$.

Since when starting from v, one will not walk through any edge of C_m before arriving u.

$\therefore h_{vu} = h'_{vu} = n^2$.

Let $R_{eff}(u, v)$ be the equivalent resistance of u, v on G.

Again, in class, we know that $h_{uv} + h_{vu} = 2(m + n) \times R_{eff}(u, v) = 2(m + n)n$.

$\therefore h_{uv} = 2(m + n)n - h_{vu} = 2(m + n)n - n^2 = 2mn + n^2$.

Problem 3. Run A for s (we will determine s later) times independently, and let the resulting $A(I)$ of the i-th time be X_i.

Let $B(I)$ be the median of $X_1, X_2, \ldots, X_s$.

We know that $p := \Pr[(1 - \epsilon)\#(I) \le X_i \le (1 + \epsilon)\#(I)] \ge \frac{3}{4}$.

Let $Y = |\{i : X_i < (1 - \epsilon)\#(I)\}|, Z = |\{i : X_i > (1 + \epsilon)\#(I)\}|$.

Let $\mu := \mathrm{E}[Y + Z]$.

$\mu = \mathrm{E}[Y+Z] = \sum_{i=1}^{s} \Pr[X_i < (1-\epsilon)\#(I) \text{ or } X_i > (1+\epsilon)\#(I)] = \sum_{i=1}^{s}(1-p) = s(1-p)$.

$\Pr[Y + Z \ge \frac{s}{2}] = \Pr[Y + Z \ge \frac{1}{2(1 - p)}\mu] = \Pr[Y + Z \ge (1 + \frac{2p - 1}{2(1 - p)})\mu] \le e^{-(\frac{2p-1}{2-2p})^2\mu/3} = e^{-\frac{(2p-1)^2}{4(1-p)}s/3} \overset{\because 2p-1 \ge \frac{1}{2}}{\le} e^{-\frac{1}{16(1-p)}s/3} \overset{\because 1-p \le \frac{1}{4}}{\le} e^{-\frac{s}{12}}$.

Take $s \ge 12 \ln \frac{1}{\delta}$, and we get $\Pr[Y + Z \ge \frac{s}{2}] \le e^{-\frac{s}{12}} \le \delta$.

$\therefore \Pr[(1-\epsilon)\#(I) \le B(I) \le (1+\epsilon)\#(I)] = 1 - \Pr[Y \ge \frac{s}{2} \vee Z \ge \frac{s}{2}] \ge 1 - \Pr[Y + Z \ge \frac{s}{2}] \ge 1 - \delta$.

Time complexity:

Since this algorithm runs algorithm A for $12 \ln \frac{1}{\delta}$ times, it is in polynomial complexity of $n, \frac{1}{\epsilon}, \ln \frac{1}{\delta}$.

Problem 4.

Lemma 4.1. If a process succeeds with probability at least p, where $p > 0$ is a constant, and each time the process runs, the results are independent, then the expected number of times running the process to get a success is at most $\frac{1}{p}$.

Proof. Let T denote the number of times running the process to get a success.
$\mathrm{E}[T] = 1 + \Pr[\text{ the process fails }]\mathrm{E}[T] \le 1 + (1-p)\mathrm{E}[T]$.
$\Rightarrow p\mathrm{E}[T] \le 1$.
$\Rightarrow \mathrm{E}[T] \le \frac{1}{p}$. ■

Lemma 4.2. The expected time complexity of step 3 is $O(i)$.

Proof. The probability that inserting j in the rebuild process fails = the probability that collisions happen on both slots $\overset{\text{Both tables have at most } j \text{ non-empty slots}}{\le} (\frac{j}{4i^{1.5}})^2 \le (\frac{i}{4i^{1.5}})^2 = \frac{1}{16i}$.
$\Rightarrow$ the probability that at least one of the above fails $\le i \times \frac{1}{16i} = \frac{1}{16}$.
$\therefore$ the probability that the rebuild process succeeds $= 1-$ the probability that at least one of the above fails $\ge \frac{15}{16}$.
By **Lemma (4.1)**, the expected number of times the rebuild process is run $\le \frac{16}{15} = O(1)$.
Since the rebuild process needs to insert at most i items, its time complexity is $O(i)$.
$\therefore$ the expected time complexity of step 3 $= O(i)O(1) = O(i)$. ■

Lemma 4.3. For any i, the expected total number of times step 3 is run when the i, $(i+1)$, ..., $(2i-1)$-th insertion arrive is $O(1)$.

Proof. If step 3 is run for 0 times, then **Lemma (4.3)** holds clearly.
Suppose that step 3 is run for the first time when the j-th insertion arrives.
For all $k = j+1, j+2, \ldots, 2i-1$, let p_k denote the expected probability that step 3 is run when the k-th insertion arrives.
$p_k =$ the expected probability that collisions happen on both slots $\overset{\text{Both tables have at most } k \text{ non-empty slots}}{\le} (\frac{k}{\text{table size}})^2 \le (\frac{k}{4j^{1.5}})^2 \le (\frac{2i}{4j^{1.5}})^2 \le (\frac{2i}{4i^{1.5}})^2 = \frac{1}{4i}$.

$\therefore$ the expected number of step 3 is called $= 1 + \sum_{k=j+1}^{2i-1} p_k \leq 1 + \sum_{k=i}^{2i-1} p_k \leq 1 + \sum_{k=i}^{2i-1} \frac{1}{4i} = \frac{5}{4} = O(1)$. ■

Let t_i denote the total expected running time when the i, $i+1$, ..., $2i-1$-th insertion arrive.

By **Lemma (4.3)**, t_i is ($O(i)$ step 1 or 2) + ($O(1)$ step 3) $\overset{\text{By } \textbf{Lemma (4.2)}}{=} O(i)O(1) + O(1)O(i) = O(i)$.

Let k be an integer such that $n \leq 2^k < 2n$.

The total expected time complexity $\leq t_1 + t_2 + t_4 + t_8 + \cdots + t_{2^k} = O(1) + O(2) + O(4) + O(8) + \cdots + O(2^k) = O(2^{k+1}) = O(4n) = O(n)$.

Problem 5.

Lemma 5.1. For all $x \geq 0$, there is $f(x) = e^{-x} + x - 1 \geq 0$.
That is, $1 - x \leq e^{-x}$.

Proof. $f'(x) = -e^{-x} + 1 \geq 0$ for all $x \geq 0$.
$\therefore f$ is increasing in $(0, \infty)$.
$\because f(0) = 1 + 0 - 1 = 0$.
$\therefore f(x) \geq 0$ for all $x \geq 0$. ■

Suppose that there are i points $P_1, P_2, \ldots, P_i$ on the circle, where the intervals between any two points are greater than $\frac{1}{n^2}$. Pick the $i+1$-th point P_{i+1} uniformly at random on the circle. Let E_j denote the event that the interval formed by P_j and the P_{i+1} is less than $\frac{1}{2n^2}$.

Since E_j, E_k are disjoint for all $j \neq k$.

$\therefore$ The probability that the intervals between any two points are still greater than $\frac{1}{n^2} \leq 1 - \Pr[E_1 \cup E_2 \cup \cdots \cup E_i] = 1 - (\Pr[E_1] + \Pr[E_2] + \cdots + \Pr[E_i]) = 1 - \frac{i}{2n^2}$.

$\therefore$ after inserting n points, the probability that the intervals between any two points are greater than $\frac{1}{n^2} \leq (1 - \frac{1}{2n^2})(1 - \frac{2}{2n^2}) \cdots (1 - \frac{n}{2n^2}) \overset{\text{By } \textbf{Lemma (5.1)}}{\leq} e^{-\frac{1}{2n^2}} e^{-\frac{2}{2n^2}} \cdots e^{-\frac{n}{2n^2}} = e^{-\frac{n(n+1)}{4n^2}} \leq e^{-\frac{n^2}{4n^2}} = e^{-\frac{1}{4}}$.

$\therefore$ the probability that the size of the smallest interval is less than $\frac{1}{n^2}$

$= 1-$ the probability that the intervals between any two points are greater than $\dfrac{1}{n^2}$
$\geq 1 - e^{-\frac{1}{4}} = \Omega(1)$.

Problem 6. Suppose that the $a_1, a_2, \ldots, a_{2n}$ are $2n$ points chosen uniform randomly in $[0,1]$.

WLOG suppose that the intervals of the first slot are $[a_1, a_1+x], [a_2, a_2+y] \pmod 1$, and WLOG suppose that $a_2 - a_1 \pmod 1 \leq \dfrac{1}{2}$.

$\Pr[x + y < \frac{1}{4n\sqrt{n}}] < \Pr[x < \frac{1}{4n\sqrt{n}} \wedge y < \frac{1}{4n\sqrt{n}}]$.

Let $A_1 := [a_1, a_1 + \frac{1}{4n\sqrt{n}}), A_2 := [a_2, a_2 + \frac{1}{4n\sqrt{n}})$.

There are two cases that $x < \frac{1}{4n\sqrt{n}} \wedge y < \frac{1}{4n\sqrt{n}}$.

Case 1: $a_2 \in A_1$.

x is guaranteed to $< \frac{1}{4n\sqrt{n}}$ in this case.

The probability that case 1 happens $= \frac{1}{4n\sqrt{n}}$.

$\Pr[y < \frac{1}{4n\sqrt{n}}] = \Pr[a_3 \in A_2 \vee \cdots \vee a_{2n} \in A_2] \leq \sum_{i=3}^{2n} \Pr[a_i \in A_2] = \sum_{i=3}^{2n} \frac{1}{4n\sqrt{n}} = \frac{2n-2}{4n\sqrt{n}} \leq \frac{2n}{4n\sqrt{n}} = \frac{1}{2\sqrt{n}}$.

Case 2: $a_2 \notin A_1$.

The probability that case 2 happens $= 1 - \frac{1}{4n\sqrt{n}}$.

By the assumption that $a_2 - a_1 \pmod 1 \leq \frac{1}{2}$, there is $A_1 \cap A_2 = \emptyset$.

$\Pr[x < \frac{1}{4n\sqrt{n}} \wedge y < \frac{1}{4n\sqrt{n}}] = \Pr[\bigcup_{i=3}^{2n} \bigcup_{j=i+1}^{2n} (a_i \in A_1 \wedge a_j \in A_2) \vee (a_i \in A_2 \wedge a_j \in A_1)] \leq \sum_{i=3}^{2n} \sum_{j=i+1}^{2n} \Pr[(a_i \in A_1 \wedge a_j \in A_2) \vee (a_i \in A_2 \wedge a_j \in A_1)] = \sum_{i=3}^{2n} \sum_{j=i+1}^{2n} 2 \times \frac{1}{4n\sqrt{n}} \times \frac{1}{4n\sqrt{n}} = \binom{2n-2}{2} \times 2 \times \frac{1}{16n^3} \leq \frac{4n^2}{16n^3} = \frac{1}{4n}$.

$\therefore \Pr[x + y < \frac{1}{4n\sqrt{n}}] \leq \frac{1}{4n\sqrt{n}} \times \frac{1}{2\sqrt{n}} + (1 - \frac{1}{4n\sqrt{n}}) \times \frac{1}{4n} \leq \frac{1}{8n^2} + \frac{1}{4n} \leq \frac{1}{2n}$.

$\Pr[$ the size of the smallest interval $< \frac{1}{4n\sqrt{n}}]$

$= \Pr[\bigcup_{i=1}^{n}($ the size of the interval of the i-th slot $< \frac{1}{4n\sqrt{n}})]$

$\leq \sum_{i=1}^{n} \Pr[$ the size of the interval of the i-th slot $< \frac{1}{4n\sqrt{n}}]$

$\leq \sum_{i=1}^{n} \frac{1}{2n} = \frac{1}{2}$.

$\therefore$ the size of the smallest interval is at least $\frac{1}{4n\sqrt{n}}$ with probability $\geq 1 - \frac{1}{2} = \frac{1}{2} = \Omega(1)$.