2020 FALL MAS583 HW1

Fanchen Bu [20194185]

September 10, 2020

1 Problem 1

1.1 Question

Prove that if there is a real p, $0 \leq p \leq 1$ such that

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t} (1-p)^{\binom{t}{2}} < 1,$$

then the Ramsey number $R(k,t)$ satisfies $R(k,t) > n$. Using this, show that

$$R(4,t) \geq \Omega(t^{3/2}/(\ln t)^{3/2}).$$

1.2 Answer

With the assumption, consider K_n and color each edge independently either red (with probability p) or blue (with probability $1-p$).
Arbitrarily choose $K \subset V(K_n)$ with $|K| = k$, let A_K denote the event that the k vertices in K form a monochromatic red K_k, clearly, $\Pr[A_K] = p^{\binom{k}{2}}$. Thus

$$\Pr[\bigvee_K A_K] \leq \sum_K \Pr[A_K] = \binom{n}{k} p^{\binom{k}{2}}.$$

Similarly, if we arbitrarily choose $T \subset V(K_n)$ with $|T| = t$, let B_T denote the event that the t vertices in T form a monochromatic blue K_t, we have

$$\Pr[\bigvee_T B_T] \leq \sum_T \Pr[B_T] = \binom{n}{t} (1-p)^{\binom{t}{2}}.$$

Then the probability of event that at least one monochromatic red K_k or one monochromatic blue K_t exists is

$$\Pr[\bigvee_K A_K \vee \bigvee_T B_T] \leq \Pr[\bigvee_K A_K] + \Pr[\bigvee_T B_T] \leq \binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t} (1-p)^{\binom{t}{2}} < 1,$$

that means with positive probability, there is a 2-coloring of K_n without any monochromatic red K_k or monochromatic blue K_t, hence $R(k,t) > n$.

Let $k = 4$, $n = c(\frac{t}{\ln t})^{3/2}$, $p = n^{-2/3}$, where c is a small constant number, say, $c \leq 3^{-3/2}$. Then with t big enough, say, $t \geq 10^6$, we have

$$\begin{aligned}
\binom{n}{4}p^6 + \binom{n}{t}(1-p)^{\binom{t}{2}} &\leq \frac{n^4p^6}{24} + (\frac{en}{t})^t e^{-pt(t-1)/2} \\
&\leq \frac{1}{24} + (\frac{cet^{1/2}}{\ln t^{3/2}})^t t^{-\frac{1}{2}(t-1)c^{-2/3}} \\
&\leq \frac{1}{24} + (\frac{cet^{1/2}}{\ln t^{3/2}})^t t^{-\frac{3}{2}(t-1)} \\
&\leq \frac{1}{24} + (\frac{cet^{1/2}}{\ln t^{3/2}})^t (100cet)^{-t} \\
&\leq \frac{1}{24} + (\frac{1}{100(\ln t)^{3/2}t^{1/2}})^t \\
&< 1
\end{aligned}$$

Thus, $R(4, t) \geq \Omega(t^{3/2}/(\ln t)^{3/2})$.

2 Problem 2

2.1 Question

Suppose $n \geq 4$ and let H be an n-uniform hypergraph with at most $4^{n-1}/3^n$ edges[1]. Prove that there is a coloring of the vertices of H by four colors so that in every edge all four colors are represented.

2.2 Answer

Arbitrarily choose an edge v in H, for the n vertices in it, we color each of them by four colors independently, with equal probability of each color at random. Let A_v denote the event that not all four colors are represented in v, clearly, $\Pr[A_v] = 3^n/4^{n-1}$. Thus

$$\Pr[\bigvee_v A_v] \leq \sum_v \Pr[A_v] < \frac{4^{n-1}}{3^n} \cdot \frac{3^n}{4^{n-1}} = 1,$$

that means with positive probability, there is a coloring of the vertices of H by four colors so that in every edge all four colors are represented.

3 Problem 3

3.1 Question

(*) Prove that for every two independent, identically distributed real random variables X and Y,

$$\Pr[|X - Y| \leq 2] \leq 3 \Pr[|X - Y| \leq 1].$$

[1] As $4^{n-1}/3^n$ is not an integer when $n \geq 4$, we can safely say that $|V(H)| < 4^{n-1}/3^n$.

3.2 Answer

The proof basically follows the paper *The 123 Theorem and Its Extensions* by Noga Alon and Raphael Yuster.
Let $T = (x_1, x_2, ..., x_m)$ be a sequence of not necessarily distinct reals. For any positive b, define

$$T_b = \{(x_i, x_j) : 1 \le i, j \le m, |x_i - x_j| \le b\}.$$

Lemma 3.1 *For any sequence T as above and for every integer $r > 1$.*

$$|T_r| < (2r-1)|T_1|.$$

Proof. We apply induction on $|T| = m$. The result is trivial for $m = 1$. Assuming it holds for $m-1$, we prove it for m (> 1). Given a sequence $T = (x_1, x_2, ..., x_m)$, let $t+1$ be the maximum number of points of T in a closed interval of length 2 centered at a member of T. Let x_i be any rightmost[2] point of T so that there are $t+1$ members of T in the interval $[x_i - 1, x_i + 1]$ and define $T' = T \setminus \{x_i\}$. The number of members of T' in the interval $[x_i - 1, x_i + 1]$ is clearly t and, hence x_i appears in precisely $2t+1$ ordered pairs of T_1 ((x_i, x_i), (x_i, x') and (x', x_i) where x' can be chosen from the t members of T' in the interval $[x_i - 1, x_i + 1]$). Thus

$$|T_1| = 2t + 1 + |T_1'|.$$

The interval $[x_i - r, x_i + r]$ is the union of the $2r - 1$ smaller intervals

$$[x_i - r, x_i - r + 1), ..., [x_i - 2, x_i - 1), [x_i - 1, x_i + 1],$$

$$(x_i + 1, x_i + 2], ..., (x_i + r - 1, x_i + r].$$

By the choice of x_i, each of these smaller intervals (the first r ones) can contain at most $t+1$ members of T, and each of the last $r-1$ ones, which lie to the right of x_i, can contain at most t members of T. Altogether there are thus at most $(r-1)(t+1) + rt$ members of T' in $[x_i - r, x_i + r]$ and, hence,

$$|T_r| \le 2(r-1)(t+1) + 2rt + 1 + |T_r'| = (2r-1)(2t+1) + |T_r'|.$$

By the induction hypothesis $|T_r'| < (2r-1)|T_1'|$ and hence

$$|T_r| \le (2r-1)(2t+1) + |T_r'| < (2r-1)(2t+1) + (2r-1)|T_1'| = (2r-1)|T_1|,$$

completing the proof. ■

Corollary 3.2 *Let X and Y be two i.i.d. real random variables. For a positive b, define $p_b = \Pr[|X - Y| \le b]$. Then for every integer r, $p_r \le (2r-1)p_1$.*

Proof. Fix an integer m, and let $S = (x_1, ..., x_m)$ be a random sequence of m elements, where each x_i is chosen, randomly and independently, according to the distribution of X. By Lemma 3.1

$$|S_r| < (2r-1)|S_1|.$$

Therefore, the expectation of $|S_r|$ is smaller than that of $(2r-1)|S_1|$. However, by the linearity of expectation it follows that the expectation of $|S_b|$ is precisely $m + m(m-1)p_b$ for

[2] Here, "rightmost" means $\forall x \in T$, if $x > x_i$, then the interval $[x-1, x+1]$ contains at most t members in T.

every positive b (there are m tuples where both elements are identical, which are trivially in S_b, and the other $m(m-1)$ tuples have a probability of p_b to be in S_b). Therefore,

$$m + m(m-1)p_r < (2r-1)(m + m(m-1)p_1),$$

implying that for every integer m, $p_r < (2r-1)p_1 + \frac{2r-2}{m-1}$. The desired result $p_r \le (2r-1)p_1$ follows, by letting m tend to infinity. ■

With Corollary 3.2, we let $r = 2$ and complete the proof.

4 Problem 4

4.1 Question

(*) Let $G = (V, E)$ be a graph with n vertices and minimum degree $\delta > 10$. Prove that there is a partition of V into two disjoint subsets A and B so that $|A| \le O(n \ln \delta/\delta)$, and each vertex of B has at least one neighbor in A and at least one neighbor in B.

4.2 Answer

Form a random vertex subset S in such a graph by including each vertex independently with probability $p = \frac{\ln \delta}{\delta}$. Given S, let T be the set of vertices outside S having no neighbors in S, and let U be the set of vertices outside $S \cup T$ whose neighbors are all in $S \cup T$. Adding T and U to S yields a desired set A. We seek the expected number of $|S \cup T \cup U|$.

Since each vertex appears in S with probability p, linearity yields $\mathbb{E}[S] = np$. The random variable $|T|$ is the sum of n indicator random variables for whether individual vertices belong to T, likewise the random variable $|U|$ is the sum of n indicator random variables for whether individual vertices belong to U. We have $v \in T$ if and only if v and its neighbors all fail to be in S. This has probability at most $(1-p)^{\delta+1}$. We have $v \in U$ if and only if (1) v fails to be in $S \cup T$; (2) all the neighbors of v are in $S \cup T$, where $\Pr[v \notin (S \cup T)] \le \Pr[v \notin S] \le 1 - p$, and for each neighbor v' of v, $\Pr[v' \in (S \cup T)] \le \Pr[v' \in S] + \Pr[v' \in T] \le p + (1-p)^{\delta+1}$, thus $\Pr[v \in U] \le (1-p)(p + (1-p)^{\delta+1})^{\delta}$. (? It looks like that the two subevents are not independent so maybe this cannot hold.) Now we have

$$\begin{aligned}
\mathbb{E}[|S| + |T| + |U|] &\le n(p + (1-p)^{\delta+1} + (1-p)(p + (1-p)^{\delta+1})^{\delta}) \\
&\le n \ln \delta/\delta + n(1-p)((1-p)^{\delta} + (p + (1-p)^{\delta+1})^{\delta}) \\
&\le n \ln \delta/\delta + n((1-p)^{\delta} + (p + (1-p)^{\delta+1})^{\delta}) \\
&\le n \ln \delta/\delta + ne^{-p\delta} + n(p + (1-p)^{\delta+1})^{\delta} \\
&\le n \ln \delta/\delta + n/\delta + n(p + (1-p)^{\delta+1}) \\
&\le 2n \ln \delta/\delta + n/\delta + n(1-p)^{\delta} \\
&\le 2n \ln \delta/\delta + 2n/\delta \\
&\le O(n \ln \delta/\delta),
\end{aligned}$$

which completes the proof.

4.3 A Failed Attempt

Form a random vertex subset S in such a graph by including each vertex independently with probability $p = \frac{\ln \delta}{\delta}$. Given S, let T be the set of vertices outside S having no neighbors in S

or having no neighbors in $(V \setminus S)$. Adding T to S yields a desired set A. (WRONG!!! Some vertices in $(V \setminus (S \cup T))$ may only have neighbors in T!!!) We seek the expected number of $|S \cup T|$.
Since each vertex appears in S with probability p, linearity yields $\mathbb{E}[S] = np$. The random variable $|T|$ is the sum of n indicator random variables for whether individual vertices belong to T. We have $v \in T$ if and only if (1) $v \notin S$, and (2) all the neighbors of v are in S or all the neighbors of v are in $(V \setminus S)$. This has probability at most $(1-p)[(1-p)^\delta + p^\delta]$, thus we have

$$\begin{aligned}\mathbb{E}[|S| + |T|] &\leq np + n(1-p)(e^{-p\delta} + p^\delta) \\ &\leq n \ln \delta/\delta + n[1/\delta + (\ln \delta/\delta)^{10}] \\ &\leq O(n \ln \delta/\delta),\end{aligned}$$

which completes the proof.

5 Problem 5

5.1 Question

(*) Let $G = (V, E)$ be a graph on $n \geq 10$ vertices and suppose that if we add to G any edge not in G then the number of copies of a complete graph on 10 vertices in it increases. Show that the number of edges in G is at least $8n - 36$.

5.2 Answer

The condition means for any two vertices, say v_1 and v_2, who are not connected with each other, there exist 8 other vertices, say $v_3, v_4, ..., v_{10}$, such that the vertices in the set $\{v_i\}_{i=1}^{10}$ are pairwise connected, except for v_1 and v_2, as we specified. We want to prove the number of such (v_1, v_2) tuples are at most $\binom{n}{2} - (8n - 36) = (n^2 - 17n + 72)/2 = \binom{n-8}{2}$. To prove this, we'll use Theorem 1.3.3 on the textbook, which goes as follows.
Let $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ be a family of pairs of subsets of the set of an arbitrary set. If $|A_i| = k$ and $|B_i| = l$ for all $1 \leq i \leq h$, $A_i \cap B_i = \emptyset$ and $A_i \cap B_j \neq \emptyset$ for all $i \neq j$, $1 \leq i, j \leq h$, then $h \leq \binom{k+l}{k}$.
For our case, we need such $\mathcal{F}$ with $k = 2$ and $l = n - 10$, which can be constructed as follows. Let $\mathcal{A}$ be the set $\{S : S \subset V, |S| = 2, \text{the two vertices in } S \text{ are not connected}\}$, for each $A_i \in \mathcal{A}$, let C_i be such 8 vertices which form a K_8 in G and all connect with both vertices in A_i, then these 8 vertices together with the two vertices in A_i can form a K_{10} by just connecting the two vertices in A_i, the condition guarantees that such vertices exist. If more than one group of such vertices exist, we randomly choose one group. Then we let B_i be $(V \setminus (A_i \cup C_i))$, and we have a $\mathcal{F}$ whose size is $|\mathcal{A}|$, for all i, $|A_i| = 2$, $|B_i| = n - 10$, clearly, $A_i \cap B_i = \emptyset$. Now we only need to check the last property, which is also obvious as if $A_i \cap B_j = \emptyset$ for some $i \neq j$, it means both vertices in A_i are in $A_j \cup C_j$, however, as we constructed, among the 10 vertices in $A_j \cup C_j$, only the two vertices in A_j are disconnected, besides, A_i and A_j cannot be same as they are two elements in a set. Therefore, $A_i \cap B_j \neq \emptyset$ holds for all $i \neq j$, then we are able to apply the theorem and complete the proof.

5.3 A Failed Attempt

The condition means any 10 vertices in G with the corresponding edges among them should form a K_{10} or only need one more additional edge to form a K_{10}. (WRONG!!!) For $n = 10$, the number of edges $|E| \geq \binom{10}{2} - 1 = 44 = 8n - 36$. Besides, clearly, if G satisfies such condition, then after removing any vertex from G (as long as the number of remaining vertices ≥ 10), the remaining graph should also satisfy the condition, which implies that the number of edges among any 10 vertices in G should be either 44 or 45. Assume the proposition holds for all $n \leq k$, then for $n = k + 1$, if we can prove that at least 8 new edges are needed, the prove will be completed by using induction. Actually, if we view the situation when $n = k + 1$ as adding a new vertex into a graph with k vertices satisfying the condition, then we arbitrarily choose 9 vertices from the original graph together with the new vertex to form a 10-vertex set, as we mentioned, at least 44 edges are needed for this 10-vertex set, while the number of edges among the 9 vertices from the original graph ≤ 36, which is the number of edges in a K_9, thus, at least 8 new edges are needed to satisfy the condition, which completes the proof.

5.4 Another Failed Attempt

It's easy to find a construction with $8n - 36$ edges which satisfies the condition. First we have a K_8, then for the remaining $n - 8$ vertices, we let each the them only have 8 neighbors, which are the 8 vertices in the K_8, then the total edge number is $\binom{8}{2} + 8(n - 8) = 8n - 36$. Intuitively, if for some vertex v in the remaining $n - 8$ vertices, we want to delete some connection between it and the K_8, say, delete m edges, then we'll need m other vertices in the remaining ones to connect with v, then the total edge number cannot decrease. However, I can hardly compose a clear mathematical proof for this idea.

6 Problem 6

6.1 Question

(*) Theorem 1.2.1 asserts that for every integer $k > 0$ there is a tournament $T_k = (V, E)$ with $|V| > k$ such that for every set U of at most k vertices of T_k there is a vertex v so that all directed arcs $\{(v, u) : u \in U\}$ are in E. Show that each such tournament contains at least $\Omega(k2^k)$ vertices.

6.2 Answer

The proof basically follows the paper *On A Problem of Schütte and Erdös* by E. Szekeres and G. Szekeres.

Consider a tournament T on the set $V = \{1, ..., n\}$, let $A = \{v_1, v_2, ..., v_{k-1}\}$ be any subset of V with size $k - 1$, and let $G(A)$ denote the set $\{v \in V : (v, v_i) \in T, \forall v_i \in A\}$ (the set of the players beat all players in A). For a single vertex v, let $G(v)$ denote $G(\{v\})$. We claim that if T has property S_k, then $|G(A)| \geq |A| + 2 = k + 1$. Otherwise, suppose that $|G(A)| \leq k$, as T has property S_k, $\exists v' \in V$ such that $v' \in G(G(A))$, i.e. $(v', v) \in T, \forall v \in G(A)$. Meanwhile, $\exists v'' \in V$ such that $v'' \in G(A \cup \{v'\})$ which implies both $v'' \in G(A)$ and $v'' \in G(v')$. However, we also have $v' \in G(G(A))$, thus we have v' and v'' beat each other, which is impossible.

We define another property $R_{k,m}$ for $m > 0, k \geq 0$. We say a tournament T on the set

$V = \{1, ..., n\}$ has property $R_{k,m}$ if for every subset $A \subset V$ with $|A| = k$, $|G(A)| \geq m$. Some special cases are described as follows: when $k = 0$, T has property $R_{0,m}$ if and only if $|V| = n \geq m$; when $m = 1$, $R_{k,1}$ is identical with S_k. With this new definition, what our previous claim says is each tournament with property S_k also has property $R_{k-1,k+1}$.

Lemma 6.1 *Consider a tournament T on the set $V = \{1, ..., n\}$. If T has property $R_{k,m}$, $(k > 0, m > 0)$, then the subtournament*[3] *on $G(v)$ for each $v \in V$ has property $R_{k-1,m}$. If T has property S_k and $R_{k-1,m}$, then the subtournament on $G(v)$ for each $v \in V$ has property S_{k-1} and $R_{k-2,m}$.*

Proof. Suppose T has property $R_{k,m}$, let v_0 be any vertex in V, and let $\{v_1, v_2, ..., v_{k-1}\}$ be any set of $k-1$ vertices of $G(v_0)$ (as we are talking about the property $R_{k,m}$ of the subtournament on $G(v_0)$, we naturally presume that $|G(v_0)| \geq k$, so such subset exists. However, I want to ask what if $|G(v_0)| < k - 1$? Or it's okay as we are going to use it only for cases where $m \geq k > 0$?). Then by $R_{k,m}$, $|G(\{v_0, v_1, ..., v_{k-1}\})| \geq m$, while $G(\{v_0, v_1, ..., v_{k-1}\}) \subset G(v_0)$ and for each $v' \in G(\{v_0, v_1, ..., v_{k-1}\})$, $v' \in G(v_i)$, $\forall 1 \leq i \leq k-1$, which implies $|G(\{v_1, v_2, ..., v_{k-1}\}| \geq m$, completing the proof of the first proposition.
Suppose T has property S_k and $R_{k-1,m}$, $m > k > 1$, let v_0 be any vertex in V. By the previous proof, the subtournament on $G(v_0)$ has property $R_{k-2,m}$, so we only need to show that it also has property S_{k-1} (here, with $m > k > 1$, $|G(v_0)| \geq m > k-1$ is guaranteed). Again, let $\{v_1, v_2, ..., v_{k-1}\}$ be any set of $k-1$ vertices of $G(v_0)$. By S_k, $\exists v' \in V$ such that $v' \in G(\{v_0, v_1, ..., v_{k-1}\})$, which implies $v' \in G(v_0)$ and $v' \in G(\{v_1, v_2, ..., v_{k-1}\})$, completing the proof. ■

Lemma 6.2 *If a tournament T on $V = \{1, ..., n\}$ has property $R_{k,m}$, then*

$$|V| = n \geq F(k, m) = 2^k(m + 1) - 1.$$

Proof. It's trivial when $k = 0$ regardless of the value of m. Suppose that $t > 0$ and the proposition is true for $k = t - 1$ for all $m \geq k$ (I modified the proof with $m \geq k$ to avoid my above problem). Suppose T has property $R_{k,m}$, let v_0 be any vertex in V, by Lemma 6.1 and the induction hypothesis $|G(v_0)| \geq F(k-1, m)$, thus the total number of edges in T is at least $nF(k-1, m)$. Hence

$$n(2^{k-1}(m + 1) - 1) \leq \binom{n}{2},$$

which means

$$n \geq 2^k(m + 1) - 1,$$

completing the proof. ■
Now, by our first claim, each tournament T on the set $V = \{1, ..., n\}$ with property S_k also has property $R_{k-1,k+1}$, then by Lemma 6.2, we have $n \geq F(k-1, k+1) = 2^{k-1}(k+2)-1 \geq \Omega(k2^k)$, completing the proof.

[3] By "subtournament on $G(v)$", we mean the subgraph of T only keeping the node set $G(V)$ and the directed edges among them, which is naturally also a tournament.

7 Problem 7

7.1 Question

Let $\{(A_i, B_i), 1 \le i \le h\}$ be a family of pairs of subsets of the set of integers such that $|A_i| = k$ for all i and $|B_i| = l$ for all i, $A_i \cap B_i = \emptyset$ and $(A_i \cap B_j) \cup (A_j \cap B_i) \ne \emptyset$ for all $i \ne j$. Prove that $h \le (k+l)^{k+l}/(k^k l^l)$.

7.2 Answer

Put $X = \bigcup_{i=1}^{h}(A_i \cup B_i)$ and consider a 2-coloring for numbers in X where each number is colored red with probability of $\frac{k}{k+l}$, colored blue with probability of $\frac{l}{k+l}$, randomly and independently. For each i, $1 \le i \le h$, let X_i be the event that all the numbers of A_i are red and all the numbers of B_i are blue. Clearly $\Pr[X_i] = k^k l^l/(k+l)^{k+l}$. Besides, we can check that the events X_i are pairwise disjoint, otherwise, say, both X_i and X_j hold, where $i \ne j$, then the given condition $(A_i \cap B_j) \cup (A_j \cap B_i) \ne \emptyset$ cannot hold as the intersection of sets with different colors is definitely $\emptyset$. Then we have

$$1 \ge \Pr[\bigvee_{i=1}^{h} X_i] = \sum_{i=1}^{h} \Pr[X_i] = hk^k l^l/(k+l)^{k+l},$$

which completes the proof.

8 Problem 8

8.1 Question

(Prefix-free codes; Kraft Inequality). Let F be a finite collection of binary strings of finite lengths and assume no member of F is a prefix of another one. Let N_i denote the number of strings of length i in F. Prove that

$$\sum_i \frac{N_i}{2^i} \le 1.$$

8.2 Answer

Let's think about the construction of the strings in F in the increasing order of length. When it comes to the strings of length j, each string should not contain any existing strings of length $< j$ as its prefix, i.e., among all the 2^j trivial choices, each previous string of length k would forbid 2^{j-k} choices, besides, as we assume no member of F is a prefix of another one, these forbidden sets are disjoint. Thus, for strings of length j, the size of feasible set is

$$2^j - \sum_{i=1}^{j-1} N_i 2^{j-i},$$

which should satisfy

$$2^j - \sum_{i=1}^{j-1} N_i 2^{j-i} \ge N_j.$$

We can rewrite it as

$$1 \geq \sum_{i=1}^{j} \frac{N_i}{2^i},$$

which completes the proof when j is the max length of strings in F.

9 Problem 9

9.1 Question

(*) (Uniquely decipherable codes; Kraft-McMillan Inequality). Let F be a finite collection of binary strings of finite lengths and assume that no two distinct concatenations of two finite sequences of codewords result in the same binary sequence. Let N_i denote the number of strings of length i in F. Prove that

$$\sum_i \frac{N_i}{2^i} \leq 1.$$

9.2 Answer

The proof basically follows a common proof of Kraft-McMillan Inequality.
Rewrite $\sum_i \frac{N_i}{2^i}$ as $\sum_{f\in F} 2^{-L(f)}$, where $L(f)$ denote the length of string f. For any positive integer k, we have

$$\begin{aligned}
(\sum_{f\in F} 2^{-L(f)})^k &= (\sum_{f_1\in F} 2^{-L(f_1)})(\sum_{f_2\in F} 2^{-L(f_2)})...(\sum_{f_k\in F} 2^{-L(f_k)}) \\
&= \sum_{f_1}\sum_{f_2}\cdots\sum_{f_k}\prod_{i=1}^{k} 2^{-L(f_i)} \\
&= \sum_{f_1,\dots,f_k} 2^{-\sum_{i=1}^{k} L(f_i)} \\
&= \sum_{f_1,\dots,f_k} 2^{-L(f_1 f_2 \dots f_k)}
\end{aligned}$$

Here, $f_1 f_2 ... f_k$ means the concatenation of strings f_1 to f_k. From the given assumption, we know $(f_1, f_2, ..., f_k)$ are the unique tuple of strings in F whose concatenation is $f_1 f_2 ... f_k$, which allows us to claim

$$(\sum_{f\in F} 2^{-L(f)})^k = \sum_{i=1}^{kL^*} \frac{\tilde{N}_i}{2^i} \leq \sum_{i=1}^{kL^*} \frac{2^i}{2^i} = kL^*,$$

where L^* is the maximum length of the strings in F, and $\tilde{N}_i$ is the number of i-length k-concatenation (the concatenation of k strings in F whose total length is k) of strings in F, which is naturally $\leq 2^i$. Thus,

$$\sum_{f\in F} 2^{-L(f)} \leq (kL^*)^{\frac{1}{k}},$$

holds for any positive k, which means

$$\sum_{f\in F} 2^{-L(f)} \leq 1,$$

by letting k tend to infinity, completing the proof.

10 Problem 10

10.1 Question

Prove that there is an absolute constant $c > 0$ with the following property. Let A be an n by n matrix with pairwise distinct entries. Then there is a permutation of the rows of A so that no column in the permuted matrix contains an increasing subsequence of length at least $c\sqrt{n}$.

10.2 Answer

Let $l = c\sqrt{n}$, where c is a constant big enough, say $c \geq 1000$. The proposition is trivial when $c\sqrt{n} > n$, i.e., when $n < c^2$, so we only need to care about $n \geq c^2 \geq 10^6$. Let π be a random permutation of rows of A, and X_i denote the event that the ith column contains an increasing subsequence of length at least l after permutation, then

$$
\begin{aligned}
\Pr[\bigvee_i X_i] &\leq \sum_i \Pr[X_i] \\
&\leq n\binom{n}{l}\frac{1}{l!} \\
&\leq n(\frac{ne}{l})^l \frac{1}{\sqrt{2\pi l}(\frac{l}{e})^l} \\
&\leq n(\frac{e}{c})^{2l}/\sqrt{2\pi l} \\
&\leq n(\frac{e}{c})^{2l}/10^3 \\
&\leq 10^{-3}n/300^{2000\sqrt{n}} \\
&< 1,
\end{aligned}
$$

which implies that with positive probability, such permutation exists when no column in the permuted matrix contains an increasing subsequence of length at least $l = c\sqrt{n}$.