Proof by Existence--AnotherExample

- Consider the following claim.
- There is a bipartite graph G = (L, R, E) such that
  - |L| = n
  - |R| = $2^{\log^2 n}$
  - Every subset of n/2 vertices of L has at least $2^{\log^2 n} - n$ neighbors in R.
  - No vertex of R has more than $12\log^2 n$ neighbors.
- We want to use the technique of proof by existence to show the above claim.

Proof by Existence--AnotherExample

- There is a bipartite graph $G = (L, R, E)$ such that
  - $|L| = n$, $|R| = 2^{\log^2 n}$. Every subset of $n/2$ vertices of L has at least $2^{\log^2 n} - n$ neighbors in R. No vertex of R has more than $12\log^2 n$ neighbors.
  - Let every vertex of L choose d neighbors in R independently and uniformly at random.
  - Choices are made with replacement.
  - Multiple edges are dropped in favor of one edge.

Proof by Existence--AnotherExample

- There is a bipartite graph G = (L, R, E) such that
  - $|L| = n$, $|R| = 2^{\log^2 n}$. Every subset of n/2 vertices of L has at least $2^{\log^2 n} - n$ neighbors in R. No vertex of R has more than $12\log^2 n$ neighbors.
  - Let every vertex of L choose d neighbors in R independently and uniformly at random.
  - Let us now estimate the degree of any vertex of R.
  - Let $|R| = r$.
  - We can think of the degree of a vertex v in R as the expectation of the random variable X that indicates how many vertices in L choose v as a neighbor.
  - Each neighbor in L makes d choices, so we have nd choices in all.

$$\Pr(X \geq 12\log^2 n) = \Pr(X \geq (1+2)\,4\log^2 n) \leq e^{-\mu\delta\ln\delta} = e^{-4\log^2 n \times 2\ln 2}$$

$\delta > 1$

$\Pr(\text{some } u \text{ has deg} \geq 12\log^2 n)$

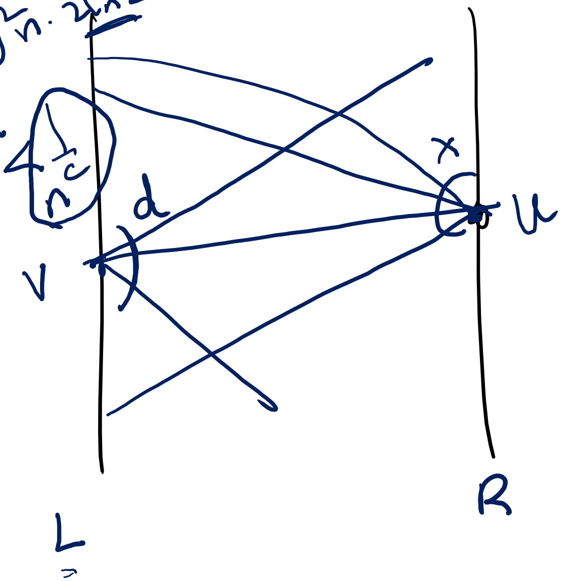
$\leq |R| \times e^{-4\log^2 n \cdot 2\ln 2}$

$\leq e^{-3\log^2 n}$

$\leq \frac{1}{n^c}$

$|R| = 2^{\log^2 n} < e^{\log^2 n}$

$\Pr\left(\bigcup_{u\in R} X_u\right)$



$\deg(u) \leq 12\log^2 n$ whp

$$X_{vi} = \begin{cases} 1 & v \text{ in } i^{th} \text{ choice picks } u \\ 0 & o/w \end{cases}$$

$$E(X_{vi}) = \frac{1}{|R|} \qquad EX = |L| \times d \times \frac{1}{|R|}$$

degree of $u \leftarrow X_u$

$$X_u = \sum_{v\in L}\sum_{i=1}^{d} X_{v,i}$$

$$|L| \times d \times \frac{1}{|R|} = 4\log^2 n$$

$$d = \frac{|R|}{|L|} \times 4\log^2 n$$

Proof by Existence--AnotherExample

- There is a bipartite graph $G = (L, R, E)$ such that
  - $|L| = n$, $|R| = 2^{\log^2 n}$, . Every subset of $n/2$ vertices of L has at least $2^{\log^2 n} - n$ neighbors in R. No vertex of R has more than $12\log^2 n$ neighbors.
  - Let every vertex of L choose d neighbors in R independently and uniformly at random.
  - Let $|R| = r$.
  - We can think of the degree of a vertex v in R as the expectation of the random variable X that indicates how many vertices in L choose v as a neighbor.
  - Each neighbor in L makes d choices, so we have nd choices in all.
  - Let Xi be a random variable if the ith choice is v.

Proof by Existence--AnotherExample

- Let $|R| = r$.
- We can think of the degree of a vertex v in R as the expectation of the random variable X that indicates how many vertices in L choose v as a neighbor.
- Each neighbor in L makes d choices, so we have nd choices in all.
- Let Xi be a random variable if the ith choice is v.
- $E[Xi] = 1/r$.
- $X = \Sigma\, Xi$ and so $E[X] = \Sigma\, E[Xi] = nd/r$.
- Pick $d = r.2\log^2 n /n$ so that $E[X] = 2\log^2 n$.
- Now apply Chernoff bounds on X for the event $X >= 12\log^2 n$.
- Use Boole's inequality to bound the probability of the bad event for every v in R.

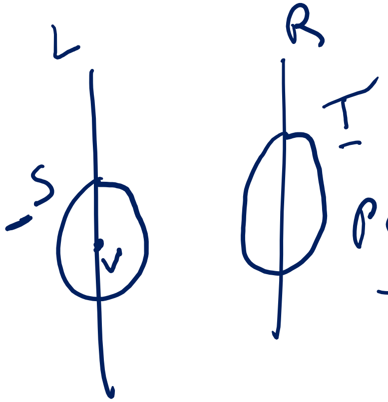
Proof by Existence--AnotherExample

- There is a bipartite graph $G = (L, R, E)$ such that
  - $|L| = n$, $|R| = 2^{\log^2 n}$, . Every subset of $n/2$ vertices of L has at least $2^{\log^2 n} - n$ neighbors in R.
  - Let every vertex of L choose d neighbors in R independently and uniformly at random.
  - We now move to property 1.
  - Let S be any subset of size $n/2$ from L.
  - Let T be any subset of R of size $2^{\log^2 n} - n$.
  - Consider the event that all the neighbors of S are in T.
  - This happens with a probability of $\left[(2^{\log^2 n} - n)/r\right]^{nd/2}$.

$|R| = 2^{\log^2 n}$

Lem: Every subset of L of size $\leq n/2$ has $\geq 2^{\log^2 n} - n$ nbrs in R.

$|T| = |R| - n$



Proof:

Pick any subset S of size $n/2$ in L

Pick any subset T of size $2^{\log^2 n} - n$ in R

$E_{ST} := \{N(S) \subseteq T\}$

$$\Pr(E_{ST}) = \left(\frac{|T|}{|R|}\right)^{d\frac{n}{2}} = \left(1 - \frac{n}{|R|}\right)^{d \cdot n/2}$$

$\#$ ways of choosing $S = \binom{n}{n/2}$

$\#$ ways of choosing $T = \binom{|R|}{|R|-n} = \binom{|R|}{n}$

$E = \bigcup_S \bigcup_T E_{ST}$

$$\Pr(E) \leq \binom{n}{n/2} \times \binom{|R|}{n} \times \left(1 - \frac{n}{|R|}\right)^{d n/2}$$

$$\leq (2e)^{n/2} \times \left(e\frac{|R|}{n}\right)^{n} \times e^{-n^2 d/2|R|}$$

$$d = \frac{|R|}{|L|} \cdot 4 \log^2 n = \frac{|R|}{n} \cdot 4 \log^2 n$$

$$\Pr(E) \leq (2e)^{n/2} \left(\frac{e|R|}{n}\right)^n \cdot e^{-n^2 d/2|R|}$$

$$= (2e)^{n/2} \left(\frac{e\,2^{\log^2 n}}{n}\right)^n e^{-\frac{n}{2} \cdot 4 \log^2 n}$$

$2^{\log^2 n} = n^{\log n} = e^{\log^2 n}$

$$\leq e^{\frac{n}{2} + \frac{n}{2} + n + n\log^2 n - n \ln n - 2n\log^2 n}$$

$n = e^{\log n}$

$$< e^{-\frac{n}{2}\log^2 n} \leq \frac{1}{n^{O(1)}}$$

$\Pr(\bar{E}) > 0$

Proof by Existence--AnotherExample

- There is a bipartite graph G = (L, R, E) such that
  - $|L| = n$, $|R| = 2^{\log^2 n}$, . Every subset of n/2 vertices of L has at least $2^{\log^2 n} - n$ neighbors in R.
  - Let S be any subset of size n/2 from L.
  - Let T be any subset of R of size $2^{\log^2 n} - n$.
  - Consider the event that all the neighbors of S are in T.
  - This happens with a probability of $\left[(r - n)/r\right]^{nd/2}$.
  - Now, consider all possible choices of S and T. The probability that for any S all its neighbors are in some T is upper bounded by: ${}^{n}C_{n/2} \,.\, {}^{r}C_{r-n} \,.\, \left[(r - n)/r\right]^{nd/2}$.
  - We will now show that the above probability is strictly less than 1.

Proof by Existence--AnotherExample

- Now, consider all possible choices of S and T. The probability that for any S all its neighbors are in some T is upper bounded by: $^{n}C_{n/2} \,.\, ^{r}C_{r-n} \,.\, \left[(r - n)/r\right]^{nd/2}$.
  - We will now show that the above probability is at most 1.
  - Use the (in)equalities
    - $^{n}C_{n-k} = {}^{n}C_{k}$ for k between 0 and n.
    - $^{N}c_{k}$ is at most $(en/k)^{k}$.
    - (1+x) is at most e^{x} for any real number x.
  - The required probability is
  - $(2e)^{n/2} . (er/n)^{n} . (e)^{-n^2 d/2r}$.
  - Recall that $d = 2\log^{2}n \,.\, r/n$.

Proof by Existence--AnotherExample

- Now, consider all possible choices of S and T. The probability that for any S all its neighbors are in some T is upper bounded by: ${}^{n}C_{n/2} \,.\, {}^{r}C_{r-n} \,.\, \left[(r - n)/r\right]^{nd/2}$.
  - We will now show that the above probability is at most 1.
  - Use the (in)equalities
    - ${}^{n}C_{n-k} = {}^{n}C_{k}$ for k between 0 and n.
    - ${}^{N}c_{k}$ is at most $(en/k)^{k}$.
    - $(1+x)$ is at most e^{x} for any real number x.
  - The required probability is
  - $(2e)^{n/2}. (er/n)^{n}. (e)^{-n^2 d/2r}$.
  - Recall that $d = 2\log^2 n \,.\, r/n$ and $\log r = \log^2 n$ to simplify.

Yet Another Example

- Consider a Boolean formula in $\boxed{\text{CNF.}}$
- In CNF, each clause is a disjunction of literals
- The formula is a conjunction of clauses.
- Another name for CNF is Product-of-Sums.

$E X_1 \geq \frac{1}{2}$ $\quad E X \geq \frac{m}{2}$

$X = \sum X_i$

P-o-S

x_1	T
x_2	F
$\vdots$	T
x_n	T

$\phi = C_1 \wedge C_2 \wedge C_3$

Product

$C_1 = x_1 \vee \bar{x}_2 \vee x_4$ compl. sums

Pure

$\phi = C_1 \wedge C_2 \wedge \dots \wedge C_m$

$\exists$ truth assignment s.t. $\geq \frac{m}{2}$ clauses are satisfied

random assignment

$X_i = T$ with prob $1/2$

$\quad\;\; F$ w.. $1/2$

$\Pr(C_1 = T) = 1 - \frac{1}{8} = 1 - \frac{1}{2^3} \geq \frac{1}{2}$

$X_1 = \begin{cases} 1 & \text{if } C_1 \text{ is } T \\ 0 & \text{o/w} \end{cases}$

Yet Another Example

- We show that for any set of m clauses, there is a truth assignment that satisfies at least m/2 clauses.
- Proof: Consider a random assignment of truth values to variables as T/F.
- Consider a clause C_i of k variables.
- C_i is not satisfied with probability 2^{-k}.
- Define a random variable Z_i that indicates the event Ci is satisfied.
- $E[Z_i] = \Pr(Ci \text{ is satisfied}) = 1 - 2^{-k}$.
- Define Z as the number of clauses satisfied. $Z = \sum Z_i$.
- $E[Z] = E[\sum Z_i] = \sum E[Z_i] = m(1 - 2^{-k}) \geq m/2$ as $k \geq 1$.