Seoul
National
University

Probabilistic Method

Name: Chong-kwon Kim

SCONE
Lab.

Conditional Expectation

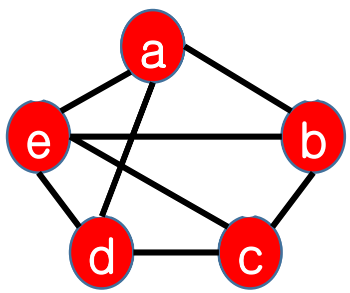
- Randomized algorithm to find a large cut
  - Random allocation of vertices to two subsets A and B
  - Repeatedly try random cuts until find one whose size $\geq m/2$

- Any (constructive) solutions that guarantee a large cut in one trial
  - Deterministic solution

- Assume that vertices are numbered 1, 2,⋯, n and we allocate vertices into A or B sequentially
- Place a vertex such that conditional cut size expectation is maximized
  - Deterministic and Constructive

Large Cut

- Suppose k vertices were already allocated into A or B sets and making a decision for the (k+1)-st vertex
  - Select A if the resulting conditional expectation after putting it to A is larger than that of allocating it to B



Allocation sequence: a-b-c-d-e
1. Allocate *a* into A
2. Given *a* is in A, where to put *b*? (assume B)
3. Given *a* in A and *b* in B, where to put *c*?

Generally, the expected cut size increase obtained by putting v_{k+1} into A, INC_A, is (# edges to {B} − # edges to {A})/2

- $E[C(A, B) \mid x_1, x_2, \cdots, x_k]$
  $= 1/2 \cdot E[C(A, B) \mid x_1, x_2, \cdots, x_k, A] + 1/2 \cdot E[C(A, B) \mid x_1, x_2, \cdots, x_k, B]$
  $\le \max\{E[C(A, B) \mid x_1, \cdots, x_k, A], E[C(A, B) \mid x_1, \cdots, x_k, B]\}$

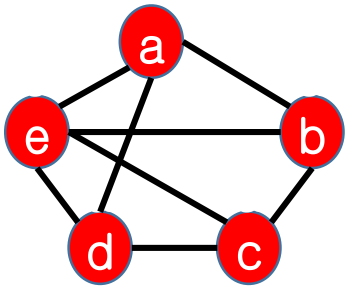
Large Cut

- $m/2 = E[C(A,B)] \le E[C(A,B) \mid x_1]$

$$\le E[C(A,B) \mid x_1, x_2]$$

$$\cdots$$

$$\le E[C(A,B) \mid x_1, x_2, \cdots, x_n]$$



a →	A
b →	B
c →	?
d →	?
e →	?

Sample & Modify

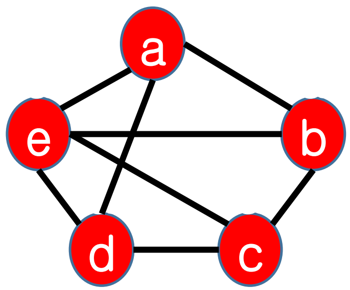
- **Conditional expectation**
  - Select wisely and deterministically

- **Sample & Modify**
  - Randomly select one
  - Modify the selection to add (or remove) a certain property (characteristics) to/from the selection

- Application of Sample & Modify
  - **Maximum Independent set** Problem
  - **Maximum Girth** Problem

Independent Set Problem

- Definition: Independent set
  - Set of vertices with no edges between them



{a,c} is an independent set while {a, d} is not

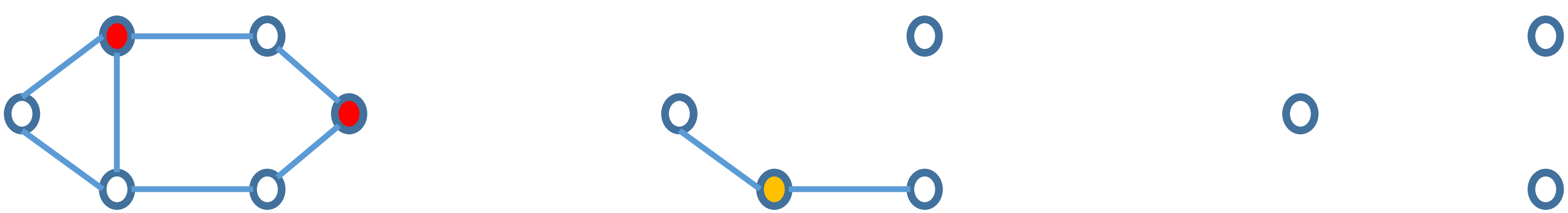
Finding an independent set is easy
How?

Finding an Max. independent set is NP-Hard

Independent Set Problem

- Theorem
  - Let G=(V,E) is a graph with n vertices and m edges, then there is an independent set with $\geq n^2/4m$ vertices

- Proof
  - # edges incident on a node
  - Let d=2m/n : Average node degree
  - Two step algorithm
    1. Sample
       Delete each vertex and its edges independently with probability 1−1/d
    2. Modify
       For any remaining edges, delete one vertex randomly



Independent Set Problem

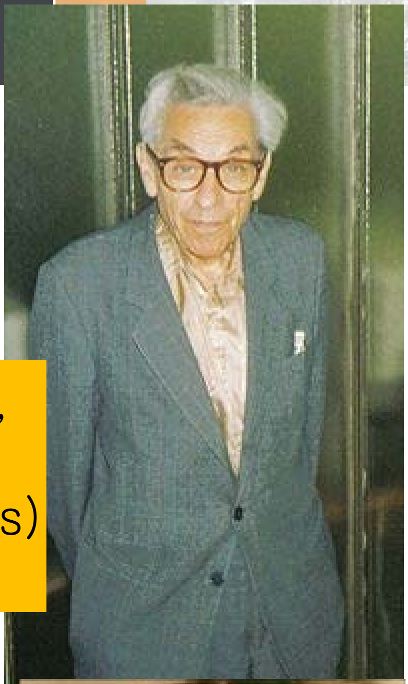
- How many vertices and edges remaining after the first step?
  - Let X, Y be # survived vertices and edges, respectively
    ➔ E[X] = n/d, because each node survival probability is 1/d
  - An edge remains if both of two end nodes survive
    ➔ E[Y] = $m/d^2 = \frac{n}{2d}$

- At the second step, at most Y vertices are removed
  ➔ Independent set size = X−Y
- E[X−Y] = $\frac{n}{2d} = n^2/4m$

Erdös–Renyi Random Graph

- **Randomly** generated networks
  - Fix number of nodes to n (parameter)
  - Random (usually undirected) edge generation

Erdös was a Hungarian mathematician, one of the most famous in the 20th C. Famous as a travelling (a.k.a. homeless) mathematician



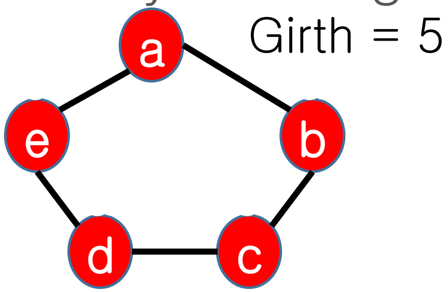
- Two models
- $G_{n,p}$:
  - Randomly generate an edge between (i, j) nodes w/ fixed probability of p
  - Average degree = (n−1)p
- $G_{n,m}$:
  - Randomly select m edges out of $\binom{n}{2}$ candidates
  - Average degree = 2m/n

Random Networks

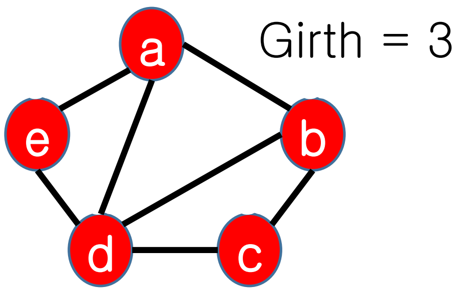
- Random network is used in
  - Graph theory
  - Sociology, economics
  - Biology, epidemiology
  - Social network analysis
  - …
- Many problems such as
  - # isolated nodes
  - Size of connected nodes (Component)
  - Appearance of giant component
  - Diameter (Six degrees of separation)
  - …

Large Girth: Example of Sample& Modify

- Definitions:
  - Cycle: A path, consists of different edges, that returns to the origin
  - Girth: Length of smallest cycle in a graph



Sparse Network



Dense Network

As more edges are added, girth decreases

- Theorem
  - For any k ≥ 3, there is a Graph with n nodes with the following properties
    - # edges m $\geq \frac{1}{4} n^{1+1/k}$
    - Girth ≥ k

Large Girth

- Proof
  - Sketch of proof
    - Sample: Generate a random graph w/ m+δ edges
    - Modify: If there are cycles of < k length, remove an edge from the cycles (Ignore efficiency)

      ➔ If Y such cycles, then # remaining edges is m+δ −Y

Sample

Select a random graph $G_{n,p}$ with p $= n^{\frac{1}{k}-1}$

Let X: # edges

$$E[X] = \binom{n}{2} \cdot p = \frac{1}{2}\left(1 - \frac{1}{n}\right) n^{\frac{1}{k}+1}$$

Large Girth

Modify

Consider a set of l ordered nodes that may create a cycle

➔ Probability that the l nodes form a cycle is p^l

How many cyclic irreversible unique sequences can be generated with l nodes?

➔ $(l-1)!/2$ **Why?**

Let Y_l: # length l cycles

$$E[Y_l] = p^l \cdot \binom{n}{l} \cdot (l-1)!/2$$

$$Y = \sum_{l=3}^{k-1} Y_l$$

$$E[Y] = \sum_{l=3}^{k-1} p^l \binom{n}{l} \cdot (l-1)!/2$$

$$\leq \sum_{l=3}^{k-1} n^l p^l < kn^{(k-1)/k}$$

Large Girth

remaining edges = X−Y

$$E[X-Y] \geq \frac{1}{2}\left(1-\frac{1}{n}\right) n^{\frac{1}{k}+1} - kn^{(k-1)/k}$$
$$\geq \frac{1}{4} n^{1+1/k}$$

Second Moment Method

- **Theorem**

  Let X be a non-negative integer-valued random variable.

  Then, Pr(X=0) $\leq \frac{Var[X]}{(E[X])^2}$

- **Proof**

  Pr(X=0) $\leq \Pr(|X-E[X]| \geq E[X])$

  $\leq \frac{Var[X]}{(E[X])^2}$

  Chebyshev's Inequality

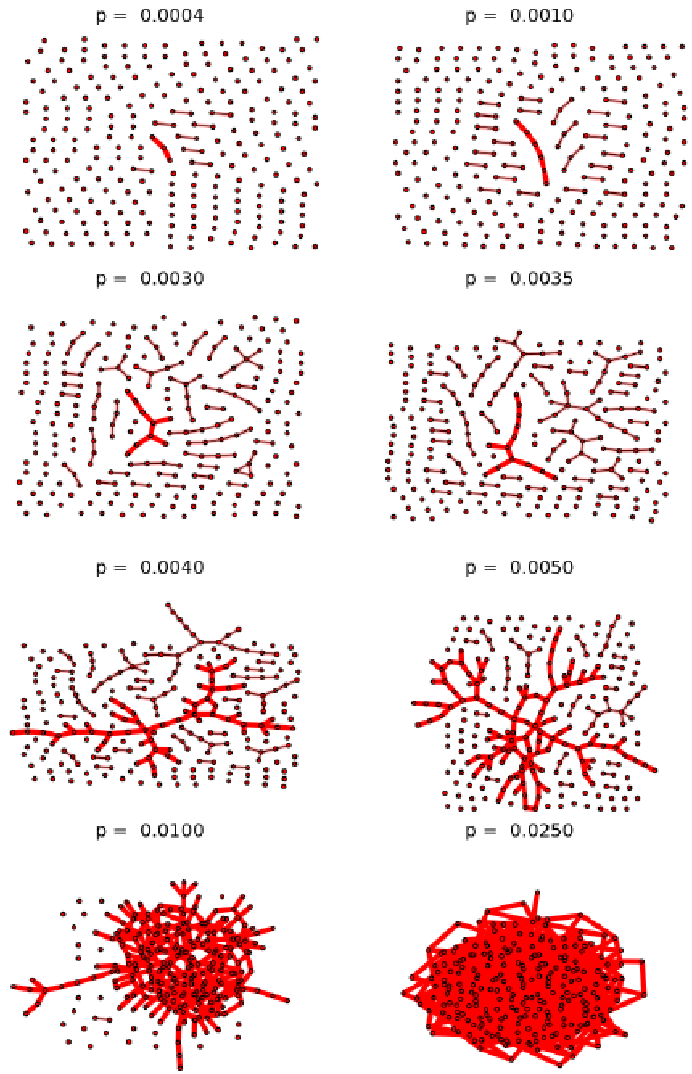
If $\frac{Var[X]}{(E[X])^2} < 1-\varepsilon$ ➔ $\Pr(X > 0) > \varepsilon$

If $\frac{Var[X]}{(E[X])^2} < \varepsilon$ ➔ $\Pr(X=0) < \varepsilon$

Cliques in Random Graphs

SCONE Lab.

- Many properties of random networks occur at threshold points (edge probability, p)
  - Giant component
  - Cliques
  - …



- Claim

  In $G_{n,p}$, let p=f(n).

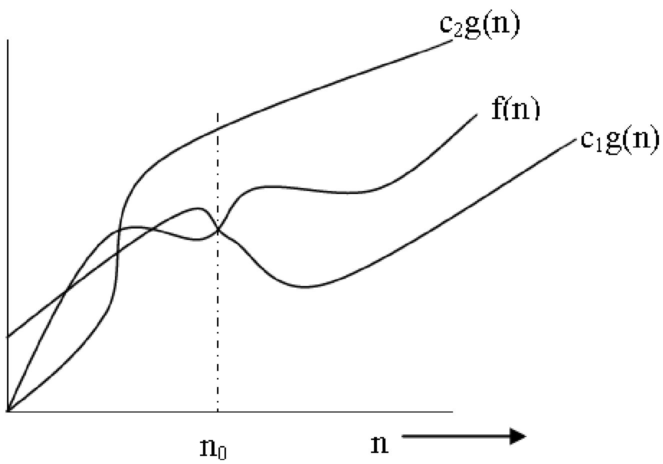
  1. If f(n)= o($n^{-2/3}$), then for large n,

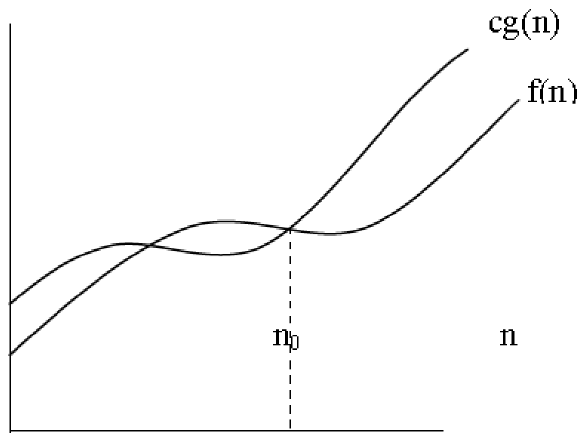
     $\Pr\left(G \in G_{n,p} \text{ has } k \geq 4 \text{ cliques}\right) < \varepsilon$

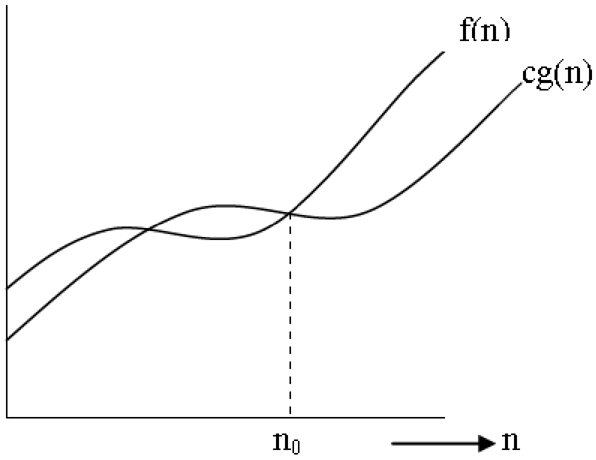
  2. If f(n)= ω($n^{-2/3}$), then for large n,

     Pr(G ∈ $G_{n,p}$ has **no** clique of k ≥ 4) < ε

Approximate Bounds







Asymptotically Tight Bound

$\Theta(g(n))$

Asymptotic Upper Bound

$O(g(n))$

Upper Bound

$o(g(n))$

Asymptotic Lower Bound

$\Omega(g(n))$

Lower Bound

$\omega(g(n))$

Big Theta

Big Oh
Small Oh

Big Omega
Small Omega

Cliques in Random Graphs

In $G_{n,p}$, let p=f(n).

1. If f(n)= o($n^{-2/3}$), then for large n, $\Pr(\mathrm{G} \in G_{n,p} \text{ has } \mathrm{k} \geq 4 \text{ cliques}) < \varepsilon$

- Proof

1. There are $\binom{n}{4}$ combinations of 4 node set

Xi = 1, if i-th combination form a 4-clique

0, ow

Probability that k random nodes form a k-clique

➔ Probability that all of $\binom{k}{2}$ edges exist

➔ $p^{\binom{k}{2}}$

X = $\sum_{i=1}^{\binom{n}{4}} X_i$

➔ E[X] = $\binom{n}{4} p^6$ = o(1)

➔ E[X] < ε, for large n

➔ Pr($X \geq 1$) $= \sum_{x=1} P_x \leq \sum_{x=0} x \cdot P_x = E[X] < \varepsilon$

What techniques are used?

Cliques in Random Graphs

- Proof of 2

2. If $f(n) = \omega(n^{-2/3})$, then for large n,
   $\Pr(G \in G_{n,p}$ has **no** clique of $k \geq 4) < \varepsilon$

- Like the proof of 1, we can prove that E[X] = $\omega(1)$

➔ $E[X] \to \infty$ as $n \to \infty$

➔ Does it guarantee the existence of 4-clique with high probability?

No! Refer to double betting

Strategy: Use Second Moment Method

$\Pr(X=0) \leq \frac{Var[X]}{(E[X])^2}$

And show that $\frac{Var[X]}{(E[X])^2} \to 0$ as $n \to \infty$

Now, derive the expectation & variance of # 4-cliques

Cliques in Random Graphs

SCONE Lab.

- First prove that
  - Let Yi, i=1,2,⋯,m be **0−1** random variables and let Y=$\sum_{i=1} Y_i$
    Then, Var[Y] ≤ E[Y] + $\sum_{i \neq j} Cov(Y_i, Y_j)$

- Proof
  - Var[Y] = $\sum_{i=1} Var[Y_i] + \sum_{i \neq j} Cov(Y_i, Y_j)$

    $Var[Y_i]$ = E[Y_i^2] − $(E[Y_i])^2$ ≤ E[Y_i^2] = E[Y_i]

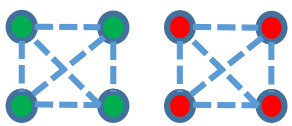
    $\leq \sum_{i=1} E[Y_i] + \sum_{i \neq j} Cov(Y_i, Y_j)$

    = E[Y] + $\sum_{i \neq j} Cov(Y_i, Y_j)$

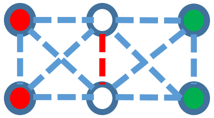
  - We have derived E[X] and need to derive Var[X]
    - Problem: Computation of Cov(Xi, Xj)

Cliques in Random Graphs

- Proof of 2
  - Cov(Xi, Xj) = E[Xi Xj] – E[Xi] E[Xj] ≤ E[Xi Xj]
  - E[Xi Xj] depends on how many edges in i-th and j-th 4 node combinations, C_i and C_j
  - Cases
    - Share No edges
      - ➔ Xi and Xj are independent
      - ➔ Cov(Xi, Xj) = 0

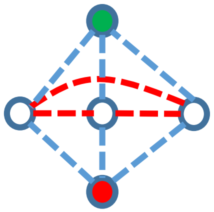


    - Share One edge:
      - ➔ 11 edges should present
      - ➔ E[Xi Xj=1] = p^{11}



$\binom{n}{6}\binom{6}{2;2;2}$cases

    - Share Three edges
      - ➔ 9 edges
      - ➔ E[Xi Xj=1] = p^{9}



$\binom{n}{5}\binom{5}{3;1;1}$cases

Cliques in Random Graphs

- Cont.
  - Var[X] $\leq \mathrm{E}[\mathrm{X}] + \sum_{i \neq j} Cov(X_i, X_j)$

    $\leq \binom{n}{4} p^6 + \binom{n}{6}\binom{6}{2;2;2} p^{11} + \binom{n}{5}\binom{5}{3;1;1} p^9$

    $= \mathrm{o}(n^8 p^{12})$ Why?

  - $(E[X])^2 = \left(\binom{n}{4} \cdot p^6\right)^2 = \Theta(n^8 p^{12})$

  ➔ Pr[X=0] $\leq \frac{Var[X]}{(E[X])^2} = \mathrm{o}(1)$

Conditional Expectation Inequality

- Special application of the Second Moment Method to the cases of **Bernoulli trials**

- Theorem

  Let X = $\sum_{i=1} X_i$ where Xi is 0−1 random variable.

  Then, $\boxed{\Pr(X > 0) \ge \sum_{i=1}^{n} \frac{\Pr(X_i=1)}{E[X \mid X_i=1]}}$

Second Moment Method

$\Pr(X=0) \le \frac{Var[X]}{(E[X])^2}$

Conditional Expectation Inequality

- Proof

Let Y= 1/X, if X > 0

0, ow

Pr(X > 0) = E[XY]

$= E[(\sum_{i=1} X_i)\ Y]$

$= \sum_{i=1} E[X_i\ Y]$

$= \sum_{i=1} E[X_i\ Y \mid X_i = 1] \Pr(X_i = 1) + E[X_i Y \mid X_i = 0] \Pr(X_i = 0)$

$= \sum_{i=1} E[Y \mid X_i = 1] \Pr(X_i = 1)$

$= \sum_{i=1} E[1/X \mid X_i = 1] \Pr(X_i = 1)$ (Jensen's Inequality, f= 1/x)

$\geq \sum_{i=1}^{n} \frac{\Pr(X_i=1)}{E[X \mid X_i=1]}$

Conditional Expectation Inequality

- Recall the Claim that
  - In $G_{n,p}$, let p=f(n).
  1. If f(n)= o($n^{-2/3}$), then for large n, $\Pr\left(\mathrm{G} \in G_{n,p} \text{ has } \mathrm{k} \geq 4 \text{ cliques}\right) < \varepsilon$
  2. If f(n)= ω($n^{-2/3}$), then for large n, Pr(G ∈ $G_{n,p}$ has **no** clique of k ≥ 4) < ε

- Proof of 2 using $\Pr(X > 0) \geq \sum_{i=1}^{n} \frac{\Pr(X_i=1)}{E[X \mid X_i=1]}$
  - Now, we compute $E[X \mid X_i = 1]$

$$E[X \mid X_i = 1] = \sum_{j=1}^{\binom{n}{4}} E\left[X_j \middle| X_i = 1\right]$$

  - Again, we consider how many vertices that Ci and Cj share

Conditional Expectation Inequality

- Cont.
  - Cases
    - No common vertex, one common vertex

      $\binom{n-4}{4}$ instances $\quad 4\binom{n-4}{3}$ instances

      $E[X_j|X_i=1] = \Pr(X_j = 1 \mid X_i = 1) = p^6$

    - Two common vertices

      $\Pr(X_j = 1 \mid X_i = 1) = p^5$

      $\binom{n-4}{2}\binom{4}{2}$ instances

    - Three common vertices

      $\Pr(X_j = 1 \mid X_i = 1) = p^3$

      $\binom{n-4}{1}\binom{4}{3}$ instances

  - $E[X \mid X_i = 1] = 1 + \binom{n-4}{4}p^6 + 4\binom{n-4}{3}p^6 + 6\binom{n-4}{2}p^5 + 4\binom{n-4}{1}p^3$

  - $\Pr(X > 0) \geq \sum_{i=1}^{n} \frac{\Pr(X_i=1)}{E[X \mid X_i=1]}$

Approaches to 1 as n increases when p= f(n)= $\omega(n^{-2/3})$

Lovasz Local Lemma (LLL)

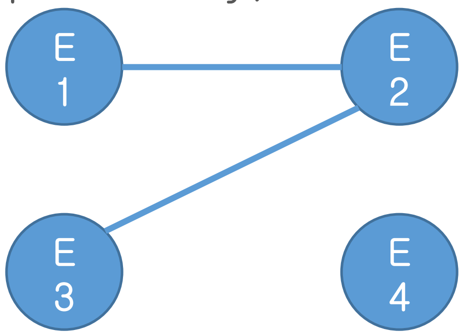
Lovasz is another Hungarian mathematician
Winner of IMO, Wolf Prize, and Kyoto Prize

- One of the most powerful tools
- Used to prove that there is an element not included in any of bad events $\bigcup_{i=1}^{n} E_i$
- Case when E_i are mutually independent
  - ➔ $\bar{E}_i$ are mutually independent also
  - ➔ $\Pr(\bigcap_{i=1}^{n} \bar{E}_i) = \prod_{i=1}^{n} \Pr(\bar{E}_i) > 0, \text{if } \Pr(E_i) < 1$ for all i

Lovasz Local Lemma expands the argument to the cases where events are not mutually independent but the dependency is limited

Lovasz Local Lemma

- Definition: Dependency Graph
  - Representation of dependency (or mutual independency) between events with a graph
  - Vertex i : Event i
  - E_i is mutually independent of events $\{E_j \mid (i, j) \notin E\}$



- Lovasz Local Lemma

  Let E1, E2, ⋯, En are events that has the following properties

  1. for all i $\Pr(E_i) < p$
  2. The degree of the dependency graph is bounded by d
  3. $4dp \le 1$

  Note: $d \ge 1$, $p \le 1/4$

  Then $\Pr(\cap_{i=1}^{n} \bar{E}_i) > 0$

Lovasz Local Lemma

- Proof of $\Pr(\cap_{i=1}^{n} \bar{E}_i) > 0$

  - $\Pr(\cap_{i=1}^{n} \bar{E}_i) = \Pr(\bar{E}_n \mid \cap_{i=1}^{n-1} \bar{E}_i) \cdot \Pr(\bar{E}_{n-1} \mid \cap_{i=1}^{n-2} \bar{E}_i) \cdot \cdots \Pr(\bar{E}_1)$
    $= \prod_{i=1}^{n} (1 - \Pr(E_i \mid \cap_{j=1}^{i-1} \bar{E}_j))$

Need to prove that
$\Pr(E_i \mid \cap_{j=1}^{i-1} \bar{E}_j) < 1$
for i=1, 2,⋯,n

In fact
$\Pr(E_i \mid \cap_{j=1}^{i-1} \bar{E}_j) \le 2p$

  - Now to prove $\Pr(E_i \mid \cap_{j=1}^{i-1} \bar{E}_j) \le 2p$

First, should be well-defined

$\Pr(\cap_{j=1}^{i-1} \bar{E}_j) > 0$

Need to prove that
$\Pr(E_j \mid \cap_{k=1}^{j-1} \bar{E}_k) < 1$
for j=1, 2,⋯,i−1

Lovasz Local Lemma

- Introduce several notations
  - S=$\{s_1, s_2, \ldots, s_s\}$ is a random permutation of a subset with s elements randomly selected from {1, 2, ⋯, n}

    ➔ For simplicity, we use 1, 2, ⋯n instead of $s_1, s_2, \ldots, s_n$
  - $F_s = \bigcap_{j=1}^{s} \bar{E}_j$ for S = {1, 2, ⋯, s}

For |S| = 1, 2, ⋯, n

① Pr($\boldsymbol{E_i} \,|\boldsymbol{F_{s-1}}) \leq \boldsymbol{2p}$

For |S| = 1, 2, ⋯, n

② Pr($\boldsymbol{F_s}$) > 0

Use Mathematical Induction

Prove for |S|=1

Assume TRUE for |S| = 1, 2, ⋯, h

➔ Prove ① for h+1

➔ Prove ② for h+1

Lovasz Local Lemma

- |S| = 1
  - ① $\Pr(E_1 \mid \emptyset) = \Pr(E_1) \leq p \leq 2p$
  - ② $\Pr(F_1) = 1 - \Pr(E_1) = 1 - p > 0$

- Assume ① and ② are true up to h

  For a particular E_k for $k \notin S$

  $S_1 = \{j \in S \mid (k, j) \text{ is an edge in the dependency graph } G\}$

  $S_2 = S - S_1$

  If $S_2 = S$

  ➔ E_k is mutually independent of events $\bar{E}_j$ for $j \in S$

  ➔ $\Pr(E_k \mid F_S) = \Pr(E_k) \leq p \leq 2p$

  If $S_2 \neq S$

  $F_S = F_{S_1} \cap F_{S_2}$

  $\Pr(E_k \mid F_S) = \Pr(E_k \mid F_{S_1} \cap F_{S_2})$

  $= \Pr(E_k \cap F_{S_1} \cap F_{S_2}) / \Pr(F_{S_1} \cap F_{S_2})$

  $= \Pr(E_k \cap F_{S_1} \mid F_{S_2}) \Pr(F_{S_2}) / \Pr(F_{S_1} \mid F_{S_2}) \Pr(F_{S_2})$

Lovasz Local Lemma

- $\Pr(E_k | F_S) = \Pr(E_k \cap F_{S_1} | F_{S_2}) / \Pr(F_{S_1} | F_{S_2})$

  $\leq \Pr(E_k | F_{S_2}) / \Pr(F_{S_1} | F_{S_2})$

  $= \Pr(E_k) / \Pr(F_{S_1} | F_{S_2})$

  $\leq p / \Pr(F_{S_1} | F_{S_2})$

- Now consider $\Pr(F_{S_1} | F_{S_2})$

  $\Pr(F_{S_1} | F_{S_2}) = \Pr(\bigcap_{j \in S1} \bar{E}_j \mid F_{S_2})$

  $= 1 - \Pr(\bigcup_{j \in S1} E_j \mid F_{S_2})$

  $\geq 1 - \sum_{j \in S1} \Pr(E_j | F_{S_2})$

  $\geq 1 - \sum_{j \in S1} 2p$

  $\geq 1 - 2pd$ (Why?)

  $\geq \frac{1}{2}$

Mathematical induction

➔ $\Pr(E_k | F_S) \leq 2p$

Lovasz Local Lemma

- We just proved ① for |S|= h+1
- Now prove ② for |S|= h+1

$\Pr(F_s) = \Pr(\bigcap_{j=1}^{h+1} \bar{E}_j)$

$= \Pr(\bar{E}_{h+1} \mid \bigcap_{i=1}^{h} \bar{E}_i) \cdot \Pr(\bar{E}_h \mid \bigcap_{i=1}^{h-2} \bar{E}_i) \cdot \cdots \Pr(\bar{E}_1)$

$= \prod_{i=1}^{h+1} (1 - \Pr(E_i \mid \bigcap_{j=1}^{h-1} \bar{E}_j))$

$\geq \prod_{i=1}^{h+1} (1-2p) > 0$

Application-Satisfiability

- Problem:
  - K-SAT problems where each literal appears at most T=$2^k/4k$ can be satisfiable

- Proof:
  - Suppose a random assignment of T/F to each literal

    Ei: Event that i- clause is not satisfied

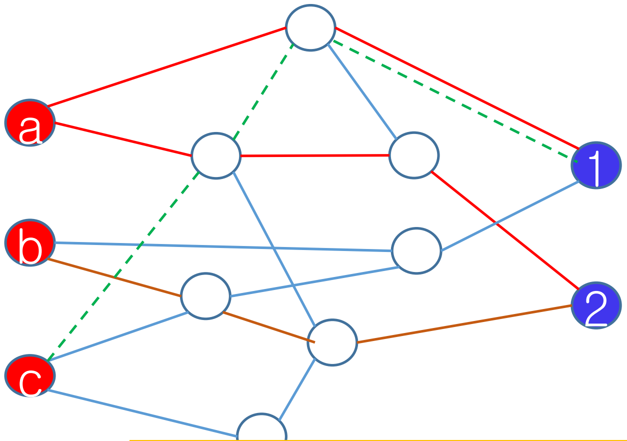
    Pr(Ei) = $1/2^k$

    ➔ $4dp = 4 \cdot k \cdot (2^k/4k) \cdot 1/2^k = 1$
  - Pr(K-SAT is satisfiable) = Pr($\bigcap_{i=1}^{n} \bar{E}_i$) > 0 because 4dp ≤ 1

Application – LLL

- Application of LLL is easy
  - Just show that $4pd \leq 1$

Critical Point:
Model proper dependency graphs



Source-destination pairs
(a,1), (a, 2), (b, 2), (c, 1)

- Example: Edge-Disjoint Paths
  - Paths that do not share same edges
- Problem: To prove that
  - Let F_j
    - Set of available paths that j-th node pair can select, $m = |F_j|$

If each path in Fi clashes with no more than k paths in Fj
➔ There exists a way to choose n edge-disjoint paths connecting n pairs when $8nk/m \leq 1$

What are the vertices (events)? d? p?

Edge-Disjoint Paths

- Proof

  $E_{i,j}$: Event that paths selected by pair i and pair j clash

  p

  ➔ $\Pr(E_{i,j}) \leq \text{k/m}$

  G: Dependency graph based on these events

  Note that $E_{i,j}$ is independent of $E_{k,l}$ for $\mathrm{i} \neq \mathrm{k}, \mathrm{j} \neq l$

  # degrees of $E_{i,j} \leq 2\mathrm{n}$

  d

  ➔ 4pd = 4 (k/m) 2n = 8kn/m ≤ 1

Application: k-SAT

- LLL can be used for **constructive solution** also
- Usually it is quite difficult to apply
  - Section 6.8 is an example of constructive application of LLL
- **Structure** of constructive application of LLL
  - In the first phase, a subset of random variables are assigned with randomly selected values
    - Such that it is possible to prove the existence of solutions to the remaining problem with LLL
    - Such that remaining problem is partitioned into several independent components that are small enough to be solved by an exhaustive search
- Note that
  - 2-SAT is polynomial
  - k-SAT for $k \geq 3$ is NP-Hard
- We will prove that Some k-SAT is polynomial

k-SAT Problem

- We already proved that
  - k-SAT is satisfiable if each literal appears at most T=$\mathbf{2^k/4k}$ times
- Now, we construct a solution in polynomial time if T=$\mathbf{2^{\alpha k}}$ for some small constant $\alpha > 0$
  - Phase 1: Random allocation of T/F to several literals
  - Phase 2: Exhaustive search with remaining literals
- Definitions and Notations
  - l literals $(x_1, x_2, \dots, x_l)$ and m clauses $(C_1, C_2, \dots, C_m)$
  - Dangerous clause: Clause that is
    - k/2 literals are fixed AND
    - Not satisfied yet
  - Surviving clause: clause unsatisfied in Phase 1
    - ➔ Surviving clause has $\geq$ k/2 unassigned literals
  - Deferred literal is a literal not used in Phase 1

k-SAT Problem

- Phase 1
  - Sequentially assign T/F values to literals
  - If x_i is not in a dangerous clause, assign it with T/F randomly

- Claim:
  - The problem remained to Phase 2 is satisfiable

- Proof
  - Prove that 4dp ≤ 1

    Let E_i be the event that i-th Phase 2 clause is not unsatisfied

    p = Pr$(E_i) \leq 2^{-k/2}$

    d $\leq k \cdot 2^{\alpha k}$

    4dp = 4 $\cdot$ $k2^{\alpha k} \cdot 2^{-k/2} \leq 1$ for small α

k-SAT Problem

- Consider clauses remained in phase 2
- Consider two dependency graphs H and H′
  - H: Graph generated from the original k-SAT formula
  - H′: Graph generated with survived clauses in Phase 2
- H (H′ also) have several (connected) components
- If the size of the largest component in H′ is O(log m), then we can find a solution in polynomial of n

- Claim
  - All components in H′ are of size O(log m) with probability 1−o(1)

k-SAT Problem

- Proof
  - First, define a tree S called 4-Tree for a component R in H as follows

    ① Any two nodes in S are at least 4 hops away each other in H

    ② Every edge in S connects two nodes that are exactly 4 hops away each other in H

    ③ Any node of R is wither in S or is within three hops from a node in S

  - Suppose R has r nodes

    ➔ How many nodes in 4-Tree generated from R? (Maximal size of the 4-Tree)

    For each vertex, there are $\leq d^3$ nodes within 3 hops

    ➔ Maximal 4-Tree from R has $\geq r/d^3$ nodes

k-SAT Problem

- Proof – Cont.
  - We need to prove that "There is no component of size r $\geq clog_2 m$ in H′ with probability 1−o(1)"

"There is 4-Tree of size $\geq r/d^3$ with probability 1−o(1)"

Let us count the number of 4-Trees of size $s = r/d^3$

➔ Ways to generate 4-Trees of size s in H

$\leq m(d^4)^{2s} = md^{8r/d^3}$ (Why?)

  - Nodes in a 4-Tree is independent and the probability that s nodes survive in H′ is $\leq ((d+1)2^{-k/2})^s$

➔ Probability that has component of size r is

$\leq md^{8r/d^3} \cdot ((d+1)2^{-k/2})^{r/d^3} \leq md^{rk/d^3(8\alpha+2\alpha-\frac{1}{2})}$ = o(1)

for r $\geq clog_2 m$

General Lovasz Local Lemma

- We proved the symmetric case of Lovasz Local Lemma
- Now, consider general case

- Theorem

  Let G be a dependency graph of BAD events {E1, E2, ⋯, En}

  If there are $x_1, x_2, \ldots, x_n \in [0,1)$ such that Pr(Ei) $\leq x_i \cdot \prod_{(i,j) in\ G}(1-x_j)$

  ➔ Pr(No BAD Events) $\geq \prod(1-x_j)$

- Proof is similar to the symmetric case

General Lovasz Local Lemma

- To prove that

  If there exist $x_1, x_2, \dots, x_n \in [0,1)$ such that

  Pr(Ei) $\leq x_i \prod_{(i,j)\in E}(1 - x_j)$, then

For |S| = 1, 2, ⋯, n

① Pr$(\boldsymbol{E_k} \mid \boldsymbol{F_{s-1}}) \leq \boldsymbol{x_k}$

For |S| = 1, 2, ⋯, n

② Pr$(\boldsymbol{F_s})$ = $\prod_{j\in S}(\boldsymbol{1 - x_j})$ > 0