Lecture Notes CS:5360 Randomized Algorithms
Lecture 20 and 21: Nov 6th and 8th, 2018
Scribe: Qianhang Sun

1 Probabilistic Method

Turning the "MaxCut proof" into an algorithm.

$$\text{Algorithm} \begin{cases} \text{Las Vegas Algorithm} \\ \text{Deterministic Algorithm} \end{cases}$$

$$\text{Derandomization} \begin{cases} \text{Pairwise Independence} \\ \text{Method of Conditional Probabilities} \end{cases}$$

2 MaxCut Proof

Theorem. Let $G = (V, E)$ be a graph with m edges. G has a cut with no less than $m/2$ edges crossing it.

Algorithm 1: LAS VEGAS ALGORITHM OF MAXCUT :

```
repeat
    Throw each vertex independently v ∈ V into A or B with prob 1/2;
    C ← edges crossing the (A, B) cut;
until |C| ≥ m/2;
```

Analyze. Let $p = Pr(|C| \geq \lfloor m/2 \rfloor)$, then

$$\begin{aligned} E[|C|] &= \sum_{c=0}^{m} c \cdot Pr(|C| = c) \\ &= \sum_{c=0}^{\lfloor m/2 \rfloor - 1} c \cdot Pr(|C| = c) + \sum_{c=\lfloor m/2 \rfloor} c \cdot Pr(|C| = c) \\ &\leq (\frac{m}{2} - 1) \cdot (1 - p) + m \cdot p \end{aligned}$$

Since $E[|C|] = \frac{m}{2}$,

$$\begin{aligned} \frac{m}{2} &\leq (\frac{m}{2} - 1) \cdot (1 - p) + m \cdot p \\ &\leq \frac{m}{2} - 1 - p \cdot (\frac{m}{2} - 1) + m \cdot p \\ p &\geq \frac{1}{1 + \frac{m}{2}} \end{aligned}$$

Thus, the expected number of repetitions is O(m), which is, total running time is polynomial in repetition.

3 Derandomization

Pairwise Independence. Let

$$X_v = \begin{cases} 1 & \text{if v falls in A} \\ 0 & \text{otherwise} \end{cases}$$

We assumed that $\{X_v | v \in V\}$ are mutually independent. Suppose $\{X_v | v \in V\}$ are only pairwise independent. i.e.

$$Pr(X_u = 1 | X_v = 0) = Pr(X_u = 1)$$
$$Pr(X_u = 1 | X_w = 1) \neq Pr(X_u = 1)$$

Is it still the case that $E[|C|] = \frac{m}{2}$? Yes. Constructing random variables that are pairwise independent. Let $m \geq 1, n = 2^m - 1$, suppose we are given m mutually independent $(0-1)$ random variables $Y_1, Y_2, Y_3 \cdots Y_m$. Let $S \subseteq 1, 2, \cdots m, S \neq \emptyset$, $X_s = XORofY_i' s, i \in S$. Since the number of such sets S is $2^m - 1$, we have $(2^m - 1)$ $(0-1)$ random variables X_s.

Claim. $\{X_s | S \; subseteq 1, 2, 3 \cdots m, S \neq \emptyset\}$ are pairwise independent. $Pr(X_s = 1) = \frac{1}{2}$

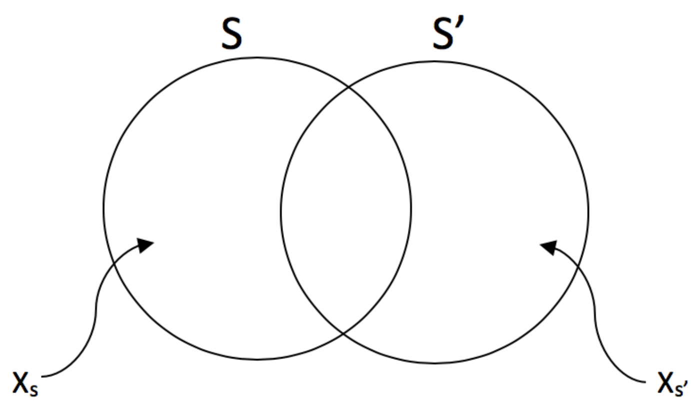

Figure 1: Pairwise Independent

We need a random variable X_v for each vertex $v \in V$. Thus we need $|V|$pairwise independent random variables $\Rightarrow \lceil \log_2 |V| \rceil$mutually independent random variables are needed. Since we only need $\Rightarrow \lceil \log_2 |V| \rceil$ random bits, we can generate all possible settings of These in $O(|V|)$ time and get the entire space. Then we explore each cut in the sample space and pick a cut of size no less than $\frac{m}{2}$, which is guaranteed to exist.

Method of Conditional Probabilities.

Claim. There exists $x_{k+1} \in \{A, B\}$, $E[C(A, B) | x_1, x_2, \cdots x_{k+1}] \geq E[C(A, B) | x_1, x_2, \cdots x_k]$

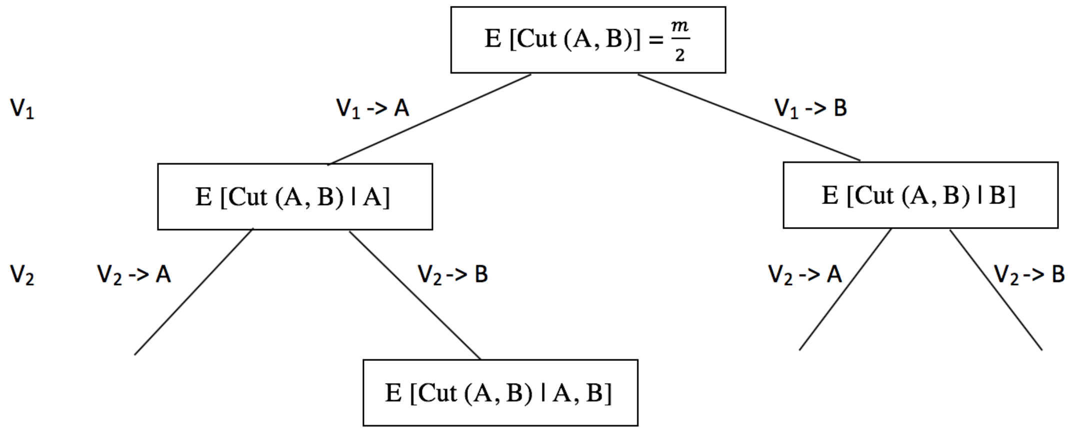

Figure 2: Note that $E[C(A,B)|x_1, x_2, \cdots x_k]$ = size of a specific cut. $\{E[C(A,B)|x_1, x_2, \cdots x_k], x_i \in \{A, B\}\}$ denotes the conditional expectation of C(A, B), the size of the cut, conditioned on v_i falling into x_i, for $i = 1, 2, 3 \cdots k$.

Proof:

$$E[C(A,B)|x_1, x_2, \cdots x_k] = \frac{1}{2} \cdot E[C(A,B)|x_1, x_2, \cdots x_k, A] + \frac{1}{2} \cdot E[C(A,B)|x_1, x_2, \cdots x_k, B]$$

Algorithm step at node $E[C(A,B)|x_1, x_2, \cdots x_k]$:
- Calculate $E[C(A,B)|x_1, x_2, \cdots x_k, A]$ and $E[C(A,B)|x_1, x_2, \cdots x_k, B]$
- Travel to the "child" node with larger expectation.

How to calculate $E[C(A,B)|x_1, x_2, \cdots x_k, A]$?
(1) Count number of edges with both end points fixed that cross the cut.
(2) The answer = count in step(1) + $\frac{1}{2}$(remaining edges)

Note that the "remaining edges" term is the same independent of which set v_{k+1} is assigned to. Thus, v_{k+1} needs to be placed in a set A or B that maximizes the number of edges crossing the cut.

Theorem: The greedy algorithm for MaxCut produces a cut of size no less than $\frac{m}{2}$

4 Lovasz Local Lemma

- "Gem" of the probabilistic method
- 1975 Lovasz & Erdos: on hypergraph coloring

We have a collection $B_1, B_2, \cdots B_n$ of "bad events". **Goal**: Pr(no "bad" event occurs) > 0. i.e. $Pr(\bigcap_{i=0}^{n} \bar{B}_i) > 0 \Rightarrow$ There exists a "good" element in the sample space. How to show $Pr(\bigcap_{i=0}^{n} \bar{B}_i) > 0$?

Approache 1: $Pr(B_i)$ is very small, then

$$Pr(\bigcap_{i=0}^{n} \bar{B}_i) = 1 - Pr(\bigcup_{i=0}^{n} B_i) \geq 1 - \sum_{i=1}^{n} Pr(B_i)$$

If $\sum_{i=1}^{n} Pr(B_i) < 1$,then $Pr(\bigcap_{i=0}^{n} \bar{B}_i) > 0$. So for example, if $Pr(B_i) < \frac{1}{n}$, then this holds.

Approache 2: Independence

$$Pr(\bigcap_{i=0}^{n} \bar{B}_i) = \prod_{i=1}^{n} Pr(\bar{B}_i) \text{ (by independence)}$$

If $Pr(\bar{B}_i) > 0$ for all i, we are done. In settings where $Pr(\bar{B}_i)$ is not too small and $\bar{B}_i$'s are not mutually independent, we use Lovasz Local Lemma.

Idea: - $Pr(B_i) \leq p$
- p is not too small, but small enough relative to the dependencies among the B_i's.

Definition: Let $B_1, B_2, \cdots B_n$ be events, A directed graph $G = (V, E)$ with $V = \{1, 2, \cdots n\}$ is a dependency graph of the events if every event B_i is mutually independent of $\{B_j | (i, j) \notin E\}$

Example: Consider an experiment in which we toss two fair, independent coins.
E_1: first coin toss is Head
E_2: second coin toss is Head
E_e: two coin tosses are identical

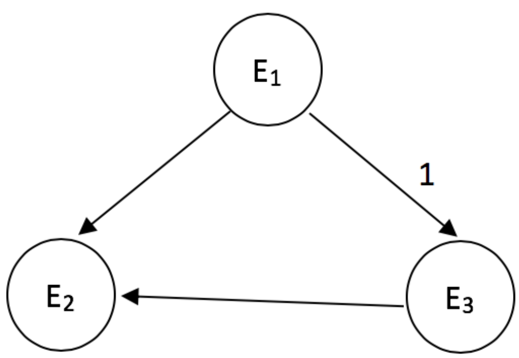

Figure 3: Relation of E_1, E_2 and E_3

Is E_1 mutually independent with respect to $\{E_2, E_3\}$? No. See edge 1. Also, dependency graph is not unique.

Lovasz Local Lemma: Let $B_1, B_2, \cdots B_n$ be events such that:
(1) $Pr(B_i) \leq p$, for i=1, 2, $\cdots n$.
(2) Maximum outdegree of a dependency graph of $B_1, B_2, \cdots B_n$ is $\leq d$.
(3) $4pd \leq 1$ (i.e. $p \leq \frac{1}{4}$)
Then $Pr(\bigcap_{i=0}^{n} \bar{B_i}) > 0)$.

Example 1: We are given a n-vertex cycle. We want to properly color the vertices, i.e. no two adjacent vertices have the same color. Then, how many colors are suffice to choose? 3 colors.

Using Lovasz Local Lemma we will show that 8 or 9 colors suffice. Let us start with a palette of c colors. Each vertex is colored uniformly random using a color from the palette. Then, good event = all pairs of adjacent vertices choose different colors.
Let $e_1, e_2 \cdots e_n$ be the edges of the cycle. B_{e_i} = both endpoints of e_i have the same color. Good event = $\bigcap_{i=1}^{n} \bar{B_{e_i}}$ and $Pr(B_{e_i}) = \frac{1}{c}$. What about dependencies among $B'_{e_i}s$?

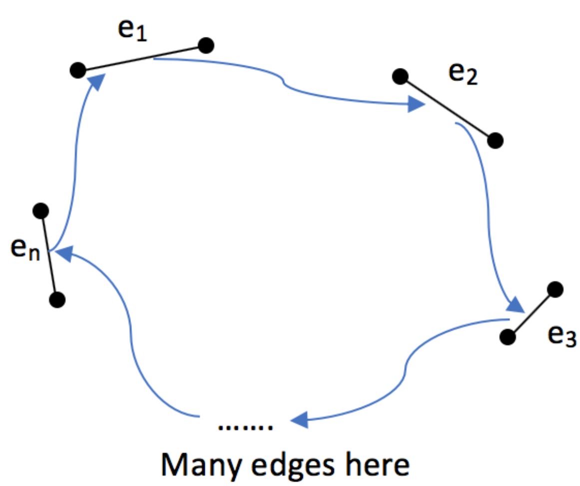

Figure 4: Relation of $e_1, e_2 \cdots e_n$. In this figure, it is easy to tell d=2.

We need

$$
\begin{aligned}
4pd &\leq 1 \\
4 \cdot \frac{1}{c} \cdot 2 &\leq 1 \\
c &\geq 8
\end{aligned}
$$

Instead of $d = 2$, we had $d = 1$, then $c \geq 4$.

Example 2: K-SAT An instance of SAT is a boolean formula in CNF. For example:

$$\underbrace{(\overbrace{\bar{X}_1}^{Literal} \vee \bar{X}_2)}_{\text{Clause}} \wedge \underbrace{(X_2 \vee \bar{X}_3 \vee \bar{X}_4)}_{\text{Clause}} \wedge \underbrace{(\bar{X}_1 \vee X_4)}_{\text{Clause}}$$

k-SAT = special case of SAT in which each clause has exactly k literals. Is the given instance of k-SAT satisfiable?

Theorem: If each variable appears at most $T := \frac{2^k}{4k}$ clauses, then the given instance of k-SAT is satisfiable.

Proof: via Lovasz Local Lemma For each variable x_i, set it independently to true or false with probability $\frac{1}{2}$ each. For each clause C, define B_c = event that C is False. $Pr(B_c) = \frac{1}{2^k}$

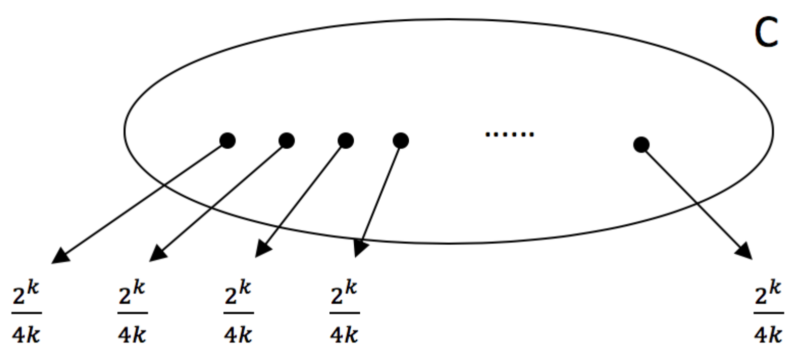

Figure 5: k-SAT

Thus, $d \leq \frac{2^k}{4}$, Then $4 \cdot \frac{1}{2^k} \cdot \frac{2^k}{4} = 1 \Rightarrow$ Lovasz Local Lemma