Advanced Algorithms (VIII)

Shanghai Jiao Tong University

Chihao Zhang

April 26, 2020

The Probabilistic Method

The Probabilistic Method

Design a probability space Ω

The Probabilistic Method

Design a probability space Ω

Show that $\Pr[\text{the object exists}] > 0$

The Probabilistic Method

Design a probability space Ω

Show that $\Pr[\text{the object exists}] > 0$

Bad events $A_1, A_2, \ldots, A_m$, each happens w.p. p_i

The Probabilistic Method

Design a probability space Ω

Show that $\Pr[\text{the object exists}] > 0$

Bad events $A_1, A_2, \ldots, A_m$, each happens w.p. p_i

Is $\Pr[\bar{A}_1 \wedge \bar{A}_2 \ldots \wedge \bar{A}_m] > 0$?

We can apply the union bound

We can apply the union bound

$$\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] = 1 - \Pr\left[\bigcup_{i\in[m]} A_i\right] \geq 1 - \sum_{i\in[m]} p_i$$

We can apply the union bound

$$\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] = 1 - \Pr\left[\bigcup_{i\in[m]} A_i\right] \geq 1 - \sum_{i\in[m]} p_i$$

So $\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] > 0$ if $\sum_{i\in[m]} p_i < 1$

We can apply the union bound

$$\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] = 1 - \Pr\left[\bigcup_{i\in[m]} A_i\right] \geq 1 - \sum_{i\in[m]} p_i$$

So $\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] > 0$ if $\sum_{i\in[m]} p_i < 1$

The union bound is tight when bad events are disjoint

On the other hand, if the bad events are mutually independent…

On the other hand, if the bad events are mutually independent…

$$\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] = \prod_{i\in[m]} (1 - p_i)$$

On the other hand, if the bad events are mutually independent…

$$\Pr\left[\bigcap_{i \in [m]} \bar{A}_i\right] = \prod_{i \in [m]} (1 - p_i)$$

So $\Pr\left[\bigcap_{i \in [m]} \bar{A}_i\right] > 0$ as long as none of $p_i = 1$

On the other hand, if the bad events are mutually independent…

$$\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] = \prod_{i\in[m]} (1 - p_i)$$

So $\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] > 0$ as long as none of $p_i = 1$

The two cases correspond to two extremes of the dependency

Lovász Local Lemma

Lovász Local Lemma

The Lovász local lemma (LLL) captures partial dependency between bad events

Lovász Local Lemma

The Lovász local lemma (LLL) captures partial dependency between bad events

Erdős and Lovász, *Infinite and Finite Sets*, 1975

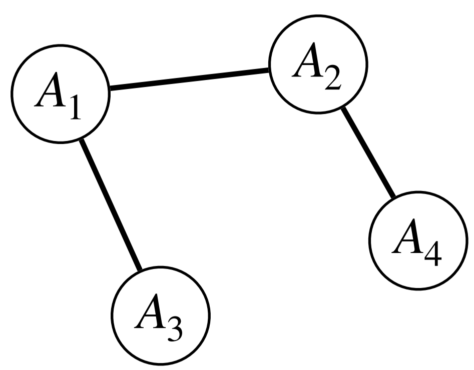
The Dependency Graph

The Dependency Graph

We describe the dependency of bad events in a graph

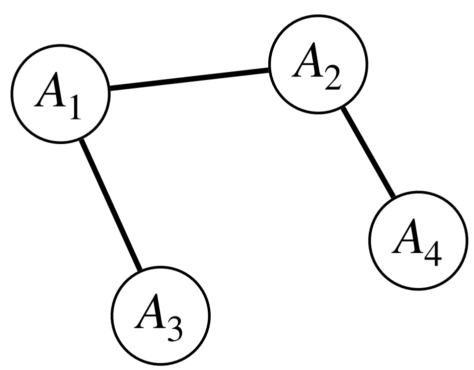
The Dependency Graph

We describe the dependency of bad events in a graph



The Dependency Graph

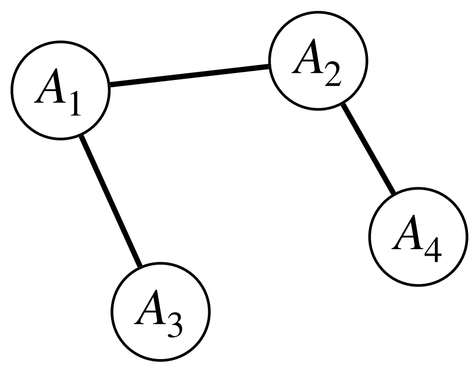
We describe the dependency of bad events in a graph



$$V = \{A_1, \ldots, A_n\}$$

The Dependency Graph

We describe the dependency of bad events in a graph

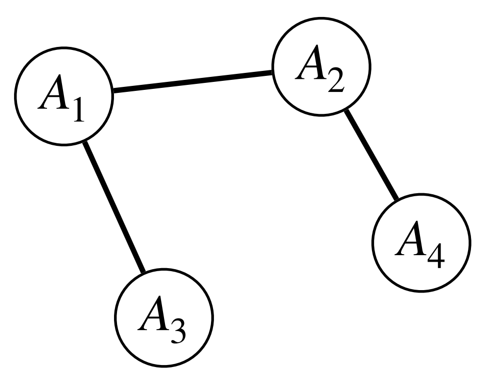


$$V = \{A_1, \ldots, A_n\}$$

$$N(A_i) = \{A_j \mid A_i \sim A_j\}$$

The Dependency Graph

We describe the dependency of bad events in a graph



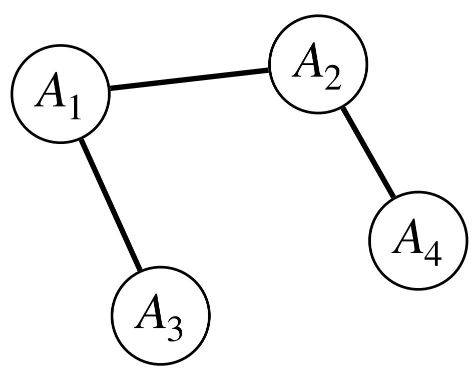
$$V = \{A_1, \ldots, A_n\}$$

$$N(A_i) = \{A_j \mid A_i \sim A_j\}$$

$$\Delta = \max_{i \in [m]} |N(A_i)|$$

The Dependency Graph

We describe the dependency of bad events in a graph



$$V = \{A_1, \ldots, A_n\}$$

$$N(A_i) = \{A_j \mid A_i \sim A_j\}$$

$$\Delta = \max_{i \in [m]} |N(A_i)|$$

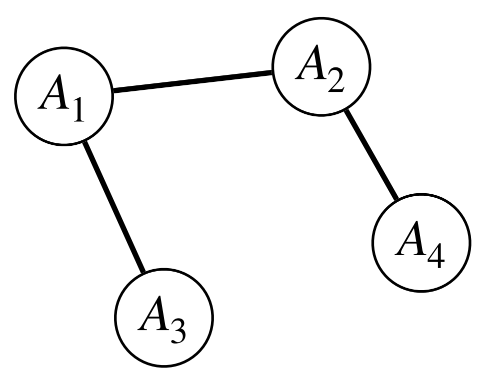
$$4\Delta p \leq 1$$

$$A_i \perp \{A_j\}_{j \notin N(A_i)}$$

$$\Pr[A_i] \leq p$$

The Dependency Graph

We describe the dependency of bad events in a graph



$$V = \{A_1, \dots, A_n\}$$

$$N(A_i) = \{A_j \mid A_i \sim A_j\}$$

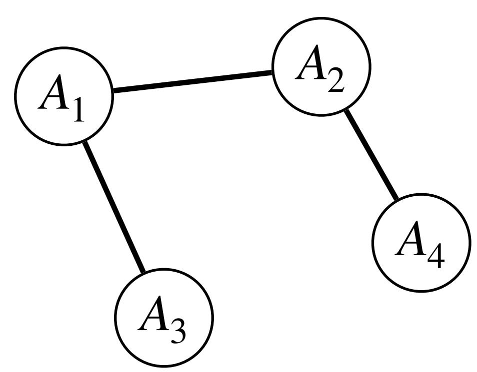
$$\Delta = \max_{i \in [m]} |N(A_i)|$$

$$4\Delta p \leq 1$$
$$A_i \perp \{A_j\}_{j \notin N(A_i)}$$
$$\Pr[A_i] \leq p$$

$$\Longrightarrow$$

The Dependency Graph

We describe the dependency of bad events in a graph



$$V = \{A_1, \ldots, A_n\}$$

$$N(A_i) = \{A_j \mid A_i \sim A_j\}$$

$$\Delta = \max_{i \in [m]} |N(A_i)|$$

$$4\Delta p \leq 1$$
$$A_i \perp \{A_j\}_{j \notin N(A_i)}$$
$$\Pr[A_i] \leq p$$

$$\implies \Pr\left[\bigcap_{i \in [m]} \bar{A}_i\right] > 0$$

Proof of (Symmetric) LLL

For $S \subseteq [m]$, we prove by induction on $|S|$ that

For $S \subseteq [m]$, we prove by induction on $|S|$ that

$$\forall i \notin S, \quad \Pr\left[A_i \mid \bigcap_{j \in S} \bar{A}_j\right] \leq 2p$$

For $S \subseteq [m]$, we prove by induction on $|S|$ that

$$\forall i \notin S, \quad \Pr\left[A_i \mid \bigcap_{j \in S} \bar{A}_j\right] \leq 2p$$

Assume $|S| = s$ and the statement holds for smaller S

For $S \subseteq [m]$, we prove by induction on $|S|$ that

$$\forall i \notin S, \quad \Pr\left[A_i \mid \bigcap_{j \in S} \bar{A}_j\right] \leq 2p$$

Assume $|S| = s$ and the statement holds for smaller S

For every $T \subseteq [m]$, we use F_T to denote the event $\bigcap_{i \in T} \bar{A}_i$

It is clear that for every $T \in \binom{[m]}{\leq s}$,

$\Pr[F_T] \geq (1 - 2p)^s > 0$

It is clear that for every $T \in \binom{[m]}{\leq s}$,

$\Pr[F_T] \geq (1 - 2p)^s > 0$

We partition S into $S = S_1 \cup S_2$ where $S_1 = \{j \mid j \sim i\}$

It is clear that for every $T \in \binom{[m]}{\leq s}$,

$\Pr[F_T] \geq (1-2p)^s > 0$

We partition S into $S = S_1 \cup S_2$ where $S_1 = \{j \mid j \sim i\}$

If $|S_2| = s$, then $\Pr[A_i \mid S] = \Pr[A_i \mid S_2] \leq p$

It is clear that for every $T \in \binom{[m]}{\leq s}$,

$\Pr[F_T] \geq (1-2p)^s > 0$

We partition S into $S = S_1 \cup S_2$ where $S_1 = \{j \mid j \sim i\}$

If $|S_2| = s$, then $\Pr[A_i \mid S] = \Pr[A_i \mid S_2] \leq p$

Otherwise,

$$\Pr[A_i \mid F_S] = \Pr[A_i \mid F_{S_1} \cap F_{S_2}] = \frac{\Pr[A_i \cap F_{S_1} \cap F_{S_2}]}{\Pr[F_{S_1} \cap F_{S_2}]}$$

$$\Pr[A_i \mid F_S] = \frac{\Pr[A_i \cap F_{S_1} \cap F_{S_2}]}{\Pr[F_{S_1} \cap F_{S_2}]} = \frac{\Pr[A_i \cap F_{S_1} \mid F_{S_2}]}{\Pr[F_{S_1} \mid F_{S_2}]}$$

$$\Pr[A_i \mid F_S] = \frac{\Pr[A_i \cap F_{S_1} \cap F_{S_2}]}{\Pr[F_{S_1} \cap F_{S_2}]} = \frac{\Pr[A_i \cap F_{S_1} \mid F_{S_2}]}{\Pr[F_{S_1} \mid F_{S_2}]}$$

$$\Pr[A_i \cap F_{S_1} \mid F_{S_2}] \leq \Pr[A_i \mid F_{S_2}] \leq p$$

$$\Pr[A_i \mid F_S] = \frac{\Pr[A_i \cap F_{S_1} \cap F_{S_2}]}{\Pr[F_{S_1} \cap F_{S_2}]} = \frac{\Pr[A_i \cap F_{S_1} \mid F_{S_2}]}{\Pr[F_{S_1} \mid F_{S_2}]}$$

$$\Pr[A_i \cap F_{S_1} \mid F_{S_2}] \leq \Pr[A_i \mid F_{S_2}] \leq p$$

$$\Pr[F_{S_1} \mid F_{S_2}] = 1 - \Pr\left[\bigcup_{j \in S_1} A_j \mid F_{S_2}\right] \geq 1 - 2dp \geq \frac{1}{2}$$

$$\Pr[A_i \mid F_S] = \frac{\Pr[A_i \cap F_{S_1} \cap F_{S_2}]}{\Pr[F_{S_1} \cap F_{S_2}]} = \frac{\Pr[A_i \cap F_{S_1} \mid F_{S_2}]}{\Pr[F_{S_1} \mid F_{S_2}]}$$

$$\left.\begin{aligned} &\Pr[A_i \cap F_{S_1} \mid F_{S_2}] \le \Pr[A_i \mid F_{S_2}] \le p \\ &\Pr[F_{S_1} \mid F_{S_2}] = 1 - \Pr\left[\bigcup_{j \in S_1} A_j \mid F_{S_2}\right] \ge 1 - 2dp \ge \frac{1}{2} \end{aligned}\right\}$$

$$\Pr[A_i \mid F_S] = \frac{\Pr[A_i \cap F_{S_1} \cap F_{S_2}]}{\Pr[F_{S_1} \cap F_{S_2}]} = \frac{\Pr[A_i \cap F_{S_1} \mid F_{S_2}]}{\Pr[F_{S_1} \mid F_{S_2}]}$$

$$\left.\begin{aligned} &\Pr[A_i \cap F_{S_1} \mid F_{S_2}] \leq \Pr[A_i \mid F_{S_2}] \leq p \\ &\Pr[F_{S_1} \mid F_{S_2}] = 1 - \Pr\left[\bigcup_{j \in S_1} A_j \mid F_{S_2}\right] \geq 1 - 2dp \geq \frac{1}{2} \end{aligned}\right\}$$

$$\Longrightarrow$$

$$\Pr[A_i \mid F_S] = \frac{\Pr[A_i \cap F_{S_1} \cap F_{S_2}]}{\Pr[F_{S_1} \cap F_{S_2}]} = \frac{\Pr[A_i \cap F_{S_1} \mid F_{S_2}]}{\Pr[F_{S_1} \mid F_{S_2}]}$$

$$\left.\begin{aligned} &\Pr[A_i \cap F_{S_1} \mid F_{S_2}] \le \Pr[A_i \mid F_{S_2}] \le p \\ &\Pr[F_{S_1} \mid F_{S_2}] = 1 - \Pr\left[\bigcup_{j \in S_1} A_j \mid F_{S_2}\right] \ge 1 - 2dp \ge \frac{1}{2} \end{aligned}\right\}$$

$$\implies \quad \Pr[A_i \mid F_S] \le 2p$$

Applications of LLL

Edge-Disjoint Paths

Edge-Disjoint Paths

n pairs of users, each has a collection of m paths F_i connecting them

Edge-Disjoint Paths

n pairs of users, each has a collection of m paths F_i connecting them

Each path in F_i shares edges with no more than k paths in F_j for any $j \neq i$

Edge-Disjoint Paths

n pairs of users, each has a collection of m paths F_i connecting them

Each path in F_i shares edges with no more than k paths in F_j for any $j \neq i$

If $8nk \leq m$, then there is a way to choose n edge-disjoint paths connecting n pairs

Define the probability space as

Define the probability space as

"Each pair of users chooses a path from its collection uniformly at random"

Define the probability space as

"Each pair of users chooses a path from its collection uniformly at random"

For every $i \neq j$, define the bad event E_{ij} as

Define the probability space as

"Each pair of users chooses a path from its collection uniformly at random"

For every $i \neq j$, define the bad event E_{ij} as

"the path chosen in F_i overlaps with the path chosen in F_j"

Define the probability space as

"Each pair of users chooses a path from its collection uniformly at random"

For every $i \neq j$, define the bad event E_{ij} as

"the path chosen in F_i overlaps with the path chosen in F_j"

So we only need to show $\Pr\left[\bigcap_{\{i,j\}\in\binom{n}{2}} \bar{E}_{ij}\right] > 0$

For each $\{i, j\} \in \binom{n}{2}$, we have $\Pr[E_{ij}] \leq \frac{k}{m}$

For each $\{i, j\} \in \binom{n}{2}$, we have $\Pr[E_{ij}] \leq \frac{k}{m}$

E_{ij} and $E_{i'j'}$ are dependent only when $\{i, j\} \cap \{i', j'\} \neq \emptyset$

For each $\{i, j\} \in \binom{n}{2}$, we have $\Pr[E_{ij}] \leq \frac{k}{m}$

E_{ij} and $E_{i'j'}$ are dependent only when $\{i, j\} \cap \{i', j'\} \neq \emptyset$

So the maximum degree of the dependency graph is at most $2n$

For each $\{i, j\} \in \binom{n}{2}$, we have $\Pr[E_{ij}] \leq \frac{k}{m}$

E_{ij} and $E_{i'j'}$ are dependent only when $\{i, j\} \cap \{i', j'\} \neq \emptyset$

So the maximum degree of the dependency graph is at most $2n$

$4\Delta p \leq 1$

$A_i \perp \{A_j\}_{j \notin N(A_i)}$

$\Pr[A_i] \leq p$

For each $\{i,j\} \in \binom{n}{2}$, we have $\Pr[E_{ij}] \leq \frac{k}{m}$

E_{ij} and $E_{i'j'}$ are dependent only when $\{i,j\} \cap \{i',j'\} \neq \varnothing$

So the maximum degree of the dependency graph is at most $2n$

$4\Delta p \leq 1$
$A_i \perp \{A_j\}_{j \notin N(A_i)}$
$\Pr[A_i] \leq p$

The LLL condition is then $8nk \leq m$

Satisfiability

Satisfiability

Recall that k-SAT problem is **NP**-hard for $k \geq 3$

Satisfiability

Recall that k-SAT problem is **NP**-hard for $k \geq 3$

On the other hand, if the formula is sparse, then it is always satisfiable

Satisfiability

Recall that k-SAT problem is **NP**-hard for $k \geq 3$

On the other hand, if the formula is sparse, then it is always satisfiable

Given $\phi = C_1 \wedge C_2 \ldots \wedge C_m$, where each $|C_i| = k$

Satisfiability

Recall that k-SAT problem is **NP**-hard for $k \geq 3$

On the other hand, if the formula is sparse, then it is always satisfiable

Given $\phi = C_1 \wedge C_2 \ldots \wedge C_m$, where each $|C_i| = k$

The degree of a variable x is the number of clauses that x or $\bar{x}$ belongs to.

Satisfiability

Recall that k-SAT problem is **NP**-hard for $k \geq 3$

On the other hand, if the formula is sparse, then it is always satisfiable

Given $\phi = C_1 \wedge C_2 \ldots \wedge C_m$, where each $|C_i| = k$

The degree of a variable x is the number of clauses that x or $\bar{x}$ belongs to.

Let d be the maximum degree of variables in ϕ

Theorem.

If $4kd \leq 2^k$, then ϕ is satisfiable

Theorem.

If $4kd \leq 2^k$, then ϕ is satisfiable

The probability space is the uniform distribution over $\{0,1\}^V$

Theorem.

If $4kd \leq 2^k$, then ϕ is satisfiable

The probability space is the uniform distribution over $\{0,1\}^V$

Each clause C_i defines a bad event $A_i :=$ "*C_i is not satisfied*"

Theorem.

If $4kd \leq 2^k$, then ϕ is satisfiable

The probability space is the uniform distribution over $\{0,1\}^V$

Each clause C_i defines a bad event $A_i :=$ "C_i *is not satisfied*"

We need to show $\Pr\left[\bigcap_{i\in[m]} \bar{A}_i\right] > 0$

Each clause C_i satisfies $\Pr[\bar{A}_i] = 2^{-k}$

Each clause C_i satisfies $\Pr[\bar{A}_i] = 2^{-k}$

Two clauses are dependent only if they share some variables

Each clause C_i satisfies $\Pr[\bar{A}_i] = 2^{-k}$

Two clauses are dependent only if they share some variables

Therefore, the maximum degree of the dependency graph is at most kd

Each clause C_i satisfies $\Pr[\bar{A}_i] = 2^{-k}$

Two clauses are dependent only if they share some variables

Therefore, the maximum degree of the dependency graph is at most kd

The LLL condition is $4kd \leq 2^k$

Asymmetric LLL

Asymmetric LLL

In many cases, bad events happen with different probabilities

Asymmetric LLL

In many cases, bad events happen with different probabilities

Assume there exist $x_1, \ldots, x_n \in [0,1]$ such that

$$\Pr[A_i] \leq x_i \prod_{j \sim i} (1 - x_j)$$

Asymmetric LLL

In many cases, bad events happen with different probabilities

Assume there exist $x_1, \ldots, x_n \in [0,1]$ such that

$$\Pr[A_i] \le x_i \prod_{j \sim i} (1 - x_j)$$

Then $\Pr\left[\bigcap_{i=1}^{n} \bar{A}_i\right] \ge \prod_{i=1}^{n} (1 - x_i)$

Algorithmic LLL

Algorithmic LLL

LLL guarantees the existence of a solution

Algorithmic LLL

LLL guarantees the existence of a solution

Can we find one efficiently?

Algorithmic LLL

LLL guarantees the existence of a solution

Can we find one efficiently?

The Gödel Prize 2020 - Laudation

The 2020 Gödel Prize is awarded to **Robin A. Moser** and **Gábor Tardos** for their algorithmic version of the Lovász Local Lemma in the paper:

"A constructive proof of the general Lovász Local Lemma," Journal of the ACM 57(2): 11:1-11:15 (2010).

The Lovász Local Lemma (LLL) is a fundamental tool of the probabilistic method. It enables one to show the existence of certain objects even though they occur with exponentially small probability. The original proof was not algorithmic, and subsequent algorithmic versions had significant losses in parameters. This paper provides a simple,