Streaming Algorithms

Meng-Tsung Tsai

03/06/2018

A Reference Book

- "The Probabilistic Method", Alon and Spencer (2004)

You may find an e-copy of this book on www.lib.nctu.edu.tw

TA

李映萱

email: s88001940@gmail.com

hours: Mon 2-3 pm

location: 電資大樓 ES 701-R6

You may contact 李映萱 for any questions about the written assignements.

Programming Issues

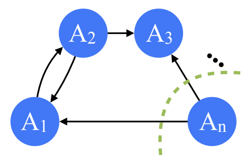
You may contact 黃書擎 for questions about the programming environment and the online judge. However, he is not a TA and please don't try to ask him hints for the problemsets.

email: frank.bozar@gmail.com

Local Lemma

Local Lemma (Symmetric Case)

Let $A_1, A_2, ..., A_n$ be events in an arbitrary probability space. Let $D = (V, E)$ be a directed dependency graph. If all the following conditions hold,



Each node has out-degree $\leq d$.

(1) each A_i is independent of a large set $S_i = \{A_j :(i, j) \notin E\}$ so that $\Pr(A_i \cap (\wedge_{A_j \in S_i} \bar{A}_j)) = \Pr(A_i)\Pr(\wedge_{A_j \in S_i} \bar{A}_j)$ and $|S| \geq n-1-d$,

(2) each event A_i happens with probability $\Pr[A_i] \leq p$,

(3) $ep(d+1) \leq 1$,

then $\Pr(\wedge_i \bar{A}_i) > 0$.

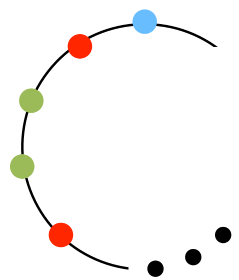
Why not use the union bound?

$\Pr(\wedge_i \bar{A}_i) = 1 - \Pr(\vee_i A_i) \geq 1 - \sum_i \Pr(A_i)$

Local Lemma gives a tigher bound than the union bound when the dependencies among the events are rare (i.e. small d).

Coloring Points on a Circle

Given a circle and 11n points on it. Let C be any coloring on the 11n points so that each point is colored with one of n colors and there are 11 points of each color.



Coloring Points on a Circle

Given a circle and 11n points on it. Let C be any coloring on the 11n points so that each point is colored with one of n colors and there are 11 points of each color.

Theorem 1. No matter what C is, one can pick n non-adjacent points whose colors are all distinct.

Coloring Points on a Circle

Given a circle and 11n points on it. Let C be any coloring on the 11n points so that each point is colored with one of n colors and there are 11 points of each color.

Theorem 1. No matter what C is, one can pick n non-adjacent points whose colors are all distinct.

Proof Strategy. For any C, randomly sample a point for each color and prove that such a random sample contains no adjacent points with positive probability.

What are bad events?

Coloring Points on a Circle

Let A_i for each $i \in [1, 11n]$ be the event that both the i-th point and the (i+1)-th point are sampled.

Thus, $\Pr[A_i] = 1/121$.

Say the i-th point has color α and the (i+1)-th point has color β. A_i has dependecies only with those events that involves a point with color α or β. Therefore, $d \leq (11*2-1)*2 = 42$.

Because $e*p*(d+1) \sim 0.966 \leq 1$, by the Local Lemma we have

$$\Pr[\wedge_i \bar{A}_i] > 0.$$

What happens if we use the union bound rather than the Local Lemma?

Coloring Points on a Circle

Let A_i for each $i \in [1, 11n]$ be the event that both the i-th point and the (i+1)-th point are sampled.

Thus, $\Pr[A_i] = 1/121$.

Say the i-th point has color α and the (i+1)-th point has color β. A_i has dependecies only with those events that involves a point with color α or β. Therefore, $d \leq (11*2-1)*2 = 42$.

Because $e*p*(d+1) \sim 0.966 \leq 1$, by the Local Lemma we have

$$\Pr[\wedge_i \bar{A}_i] > 0.$$

By the union bound, we get $\Pr[\wedge_i \bar{A}_i] \geq 1-11n/121$ which says nothing for large n.

Coloring Hypergraphs

We say a hypergraph is 2-colorable if there exists a coloring on nodes so that every edge is not monochromatic.

Recall that a hypergraph $H = (V, E)$ is defined as ordinary graphs except that each edge $e \in E$ is a subset of V. We say a hypergraph k-uniform if all edges in E has cardinality k.

Coloring Hypergraphs

We say a hypergraph is 2-colorable if there exists a coloring on nodes so that every edge is not monochromatic.

Recall that a hypergraph $H = (V, E)$ is defined as ordinary graphs except that each edge $e \in E$ is a subset of V. We say a hypergraph k-uniform if all edges in E has cardinality k.

Theorem 2. Let H be a k-uniform hypergraph and each edge in H has an non-empty intersection with at most d other edges. If

$$e(d+1) \le 2^{k-1},$$

then H is 2-colorable.

Coloring Hypergraphs

We say a hypergraph is 2-colorable if there exists a coloring on nodes so that every edge is not monochromatic.

Recall that a hypergraph $H = (V, E)$ is defined as ordinary graphs except that each edge $e \in E$ is a subset of V. We say a hypergraph k-uniform if all edges in E has cardinality k.

Theorem 2. Let H be a k-uniform hypergraph and each edge in H has an non-empty intersection with at most d other edges. If

$$e(d+1) \le 2^{k-1},$$

then H is 2-colorable.

Proof Strategy. Assign a random 2-coloring on H.

Coloring Hypergraphs

Let A_e for each $e \in H$ be the event that e is monochromatic.

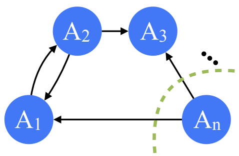
Thus, $Pr[A_e] = 1/2^{k-1}$. (Why?)

Because $e*(d+1) \le 2^{k-1}$, by the Local Lemma we have

$$Pr[\wedge_i \bar{A}_i] > 0.$$

Lovász Local Lemma (General Case)

Let $A_1, A_2, \ldots, A_n$ be events in an arbitrary probability space. Let $D = (V, E)$ be a directed dependency graph. If all the following conditions hold,



The out-degrees may be different.

(1) each A_i is independent of a large set $S_i = \{A_j : (i, j) \notin E\}$ so that $\Pr(A_i \cap (\wedge_{A_j \in S_i} \bar{A}_j)) = \Pr(A_i)\Pr(\wedge_{A_j \in S_i} \bar{A}_j)$,

(2) each event A_i happens with probability $\Pr[A_i] \leq x_i \prod_{A_j \in S_i} (1-x_j)$, where $0 \leq x_1, x_2, \ldots, x_n < 1$,

then $\Pr(\wedge_i \bar{A}_i) \geq \prod_i (1-x_i) > 0$.

Proofs

The proof of Local Lemmas can be found on Pages 64-65 in the reference book.

Why is the dependency graph directed?

For any pair of events A_i and A_j, by checking the equality

$$\Pr(A_i \cap A_j) = \Pr(A_i)\Pr(A_j)$$

one can tell whether A_i and A_j are independent.

It seems that the edges in the dependency graph are always bidirectional. Why not simply use undirected dependency graphs?

Why is the dependency graph directed?

prob	X	Y	X⊕Y
1/4	0	0	0
1/4	0	1	1
1/4	1	0	1
1/4	1	1	0

Let events $A_1 : X = 0$, $A_2 : Y = 0$, $A_3 : X \oplus Y = 0$. Observe that they are pairwise independent but not mutually independent.

Why is the dependency graph directed?

prob	X	Y	X⊕Y
1/4	0	0	0
1/4	0	1	1
1/4	1	0	1
1/4	1	1	0

Let events $A_1 : X = 0$, $A_2 : Y = 0$, $A_3 : X \oplus Y = 0$. Observe that they are pairwise independent but not mutually independent. Thus, one can draw the dependency graph as: or