

Randomly testing Boolean Functions

Lecture 1

Probability

Events, probabilities,
conditional probabilities,
union bound

Random variables,
expected value,
linearity of expectation
Markov's inequality

Algorithms

Approx. 3-SAT,

Big idea: we only want the average

$$\begin{array}{l} \text{Linearity of Expectation} \\ E[X+Y] = E[X] + E[Y] \quad (\text{stochastic variable}) \\ \text{"average sum = sum of averages"} \end{array}$$

let $Y_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{otherwise} \end{cases}$ ($i=1, \dots, m$)

$$E[Y_i] = 7/8 \quad (1 \text{ clause case})$$

$$E[\text{\# clauses satisfied}] = E\left[\sum_{i=1}^m Y_i\right] = \sum_{i=1}^m E[Y_i] = 7/8 m$$

let $Y_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{otherwise} \end{cases}$

$$Z = \sum_{i=1}^m Y_i = \text{\# clauses satisfied}$$

$$E[Z] = \frac{1}{2} E[Z | x_1 = \text{true}] + \frac{1}{2} E[Z | x_1 = \text{false}]$$

$\Rightarrow$ either $E[Z | x_1 = \text{true}]$ or $E[Z | x_1 = \text{false}]$ is $\geq E[Z]$!

QuickSort,

Random divide and conquer

① pick random 'pivot'

② Divide A into two parts



Expected running time of QuickSort

For each $i, j \in [n]$ w/ $i < j$,

$X_{ij} = \begin{cases} 1 & \text{if } i\text{th smallest element is compared to } j\text{th smallest} \\ 0 & \text{otherwise} \end{cases}$

$\sum_{i,j} X_{ij}$ = total # comparisons made by the algorithm
 $\approx$ total running time

$E[\sum_{i,j} X_{ij}]$ = average running time

$E[\sum_{i,j} X_{ij}] = \sum_{i,j} E[X_{ij}]$ Linearity of Expectation: $E[aX+bY] = aE[X] + bE[Y]$
"average sum" = "sum of averages"

now analyze $E[X_{ij}]$ in isolation

$$E[X_{ij}] = \Pr[X_{ij} = 1] = \Pr[i\text{th element is compared to } j\text{th}]$$

i th and j th elem selected before any element in between happens w/ probability $\frac{2}{j-i+1}$

$$E[X_{ij}] = \frac{2}{j-i+1}$$

by linearity of expectation

$$E[\sum_{i,j} X_{ij}] = \sum_{i,j} E[X_{ij}] = \sum_{i,j} \frac{2}{j-i+1}$$

$$\leq 2n \left(\frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \right)$$

$$\leq O(n \log n) \quad \text{with harmonic series}$$

Quickselect (HW)

Lecture 2

Probability

Hash Functions,

Universal Hash Functions

$h: [n] \rightarrow k$ s.t. for all pairs $i, j \in [n]$,
 $\Pr[h(i) = h(j)] \leq \frac{1}{k}$ (much weaker than ideal hash)

e.g. $h(x) = (ax + b) \bmod p \bmod k$

where • p prime, $> k$

• a in $\{1, \dots, p-1\}$ unif. random

• b in $\{0, \dots, p-1\}$ unif. random



Markov's Inequality

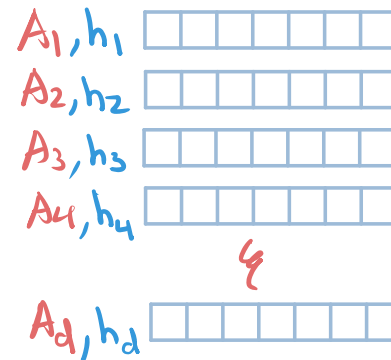
Markov's Inequality (special case)

Let $X \geq 0$ be nonnegative RV

$$\Pr[X \geq 2E[X]] \leq \frac{1}{2}$$

Algorithms

Frequency Estimation, Heavy Hitters (in streams)



estimate f_i w/
 $\min_{j=1, \dots, d} A_j[h_j(i)]$

Lecture 3

Probability

k-wise independence,
higher order moments
(generalizing universal hash
functions and Markov's inequality)

Algorithms

Hash tables w/
chaining +
linear probing

k-wise independent hash functions

$$h: [U] \rightarrow [m]$$

s.t. for any k keys $x_1, x_2, \dots, x_k \in [U]$,
 k values $y_1, y_2, \dots, y_k \in [m]$,

$$P[h(x_1)=y_1, h(x_2)=y_2, \dots, h(x_k)=y_k] = \frac{1}{m^k}$$

eg. $h(x) = a_0 + a_1x + a_2x^2 + \dots + a_{k-1}x^{k-1} \pmod{p}$

where $p \geq U$ prime,

$a_0, a_1, \dots, a_{k-1} \in [p]$ indep, uniformly random

Important lemma

k-wise independent random variables

$$P[X_1=y_1, X_2=y_2, \dots, X_k=y_k]$$

$$= P[X_1=y_1] P[X_2=y_2] \dots P[X_k=y_k]$$

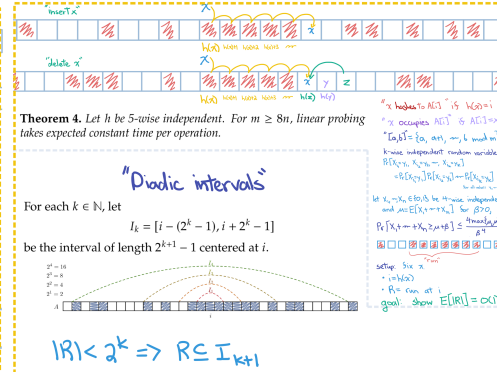
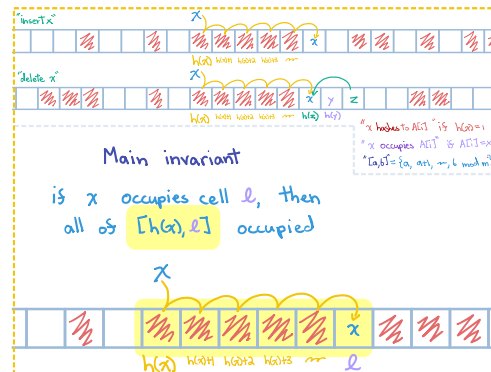
for all values $y_1, \dots, y_k$

let $X_1, \dots, X_n \in \{0, 1\}$ be 4-wise independent
and $\mu = E[X_1 + \dots + X_n]$. then for $\beta > 0$,

$$P[X_1 + \dots + X_n \geq \mu + \beta] \leq \frac{4 \max\{\mu, \mu^3\}}{\beta^4}$$

Main invariant

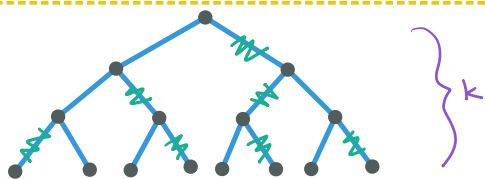
if x occupies cell i , then
all of $[h(x), i]$ occupied



Lecture 4

Probability

Galton-Watson processes



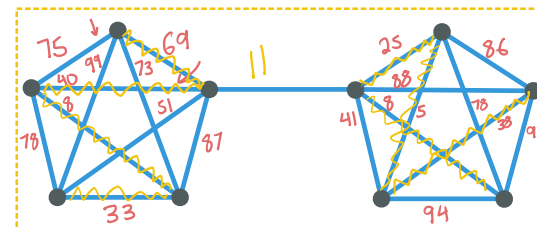
let T = complete binary tree of height k ,
w/ every edge deleted w/ prob $\leq 1/2$

Fact: $P[\text{some leaf connected to root}] \geq \frac{1}{k}$

"Galton-Watson process"

Algorithms

Random Contractions for min cut



Randomized contractions

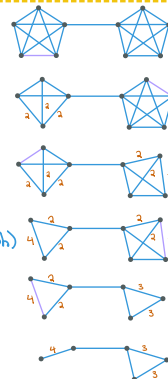
Input: $G=(V,E)$, weights $c \in \mathbb{R}_{>0}^E$
(unweighted)

① while $|E| > 1$

 Ⓐ sample $e \in E$ in proportion to c

 Ⓑ "contract" e

② return set of edges (in original graph)
 contracted to remaining edge



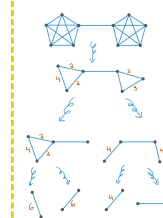
Amplification by branching

Instead of restarting at the beginning, restart partway through

High-level algorithm

① randomly contract until prob.
of ruining the min-cut is $\sim 1/2$

② Branch:
make two copies of current graph
recurse on both copies



Lecture 5

Probability

Multiplicative Chernoff Bound

Multiplicative Chernoff bound

let $X_1, \dots, X_n \in [0, 1]$ be independent

$$\mu = E[X_1 + \dots + X_n]$$

$$\epsilon \in (0, 1)$$

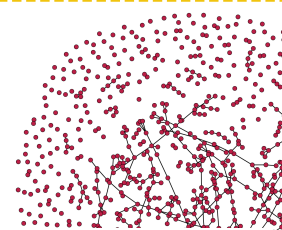
- $P[X_1 + \dots + X_n \geq (1 + \epsilon)\mu] \leq e^{-\epsilon^2 \mu / 3}$
- $P[X_1 + \dots + X_n \leq (1 - \epsilon)\mu] \leq e^{-\epsilon^2 \mu / 2}$

Algorithms (/ results)

Random graphs & the giant component

Erdős-Rényi graphs $G(n, p)$

- n vertices
- each edge sampled independently w/ probability p



Theorem 5.3. Consider a random graph $G \sim \mathcal{G}(n, p)$ for $p = c/n$, where c is a constant.

- If $c > 1$, then with high probability, there is exactly one connected component of G with $\Omega(n)$ vertices, and all other components have size $\leq O(\log n)$.
- For $c < 1$, then with high probability, all connected components of G has size $< O(\log n)$.

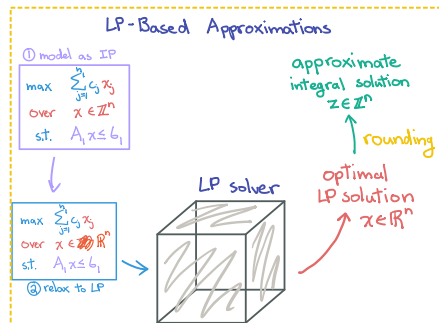
Lecture 6

Probability (& techniques)

Approximations via linear programming & randomized rounding

Linear programs (LPs)
Similar to IPs, except no integrality constraints

min/max $\sum_{j=1}^n c_j x_j$ linear objective
over $x \in \mathbb{R}^n$ linear integral variables
st. $A_1 x \leq b_1$ linear constraints
 $A_2 x = b_2$
 $A_3 x \geq b_3$



Continuous optimization vs Discrete optimization

randomization

Algorithms

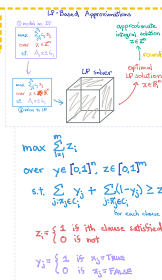
Max Sat Set Cover Covering Integer Programs

For each variable x_j , independently, randomly set $x_j = t$ with probability g_j , and $x_j = f$ otherwise.

$$E[\# \text{ sat. clauses}] \geq (1 - 1/e) \text{OPT}$$

k	oblivious	LP	average
1	1/2	$z_i \geq (3/4)z_i$	$\geq (3/4)z_i$
2	3/4	$(3/4)z_i \geq (3/4)z_i$	$\geq (3/4)z_i$
3	7/8	$(1 - 1/e)z_i \geq (3/4)z_i$	$\geq (3/4)z_i$
4	15/16	$(1 - 1/e)z_i \geq (3/4)z_i$	$\geq (3/4)z_i$
5	31/32	$(1 - 1/e)z_i \geq (3/4)z_i$	$\geq (3/4)z_i$
...	...	...	...
k	$1 - 1/2^k$	$(1 - 1/e)z_i \geq (3/4)z_i$	$\geq (3/4)z_i$

Can we get $\frac{3}{4}$ -APX?



Rounding LP solution
 $x \in [0, 1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$

For each set S_j , independently, take S_j with probability x_j .

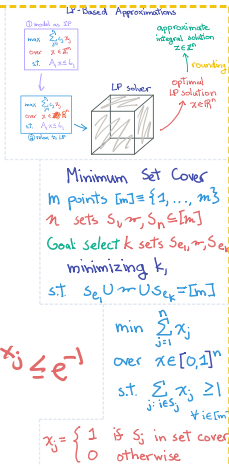
$$E[|F|] = \sum x_j \leq \text{OPT}$$

Fix a point i .

$$E[\# \text{ sets in } F \text{ covering } i] = \sum_{j \in S_i} x_j \geq z_i$$

$$P[F \text{ not cover } i] = \prod_{j \in S_i} (1 - x_j) \leq e^{-\sum_{j \in S_i} x_j} \leq e^{-z_i}$$

$$P[F \text{ cover } i] \geq 1 - e^{-1}$$



Lecture 7

Probability

Continuous Random Variables

Variance, Chebyshev's Variance Reduction

Median trick

Variance

$$\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

Chebyshev's inequality

$$P[|X - E[X]| \geq d] \leq \frac{\text{Var}(X)}{d^2}$$

"median trick", "median of means"

- ① take mean to reduce variance, get constant prob. of error
- ② take median to amplify success prob. exponentially

Algorithms

Distinct Elements



Goal: estimate # of distinct elements in the stream

Suppose you had an ideal hash function

$$h: [m] \rightarrow [0, 1]$$

and the following facts:

let $Y_1, \dots, Y_k \in [0, 1]$ independent and uniformly random

$$X = \min\{Y_1, \dots, Y_k\}$$

$$\text{then } E[X] = \frac{1}{k+1}$$

$$\textcircled{1} X \leftarrow 1$$

② for each element e

$$\textcircled{a} X \leftarrow \min\{X, h(e)\}$$

$$/* E[X] = \frac{1}{k+1}, k = (\text{\# distinct elements}) */$$

$$\textcircled{3} \text{ return } \frac{1}{X} - 1$$

$$\text{recall } E[X] = \frac{1}{k+1}$$

$$\text{goal: show } |X - E[X]| \leq \frac{\epsilon}{k+1}$$

w/ good prob.

$$P[|X - E[X]| \geq \frac{\epsilon}{k+1}]$$

$$\leq \text{Var}(X) \cdot \left(\frac{k+1}{\epsilon}\right)^2 \leftarrow \text{Chebyshev's inequality}$$

where

$$\text{Var}(X) = \frac{k}{(k+1)^2(k+2)} \approx \frac{1}{k^2}$$

Q: how to reduce $\text{Var}(X)$?

we can get $(1 \pm \epsilon)$ -APX

w/ 10% error prob. in $O(\frac{1}{\epsilon^2})$ space

what about high probability?

① make q copies $Y_1, \dots, Y_q$ of

② return median of $Y_1, \dots, Y_q$

Algo correct $\Leftrightarrow \geq \frac{1}{2}$ of Y_i 's correct

let $Z = \# \text{correct } Y_i$'s $E[Z] = \frac{q}{2}$

$$P[Z \leq (1 - \frac{\epsilon}{4})E[Z]] \leq e^{-\left(\frac{\epsilon}{4}\right)^2 E[Z]/2} \leq \frac{1}{n^{O(n)}} \text{ for } q = O(\log n)$$

ideal hash $h: [m] \rightarrow [0, 1]$

let $Y_1, \dots, Y_k \in [0, 1]$

indep. and unifs. rand.

$$X = \min\{Y_1, \dots, Y_k\}$$

$$\text{then } E[X] = \frac{1}{k+1}$$

$$\text{and } E[X^2] = \frac{2}{(k+1)(k+2)}$$

$$\text{Var}(X) = E[(X - E[X])^2]$$

$$X_1, \dots, X_q \text{ indep. copies of } X$$

$$\bar{X} = \frac{1}{q}(X_1 + \dots + X_q)$$

$$\text{Var}(\bar{X}) = \frac{1}{q} \text{Var}(X)$$

$$P[X_1 + \dots + X_q \leq (1 - \epsilon)q] \leq e^{-\epsilon^2 q/4}$$

Lecture 8

Probability

Gaussian random variables

Gaussian random variables

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

mean variance

Def: the nicest, most pleasant, convenient random variable ever

Fact let $x \sim \mathcal{N}(\mu, \sigma^2)$, $d \in \mathbb{R}$.

$$\text{then } x+d \sim \mathcal{N}(\mu+d, \sigma^2)$$

$$dx \sim \mathcal{N}(d\mu, d^2\sigma^2)$$

Fact let $x_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $x_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$

$$\text{then } x_1 + x_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Fact let $x \sim \mathcal{N}(0, 1)$, $t < \frac{1}{2}$

$$\text{then } E[e^{tx^2}] = \frac{1}{\sqrt{1-2t}}$$

Algorithms

Dimensionality reduction

Bag-of-words

literally any sentence is fine

any ↑ is literally sentence English
five encodes text in high dimensions

Lemma 8.8. Let $A \sim \mathcal{N}^{k \times d}$ be a random matrix where each coordinate A_{ij} is an independently drawn sample from $\mathcal{N}$. Let $\epsilon > 0$ be sufficiently small. For any vector x ,

$$\mathbb{P}\left[(1-\epsilon)\|x\|^2 \leq \frac{1}{k}\|Ax\|^2 \leq (1+\epsilon)\|x\|^2\right] \geq 1 - 2e^{-k/8}$$

$$k \left\{ \left[\frac{1}{k} \|Ax\|^2 \right] \leftarrow \alpha_i \right.$$

$$E[\|Ax\|^2] = \sum_{i=1}^k E[(Ax)_i^2] = k\|x\|^2$$

$$E[(\sum_{i=1}^k A_{ij} x_j)^2] = \|x\|^2$$

$$A_{ij} x_j \sim \mathcal{N}(0, x_j^2)$$

$$\sum_{j=1}^n A_{ij} x_j \sim \mathcal{N}(0, \|x\|^2)$$

$$(Ax) \sim \mathcal{N}^k(0, \|x\|^2)$$

$$\Pr[\|Ax\|^2 \geq k\|x\|^2 \cdot (1+\epsilon)] \leq e^{-\epsilon^2 k/8}$$

Dimensionality Reduction - high level

$O(\log n)$ dimensions suffice

where $n = \# \text{ points}$

$$\mathbb{R}^{\text{very large}} \xrightarrow{f(x)} \mathbb{R}^{O(\log n)}$$

One can map n points into $\mathbb{R}^{O(\log n/\epsilon^2)}$ while preserving all distances up to $(1 \pm \epsilon)$ -APX

Lecture 9

Probability

Locality-sensitive hashing

Fix $\theta \in [0, \pi]$. given q , either:

(a) output $p \in P$ w/ $\angle(p, q) \leq 4\theta$

(b) declare no $p \in P$ w/ $\angle(p, q) \leq \theta$

define $h: S^{d-1} \rightarrow \{-1, 1\}^k$ by

$$h_i(x) = \text{sign}(\langle g_i, x \rangle) \text{ w/ } g_i \sim N^d$$

$$k = \pi(\log(n)/2\theta)$$



d-dimensional sphere

when querying y :

• if there is $x \in P$ w/ $\angle(x, y) \leq \theta$, then it collides w/ probability $\geq 1/\sqrt{n}$

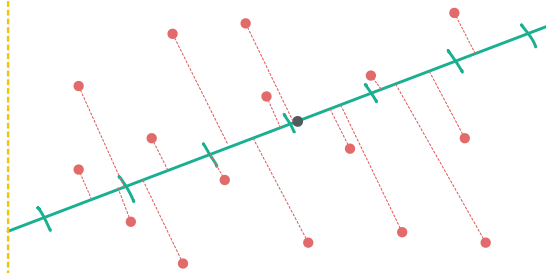
• expect to collide w/ at most one point $x \in P$ w/ $\angle(x, y) > 2\theta$

Lemma let $\bullet x, y \in S^{d-1}$
 $\bullet g \sim N^d$
 then

$$P[\text{sign}(\langle g, x \rangle) \neq \text{sign}(\langle g, y \rangle)] = \frac{\angle(x, y)}{\pi}$$

Random line embedding & Buckets

$$h(x) = \lfloor \langle g, x \rangle + d \rfloor \text{ w/ } g \sim N^d, d \in [0, 1] \text{ unifs. rand.}$$

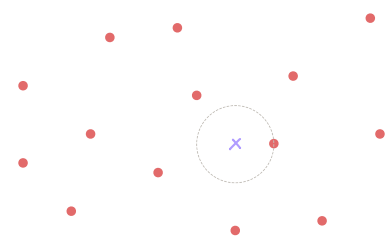


Algorithms

Approximate nearest neighbors

Nearest neighbor search

Input: n points $P \subset \mathbb{R}^d$



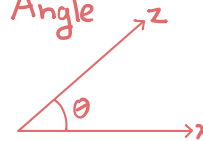
Query: given $q \in \mathbb{R}^d$, find closest point in P to q

Types of distances

• Euclidean distance

$$\|x - z\|_2 = \sqrt{\sum_i (x_i - z_i)^2}$$

• Angle



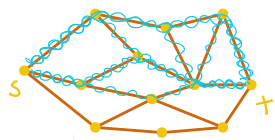
Lecture 10

Probability (and techniques)

LP duality

L_1 -metric embeddings

Max Flow Min Cut



Path Packing LP

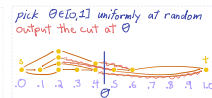
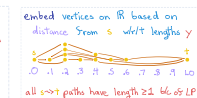
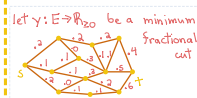
maximize $\sum_{p \in P} x_p$
over $x: P_{st} \rightarrow \mathbb{R}_{\geq 0}$
(all s-t paths)

s.t. $\sum_{p \in P} x_p \leq c_e$ for all edges e
"Ax ≤ c" (capacity of e)

Path Covering LP (i.e. cut)

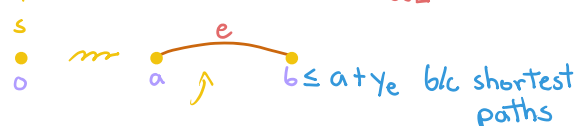
minimize $\sum_{e \in E} c_e y_e$
over $y: E \rightarrow \mathbb{R}_{\geq 0}$
(total capacity of "fractional cut")

s.t. $\sum_{e \in P} y_e \geq 1$ for all $p \in P_{st}$
"A'y ≥ b" (continuous analog of s-t cut)



$E[\text{capacity of cut}]$

$$\stackrel{\text{(linearity of expectation)}}{=} \sum_{e \in E} c_e \Pr[e \text{ cut}] \leq \sum_{e \in E} c_e y_e$$

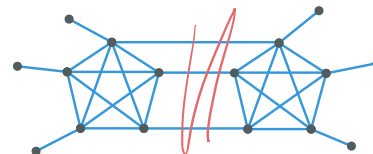


Algorithms

Sparsest Cut

Special case: uniform sparsest cut
w/ uniform demands $b(u,v) = 1 \forall u,v$

$$\text{"sparsity"} = \frac{c(\delta(s))}{|S| |T|}$$



useful primitive for graph partitioning, divide & conquer

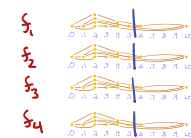
Q: how to round LP to sparse cut?

min $\sum_{e \in E} c_e d(u,v)$ over metrics $d: V \times V \rightarrow \mathbb{R}_{\geq 0}$
s.t. $\sum_{\{s,t\}} b(s,t) d(s,t) \geq 1$

Suppose $d(u,v)$ was of the form

$$d(u,v) = \sum_{i=1}^k |\xi_i(u) - \xi_i(v)| \text{ for } \xi_i: V \rightarrow [0,1]^k$$

(assume $\min \xi_i(a) = 0, \max \xi_i(a) = 1$)



" L_1 -metric"

Sparsest cut
Input: undirected graph $G=(V,E)$, edge capacities $c(e) > 0$ for $e \in E$, demands $b(s,t) > 0$ for $s,t \in V$.
Goal: compute SCF minimizing "sparsity" $\frac{c(\delta(s))}{|S| |T|}$ where S, T are edge cuts.

Lecture 11

Probability (and techniques)

Algorithms

Other metric-based problems
(sometimes via correspondence between cuts $\leftrightarrow$ metrics)

Buy-at-bulk network design
k-median clustering
metric labeling
group Steiner tree
metrical task systems
vehicle routing
hierarchical placement
distributed k-server
covering Steiner tree
⋮



these problems are easy for tree metrics

Approximation by tree metrics

Input: undirected $G=(V,E)$ w/ edge weights $w:E \rightarrow \mathbb{R}_{\geq 0}$

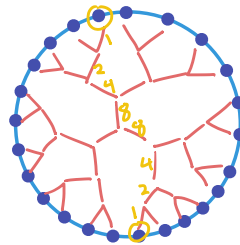
(High-level)
Goal: compute tree T (containing V as nodes)
so that shortest paths on T are close to T

$$d_T(u,v) \geq d_G(u,v)$$

$$\text{For } e=f_u,v \quad \begin{matrix} \text{(distance in } T) \\ \text{(stretch } e) = \frac{d_T(u,v)}{d_G(u,v)} \\ \text{(distance in } G) \end{matrix}$$

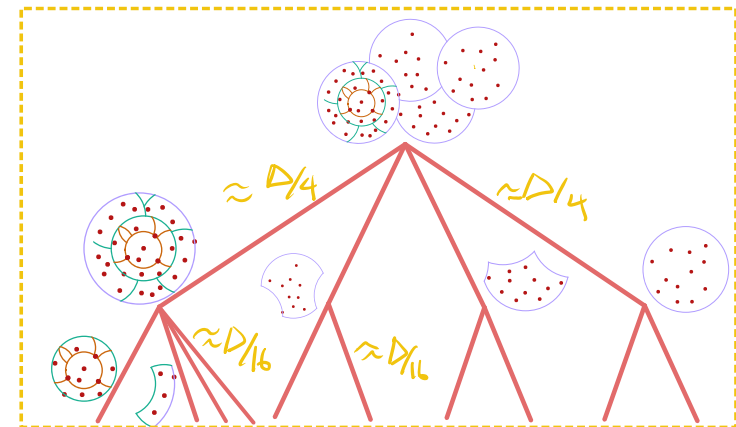
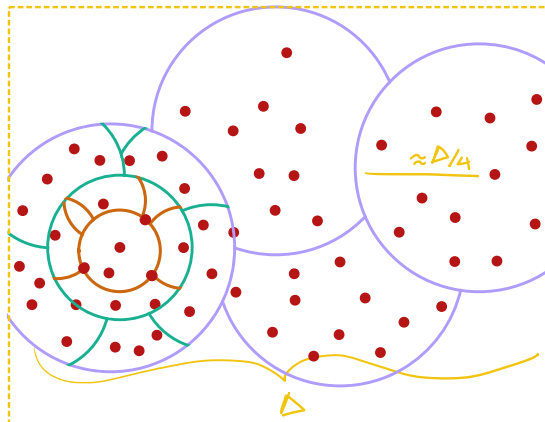
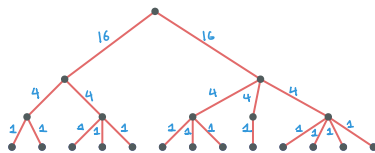
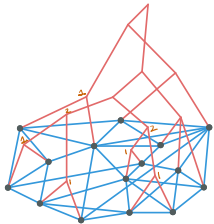
ideally

$$(\text{stretch } e) = \text{polylog}(n) \quad \forall e$$



Randomized Tree metrics

Theorem In randomized polynomial time,
randomized hierarchical dominating tree metric w/
 $E[\text{stretch } e] \leq O(\log n) \quad \forall e \in E$



Lecture 12

Probability (and techniques)

Additive Chernoff Inequality

Double-sampling

VC dimension

Additive Chernoff bound

let $X_1, \dots, X_n \in [0, 1]$ be independent

$$\mu = E[X_1 + \dots + X_n] / n$$

$\mu \in (0, 1)$

- $P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \leq e^{-2\epsilon^2 n}$
- $P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n}$

Algorithms

ϵ -samples and ϵ -nets

ϵ -sample $|M_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition
 $k = O\left(\frac{\ln(g(|P|))}{\epsilon^2}\right) \Rightarrow \epsilon$ -sample $k = O\left(\frac{\ln(g(|P|))}{\epsilon^2}\right) \Rightarrow \epsilon$ -net (w/ constant probability)

ϵ -sample theorem [VC71]
 $k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right) \Rightarrow \epsilon$ -sample (w/ constant probability)

e.g. $g(n) = O(n^d)$
 $k \geq O\left(\frac{d \ln(d/\epsilon)}{\epsilon^2}\right) \Rightarrow \epsilon$ -sample
 ↑ independent of $|P|$ (!)

$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right], P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n}$ $g(n) = \max\{|RAQ| : a \in P, |a| \leq n\}$

$\mu(r) = \frac{|P \cap r|}{|P|}$
 $M_Q(r) = \frac{|a \cap r|}{|a|}$
 $RAQ = \{r \cap a : r \in R, a \in P\}$

ϵ -sample $|M_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition
 $k = O\left(\frac{\ln(g(|P|))}{\epsilon^2}\right) \Rightarrow \epsilon$ -sample $k = O\left(\frac{\ln(g(|P|))}{\epsilon^2}\right) \Rightarrow \epsilon$ -net (w/ constant probability)

ϵ -net theorem [HW87]
 $k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right) \Rightarrow \epsilon$ -net (w/ constant probability)

e.g. $g(n) = O(n^d)$
 $k \geq O\left(\frac{d \ln(d/\epsilon)}{\epsilon^2}\right) \Rightarrow \epsilon$ -net
 ↑ independent of $|P|$ (!)

$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right], P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n}$ $g(n) = \max\{|RAQ| : a \in P, |a| \leq n\}$

$\mu(r) = \frac{|P \cap r|}{|P|}$
 $M_Q(r) = \frac{|a \cap r|}{|a|}$
 $RAQ = \{r \cap a : r \in R, a \in P\}$

Lecture 13

Probability
(and techniques)

Algorithms
(models and applications)

Consistency model

PAC learning model

Generalization error

Occam's Razor

Generalization by growth

Lecture 14

Probability
(and techniques)

Algorithms
(models and applications)

Entropy

(Optimal)
Error-Correcting
Codes

Lecture 15

Probability (and techniques)

Lovász Local Lemma

Algorithms (models and applications)

Moser's resampling
algorithm

Lecture 16

Probability (and techniques)

Algorithms (models and applications)

Random Walks

Stationary Distribution

Perron-Frobenius Theorem

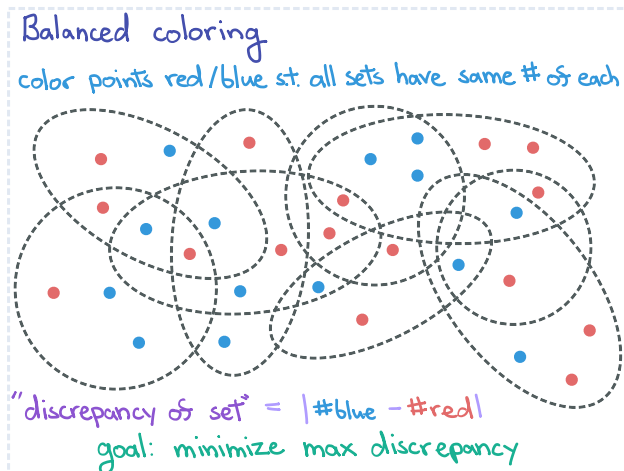
Pagerank

Lecture 17

Probability (and techniques)

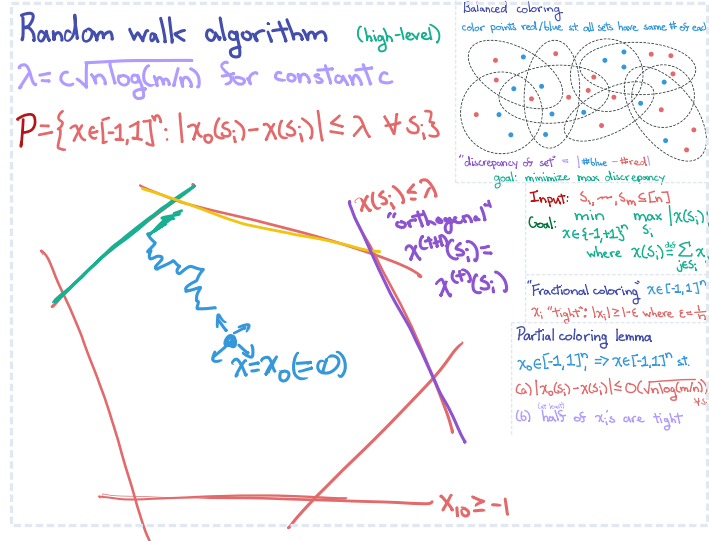
High-dimensional random walks

More fun w/ Gaussians



Algorithms (models and applications)

Discrepancy minimization



Lecture 18

Probability (and techniques)

Algorithms (models and applications)

Hitting times

Cover times

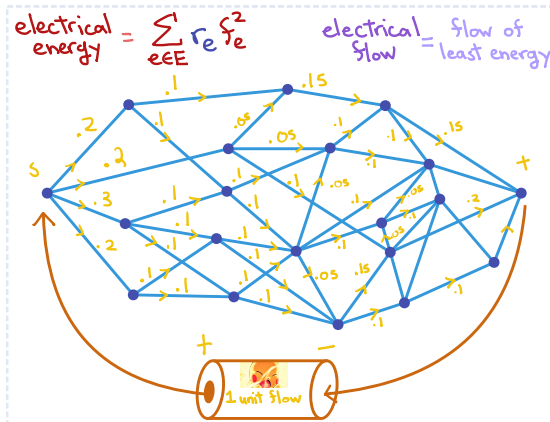
Electrical networks

Connectivity
in $O(\log n)$ space

random walk terminology:

Hitting time $H(s, t)$ = expected # steps until a random walk from s reaches t

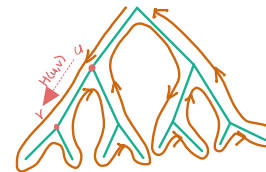
Cover time $C(s)$ = expected # steps until a random walk from s visits every vertex



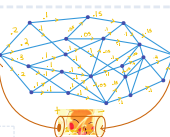
Lemma $H(u, v) + H(v, u) = 2m \cdot (\text{energy of unit flow on } (u, v))$

Theorem $C(s) \leq 2mn$

proof Six traversal of spanning tree



$$\begin{aligned} \text{total time} &= \sum_{u, v \in V} H(u, v) + H(v, u) \\ &= 2m \sum_{u, v \in V} (\text{energy of unit flow on } (u, v)) \leq 2m(n-1) \end{aligned}$$



electrical energy = $\sum_{e \in E} r_e f_e^2$
electrical flow = least energy

des $B: \mathbb{R}^E \rightarrow \mathbb{R}^V$ by
(BS) $_v$ = net flow at v
 $d_v = \begin{cases} -1 & \text{if } v = s \\ 1 & \text{if } v = t \\ 0 & \text{otherwise} \end{cases}$

min $\frac{1}{2} \sum_{e \in E} r_e f_e^2$ s.t. $BS = d$
(electrical energy) (is a typical)

effective resistance of d = optimum value/energy
Ohm's law: $f = R^{-1} B^T p$
for electric flow by electric potentials p
 $L = BR^{-1}B^T \in \mathbb{R}^{V \times V}$
(effective resist. of d) $= \langle d, L^{-1}d \rangle$

Lecture 19

Probability (and techniques)

Algorithms (models and applications)

Spectral theorem
for symmetric maps

Spectral analysis of
random walks

Lecture 20

Probability
(and techniques)

Algorithms
(models and applications)

Conductance

Cheeger's inequality

Conductance-based
mixing time

Fiedler's algorithm
for low-conductance
cuts

Lecture 21

Probability
(and techniques)

Algorithms
(models and applications)

Expanders

Zigzag product

tensors

Deterministic

Connectivity in

$O(\log n)$ space

Lecture 22

Probability
(and techniques)

classes RP and BPP

tensors of graphs

Algorithms
(models and applications)

More efficient
amplification

Efficient construction
of exponentially
large expanders

Lecture 23/24

Probability
(and techniques)

Algorithms
(models and applications)

Probabilistically
checkable proofs

PCP theorem [90's]

$$NP = PCP(\underbrace{O(\log n)}_{\text{random bits}}, \underbrace{O(1)}_{\text{queries}})$$

leads to:

Theorem. [Håstad 1997]

$\forall \epsilon > 0$, $(7/8 + \epsilon)$ -APX for 3SAT is NP-Hard

A language L is in...

P _(polytime) if there is a det. polytime decision algo

NP _(nondeterministic polytime) if there is a polytime "verifier" that takes as input $x \in \{0,1\}^n$ and $y \in \{0,1\}^{\text{poly}(n)}$ and:

- if $x \in L$, answers " $x \in L$ " for some y
- if $x \notin L$, answers " $x \notin L$ " for all y

PCP(r, q): (nonadaptive) probabilistically checkable proof w/ r random bits and q queries

Oracle model: query i th bit of proof y

Verifier:

Given $x \in \{0,1\}^n$ and oracle access to $y \in \{0,1\}^{\text{poly}(n)}$:

1. select $O(q)$ queries based on x , $O(r)$ random bits
2. make $O(q)$ queries to oracle
3. accept or reject x
 - if $x \in L$, for some y , (always) accept x
 - if $x \notin L$, for all y , accept x w/ prob. $\leq 1/2$

Graph CSP

2-ary CSP

variables = vertices

clause over $u, v \in V \approx$ edge between u and v

q -ary constraint satisfaction problems

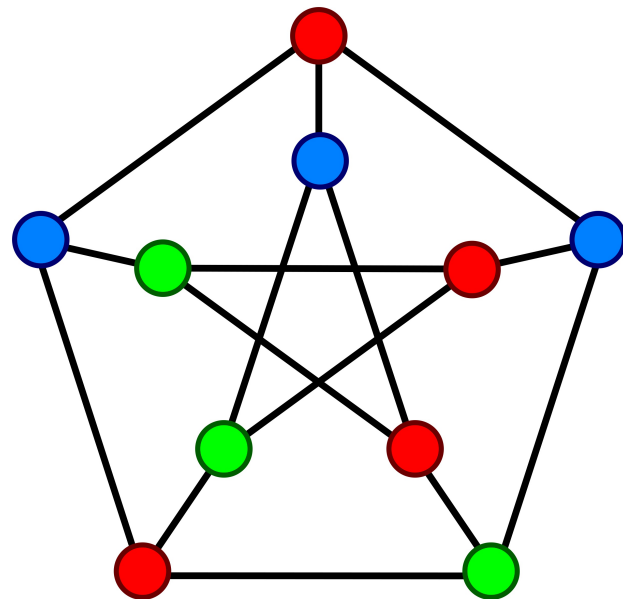
- n variables V
- finite alphabet A
- m clauses
- a clause consists of q variables $v_1, \dots, v_q \in V$ and "satisfying" set $S \subseteq A^q$

$\pi: V \rightarrow A$ satisfies clause if $(\pi(v_1), \dots, \pi(v_q)) \in S$

$V = \text{vertices}$

$A = \{\text{red, blue, green}\}$

clauses = edges



given graph-CSP G , it is NP-Hard

to distinguish:

- $UNSAT(G) = 0$
- $UNSAT(G) \geq \frac{1}{m}$

Goal: amplify $\nearrow$ to a constant

q -ary constraint satisfaction problems

- n variables V
- finite alphabet A
- m clauses
- a clause consists of q variables $v_1, \dots, v_k \in V$ and "satisfying" set $S \subseteq A^q$
- $\pi: V \rightarrow A$ satisfies clause if $(\pi(v_1), \dots, \pi(v_k)) \in S$

Graph CSP 2-ary CSP
Variables = vertices
clause over $u, v \in V$
 $\approx$ edge between u, v

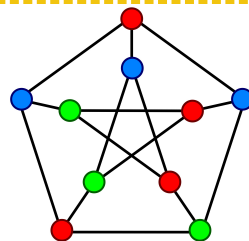
Three transformations:

1. "Expander-ification" make G a constant degree expander
2. Powering raises $UNSAT(G)$ but also increases $|A|$
3. Alphabet reduction

V = vertices

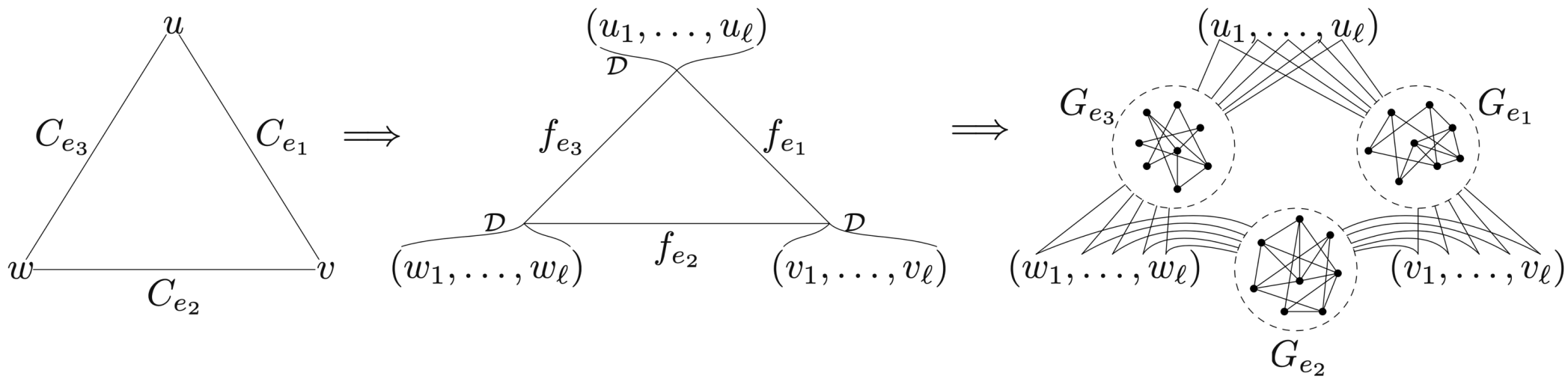
$A = \{\text{red, blue, green}\}$

clauses = edges



3. Alphabet reduction

goal: decrease alphabet size to constant



step 1: reduce to $A = \{0, 1\}$ w/ error-correcting codes

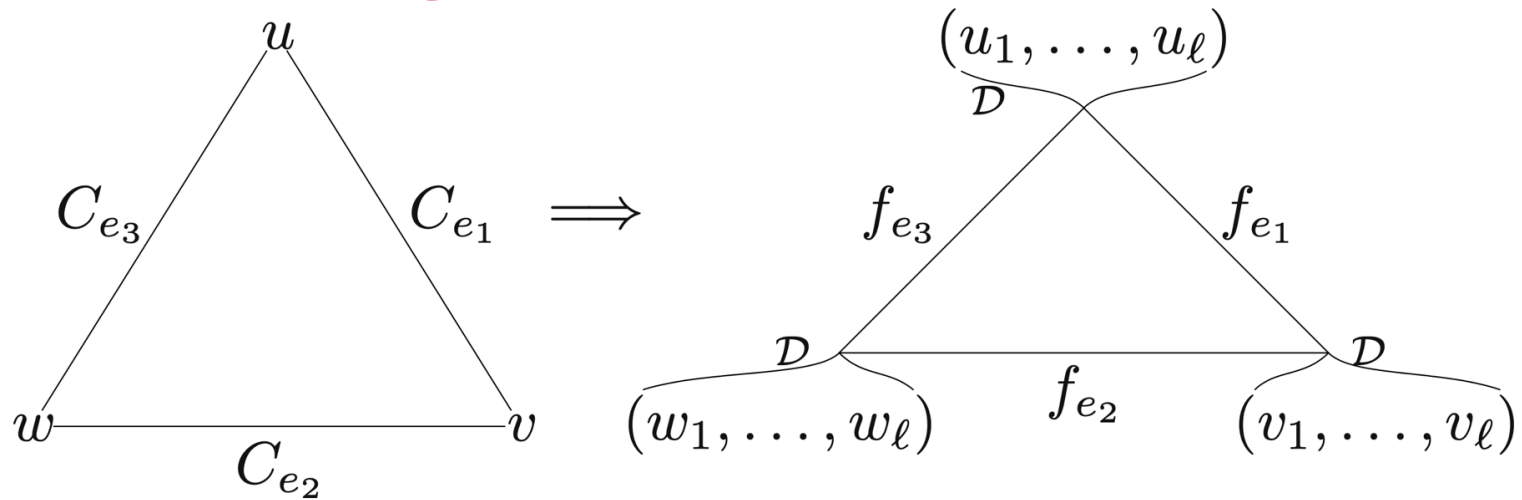
* $\uparrow$ replaces each edge w/ general constraint over $O(1)$ variables *

step 2: replace each constraint w/ graph CSP

* use techniques as black box *

step 1: reduce to $A = \{0,1\}$ w/ error-correcting codes

* $\uparrow$ replaces each edge w/ general constraint over $O(1)$ variables *



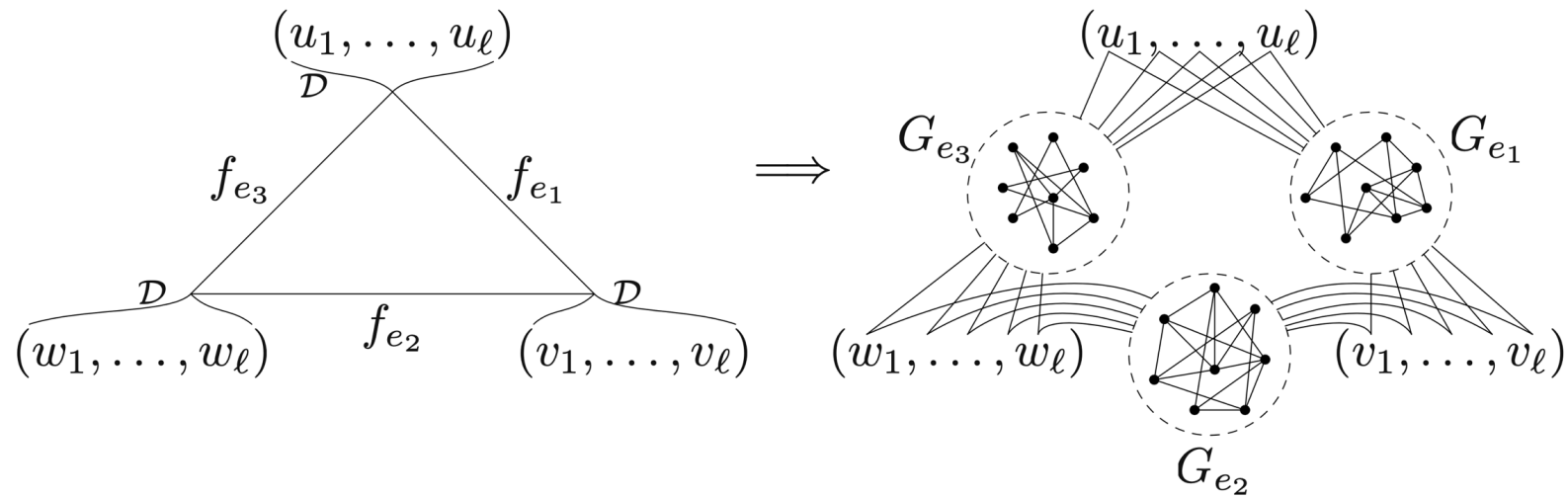
let $c: A \rightarrow \{0,1\}^\ell$ be code w/ $\ell = O(\log|A|)$ and
relative distance $\rho \in [0,1]$ ($\text{Hamming}(c(a_1), c(a_2)) \geq \rho \ell \ \forall a_1 \neq a_2$)

split each vertex u into ℓ vertices $V_u = \{v_1, \dots, v_\ell\}$

replace edge $e = \{u, v\}$ w/ formula $f_e: \{0,1\}^{V_u} \times \{0,1\}^{V_v}$ s.t.

$$f_e(x_u, x_v) = \begin{cases} 1 & \text{if } x_u = c(a_1), x_v = c(a_2), (a_1, a_2) \in C_e \\ 0 & \text{otherwise} \end{cases}$$

step 2: replace each constraint w/ graph CSP



Lemma: for $e \in E$, one can construct graph-CSP $G_e = (V_e, E_e, A_0)$ s.t.:

(a) $V_u, V_v \subseteq V_e$ (b) $|E_e|, |V_e|, |A_0| = O(1)$ and same $\forall e$

(c) completeness: if $f_e(x_u, x_v) = 1$, then one can extend (x_u, x_v) to $\sigma: V_e \rightarrow A_0$ that satisfies all of G_e

(d) soundness: if $f_e(x_u, x_v) = 0$, then $\forall$ extensions σ of (x_u, x_v) ,

$$\text{UNSAT}(\sigma|G_e) \geq \Omega\left(\min\left\{\frac{\text{Hamming}((x_u, x_v), (y_u, y_v))}{2\ell}: f(y_u, y_v) = 1\right\}\right)$$

Lecture 25

Probability (and techniques)

Discrete Fourier analysis

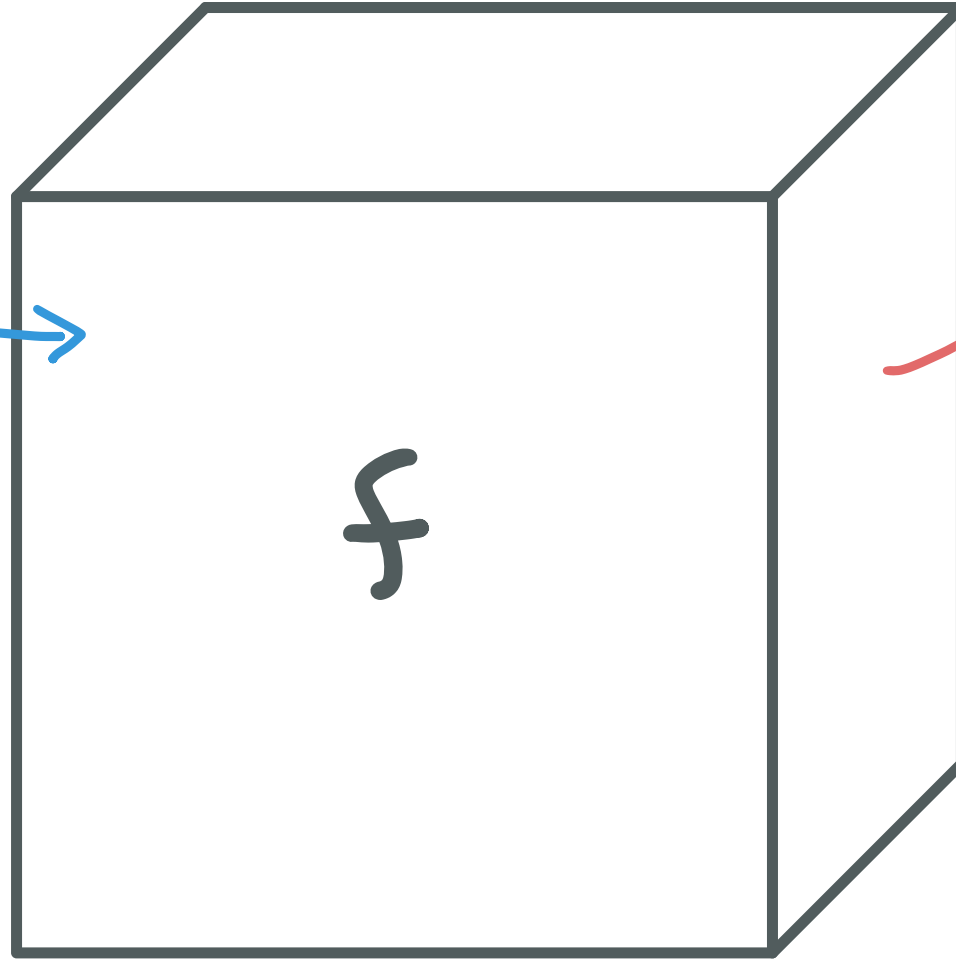
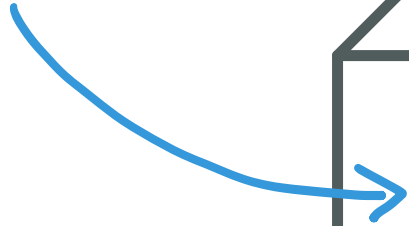
Algorithms (models and applications)

Property testing:

- Linearity test
- Dictatorship test
- Universal tester

Boolean Function as a black box

$x \in \{0,1\}^n$



$y \in \mathbb{R}$

e.g. is f a linear function?

Property testing:

given f , test (via queries $f(x)$) if f has certain properties like:

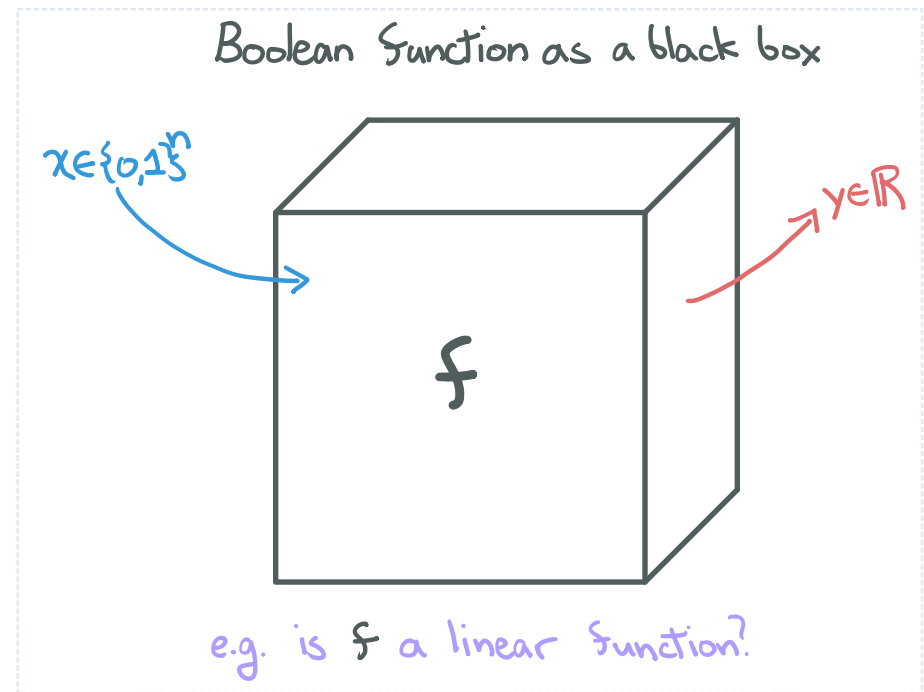
(a) linearity: $f(x+y) = f(x) + f(y)$
 $\forall x, y \in \{0,1\}^n$

(b) dictatorship: $f(x) = (-1)^{x_i}$
for some index $i \in [n]$

goal: test f w/ as few queries as possible

Impossible to be exact w/out querying all 2^n inputs

$\Rightarrow$ allow for randomization and approximations



Universal Tester

(today's main result)

let $L \subseteq \{0,1\}^n$ and $p = 2^{2^n}$

randomized algo $A_L: \{0,1\}^n \times \{0,1\}^p \rightarrow \{0,1\}$ that, given oracle access to $x \in \{0,1\}^n$ & $y \in \{0,1\}^p$, queries 3 bits of x or y and:

- If $x \in L$, $\exists y \in \{0,1\}^p$ s.t. $A_L(x,y) = 1$ always
- If $x \notin L$, $\forall y \in \{0,1\}^p$,

$$P[A_L(x,y) = 0] \geq .001 \min_{z \in L} \frac{\|x - z\|_0}{n}$$

Comparing to PCP

very large proof
(ok for $n = O(n)$)

PCP(r, q): (nonadaptive) probabilistically checkable proof
w/ r random bits and q queries

Oracle model: query i th bit of proof y

Verifier:

Given $x \in \{0,1\}^n$ and oracle access to $y \in \{0,1\}^{\text{poly}(n)}$:

1. select $O(q)$ queries based on x , $O(r)$ random bits
2. make $O(q)$ queries to oracle
3. accept or reject x
 - if $x \in L$, for some y , (always) accept x
 - if $x \notin L$, for all y , accept x w/ prob. $\leq 1/2$

Universal Tester

let $L \subseteq \{0,1\}^n$ and $p = 2^{2^n}$

randomized algo $A_L: \{0,1\}^n \times \{0,1\}^p \rightarrow \{0,1\}$ that, given oracle access to $x \in \{0,1\}^n$ & $y \in \{0,1\}^p$, queries 3 bits of x or y and:

- If $x \in L$, $\exists y \in \{0,1\}^p$ s.t. $A_L(x, y) = 1$ always

- If $x \notin L$, $\forall y \in \{0,1\}^p$,

$$P[A_L(x, y) = 0] \geq .001 \min_{y \in L} \frac{\|x - y\|_0}{n}$$

rejection probability proportional
to Hamming distance from L

Algo (high-level sketch)

let $N = 2^n$, identified w/ $\{0,1\}^n$

for $i \in [N]$, def $\chi_i: \{0,1\}^n \rightarrow \{-1,1\}$:

$$\chi_i(x) \stackrel{\text{def}}{=} (-1)^{x_i} = \begin{cases} 1 & \text{if } x_i = 0 \\ -1 & \text{if } x_i = 1 \end{cases}$$

"dictator"

let $D_L = \{\chi_i: i \in L\}$

a boolean function $f: \{0,1\}^n \rightarrow \{-1,1\}$ is a $(p=2^n)$ -bit string

for $x \in L$, the "proof" y is $\chi_x \in \{0,1\}^p$

The test: given $x \in \{0,1\}^n$ and $y \in \{0,1\}^p$:

3 queries

1. test if $y \in D_L$

2. test if $y = \chi_x$

← we will learn how to do these

Universal Tester

(today's main result)

let $L \subseteq \{0,1\}^n$ and $p = 2^n$

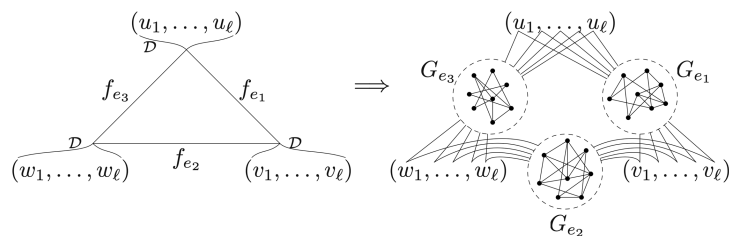
randomized algo $A_L: \{0,1\}^n \times \{0,1\}^p \rightarrow \{0,1\}$ that, given oracle access to $x \in \{0,1\}^n$ & $y \in \{0,1\}^p$, queries 3 bits of x or y and:

- If $x \in L$, $\exists y \in \{0,1\}^p$ st. $A_L(x,y) = 1$ always

- If $x \notin L$, $\forall y \in \{0,1\}^p$,

$$P[A_L(x,y) = 0] \geq 0.001 \min_{y \in L} \frac{\|x-y\|_0}{n}$$

step 2: replace each constraint w/ graph CSP



Lemma: for $e \in E$, one can construct graph-CSP $G_e = (V_e, E_e, A_e)$ st.:

(a) $\forall u, v \subseteq V_e$ (b) $|E_e|, |V_e|, |A_e| = O(\ell)$ and same $\forall e$

(c) completeness: if $f_e(x_u, x_v) = 1$, then one can extend (x_u, x_v) to $\sigma: V_e \rightarrow A_e$ that satisfies all of G_e

(d) soundness: if $f_e(x_u, x_v) = 0$, then $\forall$ extensions σ of (x_u, x_v) ,

$$\text{UNSAT}(\sigma|G_e) \geq \Omega\left(\min_{\{ \frac{\text{Hamming}((x_u, x_v), (y_u, y_v))}{2\ell} : f(y_u, y_v) = 1 \}}\right)$$

Universal Tester

(today's main result)

let $L \subseteq \{0, 1\}^n$ and $p = 2^{-n}$

randomized algo $A_L: \{0, 1\}^n \times \{0, 1\}^p \rightarrow \{0, 1\}$ that, given oracle access to $x \in \{0, 1\}^n$ & $y \in \{0, 1\}^p$, queries 3 bits of x or y and:

- If $x \in L$, $\exists y \in \{0, 1\}^p$ st. $A_L(x, y) = 1$ always
- If $x \notin L$, $\forall y \in \{0, 1\}^p$,

$$P[A_L(x, y) = 0] \geq .001 \min_{y \in L} \frac{\|x - y\|_0}{n}$$

Theorem. let $L \subseteq \{0,1\}^n$. One can construct graph-CSP $G_L = (V_L, E_L)$ s.t:

- $A = \{0,1\}^3 \equiv \{0,1,\pi,1\}$
- n vertices $v_1, \dots, v_n$ w/ labels in $\{0,1\}$ s.t.
 - if $x \in L$, then $\exists \sigma: V_L \rightarrow \{0,1\}$ s.t. $\sigma(v_i) = x_i \forall i$ and $\text{UNSAT}(\sigma|G_L) = 0$.
 - if $x \notin L$, then $\forall \sigma: V_L \rightarrow \{0,1\}$ s.t. $\sigma(v_i) = x_i \forall i$,

$$\text{UNSAT}(\sigma|G_L) \geq .001 \min_{y \in L} \frac{\|x-y\|_0}{n}$$

Universal Tester

(today's main result)

let $L \subseteq \{0,1\}^n$ and $p = 2^{-n}$

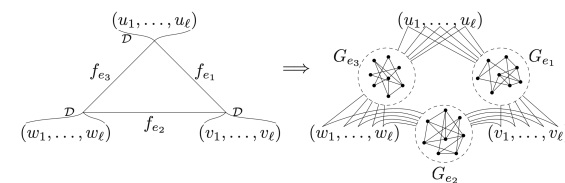
randomized algo $A_L: \{0,1\}^n \times \{0,1\}^p \rightarrow \{0,1\}$ that, given oracle access to $x \in \{0,1\}^n$ & $y \in \{0,1\}^p$, queries 3 bits of x or y and:

• If $x \in L$, $\exists y \in \{0,1\}^p$ s.t. $A_L(x,y) = 1$ always

• If $x \notin L$, $\forall y \in \{0,1\}^p$,

$$P[A_L(x,y) = 0] \geq .001 \min_{y \in L} \frac{\|x-y\|_0}{n}$$

step 2: replace each constraint w/ graph CSP



Lemma: for $e \in E$, one can construct graph-CSP $G_e = (V_e, E_e, A_e)$ s.t.:

(a) $\forall u, v \leq V_e$ (b) $|E_e|, |V_e|, |A_e| = O(1)$ and same $\forall e$

(c) completeness: if $f_e(x_u, x_v) = 1$, then one can extend (x_u, x_v) to $\sigma: V_e \rightarrow A_e$ that satisfies all of G_e

(d) Soundness: if $f_e(x_u, x_v) = 0$, then $\forall$ extensions σ of (x_u, x_v) ,

$$\text{UNSAT}(\sigma|G_e) \geq \Omega\left(\min_{\{x_u, x_v\}: f(x_u, x_v) = 0} \frac{\text{Hamming}((x_u, x_v), (y_u, y_v))}{2\ell}\right)$$

Proof. universal tester T for $L \rightarrow$

T 's input: $x \in \{0,1\}^n$, $y \in \{0,1\}^p$, $p = 2^{2^n}$

T flips finite # coins,

let Ω = all possible outcomes

for $w \in \Omega$, T queries 3 bits of x or y and accepts/rejects.

we want to simulate this w/ a graph

Universal Tester

(today's main result)

let $L \subseteq \{0,1\}^n$ and $p = 2^{2^n}$

randomized algo $A_L: \{0,1\}^n \times \{0,1\}^p \rightarrow \{0,1\}$ that, given oracle access to $x \in \{0,1\}^n$ & $y \in \{0,1\}^p$, queries 3 bits of x or y and:

- If $x \in L$, $\exists y \in \{0,1\}^p$ s.t. $A_L(x,y) = 1$ always

- If $x \notin L$, $\forall y \in \{0,1\}^p$,

$$P[A_L(x,y) = 0] \geq .001 \min_{y \in L} \frac{\|x - y\|_0}{n}$$

Theorem. let $L \subseteq \{0,1\}^n$. One can construct graph-CSP $G_L = (V_L, E_L)$ s.t:

- $A = \{0,1\}^3$

- n vertices $v_1, \dots, v_n \in \{0,1\}$ s.t.

- if $x \in L$, then $\exists \phi: V_L \rightarrow \{0,1\}$ s.t.
 $\phi(v_i) = x_i \forall i$ and $\text{UNSAT}(\phi|G_L) = 0$.

- if $x \notin L$, then $\forall \phi: V_L \rightarrow \{0,1\}$ s.t.
 $\phi(v_i) = x_i \forall i$,

$$\text{UNSAT}(\phi|G_L) \geq .001 \min_{y \in L} \frac{\|x - y\|_0}{n}$$

Vertices:

only take labels $\{0,1\}$

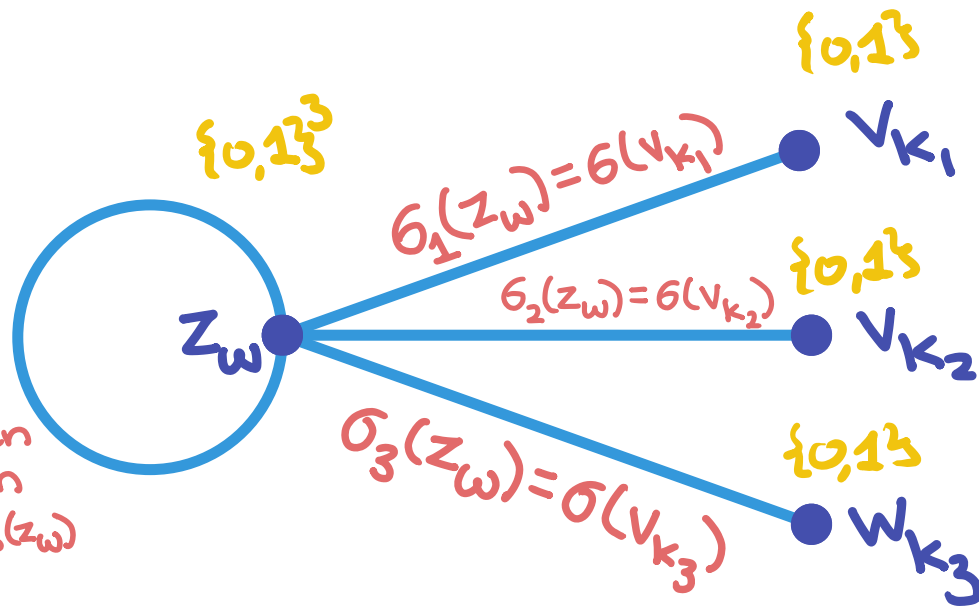
- $v_1, \dots, v_n$ modeling bits of x

- $w_1, \dots, w_p$ modeling bits of y

- z_w for each $w \in \Omega$

Edges: for $w \in \Omega$, T queries 3 bits of x or y and accepts/rejects.
 e.g. T queries k_1 th bit of x ,
 k_2 th bit of x ,
 k_3 th bit of y

4 edges:



T accepts when queries return $g_1(z_w), g_2(z_w), g_3(z_w)$

one can verify this graph-CSP satisfies the claims

Universal Tester

(today's main result)

let $L \subseteq \{0,1\}^n$ and $p = 2^{-n}$

randomized algo $A_L: \{0,1\}^n \times \{0,1\}^p \rightarrow \{0,1\}$ that, given oracle access to $x \in \{0,1\}^n$ & $y \in \{0,1\}^p$, queries 3 bits of x or y and:

• If $x \in L$, $\exists y \in \{0,1\}^p$ s.t. $A_L(x,y) = 1$ always

• If $x \notin L$, $\forall y \in \{0,1\}^p$,

$$P[A_L(x,y) = 0] \geq .001 \min_{y \in L} \frac{\|x - y\|_0}{n}$$

Theorem. let $L \subseteq \{0,1\}^n$. One can construct graph-CSP $G_L = (V_L, E_L)$ s.t:

• $A = \{0,1\}^3$

• n vertices $v_1, \dots, v_n \in \{0,1\}$ s.t.

• if $x \in L$, then $\exists g: V_L \rightarrow \{0,1\}$ s.t. $g(v_i) = x_i \forall i$ and $\text{UNSAT}(g|G_L) = 0$.

• if $x \notin L$, then $\forall g: V_L \rightarrow \{0,1\}$ s.t. $g(v_i) = x_i \forall i$,

$$\text{UNSAT}(g|G_L) \geq .001 \min_{y \in L} \frac{\|x - y\|_0}{n}$$

universal tester T for L

T 's input: $x \in \{0,1\}^n$, $y \in \{0,1\}^p$, $p = 2^{-n}$

T flips finite # coins,

let Ω = all possible outcomes

Vertices:

Only take labels $\{0,1\}$

• $v_1, \dots, v_n$ modeling bits of x

• $w_1, \dots, w_p$ modeling bits of y

• z_w for each $w \in \Omega$

Lecture 25

Probability (and techniques)

Discrete Fourier analysis

Algorithms (models and applications)

Property testing:

- Linearity test
- Dictatorship test
- Universal tester

Vector space of Boolean Functions

$$f: \{0,1\}^n \rightarrow \mathbb{R}$$

closed under sums, rescaling

identified w/ $\mathbb{R}^{2^n}$

Inner product:

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x [f^2(x)]}$$

Lemmas

- for $f, g: \{0, 1\}^n \rightarrow \{0, 1\}$

$$\|f - g\|_6^2 = \mathbb{P}_x[f(x) \neq g(x)]$$

$$\|f - g\|_6^2 = \mathbb{E}_x[(f(x) - g(x))^2]$$

Vector space of $f: \{0, 1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_6 \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0, 1\}^n} [f(x)g(x)]$$
$$\|f\|_6 \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_6} = \sqrt{\mathbb{E}_x[f^2(x)]}$$


Lemmas

- for $f, g: \{0, 1\}^n \rightarrow \{0, 1\}$

$$\|f - g\|_b^2 = \mathbb{P}_x[f(x) \neq g(x)]$$

- for $f, g: \{0, 1\}^n \rightarrow \{-1, 1\}$,

$$\frac{1}{4} \|f - g\|_b^2 = \mathbb{P}[f(x) \neq g(x)]$$

Vector space of $f: \{0, 1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0, 1\}^n} [f(x)g(x)]$$
$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x[f^2(x)]}$$

Lemmas

- for $f, g: \{0, 1\}^n \rightarrow \{0, 1\}$
 $\|f - g\|_6^2 = \mathbb{P}_x[f(x) \neq g(x)]$
- for $f, g: \{0, 1\}^n \rightarrow \{-1, 1\}$,
 $\frac{1}{4} \|f - g\|_6^2 = \mathbb{P}[f(x) \neq g(x)]$
- for $f: \{0, 1\}^n \rightarrow \{-1, 1\}$,
 $\|f\|_6^2 = 1$

Vector space of $f: \{0, 1\}^n \rightarrow \mathbb{R}$
 $\langle f, g \rangle_6 \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0, 1\}^n} [f(x)g(x)]$
 $\|f\|_6 \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_6} = \sqrt{\mathbb{E}_x[f^2(x)]}$

χ_S

for $S \subseteq [n]$, def $\chi_S: \{0,1\}^n \rightarrow \{-1,1\}$

by
$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ -1 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

Fun facts:

• $\chi_S \chi_T = \chi_{S \Delta T}$

$$\chi_S(x) \chi_T(x) = (-1)^{\sum_{i \in S} x_i + \sum_{i \in T} x_i}$$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x [f^2(x)]}$$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f - g\|_b^2 = \mathbb{P}_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$,
 $\frac{1}{4} \|f - g\|_b^2 = \mathbb{P}[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$,
 $\|f\|_b^2 = 1$

χ_S

for $S \subseteq [n]$, def. $\chi_S: \{0,1\}^n \rightarrow \{-1,1\}$

by
$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ -1 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

Fun facts:

- $\chi_S \chi_T = \chi_{S \Delta T}$

- $E[\chi_S(x)] = 0$ for $S \neq \emptyset$
 $x \in \{0,1\}^n$

$$= E[\chi_S(x) | x_n = 0] - E[\chi_S(x) | x_n = 1]$$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x[f^2(x)]}$$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f - g\|_b^2 = P_x[f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$,
 $\frac{1}{4} \|f - g\|_b^2 = P[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$,
 $\|f\|_b^2 = 1$

χ_S

for $S \subseteq [n]$, def $\chi_S: \{0,1\}^n \rightarrow \{-1,1\}$

by
$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ -1 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

Fun facts:

$$\chi_S \chi_T = \chi_{S \Delta T}$$

$$E[\chi_S(x)] = 0 \quad \text{for } S \neq \emptyset$$

$x \in \{0,1\}^n$

$$\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x[f^2(x)]}$$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f-g\|_b^2 = P_x[f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f-g\|_b^2 = P[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\langle \chi_S, \chi_T \rangle_b = E[\chi_S^{(x)} \chi_T^{(x)}]$$

$$= E[\chi_{S \Delta T}] = \begin{cases} 0 & S \neq T \\ 1 & S = T \end{cases}$$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

$$\bullet \langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$$

2^n diff. choices of S

$\{f: \{0,1\}^n \rightarrow \mathbb{R}\}$ is 2^n -dim. vector space

$\Rightarrow \{\chi_S: S \subseteq [n]\}$ is orthonormal basis

"Fourier basis"

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$
 $\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$
 $\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x [f^2(x)]}$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f-g\|_b^2 = P_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f-g\|_b^2 = P[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$

- $E_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$
- $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x [f^2(x)]}$$

✓ $\{\chi_S: S \subseteq [n]\}$ is orthonormal basis
"Fourier basis"

Fourier transform is the change of coordinates to the Fourier basis

$$f = \sum_{S \subseteq [n]} \langle f, \chi_S \rangle_b \chi_S = \sum_{S \subseteq [n]} E_x [f(x) \chi_S(x)] \chi_S$$

let $\hat{f} \in \mathbb{R}^{2^n}$ be vector w/ $\hat{f}_S = \langle f, \chi_S \rangle_b$

$f \rightarrow \hat{f}$ is the Fourier transform

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f - g\|_b^2 = P_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f - g\|_b^2 = P[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$
- $E_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$
- $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S = T \\ 0 & \text{otherwise} \end{cases}$

Fourier transform is change of coordinates to the Fourier basis

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x [f^2(x)]}$$

$$f = \sum_{S \subseteq [n]} \langle f, \chi_S \rangle_b \chi_S = \sum_{S \subseteq [n]} E_x [f(x) \chi_S(x)] \chi_S$$

let $\hat{f} \in \mathbb{R}^{2^n}$ be vector w/ $\hat{f}_S = \langle f, \chi_S \rangle_b$

$f \rightarrow \hat{f}$ is the Fourier transform

we have

$$\underline{\langle f, g \rangle_b} = \underline{\langle \hat{f}, \hat{g} \rangle} \quad \forall f, g$$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f - g\|_b^2 = P_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f - g\|_b^2 = P[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$

- $E_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$
- $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S = T \\ 0 & \text{otherwise} \end{cases}$

Linearity testing

$f: \{0,1\}^n \rightarrow \{0,1\}$ is linear if

$$f(x+y) = f(x) + f(y) \pmod{2} \quad \forall x, y$$

for technical reasons it is easier to say:

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if

$$f(x+y) = f(x)f(y) \quad \forall x, y$$

(easy to convert one to the other)

e.g. χ_S is linear

[BLR93]

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$
 $\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \in \{0,1\}^n} [f(x)g(x)]$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x [f^2(x)]}$$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f-g\|_b^2 = \mathbb{P}_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f-g\|_b^2 = \mathbb{P}[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$

- $\mathbb{E}_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$
- $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis

Fourier transform $f \rightarrow \hat{f}$

w/ $\hat{f} \in \mathbb{R}^{2^n}$, $\hat{f}_S = \langle f, \chi_S \rangle_b$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

Linearity testing

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if

$$f(x+y) = f(x)f(y) \quad \forall x,y$$

goal: test if $f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear

obvious algo:

1. sample $x, y \in \{0,1\}^n$
2. evaluate $f(x), f(y), f(x+y)$
3. Accept f if $f(x+y) = f(x)f(y)$

Theorem:

$$\min_S \mathbb{P}_x [f(x) \neq \chi_S(x)] \leq \mathbb{P}_{x,y} [f(x)f(y) \neq f(x+y)]$$

[BLR93]

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$
 $\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0,1\}^n} [f(x)g(x)]$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x [f^2(x)]}$$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f-g\|_b^2 = \mathbb{P}_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f-g\|_b^2 = \mathbb{P}[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$

- $\mathbb{E}_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$
- $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis
Fourier transform $f \rightarrow \hat{f}$

w/ $\hat{f} \in \mathbb{R}^{2^n}$, $\hat{f}_S = \langle f, \chi_S \rangle_b$
 $\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if
 $f(x+y) = f(x)f(y) \quad \forall x,y$

obvious algo:

1. sample $x, y \in \{0,1\}^n$
2. evaluate $f(x), f(y), f(x+y)$
3. Accept f if $f(x+y) = f(x)f(y)$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$
 $\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \in \{0,1\}^n} [f(x)g(x)]$
 $\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x [f^2(x)]}$

Theorem:

$$\min_S \mathbb{P}_x [f(x) \neq \chi_S(x)] \leq \mathbb{P}_{x,y} [f(x)f(y) \neq f(x+y)]$$

proof. let $P = \mathbb{P}_{x,y} [f(x)f(y) \neq f(x+y)]$

$$Z = \mathbb{1}_{[f(x)f(y) = f(x+y)]}$$

$$\begin{aligned} Z &= 1 - \frac{1}{4} (f(x)f(y) - f(x+y))^2 \\ &= 1 - \frac{1}{4} ((f^2(x)f^2(y) - 2f(x)f(y)f(x+y) + f^2(x+y))) \\ &= \frac{1}{2} (1 + f(x)f(y)f(x+y)) \end{aligned}$$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f-g\|_b^2 = \mathbb{P}_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f-g\|_b^2 = \mathbb{P}[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$

$$\mathbb{E}_{x \in \{0,1\}^n} [\chi_S(x)] = 0 \quad \text{for } S \neq \emptyset$$

$$\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis

Fourier transform $f \rightarrow \hat{f}$

$$\text{w/ } \hat{f} \in \mathbb{R}^{2^n}, \quad \hat{f}_S = \langle f, \chi_S \rangle_b$$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

$$E[Z] = \frac{1}{2} + \frac{1}{2} E_{x,y} [f(x)f(y)f(x+y)]$$

$$E_x [f(x) E_y [f(y)f(x+y)]]$$

$$\left(h(x) = E_y [f(y)f(x+y)] \right)$$

$$E_x [f(x)h(x)] = \langle f, h \rangle_b = \langle \hat{f}, \hat{h} \rangle$$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x [f^2(x)]}$$

- For $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f-g\|_b^2 = P_x [f(x) \neq g(x)]$
- For $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f-g\|_b^2 = P[f(x) \neq g(x)]$
- For $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$

- $E_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$
- $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis

Fourier transform $f \rightarrow \hat{f}$

w/ $\hat{f} \in \mathbb{R}^{2^n}$, $\hat{f}_S = \langle f, \chi_S \rangle_b$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if
 $f(x+y) = f(x)f(y) \quad \forall x, y$

$$\checkmark P = \frac{1}{2} + \frac{1}{2} \langle \hat{f}, \hat{h} \rangle$$

$$\text{claim: } \underline{\hat{h}_S} = \underline{\hat{f}_S^2} \quad \forall S$$

$$\langle h, \chi_S \rangle_b$$

$$E_x [h(x) \chi_S(x)] = E_{x,y} [f(y) f(x+y) \chi_S(x)]$$

$$\hat{h}_S = E_{x,y} [f(y) f(x+y) \chi_S(y) \chi_S(x+y)]$$

$$= E_y [f(y) \chi_S(y)] E_{x,y} [f(x+y) \chi_S(x+y)]$$

$$= \langle f, \chi_S \rangle_b^2 = \hat{f}_S^2$$

$$\text{let } P = P_{x,y} [f(x)f(y) = f(x+y)]$$

$$h(x) = E_y [f(y) f(x+y)]$$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x [f^2(x)]}$$

• for $f, g: \{0,1\}^n \rightarrow \{0,1\}$

$$\|f-g\|_b^2 = P[f(x) \neq g(x)]$$

• for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$

$$\frac{1}{4} \|f-g\|_b^2 = P[f(x) \neq g(x)]$$

• for $f: \{0,1\}^n \rightarrow \{-1,1\}$

$$\|f\|_b^2 = 1$$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

$$\chi_S \chi_T = \chi_{S \Delta T}$$

$$E_{x \in \{0,1\}^n} [\chi_S(x)] = 0 \quad \text{for } S \neq \emptyset$$

$$\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis

Fourier transform $f \rightarrow \hat{f}$

$$\text{w/ } \hat{f} \in \mathbb{R}^{2^n}, \quad \hat{f}_S = \langle f, \chi_S \rangle_b$$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if

$$f(x+y) = f(x)f(y) \quad \forall x, y$$

$$\checkmark P = \frac{1}{2} + \frac{1}{2} \langle \hat{f}, \hat{h} \rangle$$

$$\checkmark \hat{h}_S = \hat{f}_S^2 \quad \forall S$$

$$P = \frac{1}{2} + \frac{1}{2} \langle \hat{f}, \hat{f}^2 \rangle$$

$$= \frac{1}{2} + \frac{1}{2} \sum_S \hat{f}_S^3 = \frac{1}{2} + \frac{1}{2} \sum_S \hat{f}_S^2 \hat{f}_S \leq \frac{1}{2} + \frac{1}{2} \sum_S \hat{f}_S^2 \hat{f}_S^{\max}$$

$$\leq \frac{1}{2} + \frac{1}{2} \max_S \hat{f}_S$$

fix S to be max

$$\hat{f}_S \geq 2P - 1$$

$$4P[f(x) \neq \chi_S] = \|f - \chi_S\|_6^2 = \|f\|_6^2 + \|\chi_S\|_6^2 - 2\langle f, \chi_S \rangle = 2 - 2\hat{f}_S \leq 4(1 - P)$$

$$\text{let } P = \mathbb{P}_{x,y}[f(x)f(y) = f(x+y)]$$

$$h(x) = \mathbb{E}_y[f(y)f(x+y)]$$

$$\text{Vector space of } f: \{0,1\}^n \rightarrow \mathbb{R}$$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_6 \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x[f^2(x)]}$$

$$\bullet \text{ for } f, g: \{0,1\}^n \rightarrow \{0,1\}$$

$$\|f - g\|_6^2 = \mathbb{P}_x[f(x) \neq g(x)]$$

$$\bullet \text{ for } f, g: \{0,1\}^n \rightarrow \{-1,1\}$$

$$\frac{1}{4} \|f - g\|_6^2 = \mathbb{P}[f(x) \neq g(x)]$$

$$\bullet \text{ for } f: \{0,1\}^n \rightarrow \{-1,1\}$$

$$\|f\|_6^2 = 1 = \|\mathbf{1}\|_6^2$$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

$$\bullet \chi_S \chi_T = \chi_{S \Delta T}$$

$$\bullet \mathbb{E}_{x \sim \{0,1\}^n} [\chi_S(x)] = 0 \text{ for } S \neq \emptyset$$

$$\bullet \langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S = T \\ 0 & \text{otherwise} \end{cases}$$

$$\{\chi_S: S \subseteq [n]\} \text{ orthon. basis}$$

$$\text{Fourier transform } f \rightarrow \hat{f}$$

$$\text{w/ } \hat{f} \in \mathbb{R}^{2^n}, \hat{f}_S = \langle f, \chi_S \rangle_b$$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

$$f: \{0,1\}^n \rightarrow \{-1,1\} \text{ is linear if}$$

$$f(x+y) = f(x)f(y) \quad \forall x, y$$

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if
 $f(x+y) = f(x)f(y) \quad \forall x,y$

obvious algo:

1. sample $x, y \in \{0,1\}^n$
2. evaluate $f(x), f(y), f(x+y)$
3. Accept f if $f(x+y) = f(x)f(y)$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$
 $\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0,1\}^n} [f(x)g(x)]$
 $\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x [f^2(x)]}$

✓ Theorem:

$$\min_S \mathbb{P}_x [f(x) \neq \chi_S(x)] \leq \mathbb{P}_{x,y} [f(x)f(y) \neq f(x+y)]$$

Corollary All linear $f: \{0,1\}^n \rightarrow \{-1,1\}$
 are of the form χ_S for some S

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f-g\|_b^2 = \mathbb{P}_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f-g\|_b^2 = \mathbb{P}[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$
 - $\mathbb{E}_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$
 - $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$
- $\{\chi_S: S \subseteq [n]\}$ orthon. basis
 Fourier transform $f \rightarrow \hat{f}$
 w/ $\hat{f} \in \mathbb{R}^{2^n}$, $\hat{f}_S = \langle f, \chi_S \rangle_b$
 $\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$

✓ Theorem:

$$\min_S \min_x P[f(x) \neq \chi_S(x)] \leq P_{x,y}[f(x)f(y) \neq f(x+y)]$$

Corollary All linear $f: \{0,1\}^n \rightarrow \{-1,1\}$ are of the form χ_S for some S

Correcting close-to-linear functions

Theorem let $f: \{0,1\}^n \rightarrow \{-1,1\}$ be ϵ -close to χ_S . let $y \in \{0,1\}^n$ unif. random. then

$$P_y[f(y)f(x+y) = \chi_S(x)] \geq 1 - 2\epsilon$$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$
 $\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = E_{x \sim \{0,1\}^n} [f(x)g(x)]$
 $\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{E_x[f^2(x)]}$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f-g\|_b^2 = P_x[f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f-g\|_b^2 = P[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ -1 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$

- $E_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$
- $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis

Fourier transform $f \rightarrow \hat{f}$

w/ $\hat{f} \in \mathbb{R}^{2^n}$, $\hat{f}_S = \langle f, \chi_S \rangle_b$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if
 $f(x+y) = f(x)f(y) \quad \forall x, y$

Dictator functions

$\chi_{\{i\}}$ for $i \in [n]$ (abbr. χ_i)

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$
 $\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0,1\}^n} [f(x)g(x)]$
 $\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x [f^2(x)]}$

- for $f, g: \{0,1\}^n \rightarrow \{0,1\}$
 $\|f - g\|_b^2 = \mathbb{P}_x [f(x) \neq g(x)]$
- for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$
 $\frac{1}{4} \|f - g\|_b^2 = \mathbb{P}[f(x) \neq g(x)]$
- for $f: \{0,1\}^n \rightarrow \{-1,1\}$
 $\|f\|_b^2 = 1$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

- $\chi_S \chi_T = \chi_{S \Delta T}$

- $\mathbb{E}_{x \in \{0,1\}^n} [\chi_S(x)] = 0$ for $S \neq \emptyset$

- $\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S = T \\ 0 & \text{otherwise} \end{cases}$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis

Fourier transform $f \rightarrow \hat{f}$

w/ $\hat{f} \in \mathbb{R}^{2^n}$, $\hat{f}_S = \langle f, \chi_S \rangle_b$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if
 $f(x+y) = f(x)f(y) \quad \forall x, y$

Dictator functions χ_i

2-step test

① Linearity. reject if $f(x+y) \neq f(x)f(y)$

② Not-all-equal (NAE) test:

let $\Omega = \{0,1\}^3 \setminus \{(0,0,0), (1,1,1)\}$

$\Omega^n = \{(x,y,z): x,y,z \in \{0,1\}^n, (x_i,y_i,z_i) \in \Omega \forall i\}$

sample $x,y,z \in \Omega^n$.

reject f unless $(f(x), f(y), f(z)) \in \Omega$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x [f^2(x)]}$$

• for $f, g: \{0,1\}^n \rightarrow \{0,1\}$

$$\|f - g\|_b^2 = \mathbb{P}_x [f(x) \neq g(x)]$$

• for $f, g: \{0,1\}^n \rightarrow \{-1,1\}$

$$\frac{1}{4} \|f - g\|_b^2 = \mathbb{P}[f(x) \neq g(x)]$$

• for $f: \{0,1\}^n \rightarrow \{-1,1\}$

$$\|f\|_b^2 = 1$$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

$$\chi_S \chi_T = \chi_{S \Delta T}$$

$$\mathbb{E}_{x \in \{0,1\}^n} [\chi_S(x)] = 0 \text{ for } S \neq \emptyset$$

$$\langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis

Fourier transform $f \rightarrow \hat{f}$

$$\text{w/ } \hat{f} \in \mathbb{R}^{2^n}, \hat{f}_S = \langle f, \chi_S \rangle_b$$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if

$$f(x+y) = f(x)f(y) \quad \forall x, y$$

Theorem: suppose f passes both tests w/ probability $1-\epsilon$ for $\epsilon \leq .1$. Then $\exists i \in [n]$

$$\text{s.t. } \mathbb{P}_x [f(x) \neq \chi_i(x)] \leq \epsilon$$

Vector space of $f: \{0,1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle_b \stackrel{\text{def}}{=} \frac{\langle f, g \rangle}{2^n} = \mathbb{E}_{x \sim \{0,1\}^n} [f(x)g(x)]$$

$$\|f\|_b \stackrel{\text{def}}{=} \sqrt{\langle f, f \rangle_b} = \sqrt{\mathbb{E}_x [f^2(x)]}$$

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i} = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \text{ even} \\ 0 & \text{if } \sum_{i \in S} x_i \text{ odd} \end{cases}$$

$$\bullet \chi_S \chi_T = \chi_{S \Delta T}$$

$$\bullet \langle \chi_S, \chi_T \rangle_b = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{otherwise} \end{cases}$$

$\{\chi_S: S \subseteq [n]\}$ orthon. basis

Fourier transform $f \rightarrow \hat{f}$

$$\text{w/ } \hat{f} \in \mathbb{R}^{2^n}, \hat{f}_S = \langle f, \chi_S \rangle_b$$

$$\langle f, g \rangle_b = \langle \hat{f}, \hat{g} \rangle \quad \forall f, g$$

$f: \{0,1\}^n \rightarrow \{-1,1\}$ is linear if

$$f(x+y) = f(x)f(y) \quad \forall x, y$$

$$\min_x \mathbb{P}_x [f(x) \neq \chi_S(x)] \leq \mathbb{P}_{x,y} [f(x)f(y) \neq f(x+y)]$$

Dictator functions χ_i

2-step test

① Linearity. reject if $f(x+y) \neq f(x)f(y)$

② Not-all-equal (NAE) test:

let $\Omega = \{0,1\}^3 \setminus \{(0,0,0), (1,1,1)\}$

$\Omega^n = \{(x,y,z): x,y,z \in \{0,1\}, (x_i, y_i, z_i) \in \Omega \forall i\}$

sample $x,y,z \in \Omega^n$.

reject f unless $(f(x), f(y), f(z)) \in \Omega$

Theorem: suppose f passes both tests w/ probability $1-\epsilon$ for $\epsilon \leq .1$. Then $\exists i \in [n]$

$$\text{s.t. } \mathbb{P}_x [f(x) \neq \chi_i(x)] \leq \epsilon$$