Today

PCPs Course
- Introduction
- PCPs - Two Views
- FGLSS Reduction

CSS.330.1 : PCPs -
Limits of Approximation Algorithms

Lecture 01 (2023-01-27)

Instructor: Prahladh Harsha

Administrivia:

Fri - 9:30-13:00 (30-45 min break)

Grading: Problem Sets 3-4 - 60%
Class Participation 20%
Paper Presentation/Project - 20%
(No final exam).

-

PCPs: 3 decades.

1. Limits of approximation algorithms
2. Proof Checking.

Today's lecture: Statement of PCP Theorem (viewpoints).

Next 4-5 weeks: Proof of the PCP Theorem

Subsequently: Various extensions

Limits of Approximation Algorithms

Why Approximations?

NP complete problems : cope w/ hardness

- Heuristics :
- Approximation Algorithms.

Example:

① Vertex Cover:

Instance: Undirected graph $G=(V,E)$
$\omega: V \to \mathbb{R}_{\geq 0}$ (possibly)

Output: $W \subseteq V$ - cover
(i.e, $\forall (u,v) \in E$, $u \in W$ or $v \in W$)

Goal: Output a W w/ minimal cost
(i.e $\sum_{v \in W} \omega(v)$ is minimized)

Vertex Cover is NP-hard.

ω. - unweighted case.

Claim: If M is a maximal matching in G.
then $W = \{v \in V \mid v$ is an endpoint of a edge in $M\}$
satisfies W - vertex cover.

$|W| \leq 2 \cdot \text{opt-vertex cover.}$

MAX3SAT: Instance: n Boolean variables $x_1 \ldots x_n$

m- Clauses of width $= 3$.

$C_1, C_2 \ldots C_m$

$C_i = x_1 \vee x_2 \vee \overline{x_3}$

Output: Assignment $a: [n] \to \{0,1\}$

Goal: Maximize # clauses satisfied by a.

MAX3SAT is NP-hard.

Approximation Algorithm: Random assignment

C- clause w/ k variables (distinct)

$$\Pr[C \text{ is satisfied}] = 1 - \frac{1}{2^k}$$

$C_1 \ldots C_m$ - m clause

$$\mathbb{E}[\#\text{ clauses satisfied}] = \sum_{j \in [m]} \left(1 - \frac{1}{2^{k_j}}\right)$$

where $k_j = \#\text{vars}(C_j)$

$$= \frac{7}{8} m.$$

Easy exercise to derandomize the above algorithm.

Approximation Algorithm. $(\alpha \in (0,1))$

Φ - Optimization Problem (Maximization & Minimization)

$$\varphi \longrightarrow \boxed{A} \longrightarrow A(\varphi)$$

$$\alpha \cdot OPT(\varphi) < A(\varphi) \leq OPT(\varphi) \quad \text{(Maximization}$$

$$OPT(\varphi) \leq A(\varphi) < \frac{1}{\alpha} \cdot OPT(\varphi) \quad \text{(Minimization}$$

How good can we approximate a problem?

1. FPTAS: Fully-polynomial time approx scheme

   $\forall \varepsilon \in (0,1)$, there is a $(1-\varepsilon)$-approx alg runs in time $poly(n, 1/\varepsilon)$.

   eg: KNAPSACK

2. PTAS: $\forall \varepsilon \in (0,1)$, there is a $(1-\varepsilon)$-apprx alg runs in time $poly_{\varepsilon}(n)$.

   eg: MIN-MAKESPAN

(3) APX: Constant factor approximation
eg: Vertex Cover, MAX3SAT, MAXCUT

(4) log-APX: log-factor approximation
eg: SET COVER

(5) poly-APX: poly-factor apprx
eg: CLIQUE, Chromatic Number.

Question:
Given a problem, what is the best approximation one can achieve?

Work w/ a specific problem: MAX3SAT

Vertex Cover: Find VC of minimum size

VERTEX-COVER = $\{(G,k) \mid \exists$ a vertex cover $W \subseteq V(G)$ st $|W| \leq k\}$

(1) VERTEX-COVER - decision problem (0/1 answer)

(2) "Equivalent" to the original problem

(3) SAT $\leq_p$ VERTEX-COVER.

Similar to above.

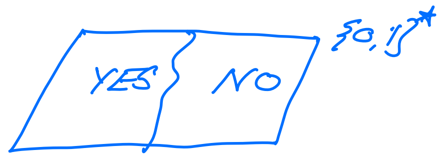
Decision-problem counterpart for "α-approximating MAX3SAT"

Gap Problems:

$(YES, NO) \subseteq \{0,1\}^*$

(i) $YES \cap NO = \emptyset$

Decision Problem (Language)



Gap problem corresponding to α-approximation MAX3SAT ($\alpha \in (0,1)$)

gap^*_α-MAX3SAT $= (YES, NO)$

$YES = \{(\varphi, k) \mid \varphi$ is a 3CNF formula & $\exists$ an assignment satisfying $\geq k$ clauses$\}$

$NO = \{(\varphi, k) \mid \varphi$ is a 3CNF formula & $\forall$ assignments $< \alpha k$ clauses are satisfied$\}$

Proposition: $\forall \alpha \in (0,1)$

an α-approximation alg for MAX3SAT exists

$\Updownarrow$

$\exists$ a ptime alg that solves gap^*_α-MAX3SAT.

Pf: ($\Downarrow$) Suppose A is α-approx alg for MAX3SAT.

B = "On input $\langle\varphi, k\rangle$

1. Run A on φ & let $k' = A(\varphi)$
2. Accept iff $k' \geq \alpha k$"

Claim: B solves gap^*_α-MAX3SAT.

Pf: $(\varphi, k) \in$ YES:

$\Rightarrow OPT(\varphi) \geq k$

$\Rightarrow k' = A(\varphi) \geq \alpha \cdot OPT(\varphi) \geq \alpha k$

$\Rightarrow B$ accepts ✓

$(\varphi, k) \in$ NO

$\Rightarrow OPT(\varphi) < \alpha k$

$\Rightarrow k' = A(\varphi) \leq OPT(\varphi) < \alpha k$

$\Rightarrow B$ rejects ✓

($\Uparrow$) Suppose B solves gap^*_α-MAX3SAT

A = "On input φ

1. Run B on $\langle\varphi, 1\rangle, \langle\varphi, 2\rangle, \ldots, \langle\varphi, m\rangle$
2. Let $k^* = \max\{k \mid B(\langle\varphi, k\rangle) = acc\}$
3. Output αk^* "

B rejects $\langle\varphi, k^*+1\rangle \Rightarrow \langle\varphi, k^*+1\rangle \notin YES \Rightarrow OPT(\varphi) \leq k^*$

accepts $\langle\varphi, k^*\rangle \Rightarrow \langle\varphi, k^*\rangle \notin NO \Rightarrow OPT(\varphi) \geq \alpha k^*$

ie, $\alpha \cdot OPT(\varphi) \leq \alpha k^* \leq OPT(\varphi)$

Hence, A is an α-approximation algorithm

<u>Qn</u>: What is the hardness of gap_α^*-MAX3SAT?

PCP Theorem: $\exists\ \alpha \in (0,1)$ and a poly time deterministic redn R from SAT to gap_α^*- MAX3SAT

ie, $\psi \in SAT \Rightarrow R(\psi) = \langle \varphi, k \rangle \in YES$

$\psi \notin SAT \Rightarrow R(\psi) = \langle \varphi, k \rangle \in NO.$

<u>Cor</u>: For the same α as in the above thm there is no α-approx for MAX3SAT unless $NP = P$

$\overline{gap}_\alpha$ - MAX3SAT

YES = $\{\langle \varphi \rangle \mid \varphi$ is a 3CNF formula & $\varphi \in SAT\}$

NO = $\{\langle \varphi \rangle \mid \varphi$ is a 3CNF formula & every assignment satisfies less than αm clauses$\}$

PCP Theorem I: $\exists\ \alpha \in (0,1)$ and a poly time deterministic redn R from SAT to gap_α-MAX3SAT

ie, $\psi \in SAT \Rightarrow R(\psi) = \varphi \in YES$

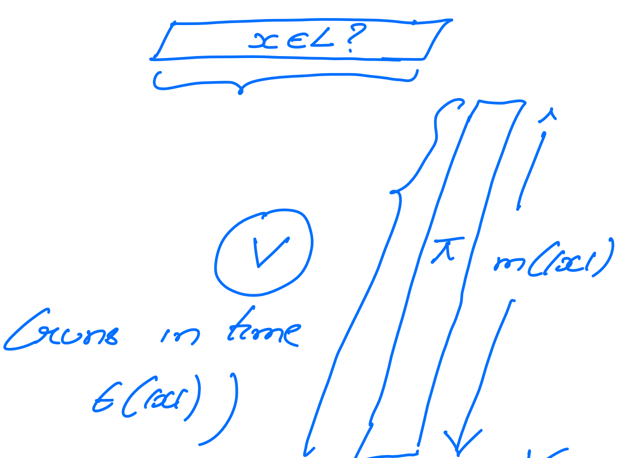
$\psi \notin SAT \Rightarrow R(\psi) = \varphi \in NO.$

Part 2:

Proof Checking.

NP: Proof-verification viewpoint of NP.

A language $L \in NP$, if there exists a deterministic verifier V and two polynomial t, m s.t.



Completeness:

$x \in L \Rightarrow \exists \pi, |\pi| = m(|x|)$

$V(x;\pi) = acc$

Soundness:

$x \notin L \Rightarrow \forall \pi, |\pi| = m(|x|)$

$V(x;\pi) = rej$

V runs in time $t(|x|)$.

Various variants of this proof verification viewpoint

- randomized
- interaction w/ prover instead of just a proof
- read only few locations of proof

— Led to notions

Interactive Proofs (IP = PSPACE)

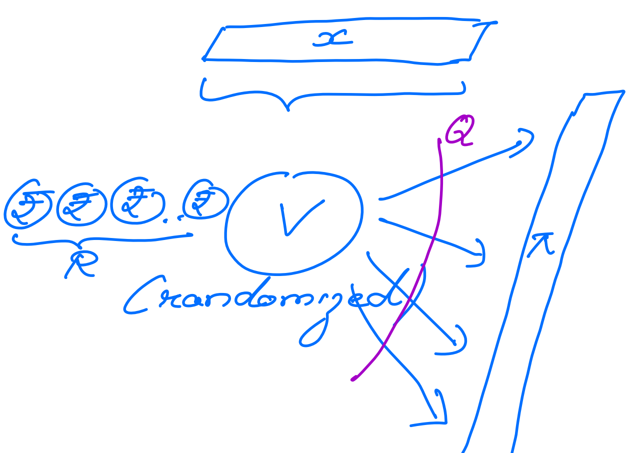
Zero knowledge

Multiprover Interactive Proofs (MIP = NEXP)

PCP Theorem.

— Restricted notion of verifier.

For $r, q, m, t : \mathbb{N} \to \mathbb{N}$, an (r, q, m, t)-restricted verifier V is a randomized algorithm that operates as follows



(0) V has explicit access to x

(a) Tosses $R \in \{0,1\}^{r(|x|)}$ coins

(b) Determines $Q = Q(x;R) \subseteq [m]$ s.t. $|Q| = q(|x|)$

(c) Predicate $C : \{0,1\}^q \to \{acc, rej\}$, $C = C(x;R)$

- V runs in time $t(|x|)$.

- V has implicit/oracle access to proof π of length m

Reads $\pi|_Q$ and acc/rej based on $C(\pi|_Q)$.

- Output is written as $V^{\pi}[x; R] \triangleq C(\pi|_Q)$

Complexity class

$PCP_{c,s}[r, q, m, t]$; $r, q, m, t : \mathbb{N} \to \mathbb{N}$

$c, s : \mathbb{N} \to [0,1]$

$c(i) \geq s(i)$, $\forall i \in \mathbb{N}$.

$L \in PCP_{c,s}[r, q, m, t]$

if $\exists$ (r, q, m, t)-restricted verifier such that

Completeness:

$x \in L \Rightarrow \exists \pi$, $|\pi| = m(|x|)$.

$$\Pr_{R \leftarrow \{0,1\}^{r(|x|)}} [V^{\pi}(x; R) = acc] \geq c(|x|).$$

Soundness

$x \notin L \Rightarrow \forall \pi$, $|\pi| = m(|x|)$

$$\Pr_{R} [V^{\pi}(x; R) = acc] < s(|x|).$$

Remarks: (1) If $t, m = poly(|x|)$, we drop these parameters.

(2) $P = PCP_{1,0}[0, 0]$

$NP = PCP_{1,0}[0, poly]$

$BPP = PCP_{3/4, 1/4}[poly, 0]$

(3) $c \in [0, 1]$

$c = 1$ — perfect completeness.

(4) Above defn is non-adaptive defn

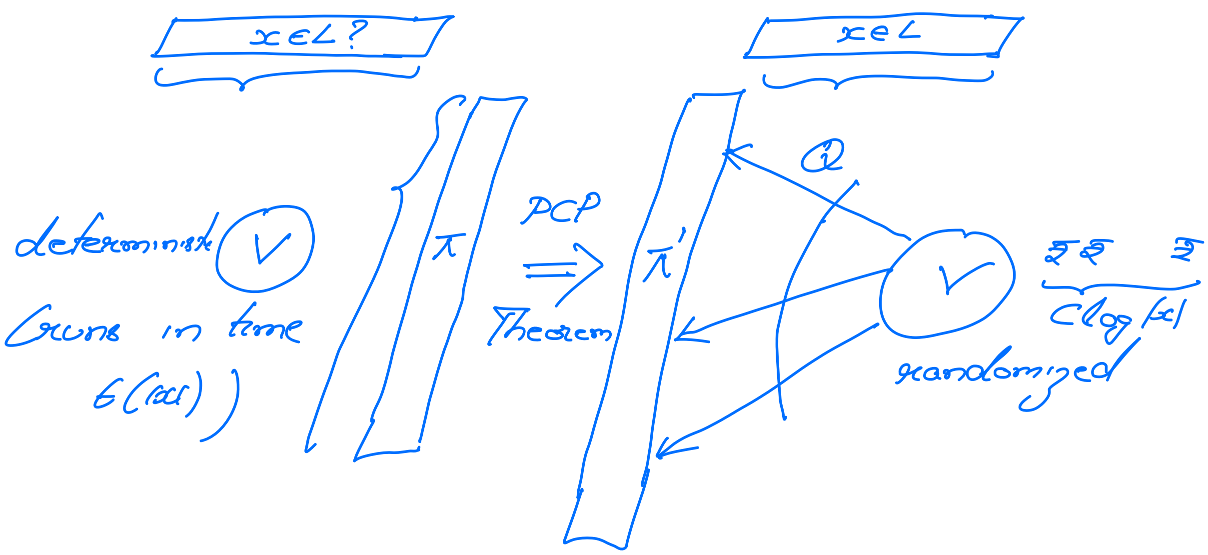
(can also define adaptive version).

PCP Theorem I:

$\exists\ Q \in \mathbb{Z}_{\geq 0}$ & $\alpha \in (0, 1)$

$\forall\ L \in NP$, $\exists\ c$

$L \in PCP_{1,\alpha}[C \log n, Q]$



Observation: PCP Theorem I $\equiv$ PCP Theorem II

(a). PCP Theorem I $\Rightarrow$ PCP Theorem II

Pf: Suppose $\exists$ a redn ρ_2 from SAT to gap$_\alpha$-MAX3SAT

Need to construct a restricted verifier for every language L in NP

$$L \xrightarrow{\rho_1} SAT \xrightarrow{\rho_2} \text{gap}_\alpha\text{-MAX3SAT}$$

$$\underbrace{\qquad\qquad\qquad}_{\rho}$$

$$x \longmapsto \psi \longmapsto \varphi$$

$V =$ "On input x & R

1. Run $\rho(x)$ to obtain 3CNF φ
2. Use R to pick a random clause C of φ
3. Set Q = vars of C
   C - predicate C."

"Expect as proof π the assignment for φ"

$x \in L \Rightarrow \varphi \in SAT \Rightarrow \exists \pi, \Pr_R[V^\pi(x;R) = acc] = 1$

$x \notin L \Rightarrow OPT(\varphi) < \alpha m \Rightarrow \forall \pi \ \Pr_R[V^\pi(x;R) = acc] < \alpha.$

(b). PCP Theorem II ⇒ PCP Theorem I.

For SAT, there is a $(C\log n, q, \text{poly}, \text{poly})$ -verifier

↓

Construct a reduction from SAT to gap_α-MAX3SAT.

: On input φ

1. For each $R \in \{0,1\}^{c\log|\varphi|}$ let h_R be the predicate of the verifier.

2. Construct

$$\Phi = \bigwedge_R h_R$$

$Var(\Phi)$ are proof bits + additional variables

(almost what we want except that Φ is not a 3CNF, but rather a q-CSP)

Observation: For every q, there exist $k(q)$, & $l(q)$ s.t. for every fn $h: \{0,1\}^q \to \{0,1\}$ there is a 3CNF formula φ_h w/

$k(q)$ clauses

$q + \ell(q)$ variables s.t

$h(x) = 1 \Rightarrow \exists z \in \{0,1\}^{\ell(q)}, \varphi_h(x, z) = 1$

$h(x) = 0 \Rightarrow \forall z \in \{0,1\}^{\ell(q)}, \varphi_h(x, z) = 0$

Modify 2 to the following

2. Construct

$$\Phi = \bigwedge_R \varphi_{h_R}$$

$\varphi \in SAT \Rightarrow \Phi \in SAT$

$\varphi \notin SAT \Rightarrow \forall \pi, \Pr_R\left[h_R\left(\pi|_{Q_R}\right) = 1\right] < \alpha$

Fix any π.

\# satisfied clauses in Φ

$$\leq \alpha \cdot 2^R \cdot k + (1-\alpha) \cdot 2^R (k-1)$$

$$= k \cdot 2^R \left(\alpha + (1-\alpha)\left(1 - \frac{1}{k}\right)\right)$$

$$= k \cdot 2^R \left(1 - (1-\alpha)\frac{1}{k}\right)$$

$$\triangleq k \cdot 2^R \, \tilde{\alpha}$$

$\boxtimes$

Inapproximability of Clique [Feige-Goldwasser-Lovasz-Safra-Szegedy]

$\alpha \in (0,1)$

gap_{α}-CLIQUE

$YES = \{\langle G,k\rangle \mid \exists$ a clique of size $\geq k$ in $G\}$

$NO = \{\langle G,k\rangle \mid$ Every clique in G is $< \alpha k\}$.

Lemma: $L \in PCP_{c,s}[r,q]$ then there is a $q \cdot 2^r$-time reduction from L to $gap_{s/c}$-CLIQUE.

Cor: $\exists \alpha \in (0,1)$, α-approximating CLIQUE is NP-hard.
(cf PCP Thm + Lemma)

Proof of Lemma:

$x \in L$ & L has a restricted verifier

$\downarrow$

(G_x, k_x)

G_x : Vertices $= \{(R, View) \mid R \in \{0,1\}^{r(|x|)}$, $View \in \{0,1\}^{q(|x|)}\}$

$\cong 2^r \times 2^q$

Edges: $(R_1, View_1) \sim (R_2, View_2)$

if (i) $(R_i, View_i)$ are accepting views for both $i = 1$ & 2 (ie $C_{R_i}(View_i) = 1$)

(ii) $View_1$ & $View_2$ must be consistent.

$x \in L \Rightarrow \Pr_R[V^\pi(x;R) = acc] \geq c.$

$W = \{(R, \pi|_{Q_R}) \mid R \in \{0,1\}^{r(|x|)}, \; C_R(\pi|_{Q_R}) = acc\}$

$|W| \geq c \cdot 2^r$

$x \notin L \Rightarrow$ If W is a clique of size $s \cdot 2^r$

$\exists \pi, \; \Pr_R[V^\pi(x;R) = acc] \geq s'$

(by constructing π by sewing together all the consistent views in W)

Hence, $s' < s$

ie any clique in G is of size $< s \cdot 2^r$.

Reduction

$$x \longmapsto \langle G_x, c2^r \rangle$$

SAT $\qquad$ $\text{gap}_{b/c}$-CLIQUE

Improving the Inapproximability:

① $PCP_{c,s}[r,q] \subseteq PCP_{c^k,s^k}[kr,kq]$
(sequential repetition)
$SAT \in PCP_{1,\alpha}[c\log n, Q] \subseteq PCP_{1,\alpha^k}[kc\log n, kQ]$

Cor: $\forall\, \alpha \in (0,1)$, gap_α-CLIQUE is NP-hard.

② By using randomness-efficient repetition
(walk on an expander graph)
$PCP_{1,1/2}[r,q] \subseteq PCP_{1,2^{-k}}[r+O(k), kq]$

Cor: $\exists\, \delta \in (0,1)$, gap_{1/n^δ}-CLIQUE is NP-hard.

③ Recycles queries.

Thm [Hås, NZ] $\forall\, \varepsilon \in (0,1)$, $\text{gap}_{1/n^{1-\varepsilon}}$-CLIQUE is NP-hard.