6.890	Lecture 12	Oct. 14, 2014

not necessarily constant — can be $c(n)$

<u>c-gap problem</u>: distinguish between
- min: $OPT \le k$ vs. $OPT > c \cdot k$ $(c > 1)$
- max: $OPT \ge k$ vs. $OPT < k/c$ $(c > 1)$
  OR: $OPT < c \cdot k$ $(c < 1)$
- promised that OPT is not in between
  i.e. don't care what algorithm does in that range
- if c-gap problem is NP-hard
  then so is $<c$-approximating optimization problem

⇒ stronger type of result

<u>(a,b)-gap SAT/CSP</u>:
- distinguish between $OPT < a \cdot \#$ clauses
  vs. $OPT \ge b \cdot \#$ clauses

⇒ gap $c = b/a$

<u>Gap-producing reductions</u>:
- output has $OPT = k$ or max: $OPT < k/c$ / min: $OPT > c \cdot k$
- simple examples:
  - Tetris: $n^{1-\varepsilon}$ gap [L3]
  - nonmetric TSP: weights $1, \infty$ or $0, 1$ (edge / nonedge)
    reduction from Ham. cycle
    ⇒ exponential or infinite gap

$PCP(O(\lg n), O(1))$ = Probabilistically Checkable Proof
- certificate of polynomial length } for YES instances
- $O(1)$-time verification algorithm } (like NP)
  given certificate & $O(\lg n)$ bits of randomness
- if YES instance: algorithm says YES
- if NO instance: $\Pr\{\text{algorithm says NO}\} \geq \Omega(1)$
- boosting: apply $\lg 1/\varepsilon$ times $\Rightarrow \Pr\{\text{incorrect}\} \leq \varepsilon$

constant < 1
- $(<1, 1)$-gap 3SAT $\in PCP(O(\lg n), O(1))$:
  - certificate = variable assignment
  - algorithm checks random clause is satisfied
  - $\Pr\{\text{wrong}\} \leq 1/\text{gap}$

$\Rightarrow$ if $(<1, 1)$-gap 3SAT is NP-hard then $NP = PCP(O(\lg n), O(1))$

(based on lecture notes by Dana Moshkovitz)

- if 3SAT $\in PCP(O(\lg n), O(1))$ then $(<1, 1)$-gap 3SAT is NP-hard ($\Rightarrow O(1)$-inapprox.)
  by gap-producing reduction:
  - PCP algorithm = $O(1)$-size formula $\rightarrow$ CNF
  - take conjunction over $n^{O(1)}$ random choices
  - if NO: $\Omega(1)$ fraction of terms false
    i.e. $\Omega(1)$ fraction of $\underbrace{\text{terms}}_{O(1) \text{ clauses}}$ have ≥ 1 false clause

    $\Rightarrow \Omega(1)$ fraction of clauses false

PCP theorem: $NP = PCP(O(\lg n), O(1))$ [Arora & Safra; Arora, Lund, Motwani, Sudan, Szegedi — FOCS 1992 / J.ACM 1998]

Gap-preserving reduction $A \to B$:

instance x of A $\xrightarrow{f}$ instance $x' = f(x)$ of B

$|x| = n$ $\qquad\qquad$ $|x'| = |n'|$

& functions $k(n)$, $k'(n')$, $c(n) \geq 1$, $c'(n') \geq 1$ satisfying:

- min: (1) $OPT_A(x) \leq k \Rightarrow OPT_B(x') \leq k'$
  (2) $OPT_A(x) \geq c \cdot k \Rightarrow OPT_B(x') \geq c' \cdot k'$
- max: (1) $OPT_A(x) \geq k \Rightarrow OPT_B(x') \geq k'$
  (2) $OPT_A(x) \leq k/c \Rightarrow OPT_B(x') \leq k'/c'$
- transitive
- gap amplifying if $c' > c$

Example: [Håstad – J.ACM 2001] & [Williamson & Shmoys book, 2010]

Max E3-X(N)OR-SAT: (linear equations, =3 terms)

- $(\frac{1}{2}+\varepsilon, 1-\varepsilon)$-gap is NP-hard $\forall \varepsilon > 0$ (PCP version)
  $\Rightarrow (\frac{1}{2}-\varepsilon)$-inapproximable
- $\frac{1}{2}$-approximation: uniform random assignment
  $\Rightarrow \Pr\{\text{right parity for equation}\} = 1/2$

Max E3SAT:

- L-reduction from Max E3-X(N)OR-SAT:
  - $x_i \oplus x_j \oplus x_k = 1 \rightarrow (x_i \vee x_j \vee x_k) \wedge (\bar{x}_i \vee \bar{x}_j \vee x_k)$
    $\wedge (x_i \vee \bar{x}_j \vee \bar{x}_k) \wedge (\bar{x}_i \vee x_j \vee \bar{x}_k)$
  - $x_i \oplus x_j \oplus x_k = 0 \rightarrow (\bar{x}_i \vee \bar{x}_j \vee \bar{x}_k) \wedge (x_i \vee x_j \vee \bar{x}_k)$
    $\wedge (\bar{x}_i \vee x_j \vee x_k) \wedge (x_i \vee \bar{x}_j \vee x_k)$
  - satisfied $\Rightarrow$ all 4 clauses satisfied
  - not satisfied $\Rightarrow$ exactly 3 clauses satisfied
  - $\Rightarrow$ additive error preserved $\Rightarrow \beta = 1$
  - $OPT_{E3SAT} = OPT_{E3X(N)OR} + 3 \cdot \#\text{equations}$
    $\leq 7 \cdot OPT_{E3X(N)OR}$ by 2-approx.
  - $\Rightarrow \alpha = 7$
- no $(1-\frac{1}{2})$-approx. for Max E3-X(N)OR-SAT
- $\Rightarrow$ no $\underbrace{(1-\tfrac{1}{2}/\alpha\beta)}_{= 1 - \frac{1}{14} = \frac{13}{14}}$-approx. for Max E3SAT
- gap argument:
  - yes instance $\Rightarrow \geq (1-\varepsilon)\cdot m$ equations satisfied
    $\rightarrow \geq (1-\varepsilon)m \cdot 4 + \varepsilon m \cdot 3$ clauses satisfied
    $= (4-\varepsilon)m$ out of $4m$
  - no instance $\Rightarrow < (\frac{1}{2}+\varepsilon)\cdot m$ equations satisfied
    $\rightarrow < (\frac{1}{2}+\varepsilon)m \cdot 4 + (\frac{1}{2}-\varepsilon)m \cdot 3$ clauses satisfied
    $= (\frac{7}{2}+\varepsilon)m$ out of $4m$
  - $\Rightarrow (\frac{7}{8}+\varepsilon, 1-\varepsilon)$-gap Max E3SAT is NP-hard
  - $\Rightarrow (\frac{7}{8}-\varepsilon)$-gap Max E3SAT is NP-hard
  - $\Rightarrow (\frac{7}{8}-\varepsilon)$-inapproximable
  - 7/8-approximation: random assignment tight!

Label Cover: Min-Rep & Max-Rep

[Arora, Babai, Stern, Sweedyk – JCSS 1997]

- given bipartite graph $G = (A \dot\cup B, E)$
  where $A = A_1 \dot\cup A_2 \dot\cup \cdots \dot\cup A_k$, $|A| = n$, $|A_i| = \frac{n}{k}$
  $B = B_1 \dot\cup B_2 \dot\cup \cdots \dot\cup B_k$, $|B| = n$, $|B_i| = \frac{n}{k}$ group
- choose $A' \subseteq A$ & $B' \subseteq B$
- __superedge__ (A_i, B_j) if ≥ 1 edge in $A_i \times B_j$
  - __covered__ if $A' \times B'$ intersects $A_i \times B_j$

Max-Rep:

- choose exactly 1 vertex from each group
  i.e. $|A' \cap A_i| = |B' \cap B_i| = 1 \quad \forall i$
- maximize # edges in $A' \times B'$ (i.e. induced by $A' \cup B'$)
  = # covered superedges

Min-Rep:

- allow >1 from each group
- cover every superedge (A_i, B_j)
- minimize $|A'| + |B'|$

Special cases:

- __regular__ = graph of superedges has uniform degree
- __star property__: each vertex $\in B_j$ adjacent to ≤ 1 vertex $\in A_i$ $\Rightarrow$ edges in $A_i \times B_j$ form disjoint stars
- __word puzzle__: A_i = set of words (e.g. "animal"), B_j = alphabet
  star = word letters in order [Moshkovitz]
- __unique game__: edges in $A_i \times B_j$ form a matching

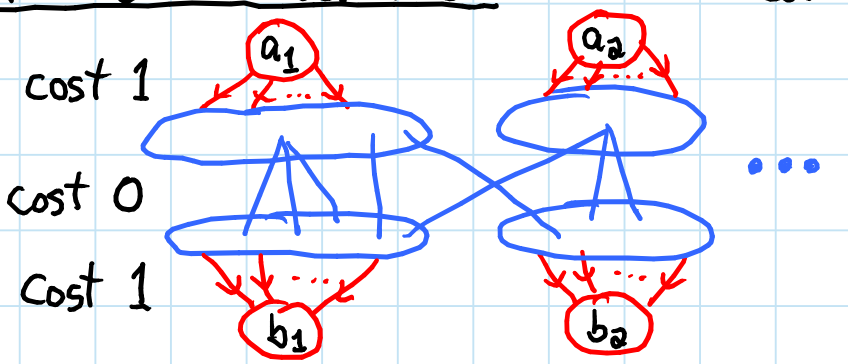
Hardness:

- (a,b)-gap Max-Rep: distinguish between a & b fraction of superedges covered
- (a,b)-gap Min-Rep: distinguish between $OPT < a(2k)$ & $OPT \geq b(2k)$ ($2k$ = super matching)
- $(\varepsilon, 1)$-gap Max-Rep is NP-hard [Raz - SICOMP 1998]
  $\forall \varepsilon > 0$ $\Rightarrow \notin APX$ (reducing to self)
  - even if $\varepsilon = 1/\log^{\varepsilon} n$ [Moshkovitz & Raz - FOCS 2008]
- $(1/p^k, 1)$-gap Max-Rep $\in P \Rightarrow NP \subseteq DTIME(n^{O(k)})$ $\forall k$
  constant - ditto $(1, p^k)$-gap Min-Rep

$\Rightarrow$ no $1/2^{\log^{1-\varepsilon} n}$-approximation algorithm unless $NP \subseteq DTIME(n^{\text{polylog } n})$ $\leftarrow$ "quasipolynomial"

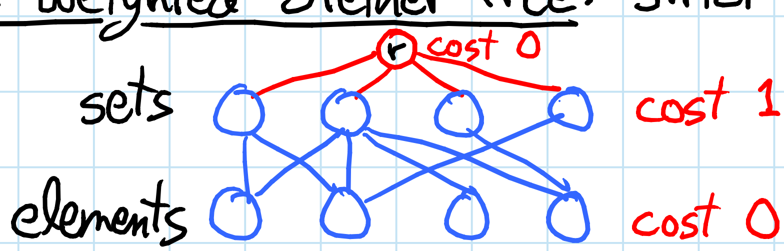
- best approx.: $\tilde{O}(n^{1/3})$ [Charikar, Hajiaghayi, Karloff - ESA 2009]
- $(O(n^{1/4}), 1)$-gap $\in P$ [Manurangsi & Moshkovitz - ESA 2013]

Directed Steiner forest: strict reduction from Min-Rep



require $a_i \to b_j$ path for each superedge (A_i, B_j)

Node-weighted Steiner tree: strict reduction from Set Cover



Unique Games Conjecture: [Khot - FOCS 2002]

$(\varepsilon, 1-\varepsilon)$-gap unique game is NP-hard $\forall \varepsilon > 0$ ($\& < \frac{1}{2}$)

$\Rightarrow$ Max 2SAT 0.940-approx., Max Cut 0.878-approx., Vertex Cover 2-approx. tight [Khot survey - CCC 2010]

- cf. 0.954, 0.941, 1.166-inapprox. via PCP [Håstad - J.ACM 2001]

- for every CSP, inapproximability factor ε less than integrality gap of natural SDP relaxation
  - $\Rightarrow$ SDP is ultimate approximation technique (here) [Raghavendra - STOC 2008]
  - e.g. for Vertex Cover, integrality gap $= 2$ for LP or SDP