Wednesday 8/27/14

Last class: For every graph $G=(V,E)$ (with $n=|V|$, $m=|E|$)

there exists a cut $S \Leftrightarrow \bar{S}$ of size $\geq \frac{m}{2}$.

How did we prove it?

Look at a random partition $S, \bar{S}$.

Let $X =$ # of edges crossing $S \Leftrightarrow \bar{S}$.

We showed that $E[X] = \frac{m}{2}$.

Hence there exists at least one cut of size $\geq \frac{m}{2}$.

Can we find such a cut?

Yes using method of conditional expectations.

X is a random variable taking values in the range $\{0, 1, 2, \ldots, m\}$

Let $V = \{v_1, v_2, \ldots, v_n\}$

In order $v_1, v_2, \ldots, v_n$, we'll assign each vertex v_i to S or $\overline{S}$.

When we assign v_i we'll use z_i to denote the assignment.

Let $z_i = \begin{cases} +1 & \text{if } v_i \in S \\ -1 & \text{if } v_i \in \overline{S} \end{cases}$

We'll assign $v_1, \ldots, v_i$ in such a way that given this assignment for $v_1, \ldots, v_i$ then for a random assignment for $v_{i+1}, \ldots, v_n$ we have that the expected cut size $\geq \frac{m}{2}$.

This means that at the end when $i=n$, there's no randomness left & we have constructed a cut of size $\geq \frac{m}{2}$.

More formally, we want to assign $z_1, \ldots, z_i$ so that:

$$E[X \mid z_1, \ldots, z_i] \geq \frac{m}{2} \qquad (*)$$

Note for $i=0$, (*) says: $E[X] \geq \frac{m}{2}$
which we showed last class.

And the case $i=1$ also holds:
v_i is assigned to S or $\overline{S}$, we might as well label the set containing v_i as S, then $z_i=+1$, and we still have:

$$E[X|z_1] \geq E[X] \geq \frac{m}{2}.$$

For $i>1$ we'll aim to assign v_i so that:

$$E[X|z_1,\ldots,z_i] \geq E[X|z_1,\ldots,z_{i-1}] \quad (**)$$

then by induction we'll have that:

$$E[X|z_1,\ldots,z_{i-1}] \geq \frac{m}{2}$$

& hence:

$$E[X|z_1,\ldots,z_i] \geq \frac{m}{2}$$

So we'll be done.

How do we assign v_i to maintain (**).

We need to factor $E[X]$ to condition on the possible values of z_i.

First, let's look at the definition of conditional expectation.

$$E[X] = \sum_{j=0}^{m} j \Pr(X=j).$$

for event ε,

$$E[X|\varepsilon] = \sum_{j=0}^{m} j \Pr(X=j|\varepsilon)$$

So for the event that $z_1=+1, z_2=-1$ then:

$$E[X|z_1,z_2] = \sum_{j=0}^{m} j \Pr(X=j|z_1,z_2)$$

Recall,

$$\Pr(X=j) = \Pr(z_1=+1)\Pr(X=j|z_1=+1) + \Pr(z_1=-1)\Pr(X=j|z_1=-1)$$
$$= \frac{1}{2}\Pr(X=j|z_1=+1) + \frac{1}{2}\Pr(X=j|z_1=-1)$$

and $$E[X] = \sum_{j=0}^{m} j\Pr(X=j) = \frac{1}{2}\sum_{j=0}^{m} j\Pr(X=j|z_1=+1) + \frac{1}{2}\sum_{j=0}^{m} j\Pr(X=j|z_1=-1)$$
$$= \frac{1}{2}E[X|z_1=+1] + \frac{1}{2}E[X|z_1=-1]$$

As we already discussed, assigning v_1 doesn't effect $E[X]$ thus:

$$E[X \mid z_1=+1] = E[X \mid z_1=-1] = E[X] \geq \frac{m}{2}.$$

But after we set z_1 what about assigning v_2?
We again have that:

$$E[X \mid z_1] = \frac{1}{2} E[X \mid z_2=+1, z_1] + \frac{1}{2} E[X \mid z_2=-1, z_1]$$

thus,

$$\max\{E[X \mid z_2=+1, z_1], E[X \mid z_2=-1, z_1]\}$$

$$\geq E[X \mid z_1]$$

& we know this $E[X \mid z_1] \geq \frac{m}{2}$.

Thus we want to assign v_2 to the best of these two

Chose z_2 to maximize $E[X \mid z_1, z_2]$

Suppose we set $z_1 = z_2 = +1$ so $v_1, v_2 \in S$.

Then the edge (v_1, v_2) if it exists doesn't count.

For all other edges, they have prob. $\frac{1}{2}$ of crossing $S \leftrightarrow \bar{S}$.

Thus, $E[X | z_1 = z_2 = +1] = \frac{m - \#\{\text{edges between } v_1, v_2\}}{2}$

Similarly for $z_1 = +1$, $z_2 = -1$,

$$E[X | z_1 = +1, z_2 = -1] = \frac{m + \#\{\text{edges between } v_1, v_2\}}{2}$$

So we can compute the assignment for v_2 to maximize $E[X | z_1, z_2]$.

More generally, if we fix the assignment for $v_1, \ldots, v_i$ (so we assign $z_1, \ldots, z_i$) then we can compute $E[X \mid z_1, \ldots, z_i]$ by:

- Counting the edges with both endpoints in $v_1, \ldots, v_i$ that are crossing the cut. ~~not~~
- plus we get every other edge with Prob. $\frac{1}{2}$

So given $z_1, \ldots, z_{i-1}$

we assign z_i by trying $z_i = +1$ & $z_i = -1$ & choosing the one maximizing the following:

Look at edges with both endpoints fixed (so both endpoints in $v_1, \ldots, v_i$)

- Keep those crossing $S \leftrightarrow \overline{S}$
- Discard those contained wholly in S or wholly in $\overline{S}$.

For unfixed edges:

add $\frac{1}{2}$ for each

Take larger of these 2 counts.

This algorithm has a simple greedy form:

Let d_i = degree of vertex v_i

Assign v_1 to S.

For $i=2 \rightarrow n$:

let j = # of neighbors of v_i assigned to S so far.

k = # of neighbors of v_i assigned to $\overline{S}$.

We know $j+k \leq d_i$

Assigning v_i to S:

we gain k edges & lose j

& get $\frac{1}{2}$ of $d_i - (k+j)$.

Assigning v_i to $\overline{S}$:

we gain j & lose k

& get $\frac{1}{2}$ of $d_i - (k+j)$.

Thus, if $j < k$ assign v_i to S

if $j \geq k$ assign v_i to $\overline{S}$.

(So we place v_i on side with fewer neighbors so far)

Greedy algorithm:

For $i = 1 \to n$:

Assign v_i to S or $\bar{S}$ to maximize its neighbors in the other set (only considering neighbors assigned earlier)

Why do we say this is a $\frac{1}{2}$-approximation algorithm?

For a graph G, let OPT denote the size of its max cut.

Let OUT denote the size of the cut output by our algorithm.

We say an algorithm is an α-approximation algorithm if

$$\min_G \frac{\text{OUT}}{\text{OPT}} \geq \alpha .$$

What about for MAX-SAT?

Formula f,

Variables $x_1, \dots, x_n$

Clauses $C_1, \dots, C_m$

For a random assignment,

let $Y = \#$ of satisfied clauses

We showed that $E[Y] \geq \frac{m}{2}$

~~For~~

Given an assignment for $x_1, \dots, x_{i-1}$

want to assign x_i to T or F so that:

$$E[Y \mid x_1, \dots, x_i] \geq E[Y \mid x_1, \dots, x_{i-1}] \geq \frac{m}{2}$$

As before:

$$\max\{E[Y \mid x_i = T, x_1, \dots, x_{i-1}], E[Y \mid x_i = F, x_1, \dots, x_{i-1}]\}$$
$$\geq E[Y \mid x_1, \dots, x_{i-1}]$$

So try both

Given an assignment to $x_1, \ldots, x_{i-1}$:

- try $x_i = T$ & $x_i = F$ and for each:
  - fix the assignments to $x_1, \ldots, x_i$ in f
  - Count the satisfied clauses
  - Drop those clauses with all literals fixed to unsatisfied
  - add $\frac{1}{2}$ for each remaining clause
- take the better of the two cases $x_i = T$ or $x_i = F$