2/2/21:

Metric TSP:

Input: Metric space (V, c)

Feasible: Hamiltonian cycle H cycle visiting all nodes once

Objective: $\min \; c(H) = \sum_{e \in H} c(e)$

Alg 1:

- Compute MST T
- Double T to get $2T$
- $2T$ Eulerian, so Eulerian tour C
- Shortcut C to get H

Thm: $2(1-\frac{1}{n})$-approx

Pf: Just like Steiner Tree!

Let H^* optimal solution,

F path from removing heaviest edge from H^*

$$\Rightarrow c(H) \underset{\text{shortcut}}{\leq} c(C) = c(2T) = 2c(T) \underset{F \text{ a spanning tree}}{\leq} 2c(F)$$

$$\leq 2(1-\tfrac{1}{n})\,c(H^*)$$

Want to do better: <u>Christofides' Algorithm</u>

why did we lose 2?

- Doubling MST

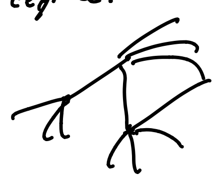
why did we do that?

- Make it Eulerian

cheaper way to make MST Eulerian?

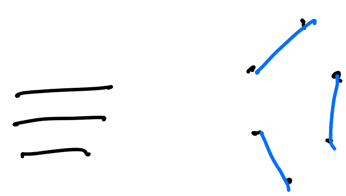
Problem: odd degree nodes

<u>Lemma</u>: Let $G=(V,E)$ be a graph. Then there are an even # nodes with odd degree.



<u>Pf</u>:

$$\sum_{v \in V} d(v) = 2|E| \quad \text{(even)}$$



<u>Def</u>: A perfect matching of $S \subseteq V$ is a matching on S of size $\frac{|S|}{2}$ (every node in S matched to other node in S)

<u>Fact</u>: Can find min-cost perfect matchings in polytime

Christofides':

- Compute MST T
- Let D be odd-degree nodes in T
- Compute min-cost perfect matching M of D
- Let C be Eulerian tour of $T+M$
- Return H = shortcutted C

<u>Claim</u>: Everything well-defined

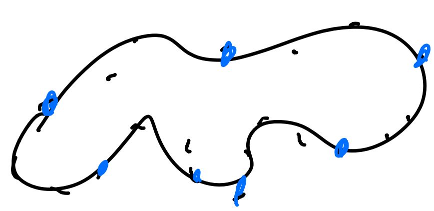
<u>Thm</u>: $\frac{3}{2}$-approximation

<u>Pf</u>: Let H^* optimal solution

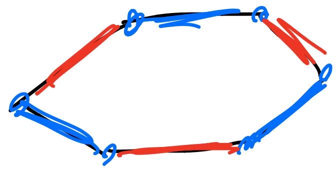
$c(T) \leq c(H^*)$ (H^* contains a spanning tree)

$c(H) \leq c(C) = c(T) + c(M) \leq c(H^*) + c(M)$

So WTS $c(M) \leq \frac{1}{2} c(H^*)$



Shortcut H^* to D, get H_D

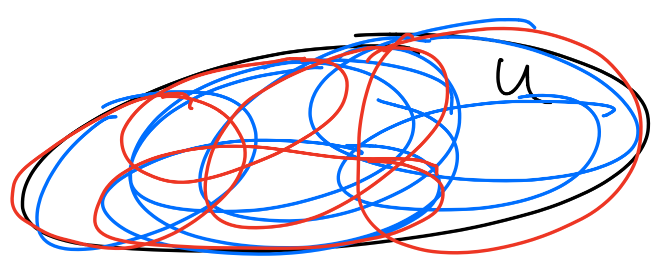


$|D|$ even, so partition into "evens" M_1 and "odds" M_2

- Each a perfect matching of D

$c(M_1) + c(M_2) = c(H_D)$

$c(M) \leq \min(c(M_1), c(M_2)) \leq \frac{1}{2} c(H_D) \leq \frac{1}{2} c(H^*)$



Set Cover:

Input: - Universe U with $|U| = n$

- Family of sets $S_1, S_2, \ldots, S_m$ with each $S_i \subseteq U$

Feasible solution: $I \subseteq [m]$ s.t. $\bigcup_{i \in I} S_i = U$

$[m] = \{1, 2, \ldots, m\}$

objective: min $|I|$

Generalizes Vertex Cover!

- $U = E$ - $S_v = N(v)$

neighborhood of v: $\{e \in E : v \in e\}$

Greedy Algorithm:

```
Init I = ∅, X = U
while (X ≠ ∅) {
    Let i ∈ [m] maximizing |X ∩ S_i|
    Add i to I, remove S_i from X
}
return I
```

Thm: If OPT uses k sets, greedy uses $\leq k(\ln \frac{n}{k} + 1)$

Cor: Greedy is $O(\log n)$-approximation

Cor: If $|S_i| \leq B$ $\forall i \in [m]$, then $O(\log B)$-approx

Pf: $k \geq \frac{n}{B} \Rightarrow$ greedy $\leq k(\ln B + 1)$

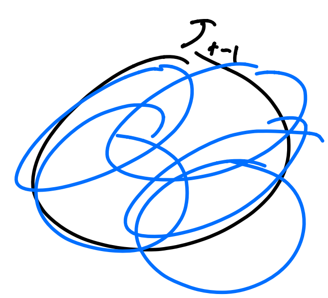
Pf of thm:

- Let I_t be sets selected in first t iterations

$J_t = U \setminus \left(\bigcup_{i \in I_t} S_i\right) =$ el'ts uncovered after t iterations

$n_t = |J_t| = n - \left|\bigcup_{i \in I_t} S_i\right|$

$\Rightarrow I_0 = \emptyset,\ J_0 = U,\ n_0 = n$



Claim: $n_t \leq (1 - \frac{1}{k}) n_{t-1}$ $\forall t$

Pf: Consider iteration t

OPT covers U with k sets

$\Rightarrow$ OPT covers J_{t-1} with $\leq k$ sets

$\Rightarrow \exists$ set covering $\geq \frac{n_{t-1}}{k}$ el'ts of J_{t-1}

$\Rightarrow$ greedy in iteration t covers $\geq \frac{n_{t-1}}{k}$ elements of J_{t-1}

$$\Rightarrow n_t \le n_{t-1} - \frac{n_{t-1}}{k} = \left(1-\frac{1}{k}\right) n_{t-1} \quad \checkmark$$

Claim $\Rightarrow n_t \le \left(1-\frac{1}{k}\right)^t n_0 = \left(1-\frac{1}{k}\right)^t n$

$\Rightarrow$ when $t = k \ln \frac{n}{k}$:

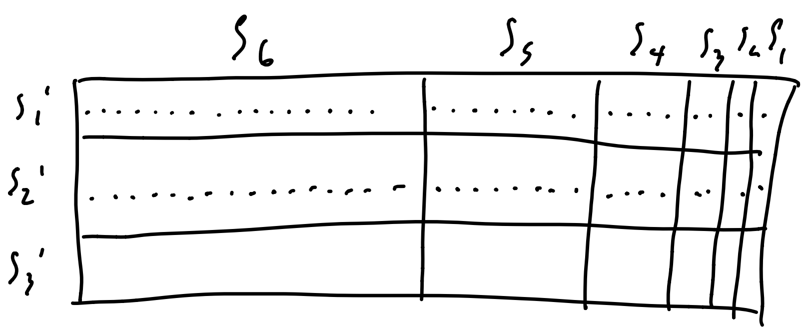
$$n_t \le \left(1-\frac{1}{k}\right)^{k \ln \frac{n}{k}} n \le \left(e^{-\frac{1}{k}}\right)^{k \ln \frac{n}{k}} n$$

$(1+x) \le e^x \quad \forall x \in \mathbb{R}$

$$= e^{-\ln \frac{n}{k}} n = \frac{k}{n} \cdot n = k$$

$\Rightarrow$ greedy takes $\le k \ln \frac{n}{k} + k = k\left(\ln \frac{n}{k} + 1\right)$ iterations

Is analysis tight, or is greedy even better?



Thm [Dinur-Steurer '14]: Unless $P = NP$, there is no $c \ln n$-approx for Set Cover for any constant $c < 1$

Weighted Set Cover:

Each set S_i also has cost c_i

$\min \sum_{i \in I} c_i$ (instead of $|I|$)

Greedy:

maximize "bang-for-the-buck": $\frac{|X \cap S_i|}{c_i}$ (density of S_i)

Thm: Greedy is an $H_n = \Theta(\log n)$-approximation

$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n}$

Pf: I_t, J_t, n_t as before.

OPT optimal solution, $C^* = \sum_{i \in OPT} c_i$

g_t index of set picked by greedy in iteration t

Claim: $\dfrac{|J_{t-1} \cap S_{g_t}|}{c_{g_t}} \geq \dfrac{n_{t-1}}{c^*}$

Pf: WTS $\exists\ i \in OPT$ s.t. $\dfrac{|J_{t-1} \cap S_i|}{c_i} \geq \dfrac{n_{t-1}}{c^*}$

(greedy choice has at least as good density)

Sps false $\Rightarrow \dfrac{c_i}{|J_{t-1} \cap S_i|} > \dfrac{c^*}{n_{t-1}} \quad \forall i \in OPT$

$$\Rightarrow c^* = \sum_{i \in OPT} c_i = \sum_{i \in OPT} \frac{c_i}{|J_{t-1} \cap S_i|} |J_{t-1} \cap S_i|$$

$$> \sum_{i \in OPT} \frac{c^*}{n_{t-1}} |J_{t-1} \cap S_i| = \frac{c^*}{n_{t-1}} \sum_{i \in OPT} |J_{t-1} \cap S_i|$$

$$\geq \frac{c^*}{n_{t-1}} |J_{t-1}| \quad \text{(OPT covers } J_{t-1}\text{)}$$

$$= c^*$$

$\Rightarrow\Leftarrow$

Let $x_t = |S_{g_t} \cap J_{t-1}|$ be # el't covered by greedy in iteration t

$\Rightarrow$ claim is $\dfrac{x_t}{c_{g_t}} \geq \dfrac{n_{t-1}}{c^*}$

$$\Rightarrow c_{g_t} \leq x_t \frac{c^*}{n_{t-1}}$$

$$c(\text{greedy}) = \sum_{t=1}^{k} c_{g_t} \leq \sum_{t=1}^{k} x_t \frac{c^*}{n_{t-1}}$$

$$= c^* \frac{x_1}{n} + c^* \frac{x_2}{n-x_1} + c^* \frac{x_3}{n-x_1-x_2} + \dots + c^* \frac{x_k}{n-x_1-\dots-x_{k-1}}$$

$$= c^* \Bigg(\underbrace{\frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}}_{x_1} + \underbrace{\frac{1}{n-x_1} + \frac{1}{n-x_1} + \dots + \frac{1}{n-x_1}}_{x_2} +$$

$$\dots + \underbrace{\frac{1}{n-x_1-\dots-x_{k-1}} + \dots + \frac{1}{n-x_1-\dots-x_{k-1}}}_{x_k} \Bigg)$$

$$\leq c^* \Big(\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \dots \frac{1}{n-x_1+1} + \frac{1}{n-x_1} + \frac{1}{n-x_1-1} + \dots + \frac{1}{n-x_1-x_2+1}$$

$$+ \frac{1}{n-x_1-x_2} + \dots + 1 \Big)$$

$$= c^* H_n$$