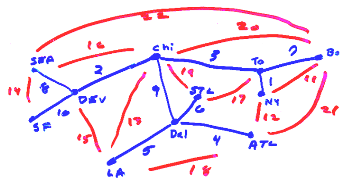
Minimum Cost Spanning Trees

Kruskal's Algorithm:

Another Example of the Greedy Method

Min Cost Spanning Tree

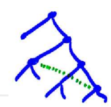
- Given: Undirected graph G = (V, E) & positive edge cost/weight $c(e) \in R$ for each $e \in E$.
- Find: connected $T \subseteq E$ minimizing $\sum_{e \in T} c(e)$



Applications

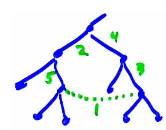
- Broadcast tree in a network
- Building roads or power lines
- Routing power & ground on a PC board
- Clustering
- …

Lemma 1: Trees and Cycles



Adding an edge to a tree creates a cycle; deleting any cycle edge gives a tree

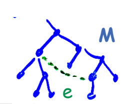
- Corollary 1: Solution to MST is a tree
- Corollary 2: Cheapest edge in E is in some MST T
- Excercises:
  - 2nd-cheapest edge also in T
  - 3rd-cheapest?



Kruskal's Algorithm

sort edges
K := ∅
while |K| < n-1 do
  e := next cheapest edge
  if $K \cup \{e\}$ is acyclic
    then $K := K \cup \{e\}$
    else discard e

Correctness

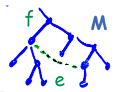


Theorem: Kruskal's algorithm builds an MST

Proof:

- Suppose Kruskal picks the tree K
- Suppose MST M maximizes $|K \cap M|$ among all MSTs
- For sake of contradiction, suppose $K \neq M$
- Let e be the cheapest edge in K – M
- Then…

Claim



$M \cup \{e\}$ has a cycle containing an edge f s.t.

(1) $f \notin K$, and

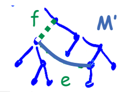
(2) $c(f) \geq c(e)$

Proof:

(1) If all the cycle edges were in K, then K wouldn't be a tree.

(2) If $c(f) < c(e)$, greedy looked at f before e. But $\{e' \in K \mid c(e') < c(e)\} \cup \{f\} \subseteq M$, hence acyclic, so f would have been picked, but it wasn't.

Correctness (cont.)



Theorem: Kruskal's algorithm builds an MST

Proof:

- Suppose Kruskal picks the tree K
- Suppose MST M maximizes $|K \cap M|$ among all MSTs
- For sake of contradiction, suppose $K \neq M$
- Let e be the cheapest edge in $K - M$
- From claim, $\exists$ f in cycle s.t. $f \notin K$, $c(f) \geq c(e)$
- Let $M' = (M \cup \{e\}) - \{f\}$
- Then M' is an MST with $|K \cap M'| > |K \cap M|$. Contradiction. QED

Implementation

Testing "if $K \cup \{e\}$ is acyclic":

Union/Find problem

Linear space

Time $n\ \alpha(n)$

$\alpha(n) < 5$ for all n < age of universe