Minimum Spanning Trees

Data Structures and Algorithms (094224)

Tutorial 12

Winter 2022/23

MST — Definitions

- Connected undirected graph $G = (V, E)$ with edge weight function $w : E \to \mathbb{R}$
- $H = (V, T)$ is a *spanning tree* ⟨עץ פורש⟩ of G if
  - H is a (spanning) subgraph of G, i.e., $T \subseteq E$; and
  - H is a tree
- Often address the spanning tree by its edge set T
  - Formally wrong!
  - Little risk for ambiguity because the vertex set is known
- $T \subseteq E$ is a *minimum spanning tree (MST)* ⟨עץ פורש מינימום⟩ if
  - T is a spanning tree of G; and
  - $w(T) \leq w(T')$ for all spanning trees T' of G
    - Recall that $w(F) = \sum_{e \in F} w(e)$
- The minimum spanning tree problem: construct an MST for a given weighted undirected graph
- Motivation: the smallest and "cheapest" subset of edges that ensures connectivity

MST — Definitions — cont.

- Edge subset $F \subseteq E$ is *good* if there exists an MST $T \supseteq F$
- Edge $e \in E - F$ is *safe* for (good) F if $F \cup \{e\}$ is still good
- Consider some vertex subset $\emptyset \subset S \subset V$
- The (bi-)partition of V into $\{S, V - S\}$ is referred to as a *cut* ⟨חתך⟩
- Let $E(S, V - S)$ be the set of edges $(u, v) \in E$ such that

$$(u \in S \wedge v \in V - S) \quad \vee \quad (v \in S \wedge u \in V - S)$$

  - I.e., $|\{u, v\} \cap S| = 1$
  - The edges in $E(S, V - S)$ are said to *cross* ⟨חוצות⟩ cut $\{S, V - S\}$
- Edge $e \in E$ is *light* ⟨קלה⟩ for cut $\{S, V - S\}$ if
  - $e \in E(S, V - S)$; and
  - $w(e) \leq w(e')$ for every edge $e' \in E(S, V - S)$
- Cut $\{S, V - S\}$ *respects* edge subset $F \subseteq E$ if $F \cap E(S, V - S) = \emptyset$
  - I.e., no edge in F crosses cut $\{S, V - S\}$

Prim's algorithm

`Prim`(G, w)

```
for all u ∈ G.V do
    u.π = NIL
    u.key = ∞
pick an arbitrary vertex r ∈ G.V
r.key = 0
Q = G.V
while Q ≠ ∅ do
    u = Extract_Min(Q)                ▷ minimum w.r.t. key
    for all v ∈ G.Adj[u] do
        if v ∈ Q and w(u, v) < v.key then
            v.π = u
            v.key = w(u, v)
```

- Run time (implementing Q using a Fibonacci heap):
  - $O(n \log n + m)$

Kruskal's algorithm

Kruskal(G, w)

1: copy the edges in $G.E$ into array $A[1 \dots m]$
2: sort A w.r.t. the edge weights
3: $T = \emptyset$
4: **for** $i = 1, \dots m$ **do**
5: $\quad e = A[i]$
6: $\quad$ **if** $(G.V, T \cup \{e\})$ is cycle free **then**
7: $\quad\quad T = T \cup \{e\}$
8: return T

- Missing details: implementation of the if condition in line 6
  - Less important since we don't focus on the algorithm's run-time

Question 1

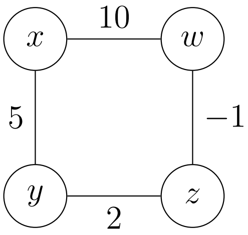
In lecture we proved the following theorem:

Let $F \subseteq E$ be a good edge subset and let $\{S, V - S\}$ be a cut that respects F. If $e \in E$ is light for $\{S, V - S\}$, then e is safe for F.

Prove/Disprove: The converse of the theorem is true, i.e.,
let $F \subseteq E$ be a good edge subset and let $\{S, V - S\}$ be a cut that respects F. If $e \in E$ is safe for F then e is light for $\{S, V - S\}$.

Solution

The claim is false.

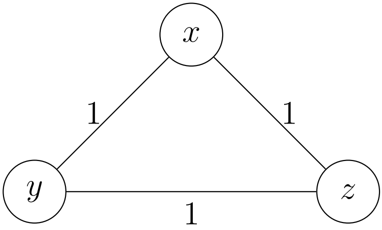


- MST of G: $T = \{(y,z),(y,x),(z,w)\}$
- $F = \{(y,z)\}$
  - F is good
- $S = \{y,z\}$
  - Cut $\{S, V-S\}$ respects F
- Edge (x,y) is safe for F but not light for cut $\{S, V-S\}$
  - Safe since $F \cup \{(x,y)\} \subseteq T$

Question 2

Give a simple example of a connected graph such that the set of edges $A = \{(u, v) \mid \text{there exists a cut } \{S, V - S\} \text{ such that } (u, v) \text{ is a light edge crossing } \{S, V - S\}\}$ does not form a minimum spanning tree.

Solution

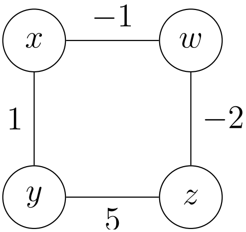


- Let $S_x = \{x\}$, $S_y = \{y\}$
- Edge (x, y) is light for cut $\{S_x, V - S_x\}$
- Edge (x, z) is light for cut $\{S_x, V - S_x\}$
- Edge (y, z) is light for cut $\{S_y, V - S_y\}$
- $A = \{(y, z), (x, z), (x, y)\}$ is not a MST since it contains a cycle

Question 3

Let $G = (V, E)$ be a connected undirected graph with weight function $w : E \to \mathbb{R}$ and $\emptyset \subset A \subset V$. Let G_1 the subgraph induced by A and G_2 the subgraph induced by $V - A$. Assume G_1 and G_2 are connected. Let T_1 be some MST of G_1 and T_2 some MST of G_2. Let e be a light edge crossing the cut $\{A, V - A\}$.
Prove/Disprove: $T = T_1 \cup T_2 \cup \{e\}$ is a MST of G.

Solution



- $T' = \{(x,y),(x,w),(w,z)\}$ is the only MST of G
- Let $A = \{x,w\}$ and $V - A = \{y,z\}$
- $G_1 = (A, E_1)$, $G_2 = (V - A, E_2)$
  - $E_1 = \{(x,w)\}$, $E_2 = \{(y,z)\}$
- $T_1 = \{(x,w)\}$, $T_2 = \{(y,z)\}$
- Edge $e = (w,z)$ is light for cut $\{A, V - A\}$
- $T = T_1 \cup T_2 \cup \{e\} \neq T'$ is not a MST of G
  - T' is the only MST of G

Question 4

Let T be a minimum spanning tree of an undirected and connected graph $G = (V, E)$ with weight function $w : E \to \mathbb{R}$ and let $e = (u, v) \in T$.
Define:

- $C_u = \{w \in V \mid \text{exists a path in } (V, T - \{e\}) \text{ from } u \text{ to } w\}$
- $C_v = \{w \in V \mid \text{exists a path in } (V, T - \{e\}) \text{ from } v \text{ to } w\}$

Prove that (u, v) is light for cut $\{C_u, C_v\}$ of G

Notice:

- $C_u \cup C_v = V,\ C_u \cap C_v = \emptyset$ and $|C_u| \geq 1,\ |C_v| \geq 1$
- $\{C_u, C_v\}$ is a cut in G
- Edge e crosses the cut $\{C_u, C_v\}$ of G
  - $e \in E(C_v, C_u)$ since $|\{u, v\} \cap C_u| = 1$

Solution

- Assume by contradiction that there exists an edge $e' \in E$ that crosses the cut $\{C_u, C_v\}$ of G and $w(e') < w(e)$
- $e' = (x, y) \notin T$ since otherwise T would contain a cycle
  - Assume w.l.o.g. that $x \in C_u$ and $y \in C_v$
    - e' crosses the cut $\{C_u, C_v\}$
  - Then there exists a cycle $C = \langle x, \cdots, u, v, \cdots, y, x\rangle$ in T
- Let $T' = T \cup \{e'\} - \{e\}$
- (V, T') is a spanning tree
  - $|T'| = |T| = |V| - 1$
  - (V, T') is connected
    - $T - \{e\}$ induces two trees
    - e' has an endpoint in each tree of $T - \{e\}$
- $w(T') = w(T) + w(e') - w(e) < w(T)$ $_{(\rightarrow\leftarrow)}$

Question 5

1. Prove: Given a connected and undirected graph $G = (V, E)$ with weight function $w : E \to \mathbb{R}$. If for every cut of G there is a unique light edge crossing the cut, then G has a unique minimum spanning tree.
2. Disprove: Given a connected and undirected graph $G = (V, E)$ with weight function $w : E \to \mathbb{R}$. If G has a unique minimum spanning tree, then for every cut of G there is a unique light edge crossing the cut.

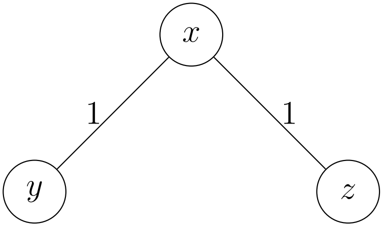
Solution — 5.1

- Let T be a minimum spanning tree of G
  - Since G is connected at least one spanning tree exists
- Assume by contradiction there exists another MST T' of G
- Since T and T' are two different spanning trees of G, there exists an edge $e = (u, v)$ such that $e \in T$ and $e \notin T'$
- Let $C_u = \{w \in V \mid \text{exists a path in } (V, T - \{e\}) \text{ from } u \text{ to } w\}$
- Let $C_v = \{w \in V \mid \text{exists a path in } (V, T - \{e\}) \text{ from } v \text{ to } w\}$
- $\{C_u, C_v\}$ is a cut in G
- Edge e is light for cut $\{C_u, C_v\}$
  - From question 4
- Let P be the unique (u, v)-path in T'
- Since P leads from $u \in C_u$ to $v \in C_v$, it must contain some edge $e' = (u', v') \neq e$ that crosses cut $\{C_u, C_v\}$
  - Otherwise T' is not a spanning tree

Solution — 5.1 — cont.

- Since every cut has a unique light edge, $w(e) < w(e')$
- Let $\hat{T} = T' \cup \{e\} - \{e'\}$
- $(V, \hat{T})$ is a spanning tree
  - $|\hat{T}| = |T'| = |V| - 1$
  - $(P - \{e'\}) \cup e$ is a (u', v')-path in $\hat{T}$ since (V, T') is connected
- $w(\hat{T}) = w(T') + w(e) - w(e') < w(T')$ $_{(\rightarrow\leftarrow)}$

Solution — 5.2



- Unique MST, $T = \{(x, y), (x, z)\}$
- The cut {{x},{y,z}} does not have a unique light edge crossing it

Question 6

Let T be a minimum spanning tree of an undirected graph $G = (V, E)$ with weight function $w : E \to \mathbb{R}$ and let L_T be the sorted list of the edge weights of T. Prove that for any other minimum spanning tree T' of G, $L_{T'} = L_T$, i.e., $L_T[i] = L_{T'}[i]$ for all $1 \leq i \leq n - 1$

Solution

Lemma

Let $G = (V, E)$ be a connected undirected graph. Let $F, F' \subseteq E$ be two edge subset such that the graphs (V, F) and (V, F') are forests and $0 \leq |F'| < |F|$. There exists an edge $e \in F$, $e \notin F'$ such that $(V, F' \cup \{e\})$ does not contain a cycle.

Solution

Proof.

- Let $k = |F|$, $k' = |F'|$ and $|V| = n$
- The graphs (V, F), (V, F') has exactly $n - k$ and $n - k'$ connected components respectively
  - A vertex with no incident edges is a connected component
- Let C_F, $C_{F'}$ be sets of vertices sets in each connected component of (V, F), (V, F') respectively
  - C_F and $C_{F'}$ are two partitions of V
- $|C_F| = n - k < |C_{F'}| = n - k'$
  - $k > k'$ thus $n - k' > n - k$
- There exists $C \in C_F$ such that for all $C' \in C_{F'}$, $C \not\subseteq C'$
  - Otherwise, $|C_F| \geq |C_{F'}|$
    - Since C_F is a partition of V, the partition $C_{F'}$ has size of at most $|C_F|$
- There exists $(u, v) \in F$ and $C' \in C_{F'}$ such that $u \in C'$ and $v \notin C'$
- Adding (u, v) to F' cannot form a cycle in $(V, F' \cup \{(u, v)\})$

Solution — cont.

- Assume by contradiction that there exist two spanning trees T and T' of G such that $L_T \neq L_{T'}$
- Let k be the first index such that $L_T[k] \neq L_{T'}[k]$
- Assume w.l.o.g. that $L_T[k] < L_{T'}[k]$
- Let F be the edges of T such that their weights correspond to $L_T[1]$ up to $L_T[k]$
- Let F' be the edges of T' such that their weights correspond to $L_{T'}[1]$ up to $L_{T'}[k-1]$
- (V, F) and (V, F') are forests
  - F, F' are subsets of tree edges
- $|F| = |F'| + 1$
- By the lemma, there exists an edge $e \in F$ such that $(V, F' \cup \{e\})$ does not contain a cycle
- Add e to T'
- $T' \cup \{e\}$ contains a cycle C
  - T' is a spanning tree

Solution — cont.

- There exists an edge $e' \in C$ such that $e' \notin F' \cup \{e\}$
  - Since $F' \cup \{e\}$ does not contain a cycle by the choice of e
- $w(e') > w(e)$
  - $e' \in T' - F'$
  - $w(e) \leq L_T[k] < L_{T'}[k] \leq w(e')$
- Let $\hat{T} = T' \cup \{e\} - \{e'\}$
- $\hat{T}$ is a spanning tree
  - $|\hat{T}| = |T'| = |V| - 1$
  - $(V, \hat{T})$ is connected since (V, T') is connected
    - Removing an edge from a cycle
- $w(\hat{T}) = w(T') + w(e) - w(e') < w(T')$ $_{(\rightarrow\leftarrow)}$

Question 7

Let $G = (V, E)$ be a connected, undirected graph with weight function $w : E \to \mathbb{R}$. Let C be a cycle in G and e the strictly heaviest edge (w.r.t. w) in C, i.e., $\forall e' \in C - \{e\}, w(e) > w(e')$.
Prove/Disprove: no MST of G contains e.

Solution (the claim is true):

- Assume by contradiction there exists a MST T of G such that $e = (u, v) \in T$
- Let $C_u = \{w \in V \mid \text{exists a path in } (V, T - \{e\}) \text{ from } u \text{ to } w\}$
- Let $C_v = \{w \in V \mid \text{exists a path in } (V, T - \{e\}) \text{ from } v \text{ to } w\}$
- $\{C_u, C_v\}$ is a cut in G
- There exists a (u, v)-path P in G that does not traverse e
  - Since u and v are part of a cycle even if e is removed from G there exists a (u, v)-path

Solution

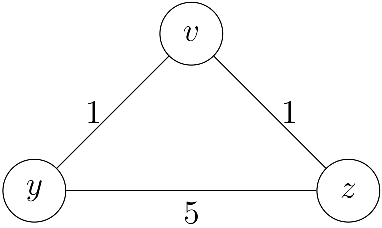
- Let P be a (u, v)-path such that $P \subseteq C - \{e\}$
- Let $e' = (x, y) \in P$ be an edge that crosses cut $\{C_u, C_v\}$ and $e' \notin T$
  - e' must exists since P leads from $u \in C_u$ to $v \in C_v$
  - If $e' \in T$, then T has a cycle
- $w(e') < w(e)$
  - e' is an edge of C
- Let $T' = T \cup \{e'\} - \{e\}$
- (V, T') is a spanning tree
  - $|T'| = |T| = |V| - 1$
  - (V, T') is connected since e' connects two connected components (subgraphs induced by C_u and C_v are trees)
- $w(T') = w(T) + w(e') - w(e) < w(T)$ $_{(\rightarrow\leftarrow)}$

Question 8

Given $G = (V, E)$ a connected undirected graph with weight function $w : E \to \mathbb{R}$, $v \in V$ and T a MST of the subgraph induced by $V - \{v\}$. Design an $O(V \log V)$-time algorithm that computes a MST of G.

Solution (Naive attempt):

- Let e be a light edge for cut $\{\{v\}, V - \{v\}\}$
- Let $T' = T \cup \{e\}$. Is T' a MST of G?
- NO!



Solution

Lemma

Given a connected and undirected graph $G = (V, E)$, a weight function $w : E \rightarrow \mathbb{R}$, $v \in V$ and T a MST of the subgraph induced by $V - \{v\}$. Let $E_v = \{(v, u) \mid u \in V\}$, and $G' = (V, E' = E_v \cup T)$. If T' is a MST of G', then T' is a MST of G.

Solution

Proof.

- Assume by contradiction there exist G, w, T, v such that T' is a MST of G' and T' is not a MST of G
- Let $\hat{T}$ be a MST of G that maximizes $\hat{T} \cap T'$
  - I.e., maximizes number of edges shared by $\hat{T}$ and T'
- $\hat{T} \not\subseteq E_v \cup T$
  - Otherwise, $\hat{T}$ is a spanning tree of $G' = (V, E_v \cup T)$
  - Implies $w(T') \leq w(\hat{T})$ since T' is a MST of $G' = (V, E_v \cup T)$
    - Cannot occur since T' is a spanning tree of G which is not a MST
- Let $\hat{e} = (x, y) \in E$ be an edge such that $\hat{e} \in \hat{T}$ and $\hat{e} \notin E_v \cup T$
- Let $C_x = \{w \in V \mid \text{exists a path in } (V, \hat{T} - \{\hat{e}\}) \text{ from } x \text{ to } w\}$
- Let $C_y = \{w \in V \mid \text{exists a path in } (V, \hat{T} - \{\hat{e}\}) \text{ from } y \text{ to } w\}$
- $\{C_x, C_y\}$ is a cut
- Edge $\hat{e}$ is light for cut $\{C_x, C_y\}$
  - From question 4

Solution — cont.

- It holds that $x, y \in V - \{v\}$
  - $\hat{e} = (x, y) \notin E_v$
- Let P be the unique (x, y)-path in T
  - P is not the edge $\hat{e}$ since $\hat{e} \notin T$
- Since P leads from $x \in C_x$ to $y \in C_y$, it must contain some edge $e \neq \hat{e}$ that crosses cut $\{C_x, C_y\}$
- Claim: $w(e) = w(\hat{e})$
- Proof of claim:
  - By question 4, $\hat{e}$ is light for cut $\{C_x, C_y\}$, thus $w(e) \geq w(\hat{e})$
  - Assume by contradiction $w(e) > w(\hat{e})$
  - Let $\tilde{T} = T \cup \{\hat{e}\} - \{e\}$
  - $(V - \{v\}, \tilde{T})$ is a spanning tree of the subgraph induced by $V - \{v\}$
    - $|\tilde{T}| = |T| = |V - \{v\}| - 1$
    - $(P - \{e\}) \cup \{\hat{e}\}$ is a (x, y)-path in $\tilde{T}$ since $(V - \{v\}, T)$ is connected
  - $w(\tilde{T}) = w(T) + w(\hat{e}) - w(e) < w(T)$ $_{(\rightarrow\leftarrow)}$

Solution — cont.

- We will show that there exists a MST of G which is closer (w.r.t number of shared edges) to T' then $\hat{T}$
- Let $\tilde{T} = \hat{T} \cup \{e\} - \{\hat{e}\}$
- $(V, \tilde{T})$ is a spanning tree
  - $|\tilde{T}| = |\hat{T}| = |V| - 1$
  - $(P - \{e\}) \cup \{\hat{e}\}$ is a (x, y)-path in $\tilde{T}$ since $(V, \hat{T})$ is connected
- $w(\tilde{T}) = w(\hat{T}) + w(e) - w(\hat{e}) = w(\hat{T})$
- $\tilde{T}$ is a MST of G and $|\tilde{T} \cap T'| > |\hat{T} \cap T'|$ $_{(\rightarrow\leftarrow)}$

Solution — cont.

Algorithm:

- Input: (G, w, v, T)
- Output: MST of G
  1. Construct $G' = (V, E' = E_v \cup T)$
  2. Run Prim's algorithm on G'

Correctness:

- From lemma and the correctness of Prim's algorithm

Run time analysis:

- $|E'| \leq 2n - 2 = O(V)$
- Construct G' – $O(V + E') = O(V)$
- Prim's algorithm on G' (implementing Q using a binary heap) – $O((V + E') \log V) = O(V \log V)$