Ελάχιστο Συνδετικό Δέντρο

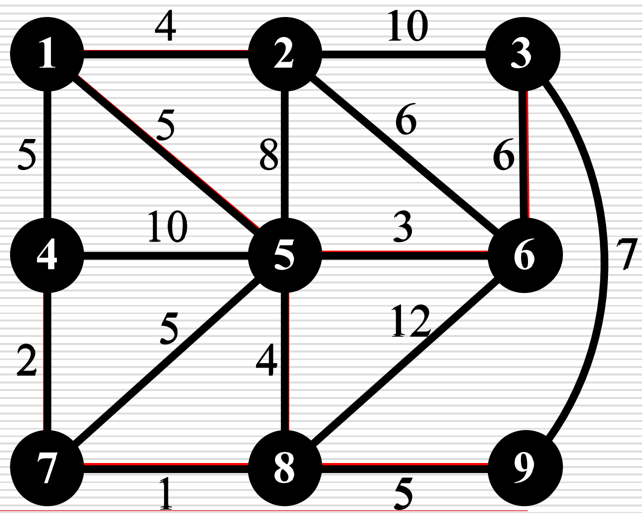
Δημήτρης Φωτάκης

Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών

Εθνικό Μετσόβιο Πολυτεχνείο

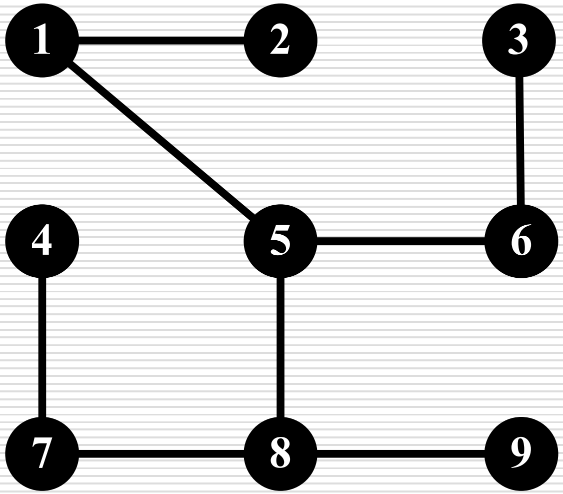
Ελάχιστο Συνδετικό Δέντρο (MST)

- Συνεκτικό μη-κατευθ. G(V, E, w) με βάρη $w : E \mapsto \mathbb{R}_{>0}$
  - Βάρος υπογραφήματος $T(V, E_T)$: $w(T) = \sum_{e \in E_T} w(e)$
- Ζητούμενο: ελάχιστου βάρους **συνεκτικό** υπογράφημα που καλύπτει όλες τις κορυφές.
  - Συνεκτικό (εξ' ορισμού) + ακυκλικό (ελάχιστο) $\Rightarrow$ Δέντρο.
  - Minimum Spanning Tree (MST, ΕΣΔ).
- Πρόβλημα συνδυαστ. βελτιστοποίησης με πολλές και σημαντικές εφαρμογές.
  - Σχεδιασμός συνδετικού δικτύου (οδικού, τηλεπ/κου, ηλεκτρικού) με ελάχιστο κόστος.



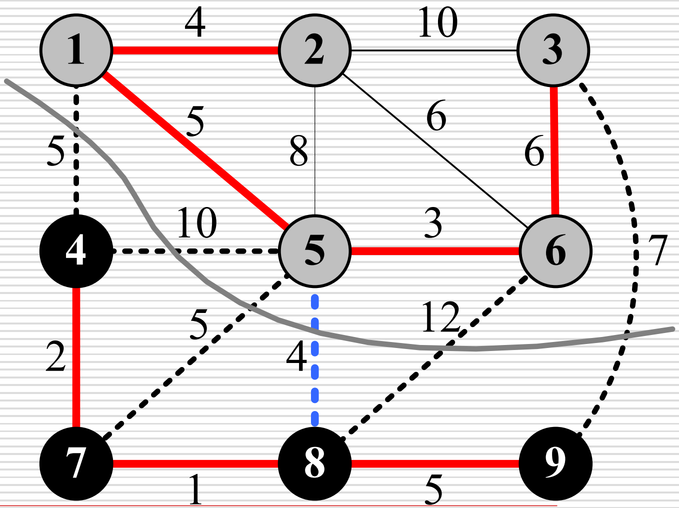
Δέντρα: Βασικές Ιδιότητες

- Δέντρο: συνεκτικό και ακυκλικό γράφημα.
- Για κάθε απλό μη-κατευθ. γράφημα T(V, E), τα παρακάτω είναι ισοδύναμα:
  - T δέντρο.
  - Κάθε ζευγάρι κορυφών ενώνεται με μοναδικό μονοπάτι.
  - T συνεκτικό και $|E| = |V| - 1$.
  - T ακυκλικό και $|E| = |V| - 1$.
  - T ελαχιστικά συνεκτικό.
  - T μεγιστικά ακυκλικό.



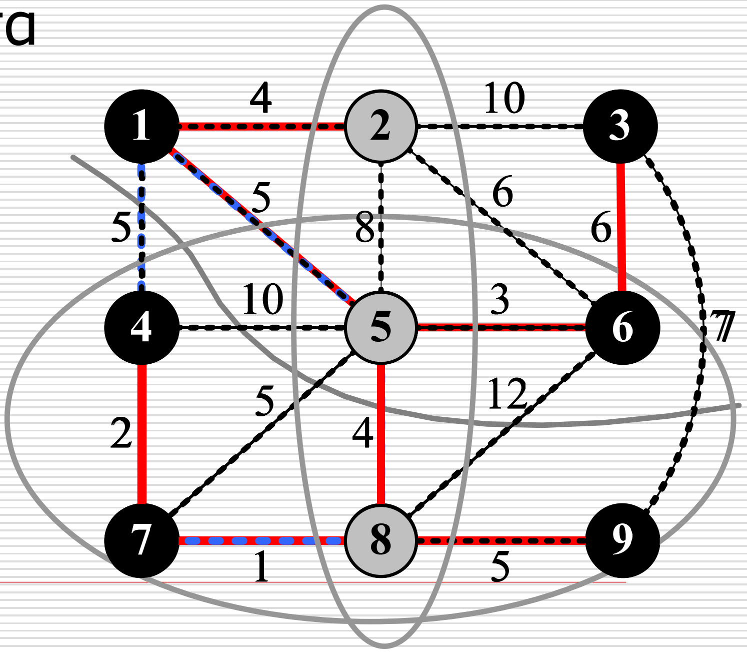
Τομές, Σύνολα Τομής, και ΕΣΔ

- Τομή (S, V \ S): διαμέριση κορυφών σε 2 σύνολα S, V \ S.
- Σύνολο τομής δ(S, V \ S): ακμές ένα άκρο στο S και άλλο άκρο στο V \ S.
  - δ(S, V \ S): όλες οι ακμές που **διασχίζουν** τομή (S, V \ S).
- Σύνολο ακμών E′ **διασχίζει** τομή (S, V \ S) αν $E' \cap \delta(S, V \setminus S) \neq \emptyset$.
- (Ε)ΣΔ ορίζεται από σύνολο ακμών (ελάχιστου) βάρους που διασχίζει όλες τις τομές.
  - Άπληστη στρατηγική: ενόσω «αγεφύρωτη» τομή, διέσχισέ την με ακμή ελάχιστου βάρους.



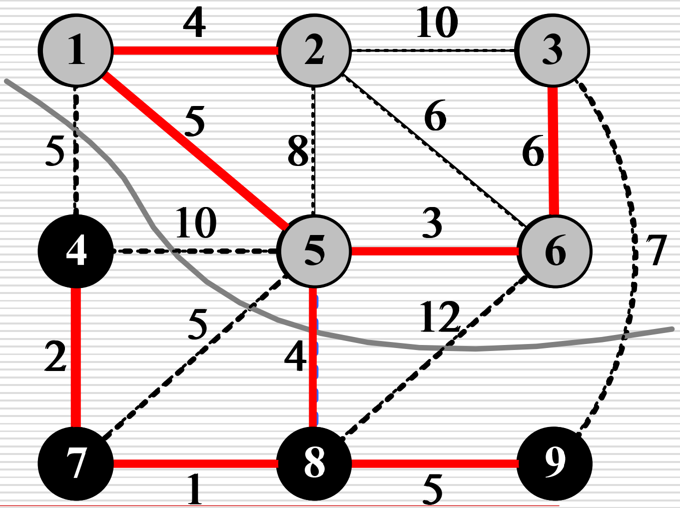
Άπληστος Αλγόριθμος για ΕΣΔ

- Έστω Δ δάσος (σύνολο ακμών χωρίς κύκλους).
- Ακμή $e \notin Δ$ είναι **ακμή επαύξησης** για Δ αν:
  - e διασχίζει μια τομή (S, V \ S) που δεν διασχίζει το Δ, και
  - e είναι ελάχιστου βάρους μεταξύ ακμών δ(S, V \ S).
- Ακμή επαύξησης για δάσος Δ συνιστά άπληστη επιλογή που σε |V|-1 βήματα οδηγεί σε ΕΣΔ:
  - Αν Δ δάσος και e ακμή επαύξησης Δ, $Δ \cup \{ e \}$ δάσος.
    - e δεν δημιουργεί κύκλο.
  - Αν $Δ \subset$ ΕΣΔ και e ακμή επαύξησης Δ, $Δ \cup \{ e \} \subseteq$ ΕΣΔ.



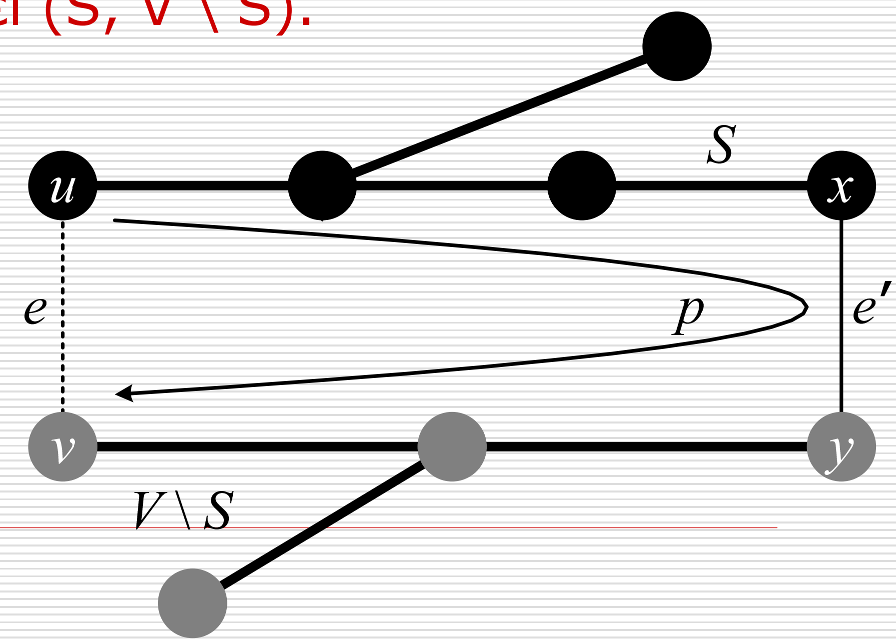
Άπληστος Υπολογισμός ΕΣΔ

- Έστω $T(V, E_T)$ ΕΣΔ για $G(V, E, w)$.
- Αφαιρώντας ακμή e από E_T προκύπτουν δύο συνεκτικές συνιστώσες, έστω S και $V \setminus S$ αντίστοιχα σύνολα κορυφών.
  - G_S και $G_{V \setminus S}$ αντίστοιχα επαγόμενα υπογραφήματα, και T_S και $T_{V \setminus S}$ αντίστοιχα υποδέντρα.
- Αρχή **βελτιστότητας**:
  - T_S αποτελεί ΕΣΔ για G_S και $T_{V \setminus S}$ αποτελεί ΕΣΔ για $G_{V \setminus S}$.
- Ιδιότητα **άπληστης επιλογής**:
  - e είναι μια ελάχιστου βάρους ακμή που διασχίζει τομή $(S, V \setminus S)$.
- **Άπληστος αλγόριθμος** για ΕΣΔ!



Άπληστη Επιλογή: Ορθότητα

- Έστω δάσος $\Delta \subset$ ΕΣΔ και e = {u, v} **ακμή επαύξησης** Δ. Τότε $\Delta \cup \{ e \} \subseteq$ ΕΣΔ.
  - (S, V \ S) τομή που δεν διασχίζει Δ και διασχίζει η ακμή e.
  - e ελάχιστου βάρους μεταξύ ακμών του δ(S, V \ S).
  - Έστω T ΕΣΔ τ.ω. $\Delta \subseteq$ T. Υποθέτουμε ότι $\Delta \cup \{e\} \not\subseteq T$
  - Έστω p μονοπάτι u – v στο T, και e′ = {x, y} ακμή T που διασχίζει (S, V \ S).
  - Αφού w(e) ≤ w(e′), και το $T' = (T \cup \{e\}) \setminus \{e'\}$ είναι **ΕΣΔ**:
    $$w(T') = w(T) + w(e) - w(e') \leq w(T)$$
  - Έχουμε ότι $\Delta \subseteq$ T και $e' \notin \Delta$. Άρα $\Delta \subseteq T \setminus \{e'\} = T' \setminus \{e\}$
  - ... και $\Delta \cup \{e\} \subseteq T'$



Άπληστος Αλγόριθμος για ΕΣΔ

MST$(G(V, E, w))$
  $\Delta \leftarrow \emptyset$;
  while $|\Delta| < |V| - 1$ **do**
    Υπολόγισε μια **ακμή επαύξησης** e για Δ;
    $\Delta \leftarrow \Delta \cup \{e\}$;
  return(Δ);

- Αρχικά $\Delta = \emptyset$ **δάσος** και **υποσύνολο** κάθε ΕΣΔ.
- Επαγωγικά, e ακμή επαύξησης για Δ:
  - $\Delta \cup \{ e \}$ **δάσος** και **υποσύνολο** κάποιου ΕΣΔ.
- Όταν $|\Delta| = |V| - 1$, Δ δέντρο, άρα και **ΕΣΔ**.

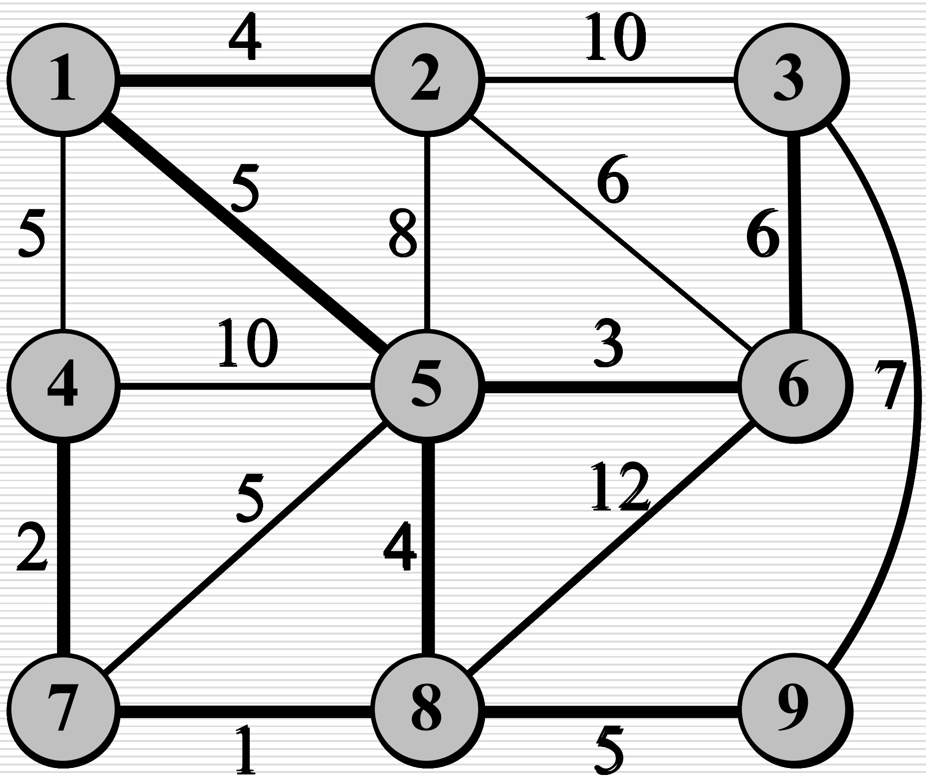
Αλγόριθμος Kruskal

MST-Kruskal$(G(V,E,w))$

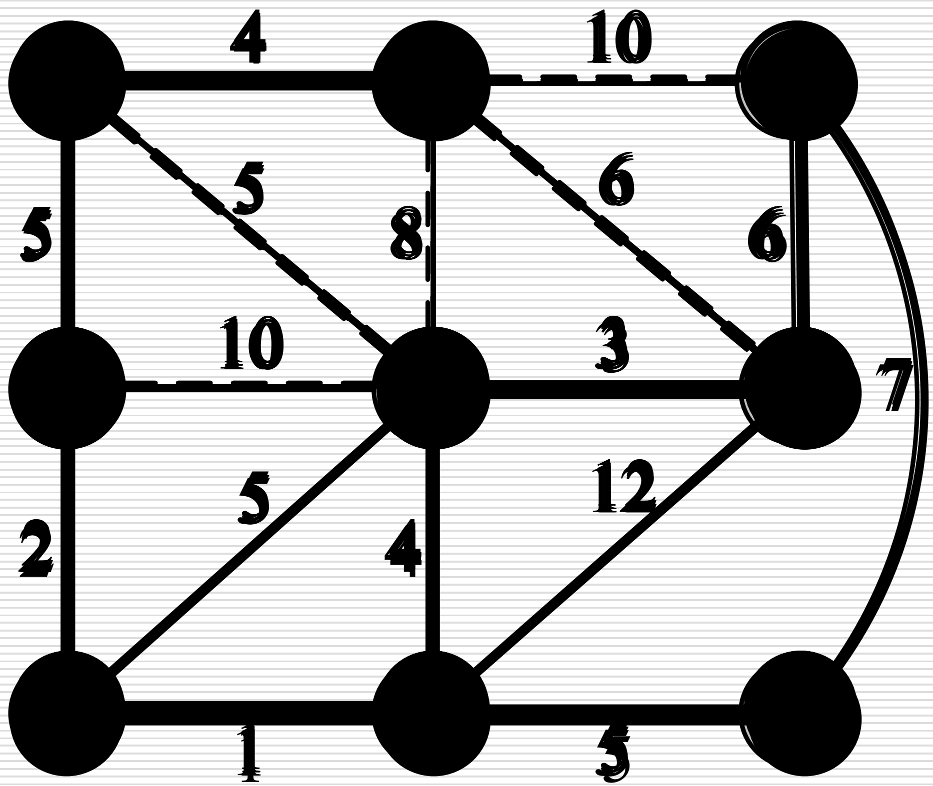
Ταξινόμησε ακμές σε αύξουσα σειρά βάρους, $w(e_1) \leq \cdots \leq w(e_m)$.
$\Delta \leftarrow \emptyset$; $i \leftarrow 1$;
while $|\Delta| < |V| - 1$ **and** $i \leq m$ **do**
  if $\Delta \cup \{e_i\}$ δεν έχει κύκλο **then**
    $\Delta \leftarrow \Delta \cup \{e_i\}$;
  $i \leftarrow i + 1$;

- Υλοποίηση: κύκλος στο $\Delta \cup \{e_i\}$ ελέγχεται με Union-Find.
  - Χρόνος εκτέλεσης: $\Theta(m \log m)$.
- **Ορθότητα**: αν e_i προστεθεί τότε ακμή επαύξησης για Δ:
  - Όχι κύκλος, άρα e_i διασχίζει μια τομή που δεν διασχίζει το Δ.
  - Αύξουσα σειρά βάρους: e_i ελάχιστου βάρους (πρώτη που ελέγχεται) από όσες ακμές διασχίζουν συγκεκριμένη τομή.

Αλγόριθμος Kruskal: Παράδειγμα



Αλγόριθμος Prim: Παράδειγμα



Αλγόριθμος Prim

- **Υλοποίηση**:
  - Ελάχιστο c[v]: ουρά προτεραιότητας.
  - Binary heap: $\Theta(m \log n)$
  - Fibonacci heap: $\Theta(m + n \log n)$
- **Ορθότητα**:
  - {v, p[v]} αποτελεί ακμή επαύξησης:
    - Διασχίζει τομή (S, V \ S).
    - Ελάχιστου βάρους μεταξύ ακμών του δ(S, V \ S).

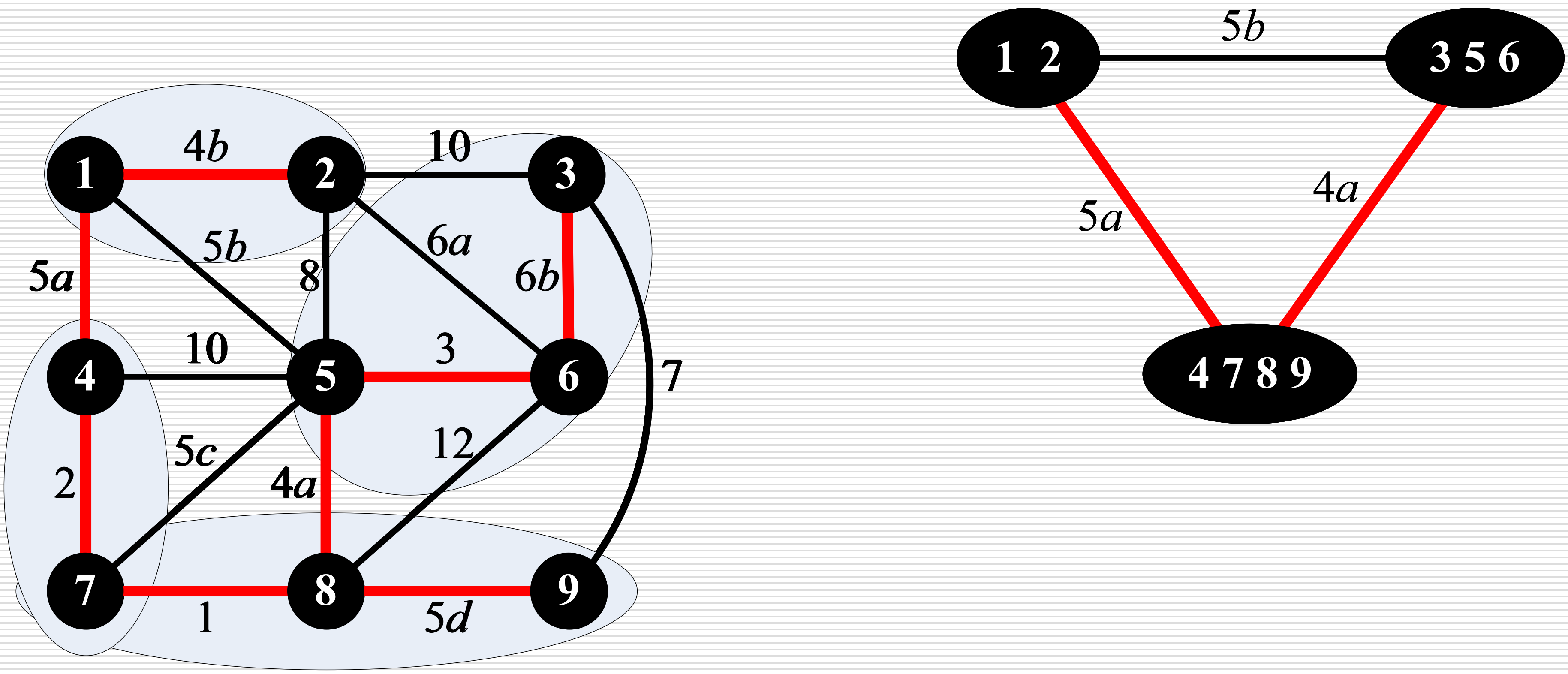
$\mathbf{MST\text{-}Prim}(G(V, E, w), s)$

- **for all** $u \in V$ **do**
  - $c[u] \leftarrow \infty;\ \ p[u] \leftarrow \mathrm{NULL};$
- $c[s] \leftarrow 0;\ \ S \leftarrow \emptyset;\ \ \Delta \leftarrow \emptyset;$
- **while** $|S| < |V|$ **do**
  - $u \notin S : c[u] = \min_{v \notin S}\{c[v]\};$
  - $S \leftarrow S \cup \{u\};$
  - **for all** $v \in \mathrm{AdjList}[u]$ **do**
    - **if** $v \notin S$ **and** $w(u, v) < c[v]$ **then**
      - $c[v] \leftarrow w(u, v);$
      - $p[v] \leftarrow u;$
  - **if** $p[u] \neq \mathrm{NULL}$ **then**
    - $\Delta \leftarrow \Delta \cup \{u, p[u]\};$

Αλγόριθμος Boruvka

- «Παράλληλη» εκδοχή γενικού άπληστου αλγόριθμου.
  - Αρχικά κάθε κορυφή αποτελεί μία συνεκτική συνιστώσα.
  - Φάση: ενόσω #συνεκτικών συνιστωσών > 1:
    - Κάθε συνεκτική συνιστώσα σ επιλέγει στο ΕΣΔ την **ελαφρύτερη** ακμή με ένα άκρο στο σ (ακμή επαύξησης).
      - Μια ακμή μπορεί να επιλεγεί και από τα δύο άκρα της.
      - Απαραίτητη ολική διάταξη των ακμών (χωρίς ισοπαλίες), διαφορετικά μπορεί να σχηματιστούν κύκλοι.
    - Συνεκτικές συνιστώσες ενημερώνονται με βάση ακμές που επιλέχθηκαν στην τρέχουσα φάση.
- Ολική διάταξη ακμών (χωρίς ισοπαλίες): υπολογίζει ΕΣΔ.
  - Κάθε ακμή που επιλέγεται, αποτελεί ακμή επαύξησης: ελάχιστου βάρους ακμή που διασχίζει μία τομή.
  - Όχι κύκλοι: **μοναδική** ελαφρύτερη ακμή διασχίζει κάθε τομή.

Αλγόριθμος Boruvka: Παράδειγμα



Αλγόριθμος Boruvka

- (Ακολουθιακή) υλοποίηση σε $O(m \log n)$.
  - Κάθε φάση σε χρόνο $O(m)$ με δύο περάσματα των ακμών.
  - 1ο πέρασμα βρίσκει ελαφρύτερη ακμή κάθε συνιστώσας.
  - 2ο πέρασμα εντάσσει ελαφρύτερες ακμές στο ΕΣΔ και ενημερώνει / συμπτύσσει συνιστώσες (π.χ., με BFS/DFS).
  - Σε κάθε φάση, #συνιστωσών μειώνεται στο μισό.
    - #φάσεων = $O(\log n)$.
- Πολλοί σύγχρονοι αλγόριθμοι (σχεδόν) γραμμικού χρόνου βασίζονται σε ιδέα Boruvka.

Κανόνες Σχηματισμού ΕΣΔ

- Ακμή e που για κάποια τομή (S, V \ S), αποτελεί ελάχιστου βάρους ακμή που διασχίζει τομή (S, V \ S):
  - e **ανήκει** σε κάποιο ΕΣΔ.
  - Γρήγορη επιλογή τέτοιων ακμών χωρίς κύκλους, και **ένταξη** σε ΕΣΔ.
- Ακμή e που για κάποιον κύκλο C αποτελεί μέγιστου βάρους ακμή κύκλου C:
  - Αν βάρος e μεγαλύτερο από βάρος άλλων ακμών του C, e **δεν ανήκει** σε κανένα ΕΣΔ.
  - Αν όλες οι ακμές του C έχουν ίδιο βάρος, e **δεν ανήκει** σε κάποιο ΕΣΔ.
  - Ενόσω υπάρχει κύκλος C, **αποκλεισμός** (μιας) βαρύτερης ακμής C.

Συζήτηση

Deterministic comparison based algorithms.

- $O(m \log n)$ [Jarník, Prim, Dijkstra, Kruskal, Boruvka]
- $O(m \log \log n)$. [Cheriton-Tarjan 1976, Yao 1975]
- $O(m\, \beta(m, n))$. [Fredman-Tarjan 1987]
- $O(m \log \beta(m, n))$. [Gabow-Galil-Spencer-Tarjan 1986]
- $O(m\, \alpha\, (m, n))$. [Chazelle 2000]

Holy grail. $O(m)$.

Notable.

- $O(m)$ randomized. [Karger-Klein-Tarjan 1995]
- $O(m)$ verification. [Dixon-Rauch-Tarjan 1992]

Συζήτηση – Ασκήσεις

- Έστω γράφημα G με διαφορετικά βάρη στις ακμές.
  - Νδο κάθε ΕΣΔ του G περιέχει την ακμή ελάχιστου βάρους.
  - Νδο G έχει μοναδικό ΕΣΔ.
  - Αληθεύει ότι η ακμή μέγιστου βάρους δεν ανήκει στο ΕΣΔ;
- Έστω γράφημα G με κύκλο C.
  - Νδο η ακμή μέγιστου βάρους του C (αν είναι μοναδική) δεν ανήκει σε κανένα ΕΣΔ του G.
- Έστω T ΕΣΔ για γράφημα G(V, E, w).
  - Να δείξετε ότι T παραμένει ΕΣΔ για G(V, E, w/2).
  - Αληθεύει ότι το T παραμένει ΕΣΔ για G(V, E, w+k);

Συζήτηση – Ασκήσεις

- (Ντετερμινιστικός) αλγόριθμος ΕΣΔ με χρόνο εκτέλεσης $O(m \log\log n)$;
- Υπολογισμός ΕΣΔ T υπό περιορισμούς ότι κάποιες ακμές πρέπει να (μην) ανήκουν στο T;
- Υπολογισμός ΣΔ T με δεύτερο μικρότερο βάρος;
- Bottleneck κόστος ΣΔ T: $b(T) = \max_{e \in T}\{w(e)\}$
  - Υπολογισμός ΣΔ με ελάχιστο bottleneck κόστος;