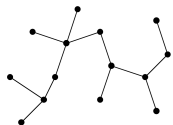


Minimum udSpændende Træer (MST)

Træer

Et (frit/u-rodet) træ er en *uorienteret* graf $G = (V, E)$ som er

- ▶ Sammenhængende: der er en sti mellem alle par af knuder.
- ▶ Acyklisk: der er ingen kreds af kanter.



(a)

Træ



(b)

Skov



(c)

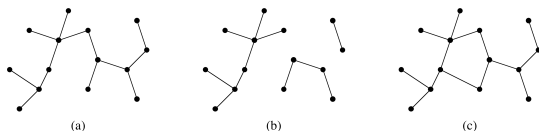
Graf med kreds (ikke træ)

(Uorienteret, acyklisk graf = skov af træer.).

Træer

Sætning (B.2): For *uorienteret* graf $G = (V, E)$ er flg. ækvivalent (gælder det ene, gælder det andet):

- ▶ G er et træ (dvs. sammenhængende og acyklisk).
- ▶ G er sammenhængende, men er det ikke hvis nogen kant fjernes.
- ▶ G er sammenhængende og $m = n - 1$.
- ▶ G er acyklisk, men er det ikke hvis nogen kant tilføjes.
- ▶ G er acyklisk og $m = n - 1$.
- ▶ Mellem alle par af knuder er der præcis een vej.



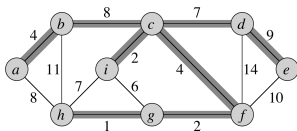
Bevis (ikke pensum): se appendix B.5.

Læs (pensum) appendix B.4 og B.5 for basale definitioner for grafer.

Minimum Spanning Tree (MST)

Udspændende træ for sammenhængende graf $G = (V, E)$:

En delgraf $T = (V, E')$, $E' \subseteq E$, som er et træ (NB: samme knudemængde V).



Iflg. sætning ovenfor har alle udspændende træer samme antal kanter ($m = n - 1$).

Minimum udSpændende Træ (MST) for en vægtet sammenhængende graf G : et udspændende træ for G som har mindst mulig sum af kantvægte (intet udspændende træ har mindre sum).

Motivation: forbind punkter i et forsyningsnetværk (elektricitet, olie, ...) billigt muligt. Kant = mulig forbindelse, vægt = pris for at etablere forbindelse. Dette var motivationen for den første algoritme for problemet (Borůvka, 1926, Østrig-Ungarn, nu Tjekkiet).

Algoritmer for MST

Grundidé: Byg MST ved at vælge kanterne een efter een. Vedligehold følgende **Invariant:** Der findes et MST som indeholder de valgte kanter A .

GENERIC-MST(G, w)

$A = \emptyset$

while A is not a spanning tree

 find an edge (u, v) that is safe for A

$A = A \cup \{(u, v)\}$

return A

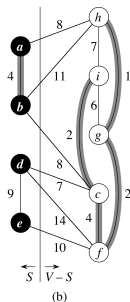
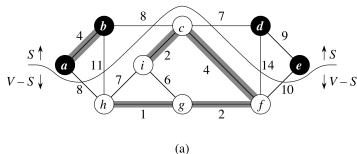
Safe kant for A : kant som kan tilføjes uden at ødelægge invarianten (mindst een må findes når invarianten gælder og $|A| < n - 1$).

- ▶ Initialisering: Enhver sammenhængende graf har et mindst eet ST (via sætningen fra B.5, punkt 2 - fjern kanter til betingelsen nås), og har derfor et MST. Dette indeholder mængden \emptyset .
- ▶ Vedligeholdelse: Per definition af safe.
- ▶ Terminering: ethvert (M)ST indeholder præcis $n - 1$ kanter. Da A vokser med een kant per iteration, giver invarianten at algoritmen terminerer, og at A da er et MST (A er indeholdt i et MST, og har samme antal kanter som dette, er derfor lig dette).

Cut-sætning

Cut: En delmængde $S \subseteq$ af knuderne.

Kan ses som en to-delning af knuderne i to mængder S og $V - S$.



Kant henover cut: en kant i $S \times (V - S)$.

Cut-sætning

Sætning:

Hvis

- ▶ der eksisterer et MST som indeholder A ,
- ▶ S er et cut som A ikke har kanter henover,
- ▶ e er en letteste kant blandt kanterne henover cuttet,

så

- ▶ er e safe for A (dvs. der der eksisterer et MST som indeholder $A \cup \{e\}$).

Cut-sætning

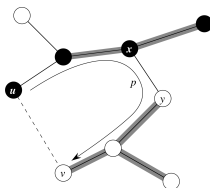
Bevis:

- ▶ Der findes et MST T som indeholder A .
- ▶ Vi skal lave et MST T' som indeholder $A \cup \{e\}$.

Lad $e = (u, v)$ være en letteste kant henover cuttet S .

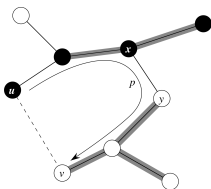
Da T er sammenhængende, må der være en sti i T mellem u og v , hvorpå der er mindst een kant (x, y) henover cuttet S .

Lad T' være T med (x, y) udskiftet til $e = (u, v)$:



Cut-sætning

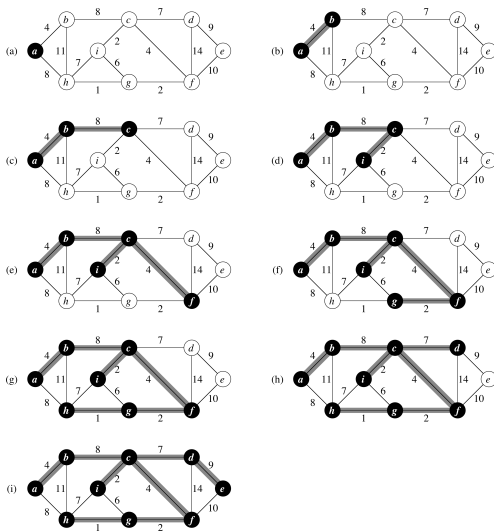
Lad T' være T med (x, y) udskiftet til $e = (u, v)$:



Som T er T' stadig sammenhængende (i alle stier kan (x, y) erstattes af resten af stien fra u til v , samt kanten (u, v)), og har T $n - 1$ kanter. Det er derfor et træ (pga. sætning tidligere). Det er klart stadig udspændende. Det kan kun være lettere end T . Det indeholder $A \cup \{e\}$ (da fjernede kant (x, y) ikke er i A).

Prim-Jarnik MST-algoritmen (Prim 1957, Jarnik 1930)

A er et træ. Bruger cut-sætningen med $S =$ alle knuder i A.



Prim-Jarnik MST-algoritmen (Prim 1957, Jarnik 1930)

A er et træ. Bruger cut-sætningen med $S =$ alle knuder i A .

Invariant: En knude $v \in V - S$ opbevarer information om sin korteste kant henover cut i $v.key$ og $v.\pi$.

$V - S$ opbevares i en (min-)prioritetskø.

```
PRIM( $G, w, r$ )
   $Q = \emptyset$ 
  for each  $u \in G.V$ 
     $u.key = \infty$ 
     $u.\pi = \text{NIL}$ 
  INSERT( $Q, u$ )
  DECREASE-KEY( $Q, r, 0$ ) //  $r.key = 0$ 
  while  $Q \neq \emptyset$ 
     $u = \text{EXTRACT-MIN}(Q)$ 
    for each  $v \in G.Adj[u]$ 
      if  $v \in Q$  and  $w(u, v) < v.key$ 
         $v.\pi = u$ 
        DECREASE-KEY( $Q, v, w(u, v)$ )
```

Korrekthed: via cut-sætningen og invarianten.

Køretid: n INSERT, n EXTRACTMIN, m DECREASEKEY på prioritetskø af størrelse $O(n)$, i alt $O(m \log n)$.

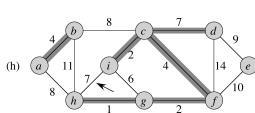
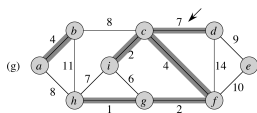
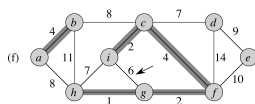
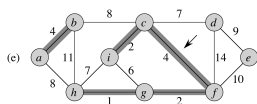
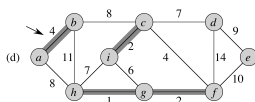
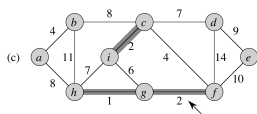
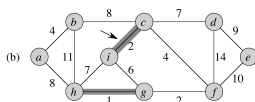
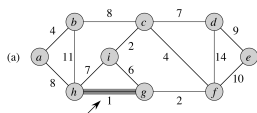
Kruskal MST-algoritmen (1956)

A er en skov. Bruger cut-sætningen med $S =$ knuderne i eet af træerne i A .

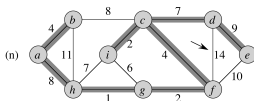
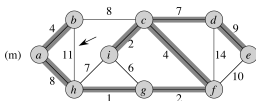
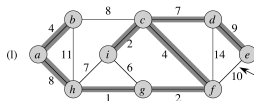
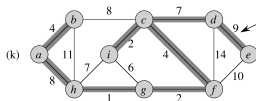
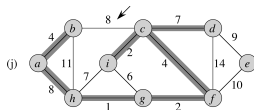
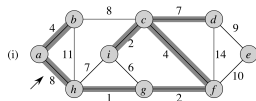
Forsøger at tilføje kanter til A i letteste-først-orden.

Tilføjer kun kant (u, v) til A hvis der ikke laves en kreds, dvs. hvis u og v ligger i forskellige træer. Hvis (u, v) tilføjes, vil disse to træer blive til eet bagefter.

Kruskal MST-algoritmen (1956)



Kruskal MST-algoritmen (1956)

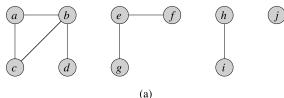


Kruskal MST-algoritmen (1956)

Tilføjer kun kant (u, v) til A hvis der ikke laves en kreds, dvs. hvis u og v ligger i forskellige træer. Hvis (u, v) tilføjes, vil disse to træer blive til eet bagefter.

Vedligeholder opdelingen i træer i A ved hjælp af en *disjoint-set* datastruktur på V :

MAKE-SET(x), UNION(x, y) FIND-SET(x)



Edge processed	Collection of disjoint sets									
initial sets	{a}	{b}	{c}	{d}	{e}	{f}	{g}	{h}	{i}	{j}
(b,d)	{a}	{b,d}	{c}		{e}	{f}	{g}	{h}	{i}	{j}
(e,g)	{a}	{b,d}	{c}		{e,g}	{f}		{h}	{i}	{j}
(a,c)	{a,c}	{b,d}			{e,g}	{f}		{h}	{i}	{j}
(h,i)	{a,c}	{b,d}			{e,g}	{f}		{h,i}		{j}
(a,b)	{a,b,c,d}				{e,g}	{f}		{h,i}		{j}
(e,f)	{a,b,c,d}				{e,f,g}			{h,i}		{j}
(b,c)	{a,b,c,d}				{e,f,g}			{h,i}		{j}

(b)

Kruskal MST-algoritmen (1956)

Lemma: der findes en datastruktur for disjoint-sets hvor

- ▶ n MAKE-SET(x)
- ▶ $n - 1$ UNION(x, y)
- ▶ m FIND-SET(x)

tager i alt $O(m + n \log n)$ tid.

Kruskal MST-algoritmen (1956)

```
KRUSKAL( $G, w$ )  
   $A = \emptyset$   
  for each vertex  $v \in G.V$   
    MAKE-SET( $v$ )  
  sort the edges of  $G.E$  into nondecreasing order by weight  $w$   
  for each  $(u, v)$  taken from the sorted list  
    if FIND-SET( $u$ )  $\neq$  FIND-SET( $v$ )  
       $A = A \cup \{(u, v)\}$   
      UNION( $u, v$ )  
  return  $A$ 
```

Invariant:

1. Der er en sti fra v til u af kanter i $A \Leftrightarrow u$ og v ligger i samme mængde i Disjoint-Set datastrukturen.
2. Når en kant forsøges tilføjet A , har enhver tidligere forsøgt kant begge sine endepunkter i samme mængde i Disjoint-Set datastrukturen.

Kruskal MST-algoritmen (1956)

Korrekthed:

Via cut-sætningen og del 2) af invarianten, som viser at en tilføjet kant (u, v) er en letteste henover et cut (nemlig cuttet givet ved knudemængden $\text{FIND-SET}(u)$), ses at A altid er indeholdt i et MST.

Når algoritmen stopper, kan der kun være een mængde tilbage i Disjoint-Set datastrukturen, af del 1) af invarianten (da den oprindelige graf er sammenhængende), ellers er der en kant, som ikke er blevet testet. Derfor er der lavet $n - 1$ unions, og dermed er $|A| = n - 1$, så A er selv dette MST.

Køretid: Sortér m kanter, lav n MAKE-SET, $n - 1$ UNION, m FIND-SET.

I alt $O(m \log m)$.