

Datenstrukturen und Algorithmen

Vorlesung 16: Minimale Spannbäume (K23)

Joost-Pieter Katoen

Lehrstuhl für Informatik 2
Software Modeling and Verification Group

<http://moves.rwth-aachen.de/teaching/ss-15/dsa1/>

12. Juni 2015



Übersicht

- 1 Spannbäume
- 2 Minimale Spannbäume
- 3 Greedy Algorithmen
- 4 Die Algorithmen von Kruskal und Prim
- 5 Implementierung und Komplexität

Übersicht

- 1 Spannbäume
- 2 Minimale Spannbäume
- 3 Greedy Algorithmen
- 4 Die Algorithmen von Kruskal und Prim
- 5 Implementierung und Komplexität

Spannbaum Probleme

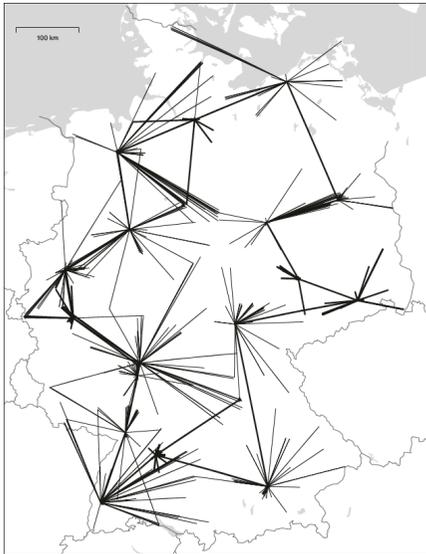
Spannbaum

Ein **Spannbaum** eines ungerichteten, zusammenhängenden Graphen G ist ein **Teilgraph** von G , der ein ungerichteter **Baum** ist und alle Knoten von G enthält.

Beispiel (Spannbaum Probleme)

- ▶ Verbinde alle Kunden durch Glasfaserkabeln
- ▶ Computernetzwerke verkabeln
- ▶ Verdrahtung von Schaltungen beim Chipdesign
- ▶ Erneure die Straßenbeläge zwischen alle Flughafenterminals
- ▶

Glasfabernetz

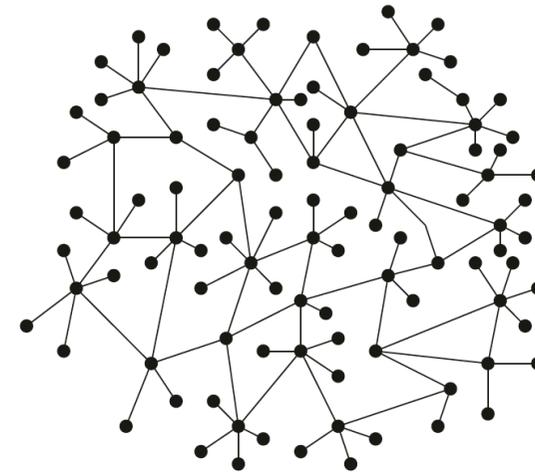


- ▶ Müssen alle Leitungen erneuert werden?
- ▶ Wie wählt man die zu erneuernden Abschnitte aus?
- ▶ Gibt es eine eindeutige Lösung?

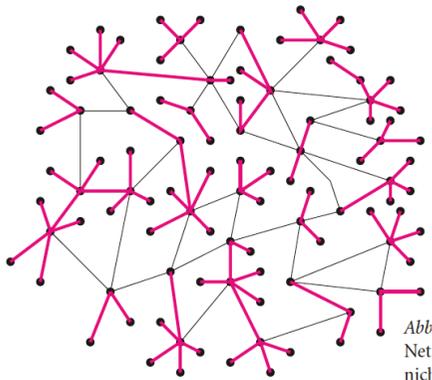
so daß

- ▶ Alle Kunden profitieren
- ▶ Jeder kann mit jedem über Glasfaser kommunizieren
- ▶ Sparsamkeit: keine doppelte Verbindungen

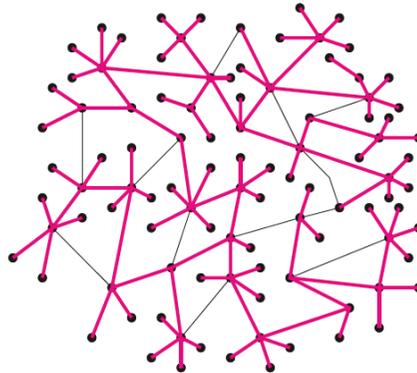
Telefonleitungsnetz



Zwei Abdeckungen



Alle Kunden sind angeschlossen



... und sind miteinander verbunden

Abb
Net
nich

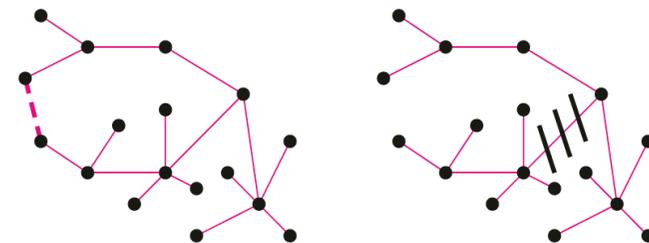
Einige Fakten bzgl. Bäume

Satz

In einem Baum sind je zwei Knoten durch genau einen Weg verbunden.

Satz

Ein Baum ist minimal zusammenhängend und maximal kreisfrei.



Fügt man eine Kante in einem Baum hinzu, entsteht ein Kreis.
Löscht man eine Kante, erhält man zwei Zusammenhangskomponenten.

Einige Fakten bzgl. Bäume

Satz

Ein Baum ist minimal zusammenhängend und maximal kreisfrei.

Beweis.

In jedem Baum sind je 2 Knoten durch einen eindeutigen Weg verbunden. Löscht man eine Kante, wird mindestens einer dieser Wege unterbrochen. Damit ist der Restgraph nicht zusammenhängend. Fügt man eine Kante (u, v) hinzu, dann bekommt man zusätzlich zu den eindeutigen Weg im Baum zwischen u und v einen weiteren Weg zwischen u und v . Das ergibt ein Kreis.

1. Jeder Baum (mit mindestens 2 Knoten) besitzt mindestens 2 Blätter
2. Ein Baum mit n Knoten hat $n-1$ Kanten.

Beweis.

Hausaufgabe. Hinweis für 2: plücken Sie Blatt um Blatt ab.

Anzahl der Spannbäume

Die Cayley Formel

Den vollständigen Graphen mit n Knoten hat n^{n-2} Spannbäume.

Beweis.

In der Vorlesung.

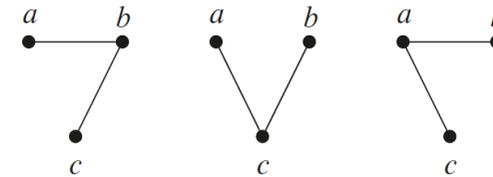
Anzahl der Spannbäume

Spannbaum

Ein **Spannbaum** eines ungerichteten, zusammenhängenden Graphen G ist ein **Teilgraph** von G , der ein **Baum** ist und alle Knoten von G enthält.

Die Cayley Formel

Den vollständigen Graphen mit n Knoten hat n^{n-2} Spannbäume.



3 Spannbäume für $n=3$ Knoten

Übersicht

- 1 Spannäume
- 2 Minimale Spannäume
- 3 Greedy Algorithmen
- 4 Die Algorithmen von Kruskal und Prim
- 5 Implementierung und Komplexität

Kosten

Beispiel

- ▶ Finde den **kostengünstigsten** Weg, um eine Menge von Flughafenterminals, Städten, ... zu verbinden
- ▶ Verdrahtung von Schaltungen mit **geringstem** Energieverbrauch
- ▶ Verbinde alle Kunden **kostengünstig** durch Glasfaserkabeln
- ▶ Computernetzwerke verkabeln

Das ist die Begründung für (kosten-) **minimale Spannäume!**

Was ist ein minimaler Spannbaum?

Kantengewichteter ungerichteter Graph

Ein **(kanten-)gewichteter** Graph G ist ein Tripel (V, E, W) , wobei:

- ▶ (V, E) ein ungerichteter Graph ist, und
- ▶ $W : E \rightarrow \mathbb{R}$ Gewichtsfunktion. $W(e)$ ist das **Gewicht** der Kante e .

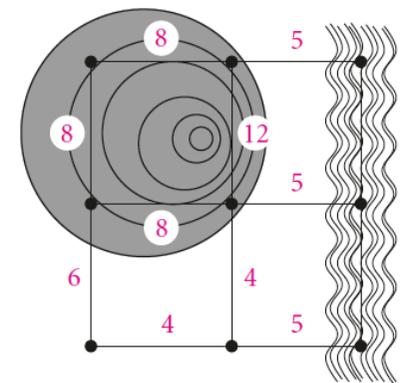
Gewicht eines Graphen

Das **Gewicht** $W(G')$ des Teilgraphen $G' = (V', E')$ vom gewichteten Graph G ist: $W(G') = \sum_{e \in E'} W(e)$.

Minimaler Spannbaum

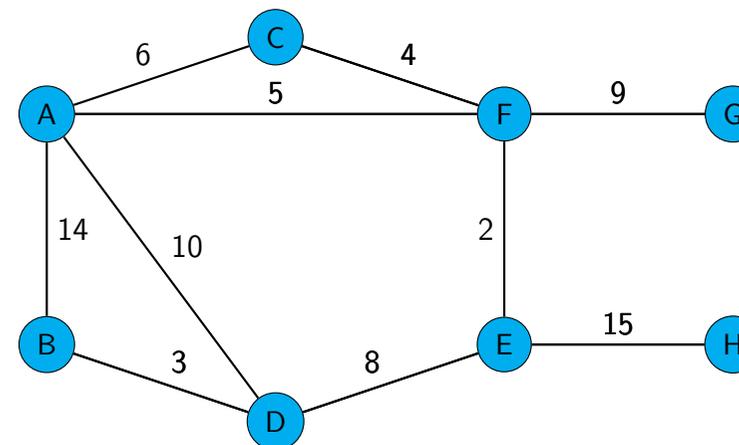
Ein Spannbaum vom (ungerichteter, gewichteter, zusammenhängen) Graphen G mit minimalem Gewicht heißt **Minimaler Spannbaum** (minimum spanning tree, MST) von G .

Leitungskosten



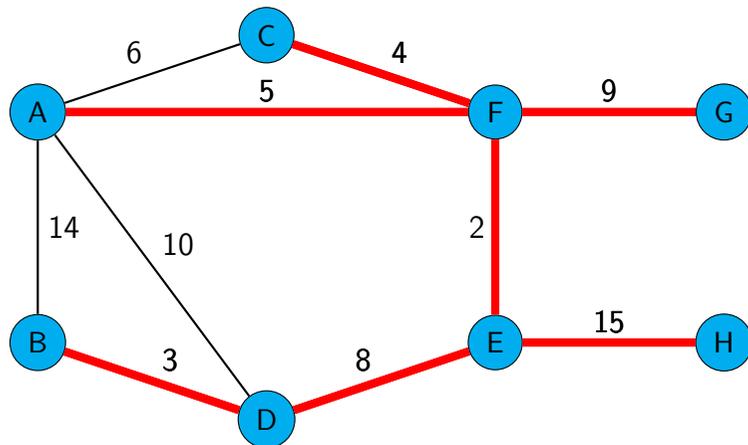
Unterschiedliche Umgebungen verursachen unterschiedliche Kosten für die Verlegung von Kabeln.

Minimaler Spannbaum – Beispiel



Was ist ein minimaler Spannbaum?

Minimaler Spannbaum – Beispiel

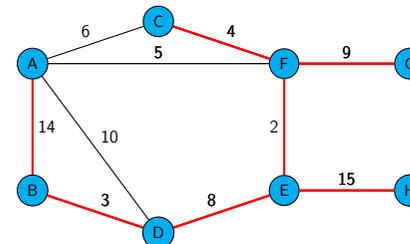


Das ist *ein* minimaler Spannbaum (mit Gesamtgewicht 46).
In diesem Fall ist es auch der einzige.

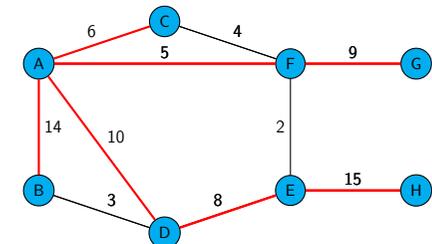
Übersicht

- 1 Spannäume
- 2 Minimale Spannäume
- 3 Greedy Algorithmen
- 4 Die Algorithmen von Kruskal und Prim
- 5 Implementierung und Komplexität

Tiefen- oder Breitensuche?



Tiefensuchbaum (von A gestartet)
Gesamtgewicht: 55



Breitensuchbaum (von A gestartet)
Gesamtgewicht: 67

Der Tiefensuchbaum und der Breitensuchbaum sind zwar Spannäume,
aber nicht notwendigerweise minimale Spannäume.

Greedy Algorithmen

Wir werden zwei Greedy-Algorithmen für MST präsentieren.

Greedy-Algorithmen („gierig“)

- ▶ Treffe in jedem Schritt eine Entscheidung, die bezüglich eines „kurzfristigen“ Kriteriums optimal ist.
- ▶ Dieses Kriterium sollte **günstig** (→ Komplexität) auswertbar sein.
- ▶ Nachdem eine Wahl getroffen wurde, kann sie **nicht mehr rückgängig** gemacht werden.

Mit Greedy-Methoden ist nicht garantiert, dass immer die beste Lösung gefunden wird, denn

- ▶ immer das lokale Optimum zu nehmen, führt nicht automatisch auch zum globalen Optimum.
- ▶ In einigen Fällen, wie dem minimalen Spannbaum und dem Kürzesten-Wege-Problem, wird aber **immer** die optimale Lösung gefunden.

Greedy?

Beispiel

Greedy kann **beliebig schlecht** werden:

- ▶ Knotenfärbungsproblem für Graphen

Greedy kann **gut** sein:

- ▶ Bin Packing ($\leq 2 \times$ Optimum)

Greedy kann **optimal** sein:

- ▶ Minimaler Spannbaum, Kürzester-Weg-Problem.

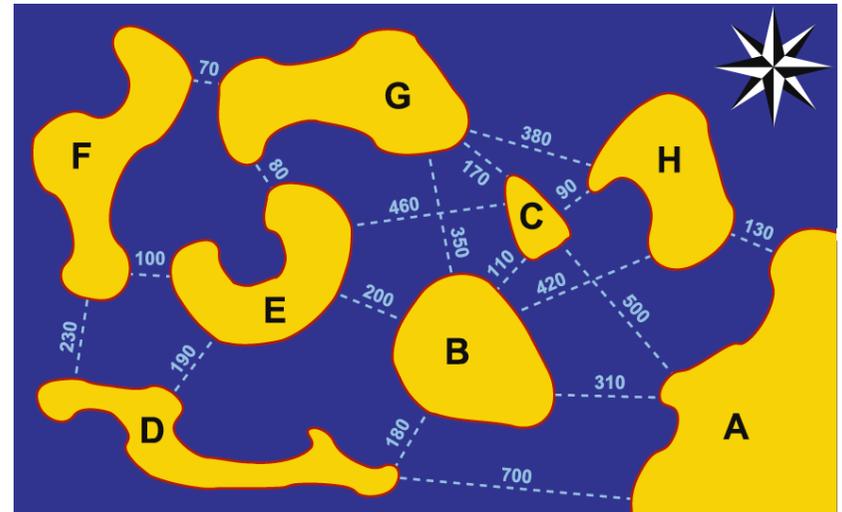
Wann ist eine greedy Lösungsstrategie optimal?

- ▶ Optimale Lösung setzt sich aus optimalen Teilproblemen zusammen
- ▶ Unabhängigkeit von anderen Teillösungen

Übersicht

- 1 Spannäume
- 2 Minimale Spannäume
- 3 Greedy Algorithmen
- 4 Die Algorithmen von Kruskal und Prim
- 5 Implementierung und Komplexität

Das Inselreich der Stamm der Algolaner



Was ist die optimale Strategie die Fährverbindungen durch Brücken zu ersetzen?

Zwei Greedy minimaler Spannbaumalgorithmen

Eingabe: ein gewichteter zusammenhängender Graph G mit n Knoten

Ausgabe: ein minimaler Spannbaum von G

Prim's Strategie

1. Wähle einen Startknoten.
2. Markiere die "billigste" vom bereits konstruierten Baum ausgehende Kante, falls sie keinen Kreis schließt
3. Wiederhole Schritt 2., so lange noch keine $n-1$ Kanten markiert sind.

Kruskal's Strategie

So lange noch keine $n-1$ Kanten markiert (d.h. selektiert) sind:

1. Wähle eine "billigste" noch unmarkierte Kante
2. Markiere sie, falls sie keinen Kreis mit anderen markierten Kanten schließt

Beispiel Prim's und Kruskal's Strategie

Das Inselreich der Algolaner

Korrektheit des Algorithmus von Kruskal

Der Algorithmus von Kruskal bestimmt einen minimalen Spannbaum.

Beweis.

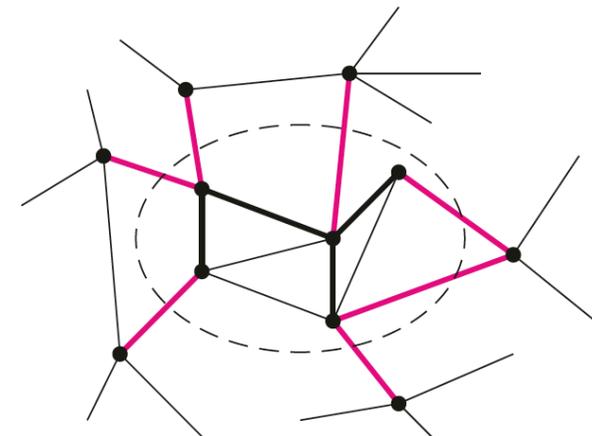
Der Algorithmus liefert ein Baum B . Zu beweisen: B ist minimal. Mit Induktion. Idee: die Mengen von Kanten die nach k Iterationen vorliegt, ist zu einem MST von G zu ergänzen. Am Ende gilt dann die Behauptung. **Basis:** die erste markierte Kante hat ein minimales Gewicht. Sie kann zu einem MST ergänzt werden. **Induktionsschritt:** sei B_k eine MST von G , die die erste k gewählte Kanten $\{e_1, \dots, e_k\}$ enthält. Betrachte die $(k+1)$. Kante e_{k+1} . Falls $e_{k+1} \in B_k$, folgt die Behauptung. Sei $e_{k+1} \notin B_k$. Das Hinzufügen von e_{k+1} zu B_k führt zu einem Kreis, da B_k ein Baum ist. Der (eindeutige!) Kreis in B_k enthält mindestens eine Kante $e \notin B$, da B ein Baum ist. Die Entfernung von e aus B_k liefert wieder ein Baum. Es folgt $W(e) \geq W(e_{k+1})$, ja sonst wäre e vom Algorithmus vor e_{k+1} gewählt worden. Damit folgt: $W(B_{k+1}) \leq W(B_k)$. Da B_k minimal ist (I.V.), ist B_{k+1} minimal. Somit ist B_{k+1} ein MST der die Kanten $\{e_1, \dots, e_{k+1}\}$ enthält. \square

Korrektheit des Algorithmus von Kruskal

Der Algorithmus von Kruskal bestimmt einen minimalen Spannbaum.

Die Beweisidee beruht auf den **Kantentausch für Kreise**: man kann aus einem Spannbaum B einen anderen Spannbaum B' konstruieren, indem man eine neue Kante hinzufügt, die—da B ein Baum ist—einen Kreis schließt, und anschließend eine andere Kante aus diesem Kreis löscht.

Beispiel: Schnitt



Der Schnitt (rote Kanten) für den bereits konstruierten Teilbaum T (dicke Kanten) besteht aus allen Kanten (u, v) mit $u \in T$ und $v \notin T$.

Korrektheit des Algorithmus von Prim

Der Algorithmus von Prim bestimmt einen minimalen Spannbaum.

Beweis.

Widerspruchsbeweis. Sei G einen Graph wofür den Algorithmus einen nicht-minimalen Spannbaum konstruiert. Betrachte die Iteration wobei zuerst eine zu teure Kante e , der "kleinste Verbrecher", gewählt wurde. Betrachte der zuvor konstruierten Teilbaum T . (Also $e \notin T$.) Der Schnitt von T enthält alle Kanten $(u, v) \in G$ mit $u \in T$ and $v \notin T$. Die Kante e gehört zum Schnitt. Der MST B_T die T enthält, enthält eine Kante e_{B_T} aus dem Schnitt, sonst ist sie nicht zusammenhängend. Es gilt $W(e_{B_T}) \geq W(e)$ da sonst der Algorithmus e nicht gewählt hätte. Andererseits, $W(e_{B_T}) \neq W(e)$, denn sonst könnte man e durch e_{B_T} ersetzen und eine MST konstruieren die günstiger ist als B_T . Da B_T ein MST ist, kann das nicht sein. Damit gilt $W(e_{B_T}) = W(e)$. Da e_{B_T} und e gleich teuer sind, liefert den Austausch von e_{B_T} durch e ein MST. Also war der Wahl des "kleinsten Verbrechers" e nicht falsch. Widerspruch. \square

Der Algorithmus von Prim – Übersicht

Wir ordnen die Knoten in drei Kategorien (BLACK, GRAY, WHITE) ein:

Baum-knoten: Knoten, die Teil vom bis jetzt konstruierten Baum sind.

Rand-knoten: Nicht im Baum, jedoch adjazent zu Knoten im Baum.

Ungesehene Knoten: Alle anderen Knoten.

Grundkonzept:

- ▶ Fange mit einem Baum aus nur einem Knoten an, indem ein beliebiger Knoten des Graphens ausgewählt wird.
- ▶ Finde die günstigste Kante (d. h. mit minimalem Gewicht), die den bisherigen Baum verlässt.
- ▶ Füge den über diese Kante erreichten (Rand-)Knoten dem Baum hinzu, zusammen mit der Kante.
- ▶ Fahre fort, bis keine weiteren Randknoten mehr vorhanden sind.

Übersicht

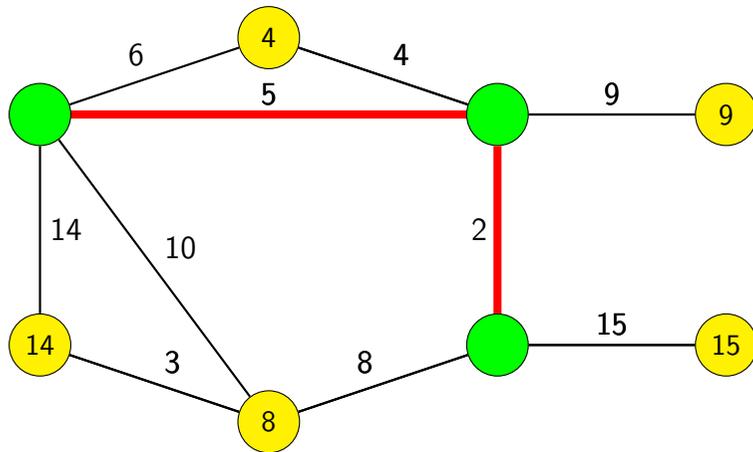
- 1 Spannäume
- 2 Minimale Spannäume
- 3 Greedy Algorithmen
- 4 Die Algorithmen von Kruskal und Prim
- 5 Implementierung und Komplexität

Prim's Algorithmus – Grundgerüst

```

1 // ungerichteter Graph G mit n Knoten
2 void primMST(Graph G, int n) {
3     initialisiere alle Knoten als ungesehen (WHITE);
4     wähle irgendeinen Knoten s und markiere ihn mit Baum (BLACK);
5     reklassifiziere alle zu s adjazenten Knoten als Rand (GRAY);
6     while (es gibt Randknoten) {
7         wähle von allen Kanten zwischen einem Baumknoten t und
8             einem Randknoten v die billigste;
9         reklassifiziere v als Baum (BLACK);
10        füge Kante (t, v) zum Baum hinzu;
11        reklassifiziere alle zu v adjazenten ungesehenen Knoten
12            als Rand (GRAY);
13    }
14 }
```

Prim's Algorithmus – Beispiel



Vorläufige Komplexitätsanalyse

Im Worst-Case:

- ▶ Jeder Knoten muss zur Prioritätswarteschlange hinzugefügt werden.
- ▶ Auf jeden Knoten muss auch wieder zugegriffen werden und er muss gelöscht werden.
- ▶ Die Priorität eines Randknotens muss nach jeder gefundenen Kante angepasst werden.

Bei einem Graph mit n Knoten und m Kanten ergibt sich:

$$T(n, m) \in O(n \cdot T(\text{insert}) + n \cdot T(\text{getMin}) + n \cdot T(\text{delMin}) + m \cdot T(\text{decrKey}))$$

wobei $T(n, m)$ die Zeitkomplexität von Prim's Algorithmus ist

Nächster Schritt:

Wähle eine geeignete Implementierung der Prioritätswarteschlange

ADT zum Vorhalten der Randknoten (I)

Die benötigten Operationen für den Algorithmus von Prim sind:

- ▶ Wähle eine billigste Kante zu einem Randknoten (Kantenkandidat).
- ▶ Reklassifiziere einen Randknoten als Baumknoten (füge den Kantenkandidat zum Baum hinzu).
- ▶ Ändere die Kosten (Randgewicht) eines Randknotens, wenn ein günstigerer Kantenkandidat gefunden wird.

Idee: Ordne die Randknoten nach ihrer Priorität (= Randgewicht).

Prioritätswarteschlange (priority queue)

- ▶ `PriorityQueue pq;`
- ▶ `pq.insert(int e, int k), int pq.getMin(), pq.delMin()`
- ▶ `void pq.decrKey(int e, int k)` setzt den Schlüssel von Element e auf k ; k muss kleiner als der bisherige Schlüssel von e sein.

⇒ Wir entscheiden uns für die Prioritätswarteschlange als Datenstruktur

Drei Prioritätswarteschlangenimplementierungen

$$T(n, m) \in O(n \cdot T(\text{insert}) + n \cdot T(\text{getMin}) + n \cdot T(\text{delMin}) + m \cdot T(\text{decrKey}))$$

Operation	Implementierung		
	unsortiertes Array	sortiertes Array	Heap
<code>pq.isEmpty()</code>	$\Theta(1)$	$\Theta(1)$	$\Theta(1)$
<code>pq.insert(e, k)</code>	$\Theta(1)$	$\Theta(n)$	$\Theta(\log n)$
<code>pq.getMin()</code>	$\Theta(n)$	$\Theta(1)$	$\Theta(1)$
<code>pq.delMin()</code>	$\Theta(n)$	$\Theta(1)$	$\Theta(\log n)$
<code>pq.getElt(k)</code>	$\Theta(n)$	$\Theta(\log n)$	$\Theta(n)$
<code>pq.decrKey(e, k)</code>	$\Theta(1)$	$\Theta(n)$	$\Theta(\log n)$
Prim	$O(n^2 + m)$	$O(n^2 + m \cdot n)$	$O(n \log n + m \log n)$

- ▶ Wir ergänzen außerdem noch zwei Operationen:

Eine vierte Prioritätswarteschlange

Prioritätswarteschlange

- ▶ `bool pq.isEmpty()` $\Theta(n)$
- ▶ `void pq.insert(int elem, VertexState &key)` $\Theta(1)$
- ▶ `float pq.getMin()` $\Theta(n)$
- ▶ `void pq.delMin()` $\Theta(n)$
- ▶ `void pq.decrKey(int elem, VertexState &newkey)` setzt den Schlüssel von `elem` auf `newkey`; `newkey.curWeight` muss kleiner als beim bisherigen Schlüssel von `elem` sein. $\Theta(1)$
- ▶ `int pq.getColor(int elem)` gibt `color` von `elem` zurück. `elem` muss dazu *nicht* in der Warteschlange sein. $\Theta(1)$
- ▶ `float pq.getWeight(int elem)` gibt `curWeight` von `elem` zurück. `elem` muss dazu *nicht* in der Warteschlange sein. $\Theta(1)$

Der Algorithmus von Prim – Implementierung

```

1 void updateFringe(PriorityQueue &pq, List adjLst[], int v) {
2   foreach (edge in adjLst[v]) {
3     // berechnet MST.
4     float newWeight = edge.weight;
5
6     if (pq.getColor(edge.w) == WHITE) { // -> GRAY
7       pq.insert(edge.w, {parent: v, curWeight: newWeight});
8     } else if (pq.getColor(edge.w) == GRAY) {
9       if (newWeight < pq.getWeight(edge.w)) {
10        // Randknoten-update: Kante von v aus ist besser
11        pq.decrKey(edge.w, {parent: v, curWeight: newWeight});
12      }
13    }
14  }
15 }

```

Der Algorithmus von Prim – Implementierung

```

1 // Ergebnis als Vorgängerbaum in .parent:
2 //  $\forall v \in V: (x, v) \in MST(V, E)$  gdw.  $x = state[v].parent, x \neq -1$ 
3 VertexState[n] primMST(List adjLst[n], int n, int start) {
4   VertexState state[n] = // (eigentlich im Konstruktor von pq)
5     { color: WHITE, parent: -1, curWeight: +inf };
6   PriorityQueue pq = VS_PriorityQueue<&VS.curWeight>(&state);
7
8   pq.insert(start, {parent: -1, curWeight: 0});
9   while (!pq.isEmpty()) { // solange es Randknoten gibt
10    int v = pq.getMin(); // günstigste Kante, bzw. Randknoten
11    pq.delMin(); // setzt auch Farbe auf BLACK
12    updateFringe(pq, adjLst, v); // update den Rand
13  }
14  return state;
15 }

```

Komplexitätsanalyse

$$T(n, m) \in O(n \cdot T(\text{insert}) + n \cdot T(\text{getMin}) + n \cdot T(\text{delMin}) + m \cdot T(\text{decrKey}))$$

- ▶ Die Schleife in `primMST` wird n mal ausgeführt.
 - ⇒ `isEmpty`, `getMin`, `delMin` und `updateFringe` wird n mal ausgeführt.
 - ▶ Beachte, dass `getMin` eine Komplexität von $\Theta(n)$ hat.
 - ▶ Die Schleife in `updateFringe` wird etwa $2m$ mal durchlaufen.
 - ⇒ `insert`, `getColor`, `getWeight` und `decrKey` werden m mal ausgeführt und sind $\Theta(1)$.
 - ▶ Der zusätzliche Speicherbedarf ist $\Theta(n)$.
 - ▶ Die untere Schranke der Zeitkomplexität ist $\Omega(m)$, da jede Kante des Graphen untersucht werden muss, um einen MST zu konstruieren.
- ⇒ Insgesamt: Worst-Case-Komplexität $O(n^2 + m) = O(n^2)$.