Datenstrukturen, Algorithmen und Programmierung 2

Amin Coja-Oghlan

May 30, 2023

Lehrstuhl Informatik 2
Fakultät für Informatik

Bäume, Wälder, Breiten- und Tiefensuche

Worum geht es?

- Bäume sind eine wichtige Klasse von Graphen
- einige Probleme, die auf allgemeinen Graphen “schwer” sind, sind auf Bäumen “leicht”
- Zusammenhang ist eine fundamentale Grapheigenschaft
- jeder zusammenhängende Graph enthält einen spannenden Baum

Bäume, Wälder, Breiten- und Tiefensuche

Definition **[Wege, Pfade]**

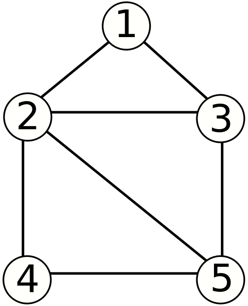
Betrachte einen Graphen $G = (V, E)$.

- ein Weg in G ist eine Folge

$$v_1, \dots, v_\ell$$

  von Knoten, so dass $v_i v_{i+1} \in E$ für $i = 1, \dots, \ell - 1$.
- dieser Weg hat Länge $\ell - 1$.
- wir sprechen von einem Weg von v_1 nach v_ℓ.
- ein *Pfad* in G ist ein Weg $v_1, \dots, v_\ell$, dessen Knoten alle verschieden sind.

Bäume, Wälder, Breiten- und Tiefensuche



Beispiel

- die Knotenfolge $1, 2, 3, 5, 2, 3$ ist ein Weg aber kein Pfad, weil 2 und 3 zweimal vorkommen.
- die Folge $1, 2, 3, 5, 4$ ist ein Pfad

Bäume, Wälder, Breiten- und Tiefensuche

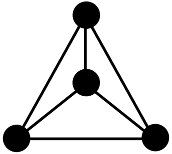
Definition **[Zusammenhangskomponenten]**

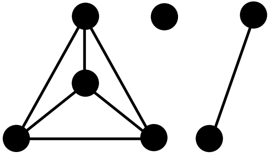
- für einen Graphen $G = (V, E)$ und $u, v \in V$ schreibe $u \sim_G v$, falls es einen Weg von u nach v gibt.
- die Relation $\sim_G$ ist eine Äquivalenzrelation, d.h.

$$
\begin{aligned}
&u \sim_G u && \text{für alle } u \in V \\
&u \sim_G v \iff v \sim_G u && \text{für alle } u, v \in V \\
&u \sim_G v \text{ und } v \sim_G w \implies u \sim_G w && \text{für alle } u, v, w \in V
\end{aligned}
$$

- die Äquivalenzklassen nennen wir Zusammenhangskomponenten von G
- zwei Knoten liegen also genau dann in derselben Zusammenhangskomponente, wenn es einen Weg vom einen zum anderen gibt

Bäume, Wälder, Breiten- und Tiefensuche

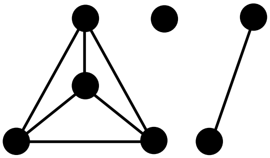




Zusammenhangskomponenten

- ein Graph ist zusammenhängend, wenn er nur eine Zusammenhangskomponente hat
- der linke Graph ist also zusammenhängend
- der rechte Graph ist unzusammenhängend
- der rechte Graph besteht aus drei Zusammenhangskomponenten

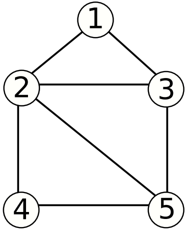
Bäume, Wälder, Breiten- und Tiefensuche

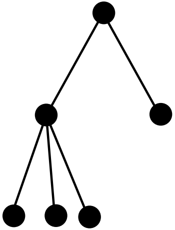


Zusammenhangskomponenten

- ein isolierter Knoten in einem Graphen ist eine Zusammenhangskomponente, die nur aus einem Knoten besteht
- eine isolierte Kante ist eine Zusammenhangskomponente, die nur eine Kante enthält

Bäume, Wälder, Breiten- und Tiefensuche

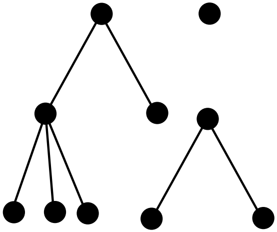


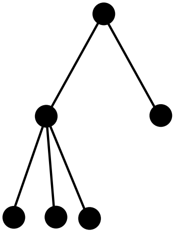


Kreise in Graphen

- ein Kreis in einem Graphen G ist eine Kopie eines Kreises C_ℓ, $\ell \geq 3$, die in G enthalten ist
- der linke Graph enthält also drei Kreise der Längen 3,4,5
- ein Graph ist kreisfrei, wenn er keine Kreise enthält
- der rechte Graph ist kreisfrei

Bäume, Wälder, Breiten- und Tiefensuche





Definition

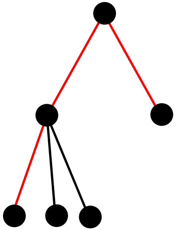
- ein kreisfreier Graph ist ein Wald
- ein zusammenhänender Wald ist ein Baum

Bäume, Wälder, Breiten- und Tiefensuche

Bemerkungen

- der kleinste Wald besteht nur aus einem Knoten
- der Graph, der nur aus einer Kante besteht, ist ebenfalls ein Baum
- jeder Pfad ist ein Baum
- ein Blatt in einem Wald ist ein Knoten vom Grad 1

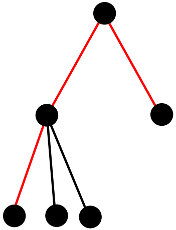
Bäume, Wälder, Breiten- und Tiefensuche



Lemma

Jeder Baum $G = (V, E)$ mit $E \neq \emptyset$ enthält mindestens zwei Blätter.

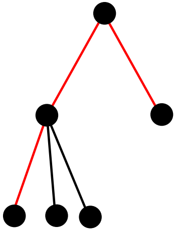
Bäume, Wälder, Breiten- und Tiefensuche



Beweis

- betrachte einen längsten Pfad $p = v_1, \ldots, v_\ell$.
- alle Nachbarn von v_1 und v_ℓ liegen auf p, weil man den Pfad sonst verlängern könnte.
- wenn v_1 zwei Nachbarn hätte, würden beide auf dem Pfad liegen; aber dann enthielte G einen Kreis

Bäume, Wälder, Breiten- und Tiefensuche



Beweis (fortgesetzt)

- genauso für v_ℓ
- also sind v_1 und v_ℓ Blätter

Bäume, Wälder, Breiten- und Tiefensuche

Proposition

Die folgenden Aussagen sind äquivalent.

1. der Graph $G = (V, E)$ ist ein Baum
2. G ist zusammenhängend und $|E| = |V| - 1$
3. G ist kreisfrei und $|E| = |V| - 1$
4. G ist kantenmaximal kreisfrei
5. G ist kantenminimal zusammenhängend
6. in G gibt es zu je zwei Knoten v, w genau einen Pfad von v nach w

Bäume, Wälder, Breiten- und Tiefensuche

Beweis $[1 \Rightarrow 2]$

- Induktion nach $n = |V|$
- für $n = 1, 2$ ist nichts zu zeigen
- für $n > 2$ finde ein Blatt v in G
- der Graph $G - \{v\}$ ist ein Baum
- nach Induktion gilt

$$|E| - 1 = |E(G - \{v\})| = |V(G - \{v\})| - 1 = |V| - 2$$

- folglich $|E| = |V| - 1$
- per Definition ist G zusammenhängend

Bäume, Wälder, Breiten- und Tiefensuche

Beweis [2 $\Rightarrow$ 3]

- Induktion nach $n = |V|$
- für $n = 1, 2$ ist nichts zu zeigen
- weil

$$\sum_{v \in V} d_G(V) = 2|E| = 2(n-1),$$

  gibt es einen Knoten $v \in V$ vom Grad 1
- der Graph $G' = G - \{v\}$ ist zusammenhängend und $|E(G')| = |V(G')| - 1$
- also ist G' nach Induktion kreisfrei
- folglich ist auch G kreisfrei
- nach Annahme gilt $|V(G)| = |E(G)| - 1$

Bäume, Wälder, Breiten- und Tiefensuche

Beweis $[3 \Rightarrow 4]$

- Induktion nach $n = |V|$; für $n = 1, 2$ ist nichts zu zeigen
- weil G kreisfrei ist, sind Anfangs- und Endknoten eines längsten Pfades Blätter
- G enthält also ein Blatt v
- der Graph $G' = G - \{v\}$ ist kreisfrei und $|E(G')| = |V(G')| - 1$
- also ist G' nach Induktion maximal kreisfrei
- somit ist G' zusammenhängend; denn sonst könnte man eine Kante zwischen zwei Komponenten von G' hinzufügen
- daher kann man keine mit v inzidente Kante hinzufügen, ohne einen Kreis zu schließen

Bäume, Wälder, Breiten- und Tiefensuche

Beweis $[4 \Rightarrow 5]$

- weil G kantenmaximal kreisfrei ist, ist G zusammenhängend
- angenommen man könnte eine Kante $e = vw$ löschen, ohne den Zusammenhang zu zerstören
- dann liegen v, w in derselben Komponente von $G' = G - \{e\}$
- also schließt e einen Kreis, Widerspruch

Bäume, Wälder, Breiten- und Tiefensuche

Beweis **[$5 \Rightarrow 6$]**

- weil G zusammenhängend ist, gibt es für alle $v, w \in V$ es einen Pfad von v nach w
- angenommen zu v, w gibt es zwei verschiedene Pfade $p \neq p'$
- dann gibt es eine Kante e auf p, die nicht auf p' vorkommt
- also ist $G - \{e\}$ zusammenhängend
- Widerspruch zur Annahme, G sei kantenminimal zusammenhängend

Bäume, Wälder, Breiten- und Tiefensuche

Beweis [$6 \Rightarrow 1$]

- Induktion nach $n = |V|$
- für $n = 1, 2$ ist nichts zu zeigen; sei also $n \geq 3$
- betrachte einen längsten Pfad $p = v_1 \cdots v_\ell$ in G
- alle Nachbarn von v_1 und v_ℓ liegen auf p
- weil p der einzige Pfad von v_1 nach v_ℓ ist, folgt, daß v_1, v_ℓ Blätter sind
- weil in $G' = G - \{v_1\}$ je zwei Knoten durch genau einen Pfad verbunden sind, ist G' ein Baum
- also ist auch G ein Baum

Bäume, Wälder, Breiten- und Tiefensuche

Definition

Ein spannender Baum eines Graphen $G = (V, E)$ ist ein Untergraph $G' = (V, E')$ von G mit derselben Knotenmenge wie G, der ein Baum ist.

Bäume, Wälder, Breiten- und Tiefensuche

Lemma

Jeder zusammenhängende Graph enthält einen spannenden Baum.

Beweis

- Induktion nach der Kantenzahl von $G = (V, E)$
- wenn G kantenminimal zusammenhängend ist, ist G bereits selbst ein Baum
- sonst gibt es eine Kante $e \in E$, so daß $G' = G - \{e\}$ zusammenhängend ist
- nach Induktion besitzt G' einen spannenden Baum

Bäume, Wälder, Breiten- und Tiefensuche

Breiten- und Tiefensuche

- wir lernen zwei Algorithmen kennen, die in einem zusammenhängenden Graphen spannende Bäume bestimmen
- darüber hinaus bestimmen diese Algorithmen die Zusammenhangskomponenten des Eingabegraphen
- die Eingabe ist jeweils ein Graph, der als Adjazenzliste gegeben ist

Bäume, Wälder, Breiten- und Tiefensuche

Breiten- und Tiefensuche

- in einem Graphen G definieren wir den Abstand von $v, w \in V(G)$ als

$$\mathrm{dist}_G(v, w) = \min_{\ell \geq 0} \exists \text{ Weg der Länge } \ell \text{ von } v \text{ nach } w$$

- falls v, w in verschiedenen Zusammenhangskomponenten liegen, verwenden wir die Konvention

$$\mathrm{dist}_G(v, w) = \infty$$

Bäume, Wälder, Breiten- und Tiefensuche

$\mathrm{BFS}(G, s)$

1. Färbe alle Knoten $v \in V(G) \setminus \{s\}$ grün und färbe s gelb.
2. Setze $d(v) = \infty$ für alle $v \in V(G) \setminus \{s\}$ und setze $d(s) = 0$.
3. Setze $p(v) = \emptyset$ für alle $v \in V$.
4. Lege eine Warteschlange Q an und füge s in Q ein.
5. Solange Q nicht leer ist,
6. entnehme v aus Q
7. färbe v rot
8. für alle $u \in \partial v$ mit Farbe grün
9. färbe u gelb, setze $p(u) = v$, $d(u) = d(v) + 1$ und füge u in Q ein

Bäume, Wälder, Breiten- und Tiefensuche

Satz

BFS hat Laufzeit $O(|V(G)| + |E(G)|)$. Bei Beendigung des Algorithmus' gilt folgendes.

(i) Die Zusammehangskomponente des Startknotens s besteht aus genau den Knoten v, für die $d(v) < \infty$.

(ii) Für alle $v \in V(G)$ gilt $d(v) = \mathrm{dist}_G(s, v)$.

(iii) Der Untergraph

$$(\{v \in V(G) : d(v) < \infty\}, \{\{v, p(v)\} : v \in V(G),\ p(v) \neq \emptyset\})$$

ist ein spannender Baum der Zusammehangskomponente von s in G.

Bäume, Wälder, Breiten- und Tiefensuche

Handschlaglemma

Für jeden Graphen G gilt

$$\sum_{v \in V(G)} d_v(G) = 2|E(G)|.$$

Bäume, Wälder, Breiten- und Tiefensuche

Lemma 1

BFS hat Laufzeit $O(|V(G)| + |E(G)|)$.

Berweis

- Laufzeit der Schritte (1)–(4) beträgt $O(|V(G|)$.
- Jeder Knoten wechselt höchstens zweimal seine Farbe, und zwar von grün auf gelb auf rot.
- Deshalb wird kein Knoten mehrmals in Q eingefügt, und folglich auch nicht wieder entnommen.
- Aus diesem Grund ist die Laufzeit der Hauptschleife (5)–(9) beschränkt durch $O(|V(G)|) + \sum_{v \in V(G)} d_G(v) = O(|V(G)|) + O(|E(G)|)$

Bäume, Wälder, Breiten- und Tiefensuche

Lemma 2

Während der gesamten Ausführung von $\mathtt{BFS}$ gilt

$$d(v) \geq \mathrm{dist}_G(s, v) \qquad \text{für alle Knoten } v.$$

Bäume, Wälder, Breiten- und Tiefensuche

Beweis

- Wir führen Induktion nach der Zahl der Iterationen der Hauptschleife (5)–(9).
- Anfangs gilt $d(v) = \infty$ für alle $v \in V(G) \setminus \{s\}$, weshalb die Behauptung trivial erfüllt ist.
- Betrachte nun den nächsten Knoten v, der aus Q entnommen wird, sowie einen grünen Nachbarn u von v
- Dann gilt $d(u) = d(v) + 1$.
- Andererseits gilt offensichtlich $\mathrm{dist}_G(s, u) \leq \mathrm{dist}_G(s, v) + 1$, weil ein Pfad von s nach v um eine Kante nach u verlängert werden kann.
- Weil $d(v) \geq \mathrm{dist}_G(s, v)$ nach Induktionsvoraussetzung, schließen wir, daß $d(u) \geq \mathrm{dist}_G(s, u)$ für alle grünen $u \in \partial_G v$.

Bäume, Wälder, Breiten- und Tiefensuche

Lemma 3

Angenommen die Warteschlange Q enthält die Knoten $q_1, \ldots, q_\ell$. Dann gilt

$$d(q_1) \leq \cdots \leq d(q_\ell) \leq d(q_1) + 1.$$

Wird ferner ein Knoten u vor einem anderen Knoten u' in Q eingefügt, so gilt

$$d(u) \leq d(u').$$

Bäume, Wälder, Breiten- und Tiefensuche

Beweis

- Wir führen Induktion nach der Zahl der Iterationen der Hauptschleife.
- Zu Beginn enthält Q nur den Knoten s, so daß nichts zu zeigen ist.
- Es ist klar, daß die Ungleichung nach einer Ausführung von Schritt (6) erhalten bleibt.
- Betrachte ferner eine Ausführung von Schritt (9).
- Dann gilt nach Induktionsvoraussetzung für den neu eingefügten Knoten v, daß $d(v) = d(v_1) + 1 \geq d(v_2)$.
- Also bleibt die Ungleichung erhalten.
- Die zweite Behauptung folgt unmittelbar aus der Ungleichung.

Bäume, Wälder, Breiten- und Tiefensuche

Beweis

- Die behauptete Laufzeit ergibt sich unmittelbar aus Lemma 1
- Wir zeigen (ii), woraus (i) direkt folgt
- Nach Lemma 2 genügt zu zeigen, daß $d(u) \leq \text{dist}_G(s, u)$ für alle u bei Beendigung von BFS
- *Beweis durch Widerspruch:* nimm an, es gäbe ein u mit

$$d(u) > \text{dist}_G(s, u)$$

Bäume, Wälder, Breiten- und Tiefensuche

Beweis

- Wähle $u \neq s$ so, daß $\mathrm{dist}_G(s, u)$ kleinstmöglich
- Da $d(u) > \mathrm{dist}_G(s, u)$ gilt $\mathrm{dist}_G(s, u) < \infty$
- Also gibt es einen kürzesten Pfad P von s nach u
- Sei v der letzte vor u auf P; dann $\mathrm{dist}_G(u) = \mathrm{dist}_G(v) + 1$
- Da $\mathrm{dist}_G(s, v) < \mathrm{dist}_G(s, u)$, folgt $d(v) = \mathrm{dist}_G(v)$
- In Schritten (8)–(12) bei Entnahme von v muß u gelb oder rot gefärbt sein
- Daher zeigt Lemma 3, daß $d(u) \leq d(v) + 1 \leq \mathrm{dist}_G(u)$
- Dies ist ein Widerspruch, also ist (ii) bewiesen
- Aussage (iii) ergibt sich direkt aus (i) und (ii)

Bäume, Wälder, Breiten- und Tiefensuche

$\texttt{DFS}(G)$

1. Färbe alle Knoten $v \in V(G)$ grün.
2. Setze $c(v) = 0$ und $p(v) = \emptyset$ für alle $v \in V$.
3. Setze $j = 1$
4. Für alle $v \in V(G)$
5.   falls v grün gefärbt ist
6.   führe $\texttt{DFSLoop}(G, v, j)$ aus.
7.   Erhöhe j um 1.

Bäume, Wälder, Breiten- und Tiefensuche

$\texttt{DFSLoop}(G, v, j)$

1. Färbe v gelb und setze $c(v) = j$.
2. Für alle $u \in \partial_G v$
3. Falls u grün ist
4. Setze $p(u) = v$.
5. Führe $\texttt{DFSLoop}(G, u, j)$ aus.
6. Färbe v rot.

Bäume, Wälder, Breiten- und Tiefensuche

Satz

- `DFS` hat Laufzeit $O(|V(G)| + |E(G)|)$.
- Die Mengen $c^{-1}(j)$ für $j \geq 1$ bilden genau die Zusammehangskomponenten von G.

Bäume, Wälder, Breiten- und Tiefensuche

Zusammenfassung

- der Zusammenhangsbegriff
- Bäume und Wälder
- Breitensuche
- Tiefensuche