Mathematische Grundlagen der Informatik

Graphentheorie

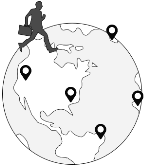
W. Gansterer, K. Schindlerová

3. Mai 2020

Graphentheorie

Herausforderungen im Alltag der Informatik

- Routing im Internet
- Reiserouten planen
- Kommunikationsnetzwerke
- Compiler und Syntaxbäume
- ...

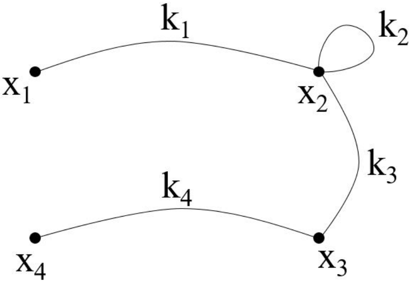


Inhaltsverzeichnis I

Ungerichtete Graphen

Definition

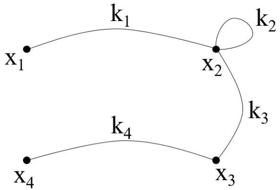
Ein (ungerichteter) Graph G besteht aus einer Menge $V = V(G)$, den Knoten von G, und aus einer Menge $E = E(G)$ von ungeordneten Paaren k=[x,y] mit $x, y \in V$, den Kanten von G. Man schreibt $G = (V, E)$.



$$V = \{x_1, x_2, x_3, x_4\}$$
$$E = \{k_1, k_2, k_3, k_4\}$$
$$k_1 = [x_1, x_2]$$
$$k_2 = [x_2, x_2]$$
$$k_3 = [x_2, x_3]$$
$$k_4 = [x_3, x_4]$$

Grundbegriffe

- Ist $k = [x, y]$ eine Kante, so heißen x und y **Endpunkte** der Kante
- x und y sind **inzident** zu k
- x und y sind **adjazent** (benachbart)
- Die Kante $k = [x, y]$ heißt **Schlinge**, wenn $x = y$
- $G' = (V', E')$ heißt **Teilgraph** von G, wenn $V' \subset V$ und $E' \subset E$
- G heißt **schlichter Graph**, wenn er weder Schlingen noch Mehrfachkanten hat (Den Zusatz „ohne Mehrfachkanten“ lässt man gewöhnlich weg und nennt Graphen mit Mehrfachkanten Multigraphen)

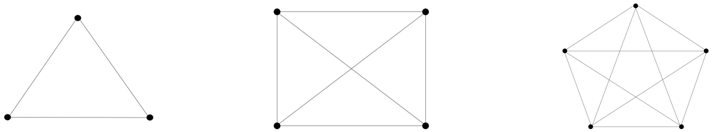


x_1 und x_2 sind adjazent
x_3 und x_4 sind Endpunkte von k_4
x_3 und x_4 sind inzident zu k_4
k_2 ist eine Schlinge

Vollständige Graphen

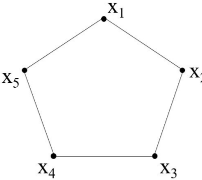
G heißt **vollständiger Graph**, wenn G ein schlichter Graph ist, in dem jedes Knotenpaar durch genau eine Kante verbunden ist.

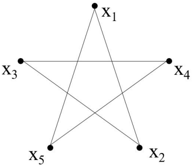
Ein vollständiger Graph mit n Knoten hat $\binom{n}{2} = \frac{n(n-1)}{2}$ Kanten



Graphische Darstellung von Graphen

Zu jedem Graphen gibt es unendlich viele graphische Darstellungen.





Schneiden sich die Kanten nur in den Knoten, so bezeichnet man die Darstellung als **geometrischer Graph**.

Jeder endliche Graph besitzt eine Realisierung als geometrischer Graph im $\mathbb{R}^3$. Graphen, die sich sogar im $\mathbb{R}^2$ überschneidungsfrei darstellen lassen, heißen **planare Graphen**.

Anwendung: Entwurf von Leiterplatten oder integrierten Schaltungen: Stromverbindungen so planen, dass sie möglichst überschneidungsfrei liegen. Im $\mathbb{R}^2$ nicht immer möglich $\Rightarrow$ oft mehrere Ebenen

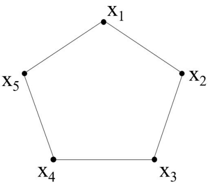
Isomorphe Graphen

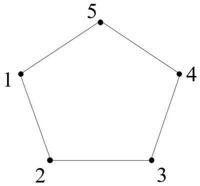
Zwei Graphen werden als „gleich“ angesehen, wenn sie sich nur in der Bezeichnung ihrer Knoten unterscheiden:

Definition

Zwei Graphen $G = (V, E)$ und $G' = (V', E')$ heißen isomorph, wenn es eine bijektive Abbildung $\phi : V \to V'$ gibt, sodass gilt:

$$[x, y] \in E \Leftrightarrow [\phi(x), \phi(y)] \in E'$$

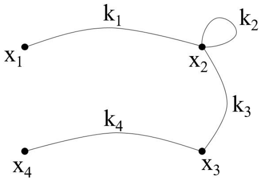




Darstellung - Adjazenzmatrix

Darstellung in Matrizenform (leicht für Computer zu verarbeiten!)

- Die Knoten werden durchnummeriert, jede Zeile/Spalte entspricht einem Knoten
- Der Eintrag der Matrix an der Stelle a_{ij} entspricht der Anzahl der Kanten zwischen den Knoten i und j. Sind die Knoten nicht adjazent, hat der Eintrag den Wert 0.
- Befindet sich im k-ten Knoten eine Schlinge, so wird an der entsprechenden Stelle auf der Diagonale eine 1 eingetragen.
- Bei *ungerichteten* Graphen ist die Adjazenzmatrix stets symmetrisch.



$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Grad

Definition

Ist x ein Knoten des Graphen G, so heißt die Anzahl der mit x inzidenten Kanten Grad von x. Dabei zählen Schlingen doppelt. Der Grad von x wird mit $d(x)$ bezeichnet.

Satz

In jedem Graphen $G = (V, E)$ gilt $\sum d(x) = 2 \cdot |E|$ ... Summe der Grade ist die doppelte Anzahl der Kanten im Graphen, sie ist also immer gerade.

Beweis: Jeder Endpunkt einer Kante liefert genau den Beitrag 1 zum Grad dieses Punktes. Jede Kante hat genau 2 Endpunkte $\Rightarrow$ sie liefert zu der linken Seite der Gleichung genau zweimal den Betrag 1.

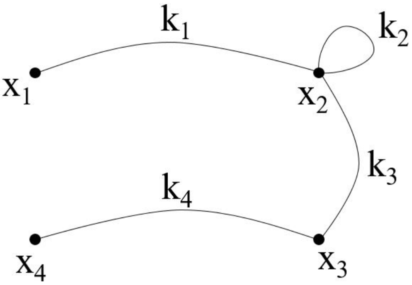
Folgerung: In jedem Graphen ist die Anzahl der Knoten mit ungeradem Grad gerade.

Beachten Sie:

Der Grad des Knotens x_n entspricht der Summe der Einträge der n-ten Zeile (bzw. Spalte) der dazugehörigen Adjazenzmatrix.

(*Achtung:* Schlingen doppelt zählen!)

Beispiel: Grad



$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$d(x_1) = 1$
$d(x_2) = 4$
$d(x_3) = 2$
$d(x_4) = 1$

Grad - Aus dem Sicht der Lineare Algebra

A die Adjazenzmatrix eines Graphen mit den Knoten $x_1, x_2, \ldots, x_n$.
Die i-te Zeile: $(a_{i1}, a_{i2}, \ldots, a_{in})$.

$a_{ij} = 1 \iff$ es eine Kante vom Knoten i zum Knoten j gibt.

Daher ist der Grad des Knotens i genau die Summe der 1er in der i-ten Zeile der Matrix (Schlingen doppelt zählen!).

In der i-ten Zeile erhaltet man den Wert $(a_{i1} + a_{i2} + \cdots + a_{in})$, es ist also

$$A \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} deg(x_1) \\ deg(x_2) \\ \vdots \\ deg(x_n) \end{pmatrix}.$$

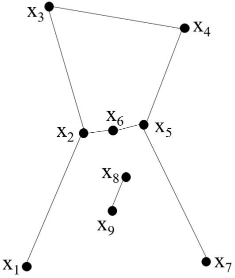
Wege in Graphen

- Eine Menge von Kanten, die den Knoten x_1 mit x_n verbindet, heißt **Kantenfolge** von x_1 nach x_n
- Eine Kantenfolge heißt **offen**, wenn $x_1 \neq x_n$, und **geschlossen**, wenn $x_1 = x_n$

- Ein **Weg** von x nach y ist eine offene Kantenfolge, in der alle Knoten verschieden sind
- Ein **Kreis** ist eine geschlossene Kantenfolge, in der bis auf Anfangs- und Endknoten alle Knoten verschieden sind

- Der Knoten y heißt **erreichbar** vom Knoten x, wenn es einen Weg von x nach y gibt.
- Ein (ungerichteter) Graph heißt **zusammenhängend**, wenn jeder Knoten von jedem erreichbar ist.

Beispiele



- $x_1x_2x_3x_4x_5x_6x_2x_1$ ist eine geschlossene Kantenfolge
- $x_1x_2x_6x_5x_7$ ist ein Weg
- $x_2x_3x_4x_5x_6x_2$ ist ein Kreis
- Der Graph ist nicht zusammenhängend, die Teilgraphen $x_1, x_2, \dots, x_7$ und x_8, x_9 sind die **Zusammenhangskomponenten** des Graphen

Zentralität von Knoten

Wer ist der Wichtigste?

In einem sozialen Netzwerk bilden Personen die Knoten und Beziehungen zwischen den Personen die Kanten eines Graphen.

- Welche Personen in einem solchen Netzwerk sind wohl die Interessantesten?
- Wer sind die zentralen Personen im Netzwerk, zu denen man selbst Kontakt aufnehmen sollte?

Zu jedem Knoten wollen wir eine Maßzahl zuordnen
- die Zentralität des Knotens.

Zentralität von Knoten

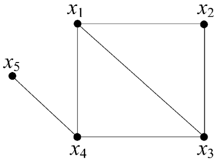
Erst überprüfen wir, wer die meisten Kontakte hat.

Die Zentralität ist dann der Grad des entsprechenden Knotens.

Es ist aber nicht nur die Anzahl der Kontakte interessant, sondern auch die Qualität der Kontakte:

Eine Person ist wichtig, wenn sie viele solche wichtige Personen kennt, die selbst auch wieder viele Kontakte haben.

Beispiel



Die Adjazenzmatrix zu dem Graphen:

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$deg(x_1) = deg(x_3) = deg(x_4) = 3, deg(x_2) = 2$

- Berücksichtigt man nur die Grade, so wären x_1, x_3, x_4 zentral.
- Bezieht man die Bedeutung der jeweiligen Nachbarn mit ein, so sollte x_4 ein geringeres Gewicht haben.

Zentralität von Knoten

Die Zentralität w_i des Knotes x_i sollte proportional zu der Summe der Zentralität aller benachbarten Knoten x_j sein.

μ .... Proportionalitätsfaktor

$$w_i = \mu \sum_{\substack{\text{alle j für die } x_j \\ \text{benachbart zu } x_i}} w_j = \mu(a_{i1}w_1 + a_{i2}w_2 + \cdots + a_{in}w_n)$$

$$a_{ij} = 1 \iff x_j \text{ benachbart zu } x_i$$

Das Ergebnis ist genau die i-te Komponente des Produkts der Matrix A mit dem Vektor $w = (w_1, w_2, \ldots, w_n)$ (siehe Folie 59, 6. Vorlesung). D.h.

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \mu A \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}.$$

Für $\lambda = 1/\mu$ gilt $Aw = \lambda w$.
D.h. w ist ein Eigenvektor zu dem Eigenwert λ der Adjazenzmatrix A des Graphen.

Gibt es für diese Gleichung überhaupt sinnvolle Lösungen?

Nicht jede Matrix hat Eigenwerte, und hier noch $\forall w_i > 0$.

Satz von Perron-Frobenius

Satz

- *Eine reelle quadratische Matrix ohne negative Einträge hat mindestens einen positiven reellen Eigenwert.*
- *Der größte dieser Eigenwerte besitzt einen Eigenvektor mit nur positiven Komponenten.*

Alle anderen Eigenvektoren haben gemischte Vorzeichen in den Komponenten. Daher ist dieser positive Eigenvektor der einzige sinnvolle Kandidat für die Gewichte der Knoten des Graphen.
Und diesen Eigenvektor gibt es immer.
Seine Komponenten heißen die Eigenvektor-Zentralität der Knoten des Graphen.

Beispiel

Die Eigenwerte berechnen wir mit z.B. Sage [1] oder Matlab. Der größte davon hat den Wert 2,64, und ein dazugehöriger Eigenvektor lautet

$$\begin{pmatrix} 2.99 \\ 2.26 \\ 2.99 \\ 2.64 \\ 1 \end{pmatrix}.$$

Die Knoten x_1 und x_3 sind jetzt wichtiger geworden als x_4. Dieser wurde herabgestuft, weil er mit dem unbedeutenden x_5 in Kontakt steht.

[1] SageMath is a free open-source mathematics software system

Anwendung - Internetsuchdienste

©Wikipedia

Z.B. die Methode PageRank;
Die Methodem benutzen unter anderem Varianten der Eigenvektor-Zentralität (eigenvector centrality), um das Ranking von Dokumenten in den Suchergebnissen zu bestimmen:

Vorne platziert werden Links auf Dokumente, auf die viele Links verweisen; diese Links werden aber selbst auch wieder von vielen anderen Seiten referenziert.

Weitere Maße für Graphen

Nähe (closeness centrality) – Die Entfernung eines Knotens zu allen anderen.

Z.B. der reziproke Wert der Summe aller Distanzen zu anderen Knoten wird genommen, um zu erreichen, dass der Wert umso höher wird, je höher die wahrgenommene Zentralität des Knotens ist.

Ähnliche Maße sind auch für gerichtete Graphen definierbar.

Länge und Gewicht

Definition

Die Anzahl der Kanten einer Kantenfolge oder eines Weges heiße **Länge** der Kantenfolge bzw. des Weges.

Definition

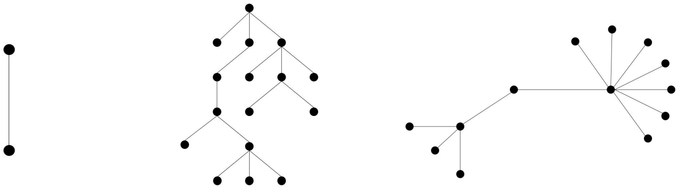
Ein Graph heißt **bewertet/gewichtet**, wenn jeder Kante $[x, y]$ ein **Gewicht** $w(x, y) \in \mathbb{R}$ zugeordnet ist. In bewerteten Graphen ist die *Länge eines Weges* von einem Knoten zu einem anderen die Summe der Gewichte aller Kanten des Weges.

Bäume

Bäume

Definition

Ein Graph, in dem je zwei Knoten durch genau einen Weg verbunden sind, heißt **Baum**. Ein Baum ist also ein zusammenhängender Graph ohne Kreise.



Charakterisierung Baum

Für einen Graphen G mit n Knoten und m Kanten sind die folgenden Aussagen äquivalent:

a) G ist ein Baum (je zwei Knoten durch genau einen Weg verbunden).

b) G ist ein zusammenhängender Graph ohne Kreise.

c) G ist zusammenhängend, aber entfernt man eine beliebige Kante, so zerfällt G in zwei Zusammenhangskomponenten.

d) G ist zusammenhängend und $n = m + 1$.
(G hat einen Knoten mehr als Kanten)

Charakterisierung Baum

Beweis.

- a) $\Rightarrow$ b): G ist natürlich zusammenhängend. Hätte G einen Kreis, so gäbe es zwei verschiedene Wege zwischen den Knoten dieses Kreises.
- b) $\Rightarrow$ c): Nehmen wir eine Kante $E = [x_1, x_2]$ weg, so erhalten wir den Graphen G', in dem x_1 und x_2 nicht mehr verbunden sind, G' ist also nicht mehr zusammenhängend (sonst hätte es in G einen Kreis gegeben).

  Sei y irgendein Knoten. Wenn y in G' nicht mehr mit x_1 verbunden ist, so enthielt der ursprüngliche Weg von y nach x_1 die Kante E und damit auch x_2. Der Weg nach x_2 existiert also noch, y ist mit x_2 verbunden. G' zerfällt also in zwei Komponenten: Die Knoten einer Komponente sind alle mit x_1 verbunden, die Knoten der anderen Komponente alle mit x_2.

■

Charakterisierung Baum

Beweis.

- c) $\Rightarrow$ a): Da G zusammenhängend ist, sind je zwei Knoten durch genau einen Weg verbunden. Gäbe es zwei verschiedene Wege von x_1 nach x_2, so könnte man aus einem dieser Wege eine Kante wegnehmen, ohne den Zusammenhang zu zerstören, was im Widerspruch zur Annahme c) steht.

$\blacksquare$

Charakterisierung Baum

Beweis.

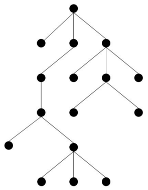
- Skizze für d) $\Leftrightarrow$ a), b), c):
  - Vorbereitung: Nach Wegnahme einer Kante aus einem Baum bilden die entstehenden Zusammenhangskomponenten wieder Bäume: Sind z. B. y und z Knoten in einer solchen Komponente, so kann es nicht mehrere verschiedene Wege von y nach z geben, da diese auch im ursprünglichen Graphen G Wege wären.
  - Zeige a) $\Rightarrow$ d) mit vollständiger Induktion nach der Anzahl der Knoten n.
  - Zeige d) $\Rightarrow$ a) mit vollständiger Induktion nach der Anzahl der Knoten n (unter Beachtung von $\sum_{x \in V} d(x) = 2 \cdot |E|$).

$\blacksquare$

Charakterisierung Baum

Zeige a) $\Rightarrow$ d) mit vollständiger Induktion nach der Anzahl der Knoten n. Der Induktionsanfang ($n = 1$ und $n = 2$) lässt sich aus dem Bild ablesen. Es soll nun für alle Bäume mit weniger als n + 1 Knoten die Behauptung erfüllt sein.

Sei $G = (V, E)$ $|V| = n + 1$. Nehmen wir eine Kante weg, so erhalten wir 2 Teilbäume mit Knotenzahl n_1 bzw. n_2, wobei $n1 + n2 = n + 1$ gilt. Für die Teilbäume ist die Kantenzahl nach Voraussetzung $n_1 - 1$ bzw. $n_2 - 1$. Damit ergibt sich für die Gesamtzahl der Kanten von $G : n_1 - 1 + n_2 - 1 + 1 = n$.



Charakterisierung Baum

Zeige d) $\Rightarrow$ a) auch mit vollständiger Induktion nach der Anzahl der Knoten n.
Induktionsanfang - das Bild.
Sei jetzt ein Graph $G = (V, E)$, $|V| = n + 1, |E| = n$. Der Grad jedes Knotens ist mindestens 1, da G zusammenhängend ist.
Dann muss es Knoten vom Grad 1 geben, also Knoten, die mit genau einer Kante inzidieren, sondern wäre für alle Knoten $d(x) \geq 2$, so würde unter Beachtung von $\sum_{x \in V} d(x) = 2 \cdot |E|$ folgen:

$$2(n+1) \leq \sum d(x) = 2|E| = 2n.$$

Charakterisierung Baum

Entfernen wir jetzt einen solchen Knoten x vom Grad 1 mit seiner Kante aus G, so entsteht ein kleinerer Graph, der wieder genau einen Knoten mehr hat als Kanten.
Dieser ist also nach Induktionsannahme ein Baum.
Nimmt man zu diesem Baum wieder x mit seiner Kante hinzu, so ist auch x mit jedem anderen Knoten des Graphen durch genau einen Weg verbunden und damit ist G ein Baum.

Charakterisierung Baum

Zusammenfassung:
Für einen Graphen G mit n Knoten und m Kanten sind die folgenden Aussagen äquivalent:

1. G ist ein Baum (je zwei Knoten durch genau einen Weg verbunden).
2. G ist ein zusammenhängender Graph ohne Kreise.
3. G ist zusammenhängend, aber entfernt man eine beliebige Kante, so zerfällt G in zwei Zusammenhangskomponenten.
4. G ist zusammenhängend und $n = m + 1$.
   (G hat einen Knoten mehr als Kanten)