Kapitel 8

Das Traveling Salesman Problem

Effiziente Algorithmen, SoSe 2018

Professor Dr. Petra Mutzel

Dr. Bernd Zey

VO 19 am 26. Juni 2018

Übersicht

I. Effiziente Graphalgorithmen

2. Starke Zusammenhangskomponenten
3. Matching-Probleme
4. Maximale Flussprobleme
5. Amortisierte Analyse
6. Minimale Schnitte

II. Approximationsalgorithmen

7. Rucksackproblem, Bin Packing Problem
8. Traveling Salesman Problem
9. Erfüllbarkeitsprobleme
10. Schnittprobleme

Design-Techniken im Verlauf der Vorlesung

1. Die Greedy-Methode: Rucksackproblem (Kap. 7)
2. Dynamische Programmierung: Rucksackproblem (FPTAS) (Kap. 7)
3. Inkrementelle Algorithmen für Partitionsprobleme: Bin Packing Problem (Kap. 7)
4. Spezielle, problemabhängige Verfahren: Traveling Salesman Problem (Kap. 8)
5. LP-basierte Verfahren, randomisiert: Max$k$SAT (Kap. 9)
6. Lokale Suchverfahren, randomisiert: Max Cut (Kap. 10)

Wiederholung

Definition Güte Sei $s \in S(w)$ (zul. Lösung) zu $w \in I$

$$r := \max\left\{\frac{v(s)}{\mathsf{OPT}(w)}, \frac{\mathsf{OPT}(w)}{v(s)}\right\}$$

heißt Güte der Lösung s

Definition Polynomialzeitalgorithmus A,
der immer Lösung mit Güte $\leq r_A$ liefert,
heißt r_A-Approximation

Definition Optimierungsproblem
mit r-Approximation für konstantes $r \geq 1$
heißt (konstant) approximierbar und
gehört zur Klasse $\mathcal{APX}$

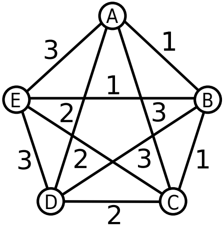
Literatur zum Traveling Salesman Problem (TSP)

- Lawler, Lenstra, Kan, Shmoys: The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization
(der Klassiker)
- Applegate, Bixby, Chvatal, Cook: The Traveling Salesman Problem: A Computational Study
(das Aktuelle)
- Website: `http://www.math.uwaterloo.ca/tsp/`
(hier findet man alles)

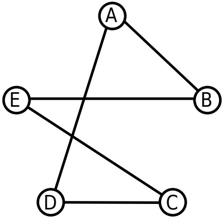
Das Traveling Salesman Problem (TSP)

Gegeben Vollständiger, ungerichteter Graph $G = (V, E), n := |V|$, mit Kantengewichten (Kosten) $w : E \to \mathbb{N} \cup \{\infty\}$

Gesucht Hamiltonkreis („Rundreise“, „Tour“) $C \subseteq E$ mit $\min \sum_{e \in C} w(e)$

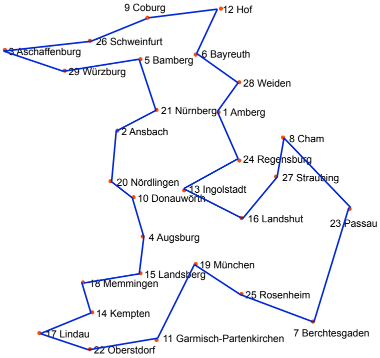


$G = (V, E)$ mit $w(e), \forall e \in E$



Tour C: A, B, E, C, D, A
Kosten $w(C) = 8$

Motivation



Viele Anwendungen, z.B.:

- Tourenplanung, Vehicle Routing, ...
- Bohrprobleme, Plotterprobleme
- Genome Sequencing (Bioinformatik)

→ Grundlegendes, gut untersuchtes Optimierungsproblem

Das Traveling Salesman Problem (TSP)

äquivalente Definition

Eingabe Symmetrische $n \times n$ Distanzmatrix $D = (d_{i,j})$ mit $d_{i,j} \in \mathbb{N} \cup \{\infty\}$

Ausgabe „Rundreise“, „Tour“:
Permutation π von $\{1, 2, \dots, n\}$
mit $\sum_{i=0}^{n-1} d_{\pi((i \bmod n)+1),\pi(((i+1) \bmod n)+1)}$ minimal

klar TSP $\in \mathcal{NPO}$

bekannt Entscheidungsvariante vom TSP ist $\mathcal{NP}$-vollständig

Wie sieht es mit Approximationen aus?

Behauptung TSP $\notin \mathcal{APX}$, sonst $\mathcal{P} = \mathcal{NP}$
d.h. Falls $\mathcal{P} \neq \mathcal{NP}$: TSP nicht approximierbar

Beweis: TSP nicht approximierbar

Annahme $\exists$ r-Approximation A für TSP

Betrachte Eingabe $G = (V, E)$ für Hamiltonkreis (HC)

Transformiere (in Polynomialzeit) G zu TSP-Eingabe: Matrix $D^{n \times n}$ mit

$$d_{i,j} = \begin{cases} 1 & \text{falls } \{i, j\} \in E \\ r \cdot n + 1 & \text{sonst} \end{cases}$$

Beobachtung:

	OPT(D)	A(D)	
falls $G \in$ HC	n	$\leq r \cdot n$	// $A(D) \leq r \cdot OPT(D)$
falls $G \notin$ HC	$> r \cdot n + 1$	$> r \cdot n + 1$	// $A(D) \geq OPT(D)$

also $\exists$ r-Approximation für TSP

$\Rightarrow$ HC $\in \mathcal{P} \Rightarrow \mathcal{P} = \mathcal{NP}$ $\square$

Eingeschränkte TSP-Varianten

Definition: Metrisches TSP:
zusätzlich Δ-Ungleichung: $\forall i, j, k\colon d_{i,k} \leq d_{i,j} + d_{j,k}$
(hier betrachtet)

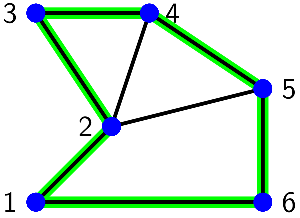
Definition: Euklidisches TSP:
Eingabe: n Punkte $x_1, x_2, \ldots, x_n \in \mathbb{R}^2$
Distanzen $d_{i,j}$ implizit gegeben durch
$d_{i,j} := \sqrt{(x_i[1] - x_j[1])^2 + (x_i[2] - x_j[2])^2}$

Vorsicht: reelle Zahlen
Aber: im Kontext der Approximation *nicht so kritisch*

Approximationen für das metrische TSP

Was hilft beim metrischen TSP?

Beobachtung „Mehrfachbesuche“ kostenlos reparierbar



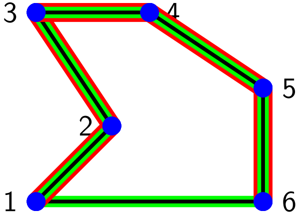
klar Knoten auslassen geht immer

aber nur beim metrischen TSP
$d_{4,5} \leq d_{4,2} + d_{2,5}$ wegen Δ-Ungleichung
also Gesamtkosten nicht größer

Idee zur Approximation

Wie finden wir günstiges Tourgerüst?

Beobachtung Touren enthalten interessante bekannte Komponenten



Beobachtung Tour enthält Spannbaum

darum OPT $\geq v(\text{MST})$

also nehme minimalen Spannbaum als „Gerüst“

„Gerüst“ $\rightsquigarrow$ Tour? mit Hilfe von Multigraph und Eulerkreis

Einschub: Multigraph, Eulerkreis

Definition 8.1

Ein ungerichteter *Multigraph* $G = (V, E)$ ist definiert durch eine endliche Knotenmenge V und eine endliche Multimenge (die Kanten) über $(V \times V) \setminus \{\{v, v\} \mid v \in V\}$.

Definition 8.2

Sei $G = (V, E)$ ein ungerichteter Multigraph. Ein *Eulerkreis* ist ein Kreis in G, der jede Kante genau einmal enthält.

Anmerkung Eulerkreise in $\mathcal{O}(|V| + |E|)$ berechenbar
Idee für Algorithmen MST $\rightsquigarrow$ Eulerkreis $\rightsquigarrow$ Tour

Lemma 8.3

Ein zusammenhängender Multigraph $G = (V, E)$ enthält einen Eulerkreis $\Leftrightarrow$ alle Knoten in G haben geraden Grad

Beweis von Lemma 8.3

„$\Rightarrow$“: Eulerkreis $\rightsquigarrow$ alle Knotengrade gerade

Berechne Knotengrade Durchlaufe Eulerkreis:
zähle jedes Betreten und Verlassen eines Knotens

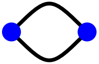
Beobachtung $\forall$Knoten: #Betreten = #Verlassen

also alle Knotengrade gerade ✓

„$\Leftarrow$“: alle Grade gerade $\rightsquigarrow$ Eulerkreis

vollständige Induktion über $m := |E|$

Induktionsanfang leer für $m \in \{0, 1\}$

für $m = 2$  ✓

Beweis von Lemma 8.3: $\forall$ Grade gerade $\rightsquigarrow$ Eulerkreis

Haben zusammenhängenden Graph $G = (V, E)$
alle Knotengrade gerade, Induktionsschritt mit $m > 2$

Konstruiere einen Kreis:

1. Starte in beliebigem Knoten v
2. Falls $\exists w$ unbesucht mit $\{v, w\}$ unbenutzt
3. Dann markiere v und $\{v, w\}$; $v \leftarrow w$; Weiter bei 2;
4. Sonst wechsle zu bereits besuchtem Nachbarn w; STOP;

Beobachtung Möglich? Ja, da alle Knotengrade gerade
Beobachtung Endstück des Pfades ist ein Kreis K
Betrachte $G' = (V, E \setminus K)$
Beobachtung bei jedem Knoten auf K ändert sich Grad um -2
also jeder Knoten in G' hat geraden Grad
darum Induktionsannahme auf G' grundsätzlich anwendbar

Lemma 8.3: G' und die Induktionsannahme

klar $G' = (V, E \setminus K)$ enthält weniger Kanten

Ist G' auch zusammenhängend?

Einsicht muss nicht sein!

Aber Induktionsannahme auf Zusammenhangskomponenten C_i von G' anwendbar
$\rightsquigarrow$ jedes C_i enthält einen Eulerkreis

Beobachtung jedes C_i berührt K

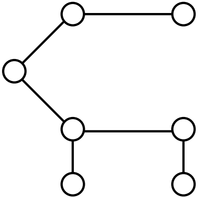
also Eulerkreise aus C_i in K einfügen $\rightsquigarrow$ Eulerkreis in G $\square$

Minimum Spanning Tree Approximation

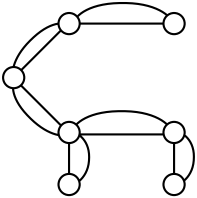
Algorithmus 8.3

1. Berechne MST $T = (V, E_T)$ auf G.
2. Erzeuge Multigraphen $G' = (V, E')$ mit $E' = \{e, e \mid e \in E_T\}$.
3. Berechne Eulerkreis K auf G'.
4. Berechne π: durchlaufe K und entferne mehrfach vorkommende Knoten.
5. Ausgabe π

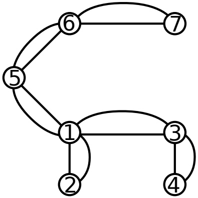
MST Approximation – Beispiel



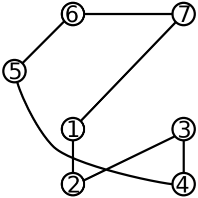
MST T



Multigraph G'



Eulerkreis K:
1,2,1,3,4,3,1,5,6,7,6,5,1



Tour π:
1,2,3,4,5,6,7,1

Minimum Spanning Tree Approximation

Algorithmus 8.3

1. Berechne MST $T = (V, E_T)$ auf G.
2. Erzeuge Multigraphen $G' = (V, E')$ mit $E' = \{e, e \mid e \in E_T\}$.
3. Berechne Eulerkreis K auf G'.
4. Berechne π: durchlaufe K und entferne mehrfach vorkommende Knoten.
5. Ausgabe π

Theorem 8.4

Algorithmus 8.3 hat Laufzeit $\mathcal{O}(n^2)$ und ist eine 2-Approximation für das metrische TSP.

Beweis von Theorem 8.4

Was ist zu zeigen?

1. Laufzeit
2. Korrektheit
3. Güte

zur Laufzeit

1. MST-Berechnung in Zeit $\mathcal{O}(n^2)$
   (Algorithmus von Prim mit Fibonacci-Heaps)
2. Multigraphberechnung in Zeit $\mathcal{O}(n)$
3. Eulerkreisberechnung in Zeit $\mathcal{O}(n)$
4. Tourberechnung in Zeit $\mathcal{O}(n)$
5. Ausgabe in Zeit $\mathcal{O}(n)$

Gesamtlaufzeit $\mathcal{O}(n^2)$

zur Korrektheit

nur kritisch: ist G' zusammenhängend, alle Knotengrade gerade?

Beobachtung wegen MST + Verdopplung gesichert ✓

Beweis von Theorem 8.4: Zur Güte

Betrachte Wert einer optimalen Lösung OPT

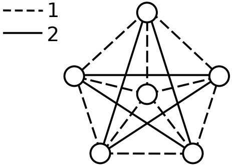
Betrachte Wert des Spannbaums $v(T) = \sum_{\{i,j\} \in E_T} d_{i,j}$

Erinnerung $v(T) \leq \mathsf{OPT}$
weil (optimale) Tour Spannbäume enthält

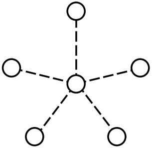
also $v(\pi) \leq \sum_{\{i,j\} \in K} d_{i,j} = \sum_{\{i,j\} \in E'} d_{i,j} = 2v(T) \leq 2\mathsf{OPT}$

also $\frac{v(\pi)}{\mathsf{OPT}} \leq 2$ □

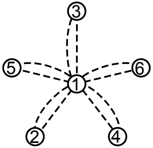
MST Approximation – Beispiel



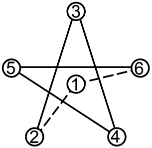
Graph G



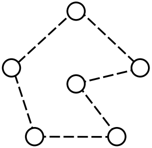
MST T



Eulerkreis K:
1,2,1,3,1,4,1,5,1,6,1



Tour π,
$v(\pi) = 10$



π_{OPT},
$v(\pi_{\mathsf{OPT}}) = 6$

$\Rightarrow$ Dieses Beispiel führt zum Worst-Case:
für n Knoten Kosten n vs. $2n - 2$
$\Rightarrow$ Güte $(2n-2)/n \rightsquigarrow 2$, d.h. Güte 2 ist *scharf*

Überlegungen zur 2-Approximation

Wir haben 2-Approximation für das metrische TSP
$\Rightarrow$ metrisches TSP $\in \mathcal{APX}$

Geht es nicht besser?

Wo verlieren wir viel?

Problem durch Spannbaumverdopplung Faktor 2

Problematisch Abschätzung des Gewinns durch Abkürzungen

Folgerung Können wir Spannbaumverdopplung einsparen?

Einsparen der Spannbaumverdopplung

Warum verdoppeln wir den Spannbaum überhaupt?

Erinnerung damit alle Knoten geraden Grad haben
für Eulerkreis-Konstruktion

Beobachtung zusätzliche Kanten nur nötig
für Knoten mit ungeradem Grad

Idee Identifiziere Knoten mit ungeradem Grad in T
$\rightarrow$ verbinde paarweise durch möglichst günstige Kanten

Erinnerung „verbinde paarweise" $\leftrightarrow$ Matching

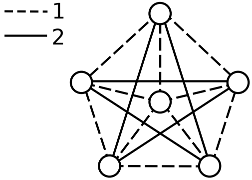
also Suche kostenminimales perfektes Matching M
$\rightarrow$ perfekt: alle Knoten sind besetzt
$\rightarrow$ kostenminimal: Summe der Kantengewichte in M ist minimal
Fakt in Zeit $O(n^3)$ berechenbar

Algorithmus von Christofides

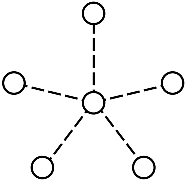
Algorithmus 8.5

1. Berechne MST $T = (V, E_T)$ auf G.
2. Berechne Menge V' der Knoten mit ungeradem Grad in T.
3. Berechne kostenmin. perfektes Matching M auf Knoten in V'.
4. Erzeuge Multigraph $G' = (V, E')$ mit $E' = E_T \cup M$.
5. Berechne Eulerkreis K auf G'.
6. Berechne π: durchlaufe K und entferne mehrfach vorkommende Knoten.
7. Ausgabe π

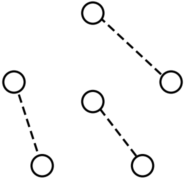
Algorithmus von Christofides – Beispiel



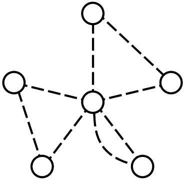
Graph G



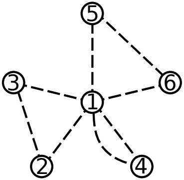
MST T



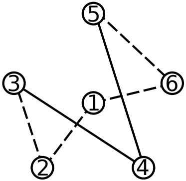
$V' = V$
$\rightarrow$ Matching M



Multigraph G'



Eulerkreis K:
1,2,3,1,4,1,5,6,1



Tour $\pi, v(\pi) = 8$

schon gesehen: $v(\pi_{\mathsf{OPT}}) = 6$

Algorithmus von Christofides

Algorithmus 8.5

1. Berechne MST $T = (V, E_T)$ auf G.
2. Berechne Menge V' der Knoten mit ungeradem Grad in T.
3. Berechne kostenmin. perfektes Matching M auf Knoten in V'.
4. Erzeuge Multigraph $G' = (V, E')$ mit $E' = E_T \cup M$.
5. Berechne Eulerkreis K auf G'.
6. Berechne π: durchlaufe K und entferne mehrfach vorkommende Knoten.
7. Ausgabe π

Theorem 8.6

Der Algorithmus von Christofides hat Laufzeit $\mathcal{O}(n^3)$ und ist eine $(3/2)$-Approximation für das metrische TSP.

Beweis von Theorem 8.6

Was ist zu zeigen?

1. Laufzeit
2. Korrektheit
3. Güte

zur Laufzeit

1. MST-Berechnung in Zeit $\mathcal{O}(n^2)$ (Algorithmus von Prim)
2. Berechnung von V' in Zeit $\mathcal{O}(n)$
3. Matchingberechnung in Zeit $\mathcal{O}(n^3)$
4. Multigraphberechnung in Zeit $\mathcal{O}(n)$
5. Eulerkreisberechnung in Zeit $\mathcal{O}(n)$
6. Tourberechnung in Zeit $\mathcal{O}(n)$
7. Ausgabe in Zeit $\mathcal{O}(n)$

Gesamtlaufzeit $\mathcal{O}(n^3)$ ✓

Beweis von Theorem 8.6: Korrektheit

Ist G' zusammenhängend und sind alle Knotengrade gerade?

Zusammenhang folgt aus Spannbaum ✓

gerade Knotengrade gerade Knoten in T: unverändert in G' ✓

ungerade Knoten in T: je 1 neue Kante in G' ✓

Perfektes Matching berechenbar? Ist $|V'|$ gerade?

$\rightarrow$ Summe aller Knotengrade in T gerade $(2 \cdot |E|)$

$\rightarrow$ Summe aller Knotengrade
der Knoten mit geradem Grad in T gerade

$\rightarrow$ Summe aller Knotengerade
der Knoten mit ungeradem Grad in T gerade

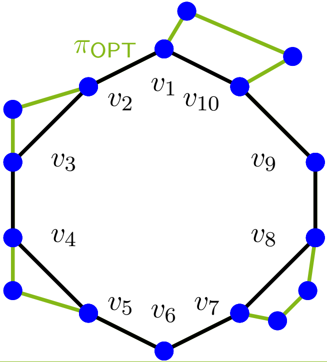
$\rightarrow$ Anzahl der Knoten mit ungeradem Grad in T gerade ✓

Beweis von Theorem 8.6: Güte

wie vorhin $v(\pi) \leq v(K) = v(T) + v(M) \leq \mathsf{OPT} + v(M)$

zu zeigen $v(\pi) \leq \frac{3}{2}\mathsf{OPT} \Rightarrow v(M) \leq \mathsf{OPT}/2$

Betrachte Knoten aus V' in optimaler Tour π_{OPT}:
Benenne so, dass Knoten in π_{OPT}
in Reihenfolge $v_1, v_2, \ldots, v_{|V'|}$ durchlaufen werden



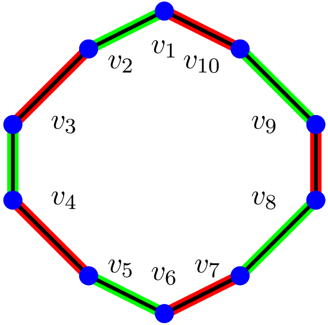
Betrachte Kreis $C = (v_1, v_2, \ldots, v_{|V'|}, v_1)$

Beobachtung $v(C) \leq v(\pi_{\mathsf{OPT}}) = \mathsf{OPT}$
(wegen Δ-Ungleichung)

Beweis von Theorem 8.6: $v(M) \leq \mathsf{OPT}/2$

Wir haben

- kostenminimales perfektes Matching M auf Knoten in V'
- $V' = \{v_1, v_2, \ldots, v_{|V'|}\}$
- Kreis $C = (v_1, v_2, \ldots, v_{|V'|})$ mit $v(C) \leq v(\pi_{\mathsf{OPT}}) = \mathsf{OPT}$



Matching $M_1 = \{\{v_1, v_2\}, \{v_3, v_4\}, \ldots\}$
Matching $M_2 = \{\{v_2, v_3\}, \{v_4, v_5\}, \ldots\}$

Beobachtung $v(M_1) + v(M_2) = v(C)$
Beobachtung $\min\{v(M_1), v(M_2)\} \leq v(C)/2$
Beobachtung $v(M) \leq \min\{v(M_1), v(M_2)\}$
also $v(M) \leq v(C)/2 \leq \mathsf{OPT}/2$ $\square$

Anmerkungen zum metrischen TSP

- Algorithmus von Christofides beste bekannte Approximation seit 1976 (!)
- obere Schranke scharf: entsprechende Worst-Case-Instanzen bekannt
- Existenz besserer Approximationen offen
- bekannt: $\mathcal{P} \neq \mathcal{NP} \Rightarrow$ metrisches TSP $\notin \mathcal{PTAS}$
- aber: für euklidisches TSP existiert ein PTAS
  (siehe Skript, aber nicht Teil unserer Vorlesung)

Design-Techniken für Approximationsalgorithmen im Verlauf der Vorlesung

1. Die Greedy-Methode: Rucksackproblem (Kap. 7)
2. Dynamische Programmierung: Rucksackproblem (FPTAS) (Kap. 7)
3. Inkrementelle Algorithmen für Partitionsprobleme: Bin Packing Problem (Kap. 7)
4. Spezielle, problemabhängige Verfahren: Traveling Salesman Problem (Kap. 8)
5. LP-basierte Verfahren, randomisiert: Max$k$SAT (Kap. 9)
6. Lokale Suchverfahren, randomisiert: Max Cut (Kap. 10)