c.) BRANCH & BOUND für Δ-TSP (metrisches TSP)

Eingabe: n Städte $N = \{1, 2, 3, \ldots, n\}$
und Distanzen $d(i,j)$ zwischen je zwei Städten i, j
die einer Metrik entsprechen

Ausgabe: eine kürzeste Rundreise

(Sei $K_n(N, E)$ der vollständige ungerichtete Graph über N)

Überlegungen:

→ Was seien die Knoten des B&B Baums?

Wir legen Teillösungen fest, (und schränken somit den Lösungsraum ein), wie üblich:

- die Knoten v des Baums werden durch Paare (S, T) beschrieben: $S \subseteq E$ ist eine Menge erzwungener Kanten, die müssen Teil der Rundreise sein

  $T \subseteq E$ ist eine Menge verbotener Kanten
- das Ausgangsproblem ist $(\emptyset, \emptyset)$. (Wurzel)
  Vom Branching-Operator werden natürlich jeweils weitere Kanten entweder erzwungen oder verboten (siehe später).

→ Wie findet man gute Anfangslösung y_0?

Wir kennen gute Heuristiken für Δ-TSP, wie Christofides' Algorithmus, Farthest-Insertion, oder Nearest-Neighbor Heuristik.

Wir haben hier ein Minimierungsproblem, und für ein B&B Algorithmus brauchen wir gute untere Schranke $unten(v)$ für die Teillösungen $v = (S, T)$

→ Die Bestimmung einer guten unteren Schranke unten(v)

Der Einfachheit halber wird unten(v) für die Wurzel $v = (\emptyset, \emptyset)$ (d.h. für das unbeschränkte Problem) gesucht

Untere Schranke mit 1-Bäumen

Wir relaxieren das Δ-TSP Problem, und erhalten ein effizient lösbares (optimierbares) Problem, so dass alle Rundreisen, aber auch andere Teilgraphen Objekte Lösungen dieses relaxierten Problems sind:
Wir beobachten, dass jede Rundreise aus einem Spannbaum und noch einer Kante besteht. Wir werden über solche Objekte optimieren. Präziser: Jede Rundreise besteht aus einem Spannbaum über die Städte $\{2, 3, 4, \ldots, n\}$ und zwei Kanten inzident mit der Stadt 1.
Einen solchen Teilgraphen nennen wir 1-Baum (OneTree), und wir optimieren über alle 1-Bäume. (Beachte, dass ein 1-Baum ein Spanngraph mit genau einem Kreis (über Stadt 1) ist.)

Definition: Ein 1-Baum besteht aus einem Spannbaum über der Städte $\{2, 3, 4, \ldots, n\}$ und zwei mit 1 inzidenten Kanten.

Beobachtung 1. Bezeichne OneT(d) die minimale Länge über 1-Bäumen, und TSP(d) die minimale Länge über Rundreisen für die Distanzfunktion $d()$.
Dann gilt $\text{OneT}(d) \leq \text{TSP}(d)$

Warum? Da eine minimale Rundreise selbst ein 1-Baum ist, kann sie nicht kürzer sein als OneT(d) die minimale 1-Baum Länge.

BB13.

Somit ist OneT(d) eine untere Schranke!

Beobachtung 2: OneT(d) die minimale 1-Baum-Länge kann effizient berechnet werden.

(Warum? Mit dem Greedy Algorithmus von Kruskal berechnen wir einen minimalen Spannbaum über $\{2, 3, 4, \ldots, n\}$; dann verbinden wir Stadt 1 mit diesem Spannbaum über seine beiden kürzesten inzidenten Kanten. Für beide Teile des 1-Baums haben wir jetzt eine minimale Lösung (für den Spannbaum über $N \setminus \{1\}$ und für die zwei Verbindungskanten), die gemeinsam einen 1-Baum ergeben. – besser geht's nicht!)

Somit ist OneT(d) eine schnell berechenbare untere Schranke!

Bemerkung: **Die Festlegung** der Stadt 1 in der Definition von 1-Bäumen macht die Definition asymmetrisch. Wir könnten versuchen OneT(d) für andere fixierte Städte auch zu berechnen, und die höchste solche untere Schranke nehmen. Mit der folgenden Methode erhält man jedoch eine bessere Schranke, und für diese Methode darf man die Stadt 1 in der Definition von OneT(d) ohne Beschränkung der Allgemeinheit festlegen.

Die Held-Karp Schranke

Wir verfeinern die untere Schranke OneT(d) wie folgt:

- wir führen ein Parameter π_i für jede Stadt $i \in N$ ein (π_i wird auch Strafe oder Preis für den „Besuch" dieser Stadt genannt)

Sei $\pi = (\pi_1, \pi_2 \, . \, . \, . \, . \, \pi_n)$ der Vektor dieser Parameter.

- wir modifizieren alle Distanzwerte d() und erhalten eine neue Metrik d^π (neue Distanzen)

$$d^\pi(i,j) = d(i,j) + \pi_i + \pi_j.$$

- Für die Metrik d^π gilt natürlich auch, dass ein kürzester 1-Baum nicht länger ist als eine kürzeste Rundreise

$$\mathrm{OneT}(d^\pi) \le \mathrm{TSP}(d^\pi) \qquad (1)$$

- weiterhin gilt

$$\mathrm{TSP}(d^\pi) = \mathrm{TSP}(d) + 2\sum_{i=1}^{n} \pi_i \qquad (2)$$

Warum? Diese Gleichung gilt für die Länge jeder konkreten <u>Rundreise</u> in den Metriken d bzw. d^π, weil jeder Knoten i mit genau zwei inzidenten Kanten teilnimmt und so genau $2\pi_i$ plus Länge verursacht in d^π im Vergleich zu d. Dies gilt <u>nicht</u> für <u>1-Bäume</u>! Da die Länge jeder <u>Rundreise</u> um den gleichen Wert $2\sum\pi_i$ verschoben ist, ist eine kürzeste Rundreise in d auch kürzeste in d^π, und für ihre Länge TSP(d) bzw. $\mathrm{TSP}(d^\pi)$ gilt die Gleichung (2).

Aus (1) und (2) folgt

$$\mathrm{OneT}(d^\pi) - 2\sum_{i=1}^{n} \pi_i \le \mathrm{TSP}(d)$$

BB15.

Für beliebige π, $OneT(d^{\pi}) - 2\sum_{i=1}^{n} \pi_i$ ist also eine untere Schranke

für $TSP(d)$, und (für fixierten $\pi = (\pi_1 \pi_2 \ldots \pi_n)$)

es ist leicht zu berechnen, weil der minimale

1-Baum leicht zu berechnen ist, mit Länge $OneT(d^{\pi})$.

Eine gute untere Schranke ist hoch, also nehmen

wir das Maximum dieser Schranken über alle

Vektoren π

$$\max_{\pi} \left[OneT(d^{\pi}) - 2\sum_{i=1}^{n} \pi_i \right]$$

Diese Schranke heißt Held-Karp Schranke, und

ist effizient berechenbar. (Die Schwierigkeit dieser

Berechnung besteht in der Maximierung über alle π.)

Der Branching Operator für die Wurzel $(\emptyset, \emptyset)$:

Sei π^* der Vektor der die Held-Karp Schranke

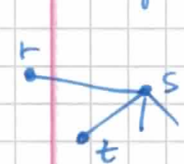
(für $(\emptyset, \emptyset)$) bestimmt, und T^* ein minimaler

1-Baum für die Distanzen d^{π^*}.

→ Falls in T^* alle Knoten i Grad $= 2$ besitzen, dann

ist T^* eine minimale Rundreise, und wir sind fertig.

→ Sonst gibt es eine Stadt s mit Grad ≥ 3. Seien



$\{r, s\}$ und $\{t, s\}$ die längsten,

mit s inzidenten Kanten in T^*

[weil $w(\pi^*) \overset{\text{Länge}}{=} d(T^*) + 0 \leq TSP(d)$, siehe später]

Die Wurzel $v = (\emptyset, \emptyset)$ des B&B Baums hat dann 3 Kinder:

- in v_1 wird die Benutzung der Kante $\{r, s\}$ verboten
- in v_2 wird $\{r, s\}$ erzwungen aber $\{t, s\}$ verboten
- in v_3 werden $\{r, s\}$ und $\{t, s\}$ erzwungen, alle andere verboten
- das erfolgversprechendste Blatt wird stets mit Tiefensuche gewählt; (bessere Lösungen ergeben sich auch während der Tiefensuche)
- für ein aktiviertes Blatt $v = (S, T)$, um die Held-Karp Schranke zu berechnen, betrachtet man entsprechend 1-Bäume die alle Kanten aus S enthalten aber alle Kanten aus T verbieten.

Weitere Erklärung zur Held-Karp Schranke

Die Held-Karp Schranke ist $\max\limits_{\pi} w(\pi)$ für

$$\rightarrow w(\pi) = \mathrm{OneT}(d^{\pi}) - 2\sum_{i=1}^{n} \pi_i = \min_{T:\,1\text{-Baum}} \left[\text{Länge}_T(d^{\pi}) - 2\sum_{i=1}^{n} \pi_i \right] =$$

$$= \min_{T:\,1\text{-Baum}} \left[\text{Länge}_T(d) + \sum_{i=1}^{n} (\deg_i^T - 2)\pi_i \right]$$

wobei $\text{Länge}_T()$ ist die Länge des 1-Baums T in der gegebenen Metrik, und $\deg_i^T$ der Grad von Knoten (Stadt) i im 1-Baum T.

BB 17.

Beachte, dass falls T eine Rundreise ist $T = R$, dann ist der Ausdruck die Länge der Rundreise und unabhängig von Π.

$$\text{Länge}_R(d) + \sum_{i=1}^{n} (2-2) \cdot \Pi_i = \text{Länge}_R(d)$$

→ man kann Π mit einer Art „lokaler Suche" (Gradienten-Verfahren) Schritt für Schritt modifizieren um $w(\Pi)$ zu maximieren. Die Veränderung von Π beeinflusst den Wert für 1-Bäume die Rundreisen sind, nicht.

→ Für einzelne (!) 1-Bäume T, die keine Rundreisen sind, kann der Wert erhöht werden, indem Städte i mit $\deg_i^T \geq 3$ immer höhere Strafe Π_i zugeordnet wird.

→ Wenn durch die Änderung in Π ein anderer 1-Baum T' minimalen Wert hat, dann ändert man Π auch entsprechend gemäß diesem Baum T'

→ Dies hilft jedoch nicht weiter, wenn mit beliebiger Änderung von Π irgendein 1-Baum seinen Wert senken würde (oder eine Rundreise als 1-Baum mit minimalem Wert erreicht wurde → dann ergibt die Held-Karp Schranke die optimale Länge einer Rundreise

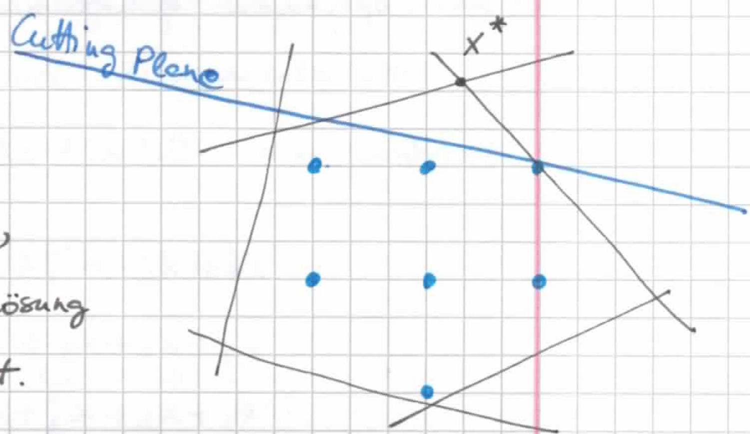
d.) Cutting Planes (Schnittebenen)

Untere Schranken für Minimierungsprobleme, bzw. obere Schranken für Maximierungsprobleme (gegeben in IP-Formulierung)

Grobstruktur

Aufgabe: Minimiere $c^T \cdot x$ so dass $A \cdot x = b$ | oder $Ax \geq b$
$x \geq 0$
x ganzzahlig

- löse die LP-Relaxierung, sei x^* eine optimale fraktionale Lösung
- falls x^* ganzzahlig, halt
- sonst füge eine neue Nebenbedingung (Ungleichung) ein, die von jeder ganzzahligen Lösung erfüllt wird, von x^* aber nicht.



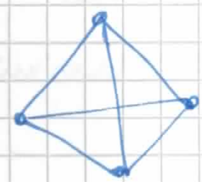
Diese Ungleichung heißt Schnittebene (Cutting Plane)
(sie bestimmt eine neue Hyperebene als Seitenfläche des Lösungspolyeders)
- (Schnittebenen werden ggf. solange eingeführt, bis ein ganzzahliges Optimum berechnet wird.)

Cutting-Plane Methoden unterscheiden sich in der Wahl der Schnittebene, die x^* aus der Lösungsmenge ausschneidet.

BB19.

Beispiel/ Illustration: max-INDEPENDENT SET auf dem vollständigen Graphen (ungewichtet)

Sei $V = \{1, 2, \dots, n\}$



IP: maximiere $\sum_{i=1}^{n} x_i$

so dass $x_i + x_j \leq 1 \quad \forall i,j \; i \neq j$

$x_i \in \{0,1\}$

- $x_i = \frac{1}{2} \; \forall i$ ist eine fraktionale Lösung mit Zielwert $\frac{n}{2}$

- die optimale ganzzahlige Lösung hat Wert 1

Wir werden die Schnittebenen

$$x_1 + x_2 \leq 1$$
$$x_1 + x_2 + x_3 \leq 1$$
$$x_1 + x_2 + x_3 + x_4 \leq 1$$
$$\vdots$$
$$x_1 + x_2 + x_3 + x_4 + \dots + x_n \leq 1$$

Schnitt für Schnitt einfügen können, d.h. zeigen, mit Induktion, dass eine ganzzahlige Lösung sie erfüllen soll.

Die letzte $\sum_{i=1}^{n} x_i \leq 1$ hat eine ganzzahlige optimale Lösung $(1, 0, 0, \dots, 0)$

Angenommen, die Ungleichung $\sum_{i=1}^{k-1} x_i \le 1$ wurde eingefügt, wir zeigen, dass $\sum_{i=1}^{k} x_i \le 1$ von einer ganzzahligen Lösung erfüllt werden soll:

Wir haben die Bedingungen

$$
\begin{array}{r|l}
1\cdot & x_1 + x_k \le 1 \\
1\cdot & x_2 + x_k \le 1 \\
 & \ddots \\
1\cdot & x_{k-1} + x_k \le 1 \\
(k-2)\cdot & x_1 + x_2 + \dots + x_{k-1} \le 1 \\
\hline
\end{array}
$$

Wir multiplizieren diese Ungleichungen durch $1, 1, 1 \dots$ bzw. $(k-2)$ und summieren sie.

wir erhalten:

$$(k-1)\cdot(x_1 + x_2 + \dots + x_{k-1} + x_k) \le k-1 + k-2$$

$$x_1 + x_2 + \dots + x_{k-1} + x_k \le 1 + \frac{k-2}{k-1}$$

Für <u>ganzzahlige</u> Lösungen soll also gelten

$$x_1 + x_2 + \dots + x_{k-1} + x_k \le 1$$ D

Das ist ein Spezialfall von <u>Gomory-Chvatal</u> Schnitten

Solving hard problems in practice

(aus: The Nature of Computation, Sec. 9.10)

TSP-Instanz: 42 Großstädte der USA
$N = \{1, 2, 3, \ldots, 42\}$

Wir versuchen obere bzw. untere Schranken für die Länge einer kürzesten Rundreise nah aneinander bringen

obere Schranke: die Länge einer bislang besten gefundenen Rundreise

untere Schranke: der optimale Wert einer relaxierten, effizient berechenbaren Version des TSP Problems

1. erste approximative Lösungen

→ man findet eine Rundreise mit der Greedy Nearest-Neighbor Heuristik, wobei in jedem Schritt die nächstmögliche noch unbesuchte Stadt besucht wird

→ man verbessert diese Tour mit lokaler Suche:

Vertex-Insertion ändert die Position einer Stadt in der Rundreise; es gibt n Positionen und $\approx n \cdot n = n^2$ mögliche Änderungen (so viele benachbarte Rundreisen zu einer gegebenen Rundreise)

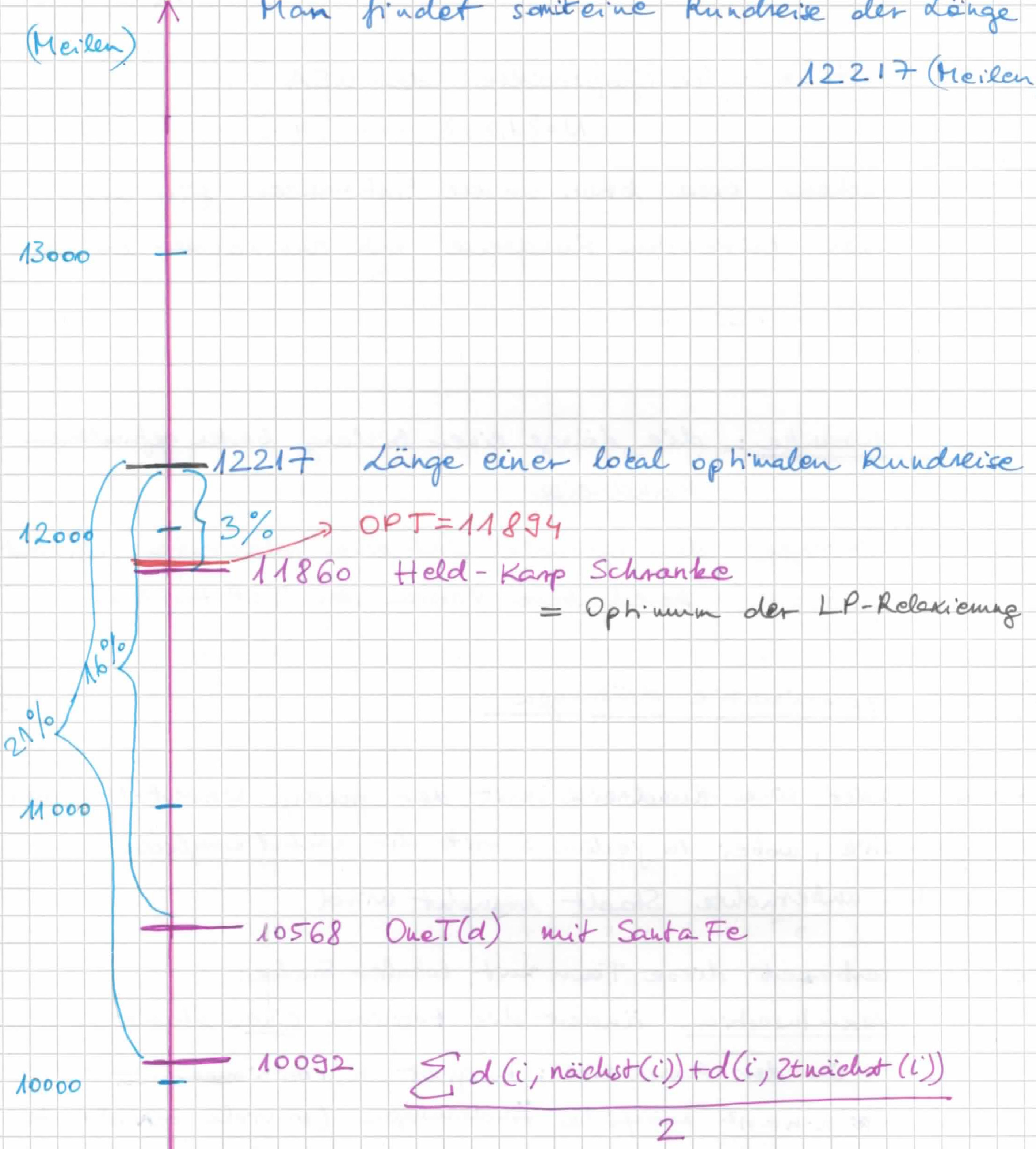
jeweils die beste Nachbarlösung wird gewählt

2-opt nimmt 2 Kanten heraus, und fügt sie andersrum in die beiden Teilrundreisen ein.

Wieder gibt es jeweils $\approx n^2$ Nachbarlösungen.

Diese Schnitte werden iteriert, bis die Tour **lokal** optimal wird, d.h. alle benachbarten Touren sind länger.

Man findet somit eine Rundreise der Länge 12217 (Meilen)



2. Untere Schranken

→ Sei $\sigma_1 \sigma_2 \ldots \sigma_n$ eine Permutation entsprechend einer optimalen Rundreise $\sum_{i=1}^{n} d(\sigma_i, \sigma_{i+1}) = OPT$

$$\sum_{i=1}^{n} \left(d(\sigma_i, \sigma_{i-1}) + d(\sigma_i, \sigma_{i+1}) \right) = 2\,OPT$$

$$\sum_{i=1}^{n} \text{Distanz von } i \text{ zu den beiden nächsten Nachbarn} \leq 2\,OPT$$

→ diese untere Schranke ergibt: 10092 Meilen

→ 1-Baum (mit Stadt 21 Santa Fe als Knoten 1)

minimale Länge: 10568

→ Held-Karp Schranke: 11860

3. LP-Relaxierung

Man möchte eine LP-Relaxierung des TSP lösen, um noch bessere untere Schranke zu erhalten

TSP hat (unter anderen) die folgende IP-Formulierung:

Sei x_e die Variable für Kante e im vollständigen Graphen über alle Städte in N, und d_e ihre Länge

minimiere $\sum_e d_e \cdot x_e$

$x_e \in \{0,1\}$

so dass $\sum_{e \in \delta(i)} x_e = 2$ für jede Stadt $i \in N$ (IP)

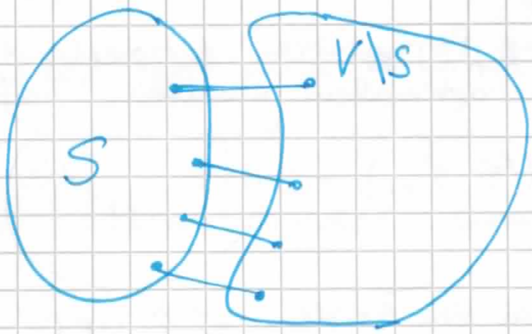
Schnitt-Ungleichung / Bedingungen (cut-constraints) $\sum_{e \in \delta(S)} x_e \geq 2 \quad \forall\, S \subset \{1, 2, 3, \ldots, 42\}$

$\delta(i)$: Menge der mit i inzidenten Kanten

$\delta(S)$: Menge der kreuzenden Kanten in $(S, V \setminus S)$

BB24.

eine Rundreise ist zweifach zusammenhängend:



Für jeden Schnitt $(S, V\setminus S)$ sollen mindestens 2 kreuzende Kanten in der Rundreise sein

(LP) Relaxierung: $x_e \in [0,1]$ statt $x_e \in \{0,1\}$

+ sonst dieselben Bedingungen

[Beachte, dass hier „Schnitt" in einem anderen Sinn benutzt wird als bei den Schnittebenen, nämlich als Zweiteilung eines Graphen!]

Die Lösungsmenge dieses LP ist das sog. Cut-Polytop

Aber! dieses LP hat exponentielle-viele Bedingungen!

somit ist es nicht in polynomieller Zeit optimierbar mit LP-Solver (also nicht mal die LP-relaxierung!)

Wir machen etwas ähnliches wie bei der Cutting-Plane Methode:

→ das LP ohne Schnitt-Bedingungen wird gelöst:

$$\text{minimiere } \sum_e d_e x_e$$
$$\text{s.d. } \sum_{e \in \delta(i)} x_e = 2 \quad \forall i \in N$$
$$0 \le x_e \le 1$$

→ eine optimale Lösung (x_e^*) entspricht einem Graphen mit Kantengewichten x_e^*

→ ein minimaler Schnitt in diesem Graphen kann effizient berechnet werden (siehe Max-Flow-Min-Cut Theorem)

→ FALLS MinCut ≥ 2 $\Rightarrow$ die Schnittbedingungen werden erfüllt (x_e^*) ist eine fraktionale Rundreise

→ SONST füge die Bedingung (Cut-Constraint o. Schnitt-Ungleichung als Schnittebene (!))

$$\sum_{e \in \delta(S_0)} x_e \geq 2$$

für den minimalen Schnitt $(S_0, V \setminus S_0)$ zu den LP-Bedingungen

→ in unserem Beispiel erhält man nach dem Einfügen von 6 Schnitt-Bedingungen eine minimale fraktionale Lösung von TSP (d.h. die alle Bedingungen von (LP) erfüllt)

ihre Länge ist 11860 — genau die Held-Karp-Schranke!

Warum?

Die Held-Karp Schranke (mit den Variablen $\pi_1\ \pi_2 \ldots \pi_n$)

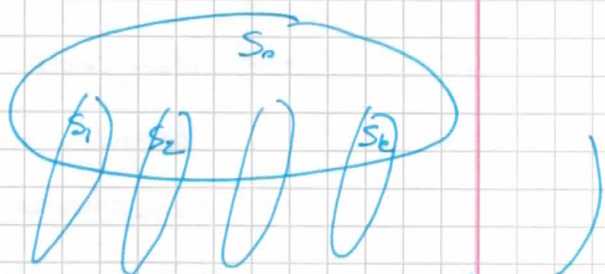
ist das duale der LP-Relaxierung!

(4.) Ende I: man sollte die untere oder die obere Schranke weiter verbessern.

Es gibt weitere lineare Bedingungen, die eine ganzzahlige Lösung erfüllen soll.

(Kamm-Ungleichungen

die Summe der Schnitte $k+1$ solcher Mengen soll $\geq 3k+1$ sein



Nach dem Einfügen von 2 Kamm-Ungleichungen (als Cutting Planes)
wird die optimale Lösung des LP (in unserem Beispiel) eine ganzzahlige Rundreise also eine optimale Rundreise

BB26.

⑤ Ende II.

Eine optimale Lösung mit nur Cutting planes braucht zu viel Kreativität oder ist sogar nicht möglich zu finden im Allgemeinen

→ Man kann die Cutting-Plane Methode mit Branch & Bound kombinieren

→ Die optimale Lösung vom (LP) enthält fraktionale Kanten mit $x_e = \frac{1}{2}$

Man fixiert eine dieser fraktionalen Kanten, und prüft separat zwei Möglichkeiten: $x_e = 1$ oder $x_e = 0$

das ist ein Branching-Operator

In beiden Fällen (Zweigen) hat man eine eingeschränkte TSP-Instanz, und kann für diese die LP-Relaxierung optimieren. (ohne Schnitt-Ungleichungen, oder mit manchen von ihnen)

In der konkreten Instanz erhält man die folgenden Ergebnisse:

→ Fall 1. $x_e = 1$ (für die konkret ausgewählte Kante)

ergibt die Optimierung des LP eine ganzzahlige Rundreise der Länge: 11 894

→ Fall 2. $x_e = 0$

das LP ergibt eine optimale Lösung, (die keine Rundreise ist), mit Wert 11 928

→ da jede weitere Bedingung, bzw. Branching-Operator die Lösungsmenge einschränkt, wird das Optimum in diesem Zweig mindestens 11928

diese ist eine untere Schranke

⇒ diesen Zweig (Lösungen mit $x_e = 0$) braucht man nicht weiter zu prüfen.

⇒ In der Praxis funktioniert Branch & Bound oft, selbst wenn es im Worst-Case kein Polynomialzeit Algorithmus ist. Gute untere Schranken zu finden ist wesentlich!

→ Die Methode wird Branch and Cut genannt, falls gute untere Schranken in den Zweigen (Kinder-Knoten) mit Hilfe von Cutting Planes berechnet werden (d.h. für die beste ganzzahlige Lösung in diesem Teilbaum)

(TSP wurde mit Branch & Cut gelöst für alle 24978 Städte in Schweden, und für alle 15112 Städte in Deutschland.

Für 1904711 Orte der Welt hat man mit Branch and Cut eine Lösung mit höchstens 0,05 % Fehler)