Approximationsalgorithmen

5. Vorlesung: SetCover via LP-Runden,
via Primal-Dual-Schema und via Dual Fitting

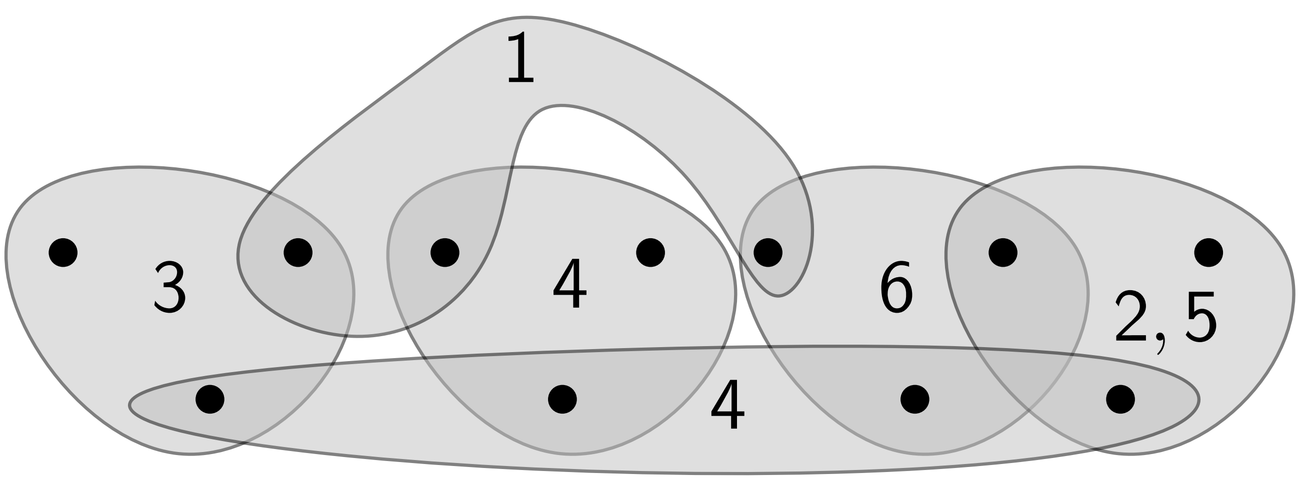
– Folien von J. Spoerhase und A. Wolff –

Alexander Wolff Lehrstuhl für Informatik I WS 2018/19

SetCover – Formulierung als ILP

Minimiere	$\sum_{S \in \mathcal{S}} c_S x_S$	
unter den NB.	$\sum_{S \ni e} x_S \geq 1$	$e \in U$
	$x_S \in \{0,1\}$	$S \in \mathcal{S}$

- Grundmenge U
- Familie $\mathcal{S} \subseteq 2^U$ mit $\bigcup \mathcal{S} = U$
- Kosten $c\colon \mathcal{S} \to \mathbb{Q}^+$



Finde Überdeckung $\mathcal{S}' \subseteq \mathcal{S}$ von U mit minimalen Kosten.

SetCover – LP-Relaxierung

$$\begin{array}{lll}
\text{Minimiere} & \displaystyle\sum_{S \in \mathcal{S}} c_S x_S & \\
\text{unter den NB.} & \displaystyle\sum_{S \ni e} x_S \geq 1 & e \in U \\
 & x_S \geq 0 & S \in \mathcal{S}
\end{array}$$

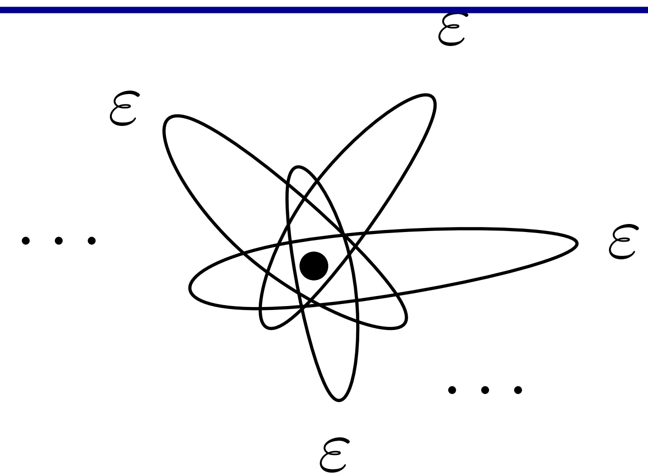
LP-Runden: Ansatz I

$$\begin{array}{lll} \text{Minimiere} & \displaystyle\sum_{S \in \mathcal{S}} c_S x_S & \\ \text{unter den NB.} & \displaystyle\sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{array}$$

LP-Runden-Eins$(U, \mathcal{S}, c)$
Ermittle optimale Lösung $\mathbf{x}$ für LP-Relaxierung.
Runde jedes x_S mit $x_S > 0$ auf 1.

- generiert zulässige Lösung
- Skalierungsfaktor beliebig groß

Benutze Häufigkeit h



LP-Runden: Ansatz II

$$\begin{array}{lll} \text{Minimiere} & \sum_{S \in \mathcal{S}} c_S x_S & \\ \text{unter den NB.} & \sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{array}$$

LP-Runden-Zwei($U, \mathcal{S}, c$)
Ermittle optimale Lösung $\mathbf{x}$ für LP-Relaxierung.
Runde jedes x_S mit $\boxed{x_S \geq 1/h}$ auf 1; restliche auf 0.

Satz. Obiger Algorithmus ist ein Faktor-h-Approximationsalgorithmus für $\textsc{SetCover}$.

Komplementärer Schlupf

Minimiere $\sum_{j=1}^{n} c_j x_j$	Maximiere $\sum_{i=1}^{m} b_i y_i$
unter den NB. $\sum_{j=1}^{n} a_{ij} x_j \geq b_i$	unter den NB. $\sum_{i=1}^{m} a_{ij} y_i \leq c_j$
$x_j \geq 0$	$y_i \geq 0$

Satz. Seien $\mathbf{x} = (x_1, \ldots, x_n)$ und $\mathbf{y} = (y_1, \ldots, y_m)$ zulässige Lösungen für das primale bzw. duale Programm.
Die Lösungen $\mathbf{x}$ und $\mathbf{y}$ sind genau dann optimal, wenn die folgenden Bedingungen erfüllt sind:

Primaler KS:
Für jedes $j = 1, \ldots, n$: entw. $x_j = 0$ od. $\sum_{i=1}^{m} a_{ij} y_i = c_j$.

Dualer KS:
Für jedes $i = 1, \ldots, m$: entw. $y_i = 0$ od. $\sum_{j=1}^{n} a_{ij} x_j = b_i$.

Relaxierter komplementärer Schlupf

Minimiere	$\sum_{j=1}^{n} c_j x_j$	Maximiere	$\sum_{i=1}^{m} b_i y_i$
unter den NB.	$\sum_{j=1}^{n} a_{ij} x_j \geq b_i$	unter den NB.	$\sum_{i=1}^{m} a_{ij} y_i \leq c_j$
	$x_j \geq 0$		$y_i \geq 0$

~~**Primaler KS**~~: relaxierter primaler KS

Für jedes $j = 1, \ldots, n$: Entweder $x_j = 0$ oder $\cancel{\sum_{i=1}^{m} a_{ij} y_i = c_j}$

$$c_j/\alpha \leq \sum_{i=1}^{m} a_{ij} y_i \leq c_j$$

~~**Dualer KS**~~: relaxierter dualer KS

Für jedes $i = 1, \ldots, m$: Entweder $y_i = 0$ oder $\cancel{\sum_{j=1}^{n} a_{ij} x_j = b_i}$

$$b_i \leq \sum_{j=1}^{n} a_{ij} x_j \leq \beta \cdot b_i$$

$$\cancel{\Leftrightarrow \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i} \qquad \Rightarrow \sum_{j=1}^{n} c_j x_j \leq \alpha\beta \sum_{i=1}^{m} b_i y_i \leq \alpha\beta \cdot \mathrm{OPT}_{\mathsf{LP}}$$

Primal-Dual-Ansatz

- Beginne mit zulässiger Dual- und unzulässiger Primal-Lösung (häufig trivial).
- „Verbessere“ die Zulässigkeit der Primal-Lösung...
- ...und simultan den Zielwert der Dual-Lösung.
- Gewährleiste dabei die relaxierten KS-Bedingungen.
- Erweitere Primal-Lösung immer nur ganzzahlig.
- Zulässigkeit der Primal-Lösung und relaxierte CS-Bedingungen führen zu Approximationsgüte.

Relaxierter KS für SetCover

Minimize $\sum_{S \in \mathcal{S}} c_S x_S$ subject to $\sum_{S \ni e} x_S \geq 1 \quad e \in U$ $x_S \geq 0 \quad S \in \mathcal{S}$	Maximize $\sum_{e \in U} y_e$ subject to $\sum_{e \in S} y_e \leq c_S \quad S \in \mathcal{S}$ $y_e \geq 0 \quad e \in U$
Primales LP	Duales LP

kritische Menge

(Unrelaxierter) primaler KS: $x_S \neq 0 \Rightarrow \sum_{e \in S} y_e = c_S$

wähle nur kritische Mengen

trivial für binäres $\mathbf{x}$

Relaxierter dualer KS: $y_e \neq 0 \Rightarrow 1 \leq \sum_{S \ni e} x_S \leq 1 \cdot h$

Primal-Dual-Schema für SetCover

PrimalDualSetCover($U, \mathcal{S}, c$)

$\mathbf{x} \leftarrow 0, \mathbf{y} \leftarrow 0$

repeat

Wähle ein noch unüberdecktes Element e.

Erhöhe y_e, bis eine Menge S kritisch ($\sum_{e' \in S} y_{e'} = c_S$).

Wähle alle kritischen Mengen und aktualisiere $\mathbf{x}$.

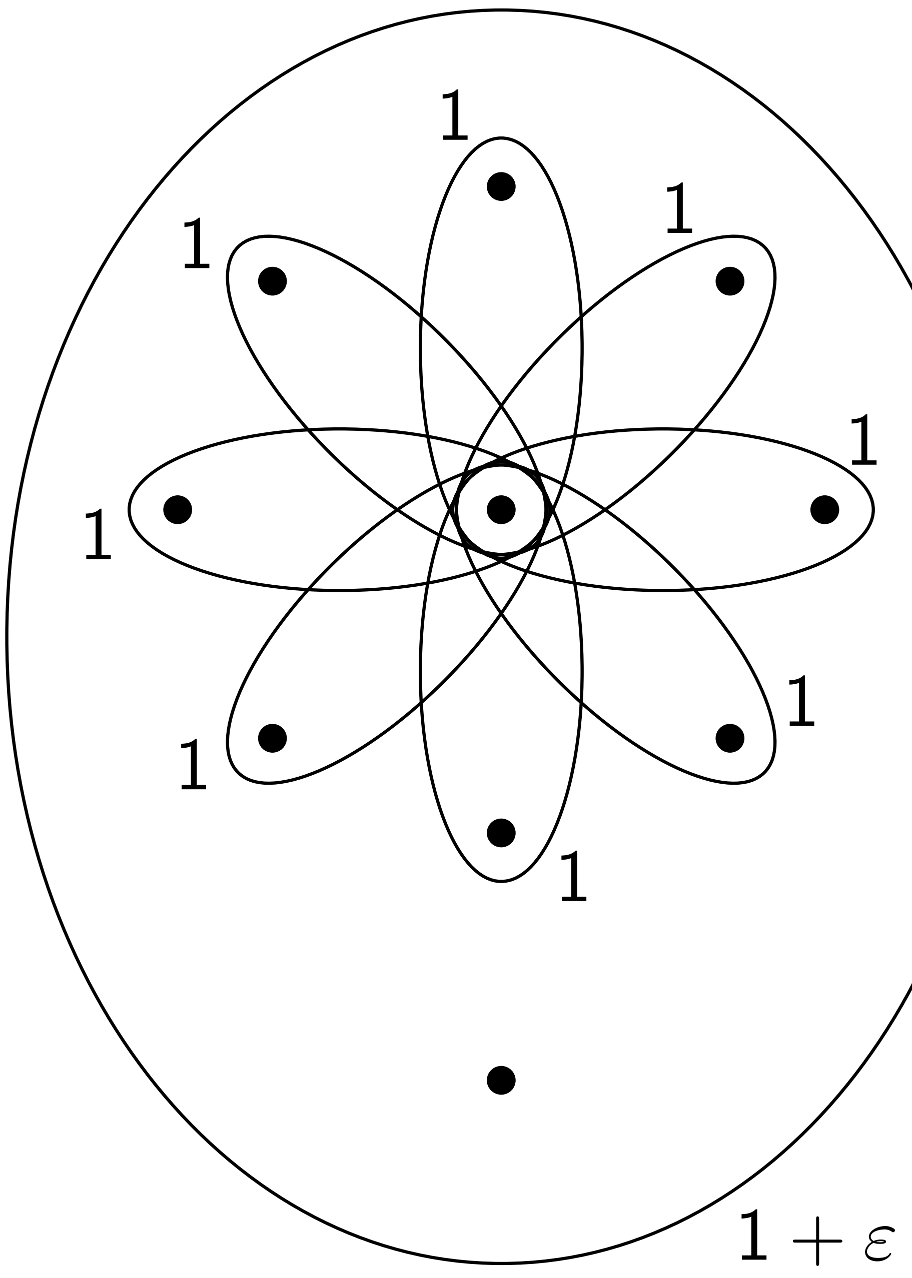
Markiere alle Elemente in diesen Mengen als überdeckt.

until alle Elemente sind überdeckt.

return $\mathbf{x}$

Satz. Obiger Algorithmus ist ein Faktor-h-Approximationsalgorithmus für SetCover. Diese Schranke ist scharf.

Scharfes Beispiel



Dual Fitting für SetCover

Kombinatorischer (Greedy-) Algorithmus (s. 2. Vorlesung):

GreedySetCover($U, \mathcal{S}, c$)
  $C \leftarrow \emptyset$
  $\mathcal{S}' \leftarrow \emptyset$
  while $C \neq U$ **do**
    $S \leftarrow$ Menge aus $\mathcal{S}$, die $\frac{c(S)}{|S \setminus C|}$ minimiert
    foreach $e \in S \setminus C$ **do**
      preis(e) $\leftarrow \frac{c(S)}{|S \setminus C|}$
    $C \leftarrow C \cup S$
    $\mathcal{S}' \leftarrow \mathcal{S}' \cup \{S\}$
  return $\mathcal{S}'$ // Überdeckung von U

Zur Erinnerung: $\sum_{e \in U}$ preis(e) bezahlt $\mathcal{S}'$ vollständig.

Neu: LP-basierte Analyse

Beob. Für jedes $e \in U$ ist preis(e) eine duale Variable.
Aber diese duale Lösung ist i.A. nicht zulässig.

ÜA: Konstruiere Instanz, wo manche S um den Faktor $\approx H_{|S|}$ überpackt werden.

Dual-Fitting-Trick:

Skaliere duale Variable, so dass keine Menge überpackt wird.

Nimm $y_e = \text{preis}(e)/H_n$.

Der Greedy-Alg. benützt *diese* dualen Variablen als untere Schranke für OPT.

Lemma.
Der Vektor $(y_e)_{e \in U}$ ist eine zulässige Lösung für das duale LP.

Duales LP

Maximize	$\sum_{e \in U} y_e$	
subject to	$\sum_{e \in S} y_e \le c_S$	$S \in \mathcal{S}$
	$y_e \geq 0$	$e \in U$

Beweis. Zu zeigen: Keine Menge wird durch $\mathbf{y}$ überpackt.
Sei $S \in \mathcal{S}$ und $k = |S|$.

Seien $e_1, \dots, e_k$ die Elemente von S –
in der Reihenfolge, in der Greedy sie überdeckt.

Betrachte Iteration, in der e_i überdeckt wird.

Kurz davor sind $\geq k - i + 1$ Elem. von S unüberdeckt.

Also ist $\text{preis}(e_i) \leq c(S)/(k - i + 1)$.

$$\Rightarrow y_{e_i} \leq \frac{c(S)}{H_n} \cdot \frac{1}{k-i+1} \Rightarrow \sum_{i=1}^{k} y_{e_i} \leq \frac{c(S)}{H_n} \cdot \overbrace{\left(\frac{1}{k} + \dots + \frac{1}{1}\right)}^{H_k} \leq c(S)$$

$\square$

Lemma.
Der Vektor $\mathbf{y} = (y_e)_{e \in U}$ ist eine zulässige Lösung für das duale LP.

Duales LP

Maximize	$\sum_{e \in U} y_e$	
subject to	$\sum_{e \in S} y_e \leq c_S$	$S \in \mathcal{S}$
	$y_e \geq 0$	$e \in U$

Resultat von Dual Fitting

Satz. Der Approximationsfaktor von GreedySetCover ist H_n, wobei $n = |U|$.

Beweis.
$$\begin{aligned}\mathsf{ALG} = c(\mathcal{S}') \le \textstyle\sum_{e \in U} \mathsf{price}(e) = H_n \cdot \textstyle\sum_{e \in U} y_e &\le \\ &\le H_n \cdot \mathsf{OPT}_f \\ &\le H_n \cdot \mathsf{OPT}\end{aligned}$$
$\square$