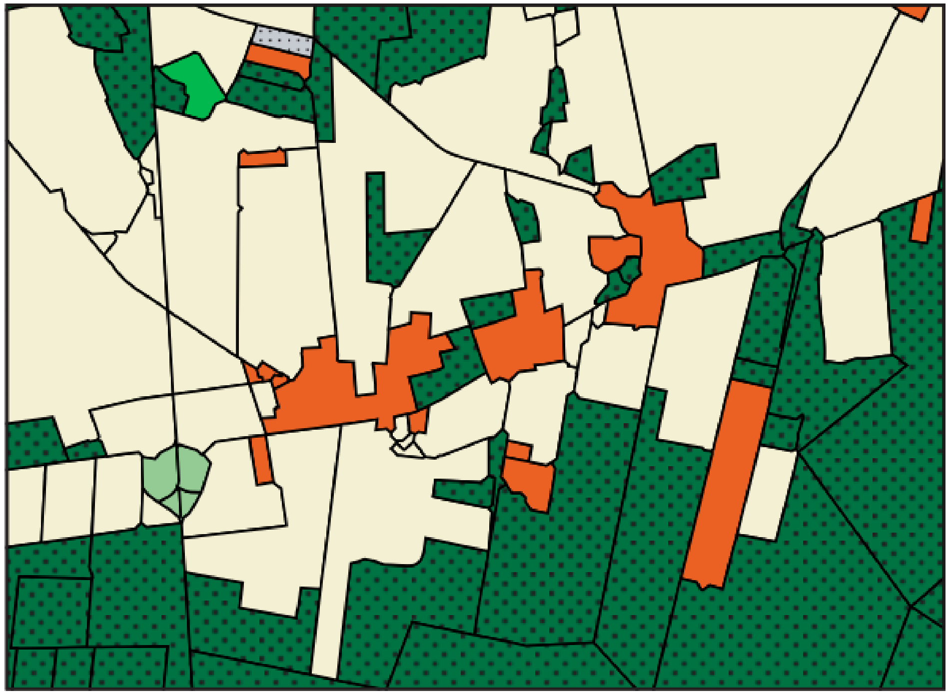
Vorlesung Algorithmische Kartografie
Flächenaggregation

LEHRSTUHL FÜR ALGORITHMIK I · INSTITUT FÜR THEORETISCHE INFORMATIK · FAKULTÄT FÜR INFORMATIK

Benjamin Niedermann · Martin Nöllenburg

30.04.2015/05.05.2015

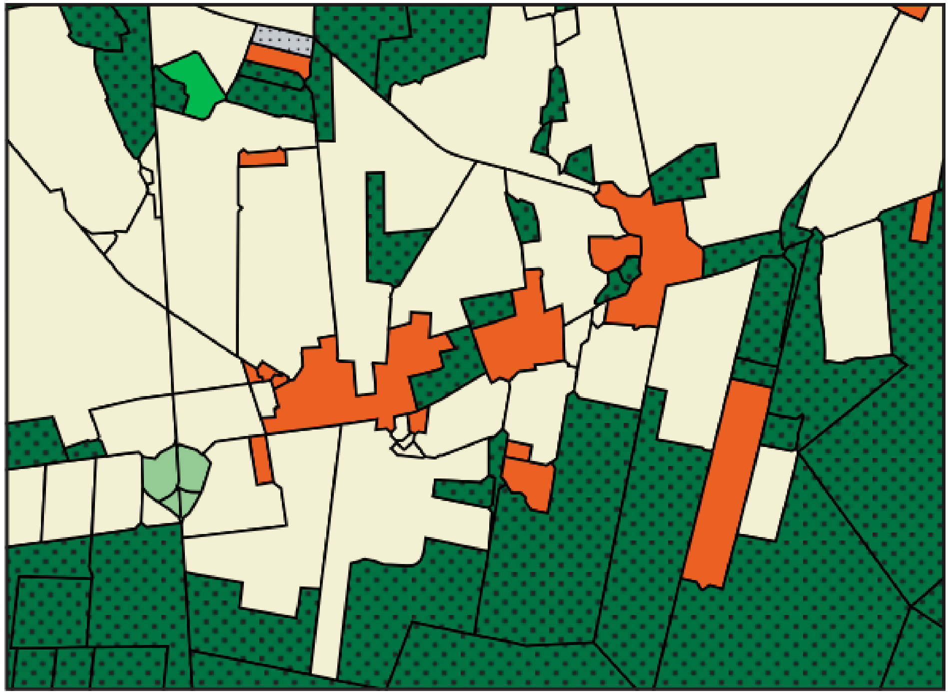
Flächenaggregation



Flächennutzung Maßstab 1:50.000

- Grassland
- Forest
- Farmland
- Settlement
- Industry

Flächenaggregation



Flächennutzung Maßstab 1:50.000

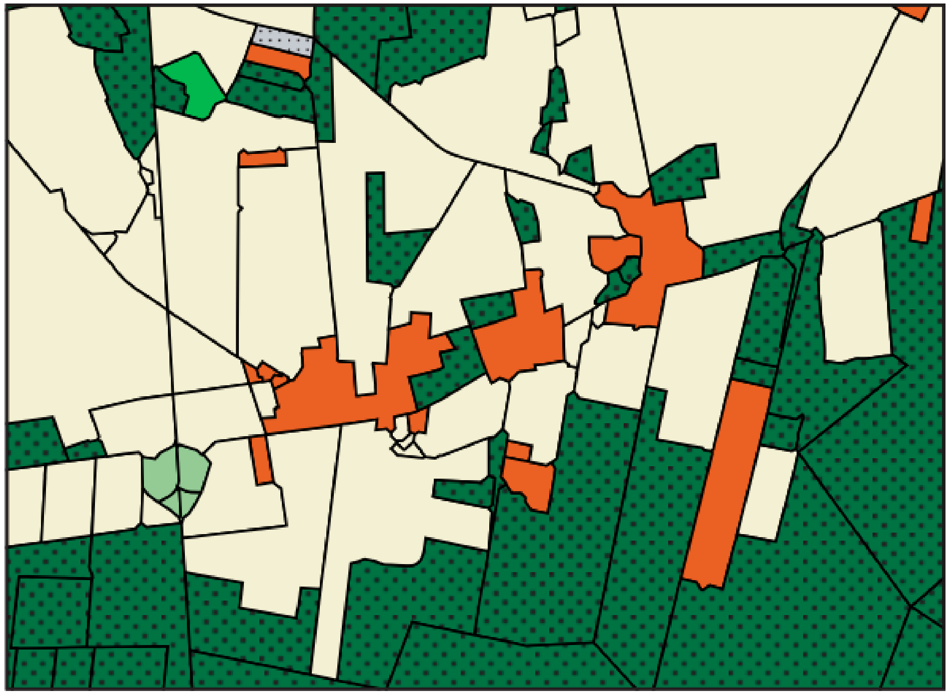
- Grassland
- Forest
- Farmland
- Settlement
- Industry

direktes Skalieren

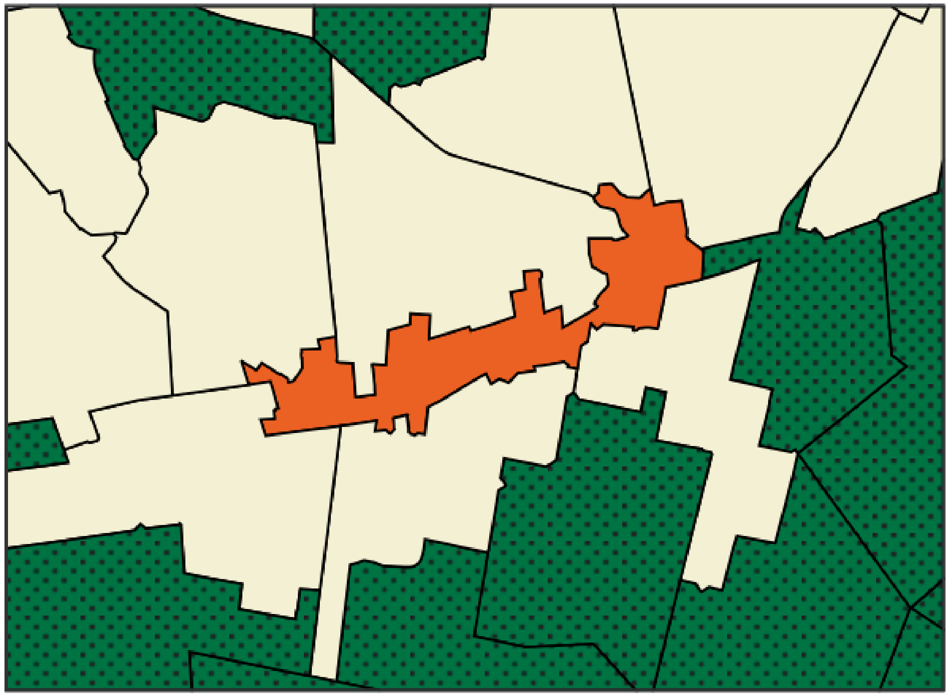
?

Flächennutzung Maßstab 1:250.000

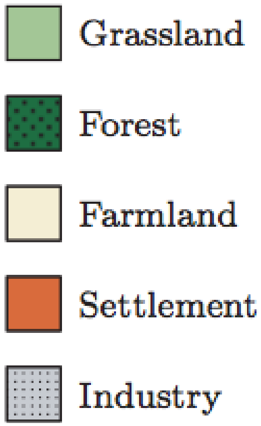
Flächenaggregation



Aggregation

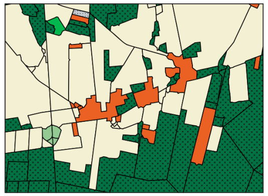


Flächennutzung Maßstab 1:50.000



direktes Skalieren

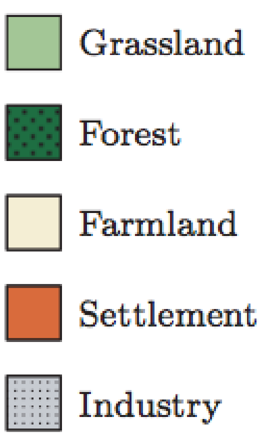
?



Flächennutzung Maßstab 1:250.000

Flächenaggregation

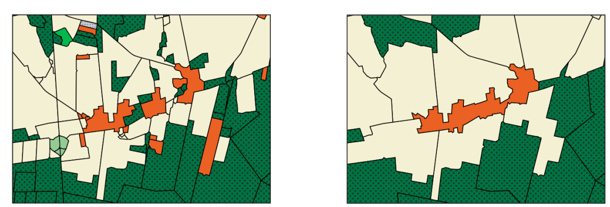
Flächennutzung Maßstab 1:50.000



direktes Skalieren

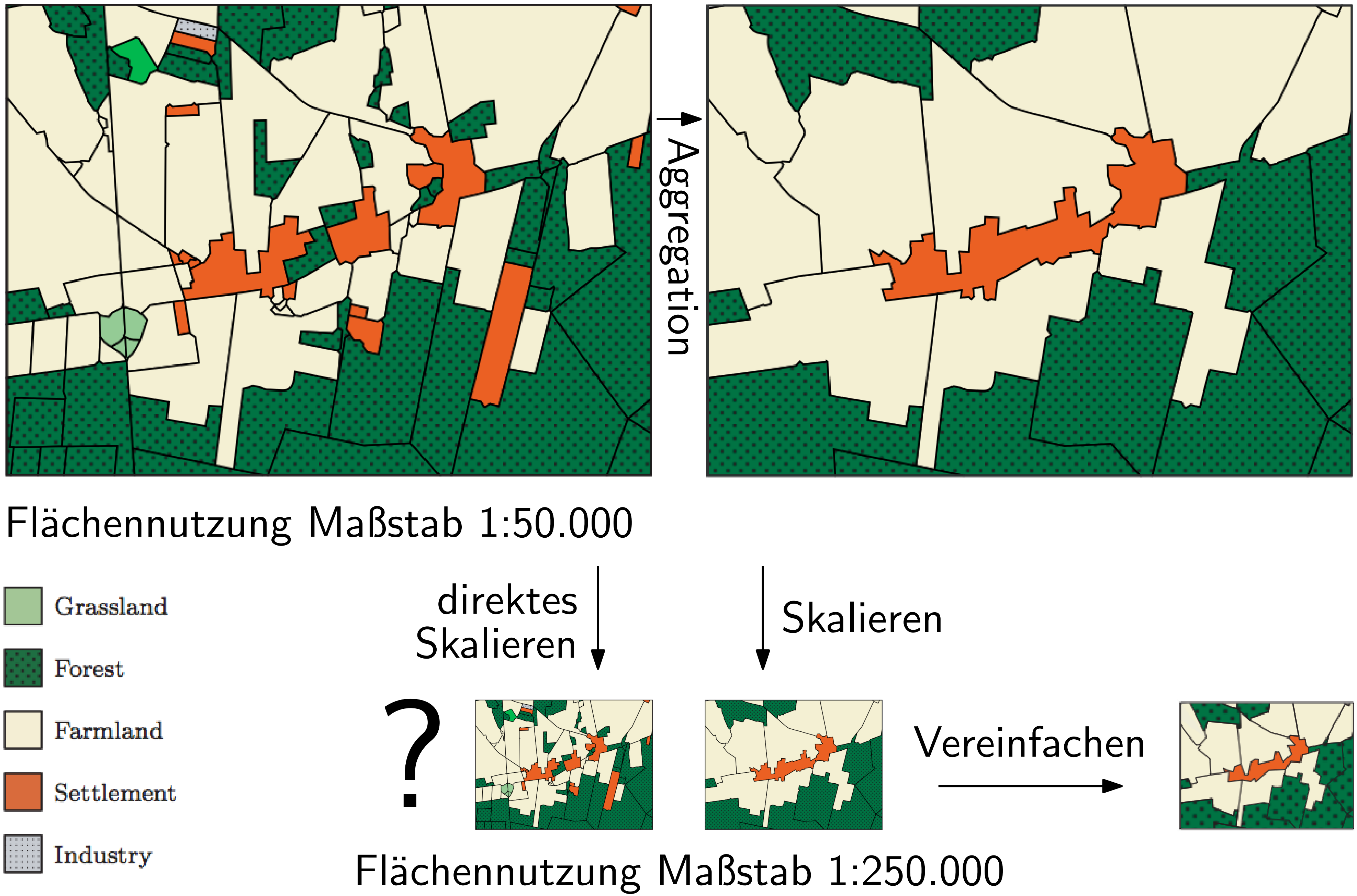
Skalieren

?

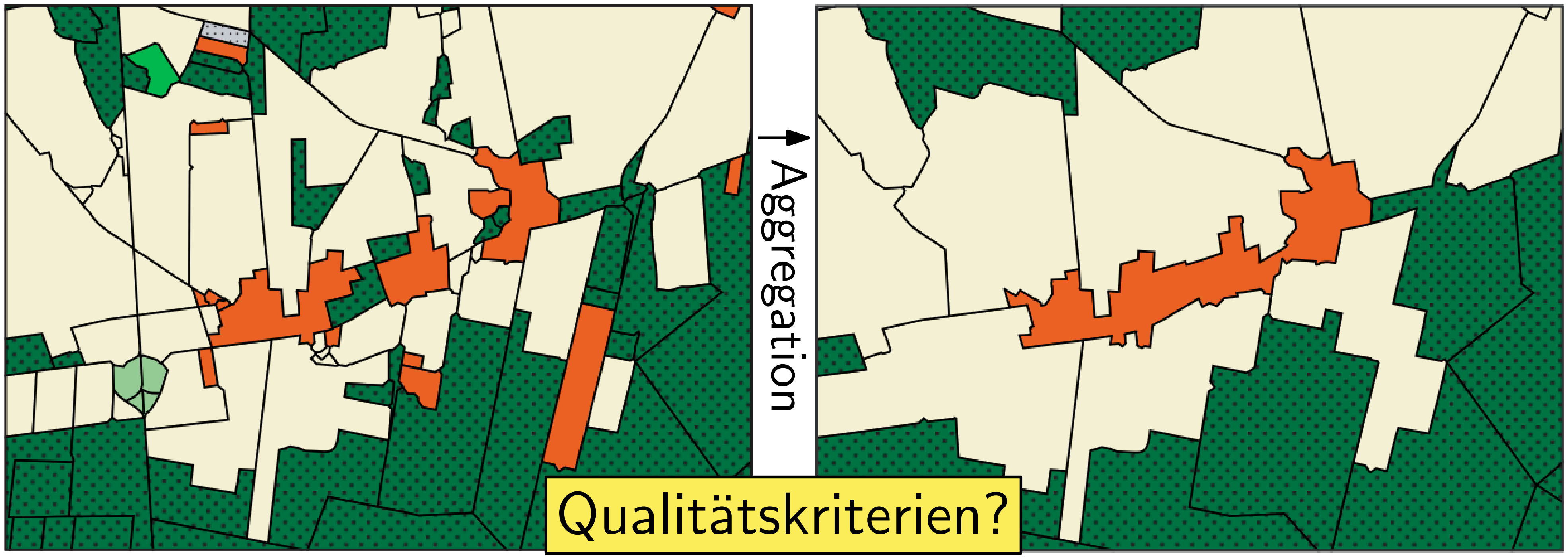


Flächennutzung Maßstab 1:250.000

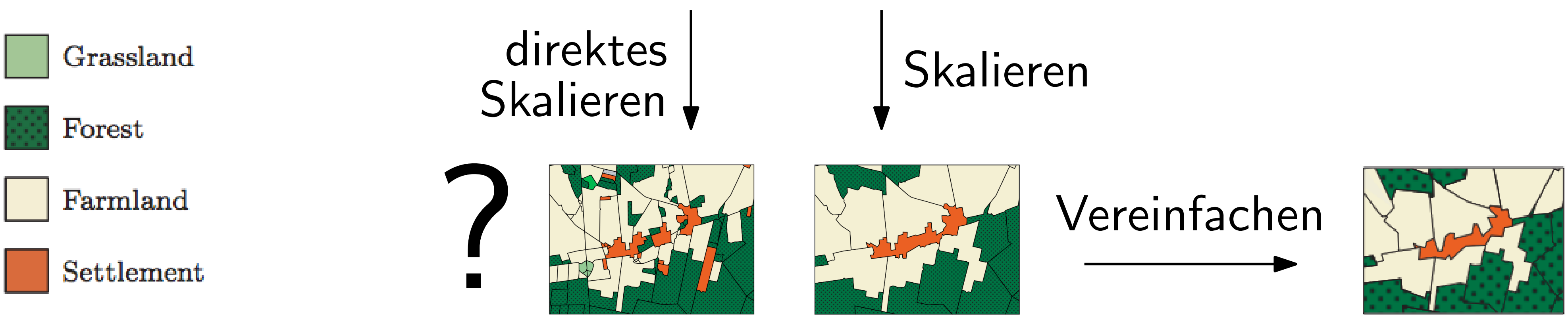
Flächenaggregation



Flächenaggregation



Flächennutzung Maßstab 1:50.000



Flächennutzung Maßstab 1:250.000

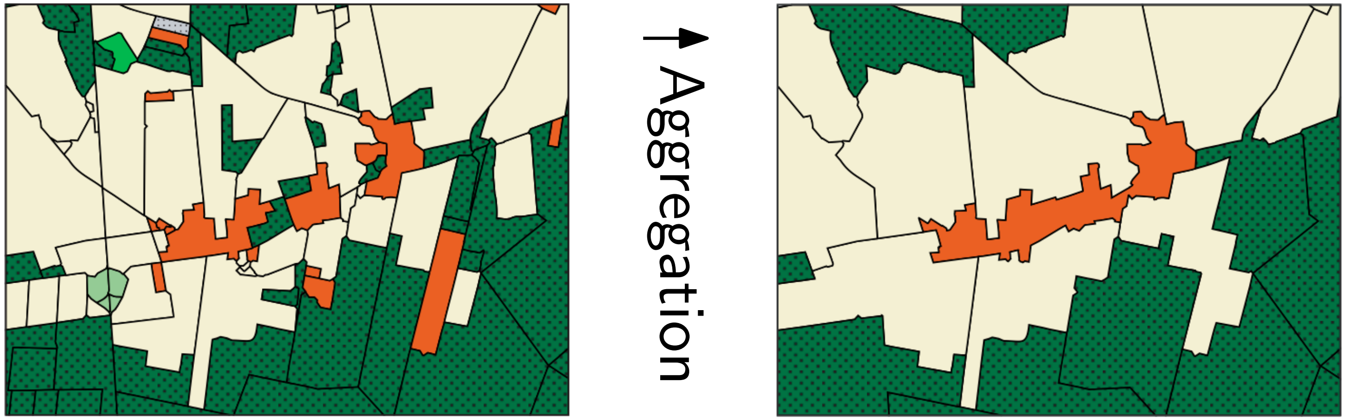
Aufgabenstellung

Daten: topographische Datenbank mit Flächennutzungsdaten

- Unterteilung der Ebene in Polygone bestimmter Nutzungskategorien
- z.B.: Siedlung, Gewerbe, Ackerland, Wald, ...

Ziel: generiere *sinnvolle* Karte in Maßstab 1:Z

- dargestellte Flächen überschreiten Mindestgröße
- dazu Aggregation benachbarter Flächen nötig
- möglichst kleine Kategorienänderungen
- möglichst kompakte Flächen



Typische Heuristik

Algorithmus RegionGrowing

$S \leftarrow$ Menge der Regionen kleiner als Mindestfläche
while $S \neq \emptyset$ **do**
  $a \leftarrow$ kleinste Region in S
  verschmelze a mit passendstem Nachbarn
  update S

Typische Heuristik

Algorithmus RegionGrowing

$S \leftarrow$ Menge der Regionen kleiner als Mindestfläche
while $S \neq \emptyset$ **do**
  $a \leftarrow$ kleinste Region in S
  verschmelze a mit passendstem Nachbarn
  update S

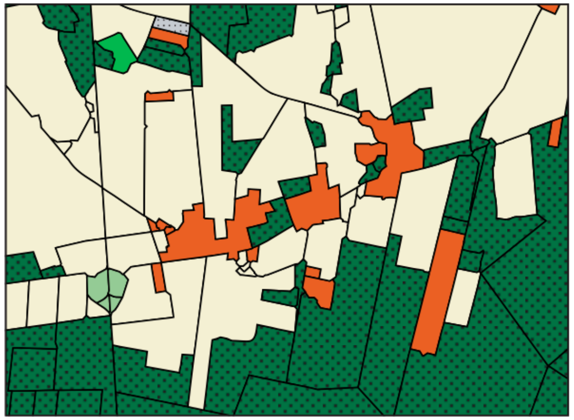
Vor- und Nachteile?

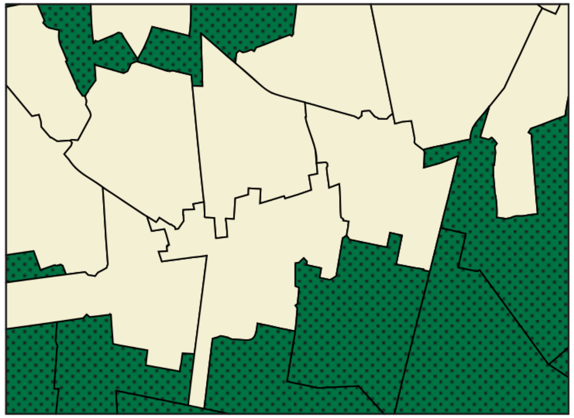
Typische Heuristik

Algorithmus RegionGrowing

$S \leftarrow$ Menge der Regionen kleiner als Mindestfläche
while $S \neq \emptyset$ **do**
  $a \leftarrow$ kleinste Region in S
  verschmelze a mit passendstem Nachbarn
  update S

Vor- und Nachteile?





- einfach zu implementieren
- Mindestflächen werden eingehalten
- nur lokale greedy Entscheidungen
- wichtige Features können verschwinden
- keine Qualitätsgarantie für Änderung der Kategorien

Qualitätskriterien

1) Logische Konsistenz

- Ausgabe ist Unterteilung der Ebene
- Einhaltung von Mindestflächen, ggf. kategorienabhängig (z.B. Wald $\geq$ 40 ha für 1:250K)

Qualitätskriterien

1) Logische Konsistenz

- Ausgabe ist Unterteilung der Ebene
- Einhaltung von Mindestflächen, ggf. kategorienabhängig (z.B. Wald $\geq$ 40 ha für 1:250K)

2) Vollständigkeit

- jede Eingaberegion, die groß genug ist muss korrekt in ihrer Kategorie dargestellt sein

Qualitätskriterien

1) Logische Konsistenz

- Ausgabe ist Unterteilung der Ebene
- Einhaltung von Mindestflächen, ggf. kategorienabhängig (z.B. Wald $\geq$ 40 ha für 1:250K)

2) Vollständigkeit

- jede Eingaberegion, die groß genug ist muss korrekt in ihrer Kategorie dargestellt sein

3) Semantische Genauigkeit

- semantische Distanz $d\colon \Gamma \times \Gamma \to \mathbb{R}_0^+$ zwischen Kategorien
- minimiere

$$\overline{d} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v)) / \sum_{v \in V} w(v)$$

Qualitätskriterien

1) Logische Konsistenz

- Ausgabe ist Unterteilung der Ebene
- Einhaltung von Mindestflächen, ggf. kategorienabhängig (z.B. Wald $\geq$ 40 ha für 1:250K)

2) Vollständigkeit

- jede Eingaberegion, die groß genug ist muss korrekt in ihrer Kategorie dargestellt sein

Γ = Menge der Kategorien

3) Semantische Genauigkeit

- semantische Distanz $d\colon \Gamma \times \Gamma \to \mathbb{R}_0^+$ zwischen Kategorien
- minimiere

$$\overline{d} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v)) / \sum_{v \in V} w(v)$$

Qualitätskriterien

1) Logische Konsistenz

- Ausgabe ist Unterteilung der Ebene
- Einhaltung von Mindestflächen, ggf. kategorienabhängig (z.B. Wald $\geq$ 40 ha für 1:250K)

2) Vollständigkeit

- jede Eingaberegion, die groß genug ist muss korrekt in ihrer Kategorie dargestellt sein

Γ = Menge der Kategorien

3) Semantische Genauigkeit

- semantische Distanz $d\colon \Gamma \times \Gamma \to \mathbb{R}_0^+$ zwischen Kategorien
- minimiere

$$\overline{d} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v)) / \sum_{v \in V} w(v)$$

V = Menge der Regionen

Qualitätskriterien

1) Logische Konsistenz

- Ausgabe ist Unterteilung der Ebene
- Einhaltung von Mindestflächen, ggf. kategorienabhängig (z.B. Wald $\geq$ 40 ha für 1:250K)

2) Vollständigkeit

- jede Eingaberegion, die groß genug ist muss korrekt in ihrer Kategorie dargestellt sein

Γ = Menge der Kategorien

3) Semantische Genauigkeit

- semantische Distanz $d\colon \Gamma \times \Gamma \to \mathbb{R}_0^+$ zwischen Kategorien
- minimiere

$$\overline{d} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v)) / \sum_{v \in V} w(v)$$

V = Menge der Regionen

$\gamma\colon V \to \Gamma$ = Kategorie einer Region

Qualitätskriterien

1) Logische Konsistenz

- Ausgabe ist Unterteilung der Ebene
- Einhaltung von Mindestflächen, ggf. kategorienabhängig (z.B. Wald $\geq$ 40 ha für 1:250K)

2) Vollständigkeit

- jede Eingaberegion, die groß genug ist muss korrekt in ihrer Kategorie dargestellt sein

Γ = Menge der Kategorien

3) Semantische Genauigkeit

- semantische Distanz $d \colon \Gamma \times \Gamma \to \mathbb{R}_0^+$ zwischen Kategorien
- minimiere

$$\overline{d} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v)) / \sum_{v \in V} w(v)$$

V = Menge der Regionen

$w \colon V \to \mathbb{R}_0^+$ = Fläche einer Region

$\gamma \colon V \to \Gamma$ = Kategorie einer Region

Qualitätskriterien

1) Logische Konsistenz

- Ausgabe ist Unterteilung der Ebene
- Einhaltung von Mindestflächen, ggf. kategorienabhängig (z.B. Wald $\geq$ 40 ha für 1:250K)

2) Vollständigkeit

- jede Eingaberegion, die groß genug ist muss korrekt in ihrer Kategorie dargestellt sein

3) Semantische Genauigkeit

- semantische Distanz $d\colon \Gamma \times \Gamma \to \mathbb{R}_0^+$ zwischen Kategorien
- minimiere

$$\overline{d} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v)) / \sum_{v \in V} w(v)$$

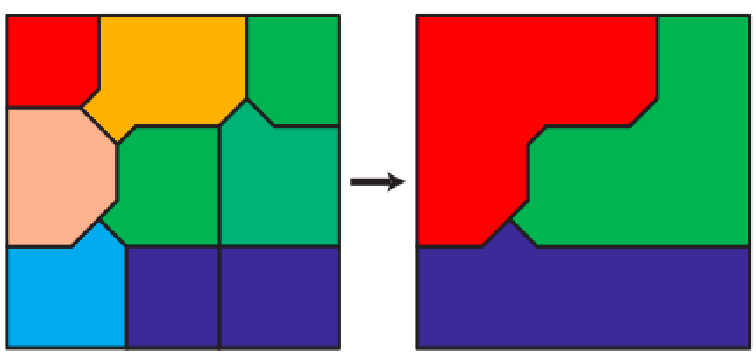
4) Kompaktheit

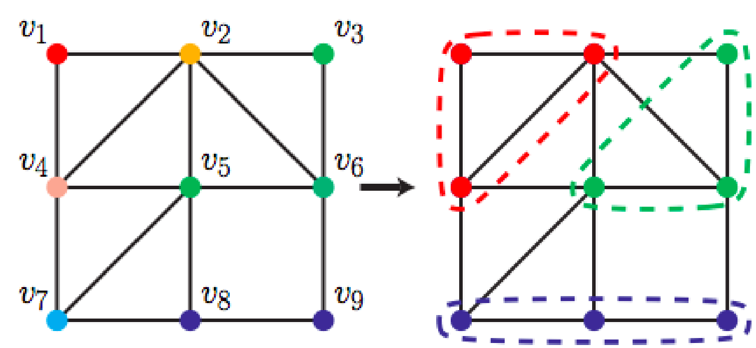
- aggregierte Regionen sollen möglichst kompakt sein, gemessen an einem Kompaktheitsmaß $c\colon 2^V \times \Gamma \to \mathbb{R}_0^+$

Formale Problemstellung AreaAggregation

Geg:
- planarer Graph $G = (V, E)$ mit Knotengewichten $w \colon V \to \mathbb{R}^+$ und Knotenfärbung $\gamma \colon V \to \Gamma$
- Mindestgewicht jeder Farbe $\theta \colon \Gamma \to \mathbb{R}^+$
- semantischer Abstand $d \colon \Gamma \times \Gamma \to \mathbb{R}_0^+$
- (Nicht-)Kompaktheit $c \colon 2^V \times \Gamma \to \mathbb{R}_0^+$
- Parameter $s \in [0, 1]$

Ges: Knotenfärbung $\gamma' \colon V \to \Gamma$ und Partitionierung $P = \{V_1, \dots, V_p\}$ von V mit
- $\gamma'(v) = \gamma'_i$ für alle $v \in V_i$
- $\sum_{v \in V_i} w(v) \geq \theta(\gamma'_i)$
- induzierter Graph $G[V_i]$ ist zusammenhängend
- $\exists$ Knoten $v \in V_i$, der seine Farbe behält, d.h. $\gamma(v) = \gamma'(v)$



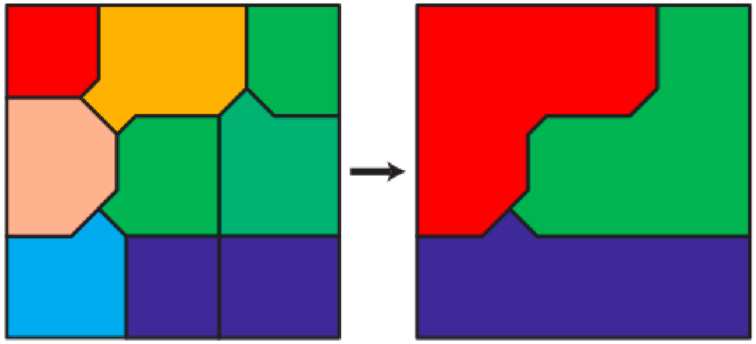


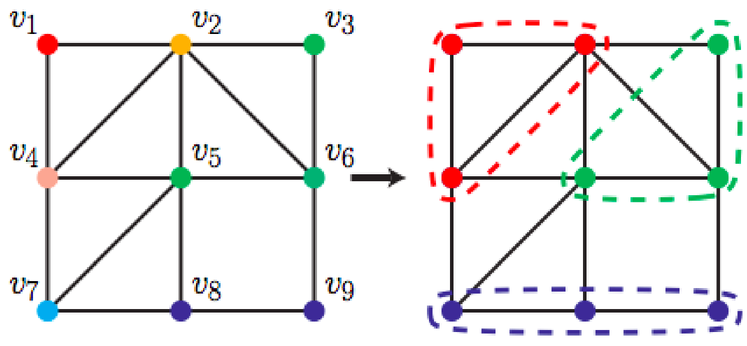
Formale Problemstellung AREAAGGREGATION

Geg:
- planarer Graph $G = (V, E)$ mit Knotengewichten $w: V \to \mathbb{R}^+$ und Knotenfärbung $\gamma: V \to \Gamma$
- Mindestgewicht jeder Farbe $\theta: \Gamma \to \mathbb{R}^+$
- semantischer Abstand $d: \Gamma \times \Gamma \to \mathbb{R}_0^+$
- (Nicht-)Kompaktheit $c: 2^V \times \Gamma \to \mathbb{R}_0^+$
- Parameter $s \in [0, 1]$

Ges: Knotenfärbung $\gamma': V \to \Gamma$ und Partitionierung $P = \{V_1, \dots, V_p\}$ von V mit
- $\gamma'(v) = \gamma'_i$ für alle $v \in V_i$
- $\sum_{v \in V_i} w(v) \geq \theta(\gamma'_i)$
- induzierter Graph $G[V_i]$ ist zusammenhängend
- $\exists$ Knoten $v \in V_i$, der seine Farbe behält, d.h. $\gamma(v) = \gamma'(v)$

Alle Knoten einer Partition gehören zur selben Kategorie.



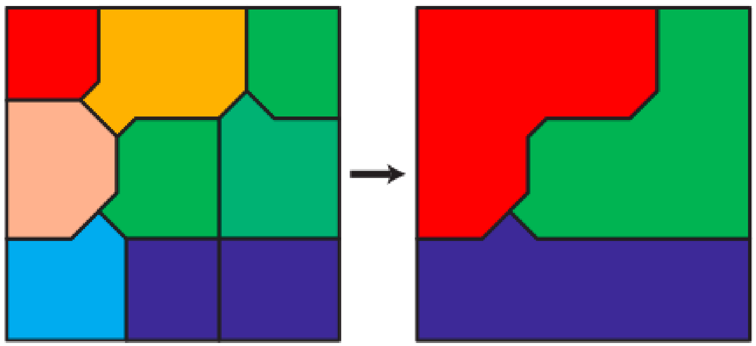


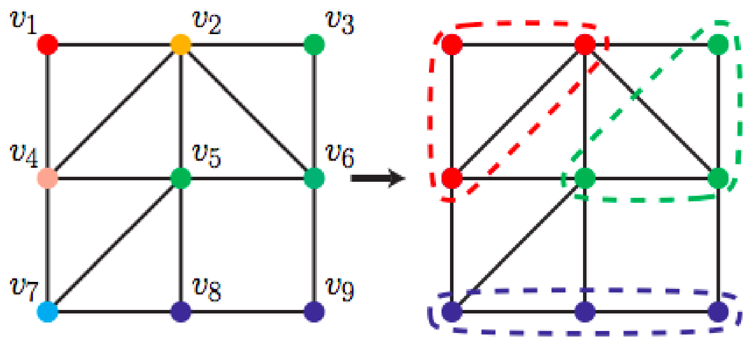
Formale Problemstellung AreaAggregation

Geg:
- planarer Graph $G = (V, E)$ mit Knotengewichten $w \colon V \to \mathbb{R}^+$ und Knotenfärbung $\gamma \colon V \to \Gamma$
- Mindestgewicht jeder Farbe $\theta \colon \Gamma \to \mathbb{R}^+$
- semantischer Abstand $d \colon \Gamma \times \Gamma \to \mathbb{R}_0^+$
- (Nicht-)Kompaktheit $c \colon 2^V \times \Gamma \to \mathbb{R}_0^+$
- Parameter $s \in [0, 1]$

Ges: Knotenfärbung $\gamma' \colon V \to \Gamma$ und Partitionierung $P = \{V_1, \ldots, V_p\}$ von V mit
- $\gamma'(v) = \gamma'_i$ für alle $v \in V_i$
- $\sum_{v \in V_i} w(v) \geq \theta(\gamma'_i)$
- induzierter Graph $G[V_i]$ ist zusammenhängend
- $\exists$ Knoten $v \in V_i$, der seine Farbe behält, d.h. $\gamma(v) = \gamma'(v)$

Partition bestitzt geforderte Mindestgewicht (Fläche).





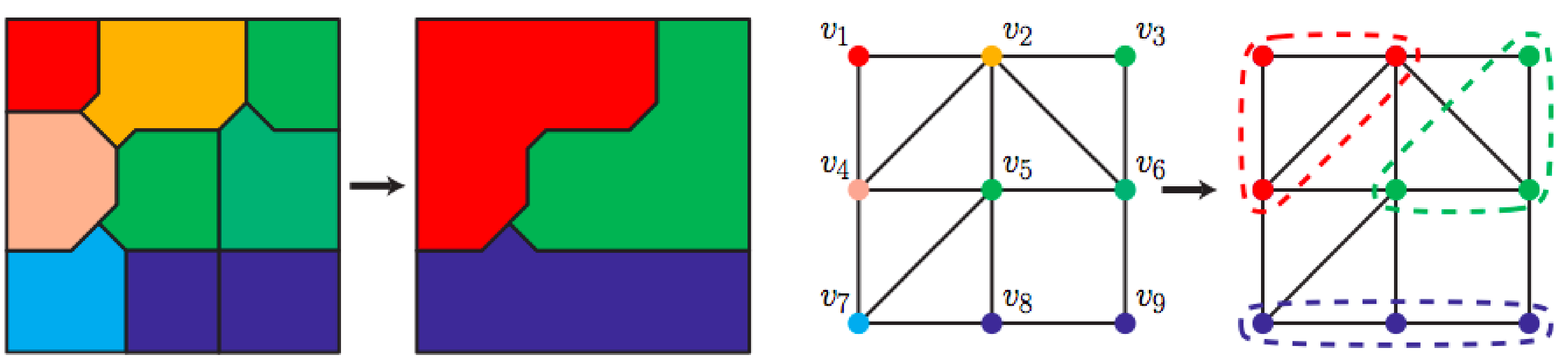
Formale Problemstellung AreaAggregation

Karlsruhe Institute of Technology

Geg:
- planarer Graph $G = (V, E)$ mit Knotengewichten $w\colon V \to \mathbb{R}^+$ und Knotenfärbung $\gamma\colon V \to \Gamma$
- Mindestgewicht jeder Farbe $\theta\colon \Gamma \to \mathbb{R}^+$
- semantischer Abstand $d\colon \Gamma \times \Gamma \to \mathbb{R}_0^+$
- (Nicht-)Kompaktheit $c\colon 2^V \times \Gamma \to \mathbb{R}_0^+$
- Parameter $s \in [0, 1]$

Ges: Knotenfärbung $\gamma'\colon V \to \Gamma$ und Partitionierung $P = \{V_1, \dots, V_p\}$ von V mit
- $\gamma'(v) = \gamma'_i$ für alle $v \in V_i$
- $\sum_{v \in V_i} w(v) \geq \theta(\gamma'_i)$
- induzierter Graph $G[V_i]$ ist zusammenhängend
- $\exists$ Knoten $v \in V_i$, der seine Farbe behält, d.h. $\gamma(v) = \gamma'(v)$

Partition repräsentiert zusammenhängende Fläche.

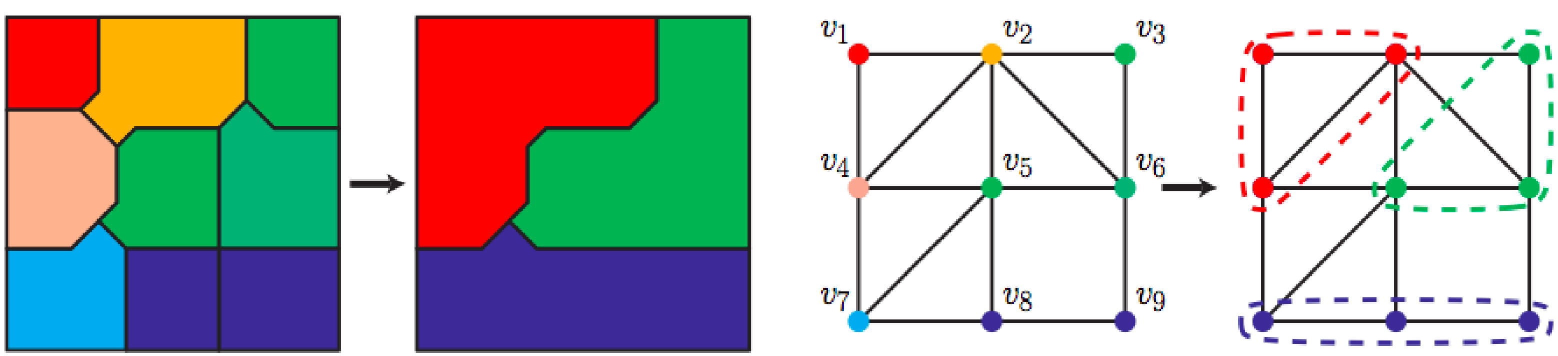


Formale Problemstellung AREAAGGREGATION

Geg:
- planarer Graph $G = (V, E)$ mit Knotengewichten $w \colon V \to \mathbb{R}^+$ und Knotenfärbung $\gamma \colon V \to \Gamma$
- Mindestgewicht jeder Farbe $\theta \colon \Gamma \to \mathbb{R}^+$
- semantischer Abstand $d \colon \Gamma \times \Gamma \to \mathbb{R}_0^+$
- (Nicht-)Kompaktheit $c \colon 2^V \times \Gamma \to \mathbb{R}_0^+$
- Parameter $s \in [0, 1]$

Ges: Knotenfärbung $\gamma' \colon V \to \Gamma$ und Partitionierung $P = \{V_1, \ldots, V_p\}$ von V mit
- $\gamma'(v) = \gamma'_i$ für alle $v \in V_i$
- $\sum_{v \in V_i} w(v) \geq \theta(\gamma'_i)$
- induzierter Graph $G[V_i]$ ist zusammenhängend
- $\exists$ Knoten $v \in V_i$, der seine Farbe behält, d.h. $\gamma(v) = \gamma'(v)$

Partition besitzt Region, die nicht Kategorie gewechselt hat.



Formale Problemstellung AreaAggregation

Geg:
- planarer Graph $G = (V, E)$ mit Knotengewichten $w\colon V \to \mathbb{R}^+$ und Knotenfärbung $\gamma\colon V \to \Gamma$
- Mindestgewicht jeder Farbe $\theta\colon \Gamma \to \mathbb{R}^+$
- semantischer Abstand $d\colon \Gamma \times \Gamma \to \mathbb{R}_0^+$
- (Nicht-)Kompaktheit $c\colon 2^V \times \Gamma \to \mathbb{R}_0^+$
- Parameter $s \in [0, 1]$

Ges: Knotenfärbung $\gamma'\colon V \to \Gamma$ und Partitionierung $P = \{V_1, \ldots, V_p\}$ von V mit
- $\gamma'(v) = \gamma'_i$ für alle $v \in V_i$
- $\sum_{v \in V_i} w(v) \geq \theta(\gamma'_i)$
- induzierter Graph $G[V_i]$ ist zusammenhängend
- $\exists$ Knoten $v \in V_i$, der seine Farbe behält, d.h. $\gamma(v) = \gamma'(v)$

minimiere dabei
- Kosten $f = s \cdot f_{\text{col}} + (1 - s) \cdot f_{\text{comp}}$
- $f_{\text{col}} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v))$
- $f_{\text{comp}} = \sum_{V_i \in P} c(V_i, \gamma'_i)$

Formale Problemstellung AreaAggregation

Geg:
- planarer Graph $G = (V, E)$ mit Knotengewichten $w \colon V \to \mathbb{R}^+$ und Knotenfärbung $\gamma \colon V \to \Gamma$
- Mindestgewicht jeder Farbe $\theta \colon \Gamma \to \mathbb{R}^+$
- semantischer Abstand $d \colon \Gamma \times \Gamma \to \mathbb{R}_0^+$
- (Nicht-)Kompaktheit $c \colon 2^V \times \Gamma \to \mathbb{R}_0^+$
- Parameter $s \in [0, 1]$

Ges: Knotenfärbung $\gamma' \colon V \to \Gamma$ und Partitionierung $P = \{V_1, \dots, V_p\}$ von V mit
- $\gamma'(v) = \gamma'_i$ für alle $v \in V_i$
- $\sum_{v \in V_i} w(v) \geq \theta(\gamma'_i)$
- induzierter Graph $G[V_i]$ ist zusammenhängend
- $\exists$ Knoten $v \in V_i$, der seine Farbe behält, d.h. $\gamma(v) = \gamma'(v)$

minimiere dabei
- Kosten $f = s \cdot f_{\text{col}} + (1 - s) \cdot f_{\text{comp}}$
- $f_{\text{col}} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v))$
- $f_{\text{comp}} = \sum_{V_i \in P} c(V_i, \gamma'_i)$

Bestrafe Kategorienwechsel.

Formale Problemstellung AreaAggregation

Geg:
- planarer Graph $G = (V, E)$ mit Knotengewichten $w\colon V \to \mathbb{R}^+$ und Knotenfärbung $\gamma\colon V \to \Gamma$
- Mindestgewicht jeder Farbe $\theta\colon \Gamma \to \mathbb{R}^+$
- semantischer Abstand $d\colon \Gamma \times \Gamma \to \mathbb{R}_0^+$
- (Nicht-)Kompaktheit $c\colon 2^V \times \Gamma \to \mathbb{R}_0^+$
- Parameter $s \in [0, 1]$

Ges: Knotenfärbung $\gamma'\colon V \to \Gamma$ und Partitionierung $P = \{V_1, \dots, V_p\}$ von V mit
- $\gamma'(v) = \gamma'_i$ für alle $v \in V_i$
- $\sum_{v \in V_i} w(v) \geq \theta(\gamma'_i)$
- induzierter Graph $G[V_i]$ ist zusammenhängend
- $\exists$ Knoten $v \in V_i$, der seine Farbe behält, d.h. $\gamma(v) = \gamma'(v)$

minimiere dabei
- Kosten $f = s \cdot f_{\text{col}} + (1 - s) \cdot f_{\text{comp}}$
- $f_{\text{col}} = \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v))$
- $f_{\text{comp}} = \sum_{V_i \in P} c(V_i, \gamma'_i)$

Bestrafe nicht-kompakte Partitionen

Kompaktheit

Umfangsbasiert

- je kleiner der Umfang einer Region desto kompakter
- $c_p(V') = \text{Umfang}(V')$
- präferiert in der Summe über P größere Regionen
- Alternativen wie $c_p/(2\sqrt{\pi\text{Fläche}})$ oder $\sqrt{\text{Fläche}} \cdot c_p/2\sqrt{\pi}$ sind problematisch für lineare Optimierung

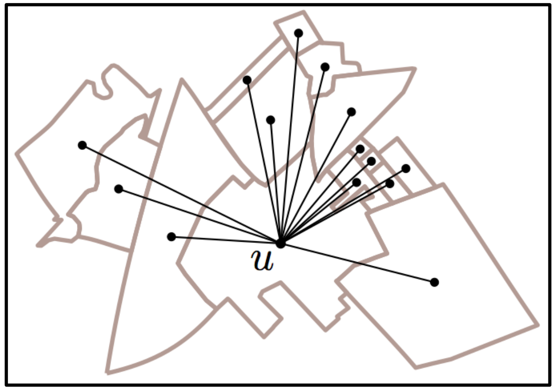
Kompaktheit

Umfangsbasiert

- je kleiner der Umfang einer Region desto kompakter
- $c_p(V') = \text{Umfang}(V')$
- präferiert in der Summe über P größere Regionen
- Alternativen wie $c_p/(2\sqrt{\pi \text{Fläche}})$ oder $\sqrt{\text{Fläche}} \cdot c_p/2\sqrt{\pi}$ sind problematisch für lineare Optimierung

Distanz zum Zentrum

- sei $\delta \colon V \times V \to \mathbb{R}_0^+$ Euklidischer Abstand zwischen Schwerpunkten
- $c_d(V', k) = \min\{\sum_{v \in V'} w(v) \cdot \delta(v, u) \mid u \in V' \text{ und } \gamma(u) = k)\}$
- Flächen sammeln sich rund um Zentrum mit unverändertem Typ
- präferiert kleinere Regionen, jedoch beschränkt durch θ
- Kombination mit c_p möglich



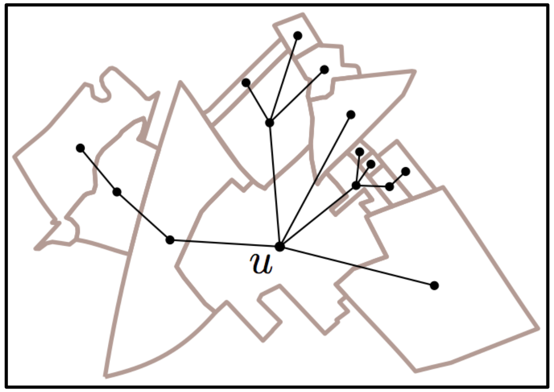
Kompaktheit

Umfangsbasiert

- je kleiner der Umfang einer Region desto kompakter
- $c_p(V') = \text{Umfang}(V')$
- präferiert in der Summe über P größere Regionen
- Alternativen wie $c_p/(2\sqrt{\pi \text{Fläche}})$ oder $\sqrt{\text{Fläche}} \cdot c_p/2\sqrt{\pi}$ sind problematisch für lineare Optimierung

Distanz zum Zentrum

- sei $\delta \colon V \times V \to \mathbb{R}_0^+$ Euklidischer Abstand zwischen Schwerpunkten
- $c_d(V', k) = \min\{\sum_{v \in V'} w(v) \cdot \delta(v, u) \mid u \in V' \text{ und } \gamma(u) = k)\}$
- Flächen sammeln sich rund um Zentrum mit unverändertem Typ
- präferiert kleinere Regionen, jedoch beschränkt durch θ
- Kombination mit c_p möglich
- nutze alternativ kürzeste Wege $\delta'_{V'}(u, v)$ in G, die innerhalb V' verlaufen $\to c_{sp}(V', k)$



Komplexität [Haunert, Wolff 2010]

Satz 1: Für eine gegeben Instanz von $\textsc{AreaAggregation}$ und einen Wert $C \in \mathbb{Z}$ ist es NP-schwer zu entscheiden, ob eine Lösung mit Kosten $\leq C$ existiert, selbst wenn $|\Gamma| = 2$, $w \equiv 1$, θ für alle Farben gleich, Distanz $d \in \{0, 1\}$ und $s = 1$.

Beweis: Reduktion von planarem Vertex Cover (s. Tafel)

$\textsc{Planares Vertex Cover}$

Geg: Planarer Graph $G = (V, E)$, Parameter $C \in \mathbb{N}$

Ges: Vertex Cover $V \subseteq V'$ mit $|V'| \leq C$.

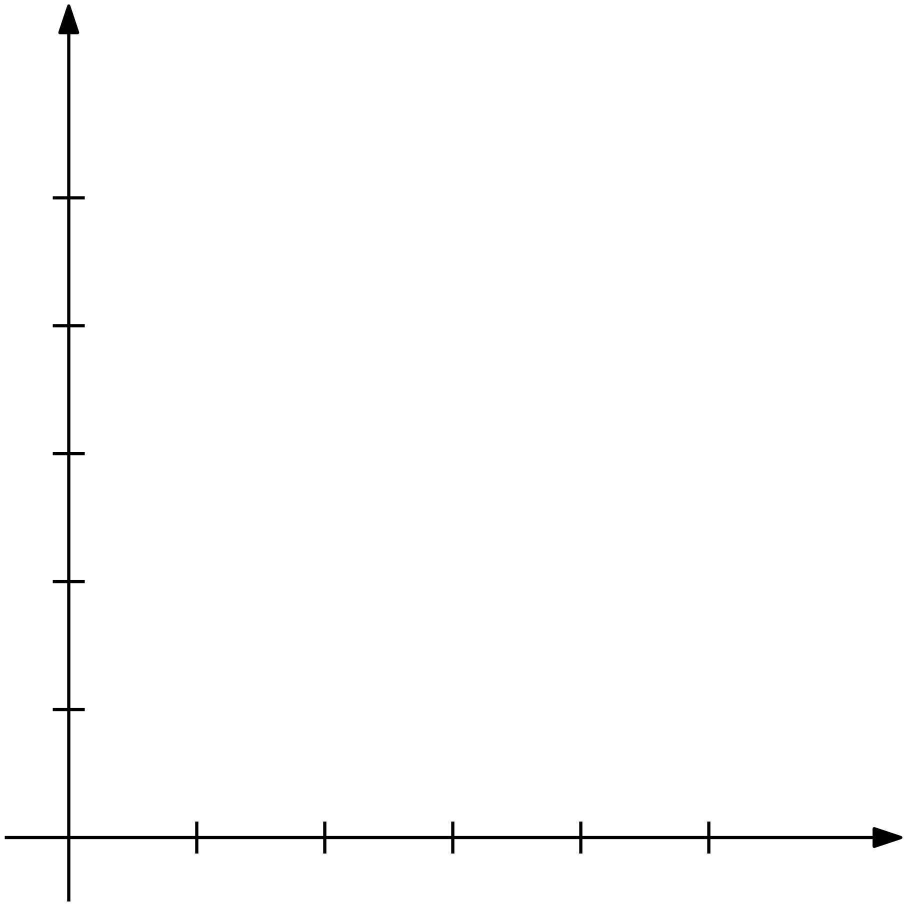
$V' \subseteq V$ heißt *Vertex Cover*, falls für jede Kante $\{u, v\} \in E$ gilt $u \in V'$ oder $v \in V'$.

(Gemischt) Ganzzahlige Programmierung

Lineare Programmierung (LP) ist ein effizientes Optimierungsverfahren für

- lineare Nebenbedingungen
- lineare Zielfunktion
- reellwertige Variablen

Beispiel:
maximiere $x + 2y$ sodass
$y \leq 0,9x + 1,5$
$y \geq 1,4x - 1,3$

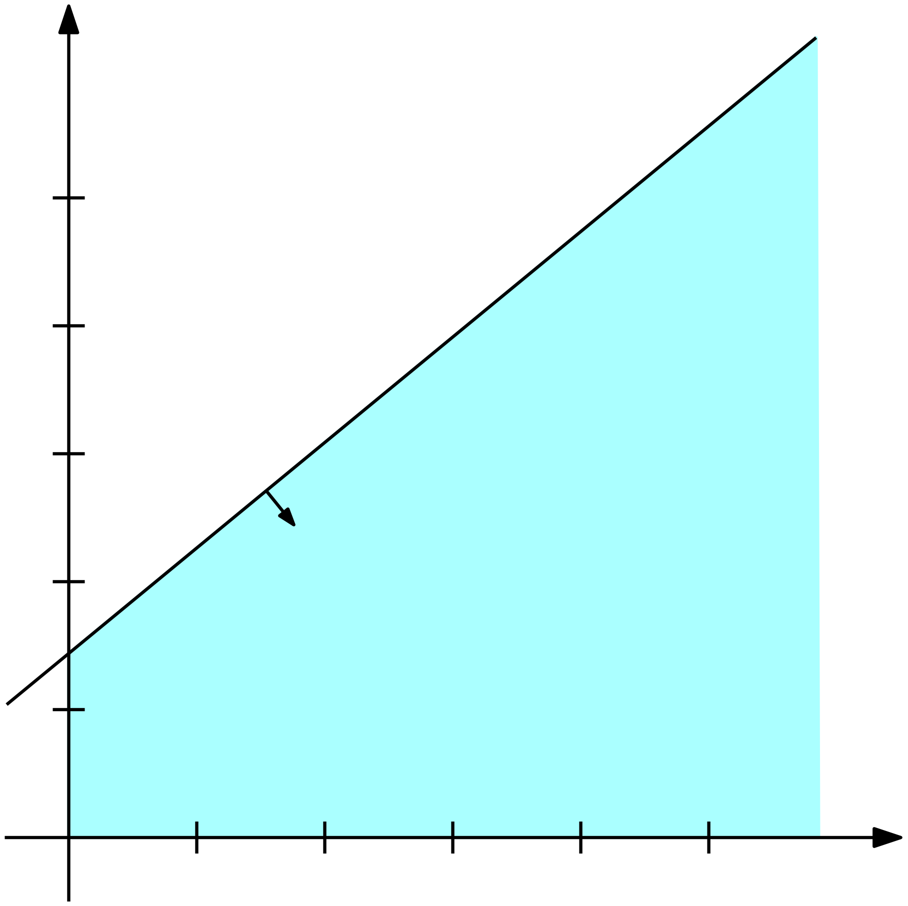


(Gemischt) Ganzzahlige Programmierung

Lineare Programmierung (LP) ist ein effizientes Optimierungsverfahren für

- lineare Nebenbedingungen
- lineare Zielfunktion
- reellwertige Variablen

Beispiel:
maximiere $x + 2y$ sodass
$y \leq 0,9x + 1,5$
$y \geq 1,4x - 1,3$

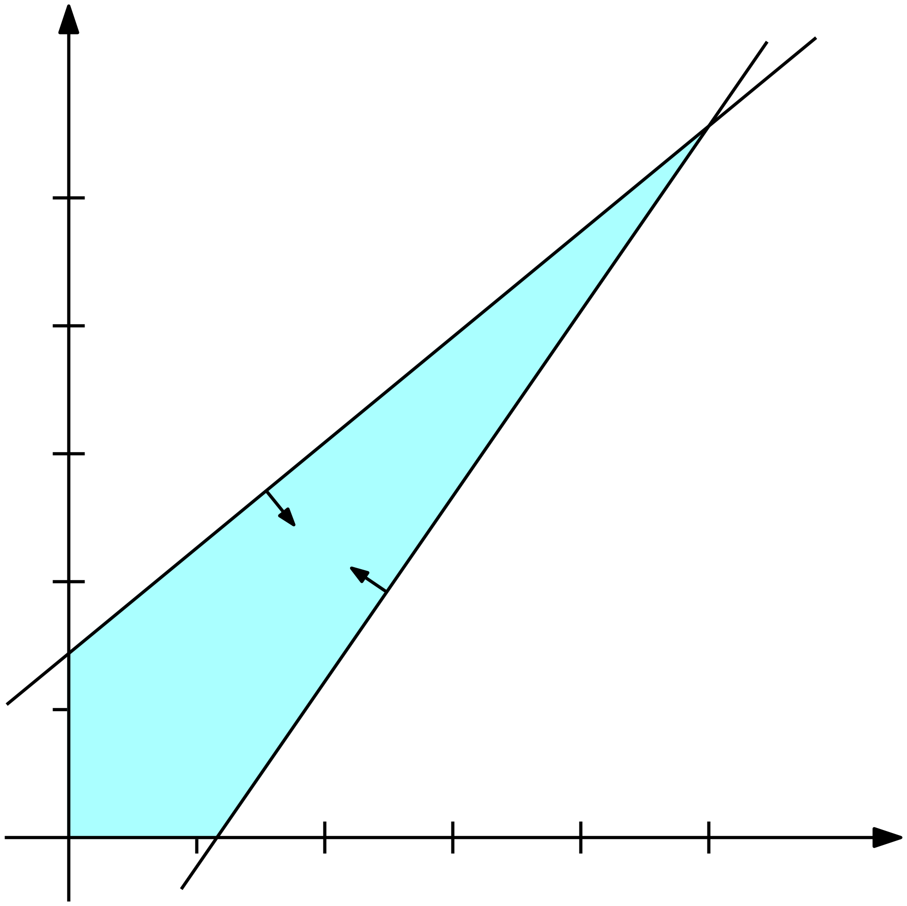


(Gemischt) Ganzzahlige Programmierung

Lineare Programmierung (LP) ist ein effizientes Optimierungsverfahren für

- lineare Nebenbedingungen
- lineare Zielfunktion
- reellwertige Variablen

Beispiel:
maximiere $x + 2y$ sodass
$y \leq 0,9x + 1,5$
$y \geq 1,4x - 1,3$

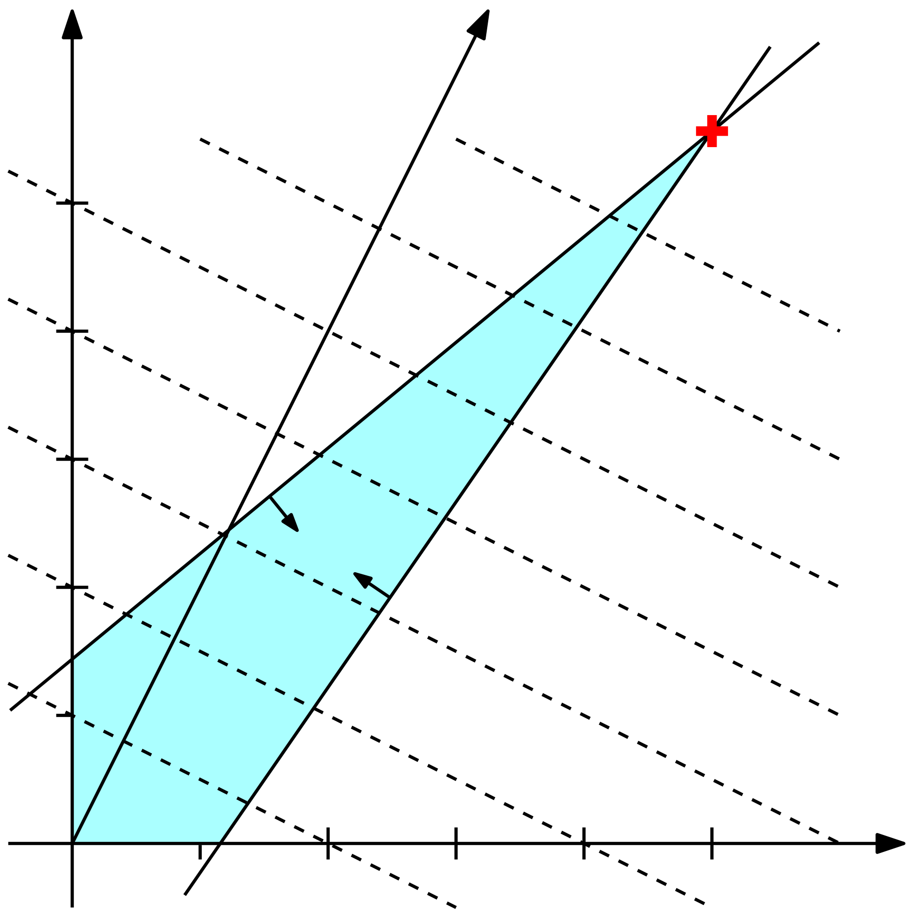


(Gemischt) Ganzzahlige Programmierung

Lineare Programmierung (LP) ist ein effizientes Optimierungsverfahren für

- lineare Nebenbedingungen
- lineare Zielfunktion
- reellwertige Variablen

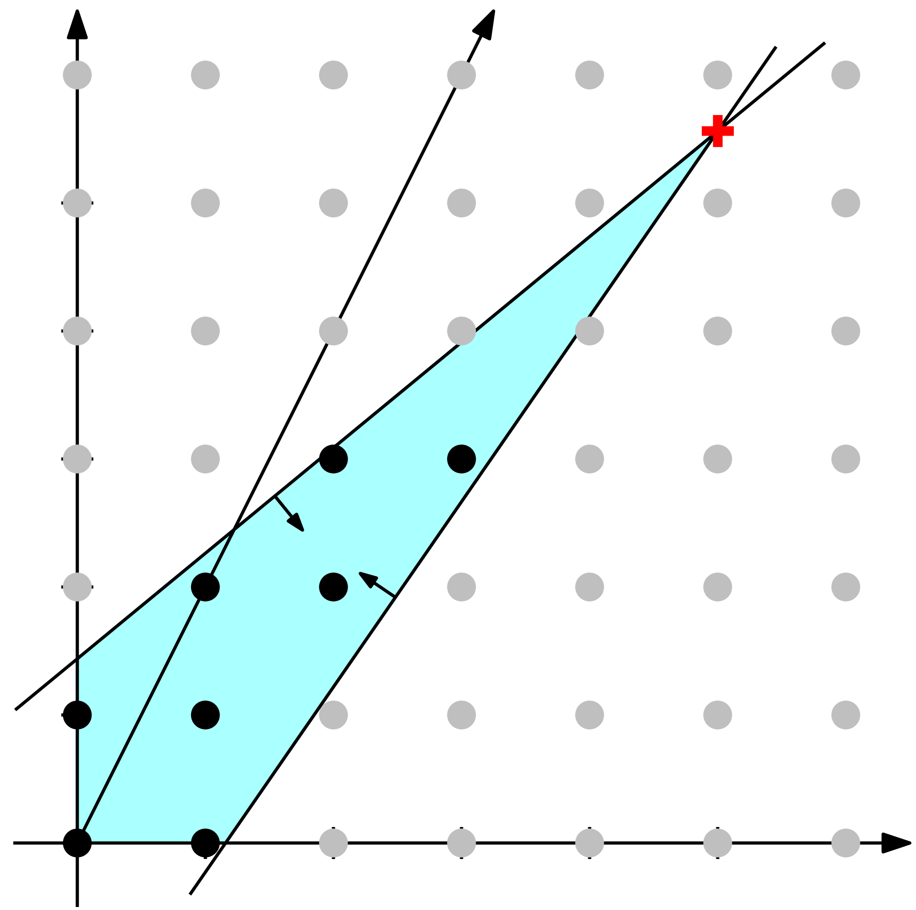
Beispiel:
maximiere $x + 2y$ sodass
$y \leq 0,9x + 1,5$
$y \geq 1,4x - 1,3$



(Gemischt) Ganzzahlige Programmierung

Lineare Programmierung (LP) ist ein effizientes Optimierungsverfahren für

- lineare Nebenbedingungen
- lineare Zielfunktion
- reellwertige Variablen



Beispiel:
maximiere $x + 2y$ sodass
$y \leq 0,9x + 1,5$
$y \geq 1,4x - 1,3$

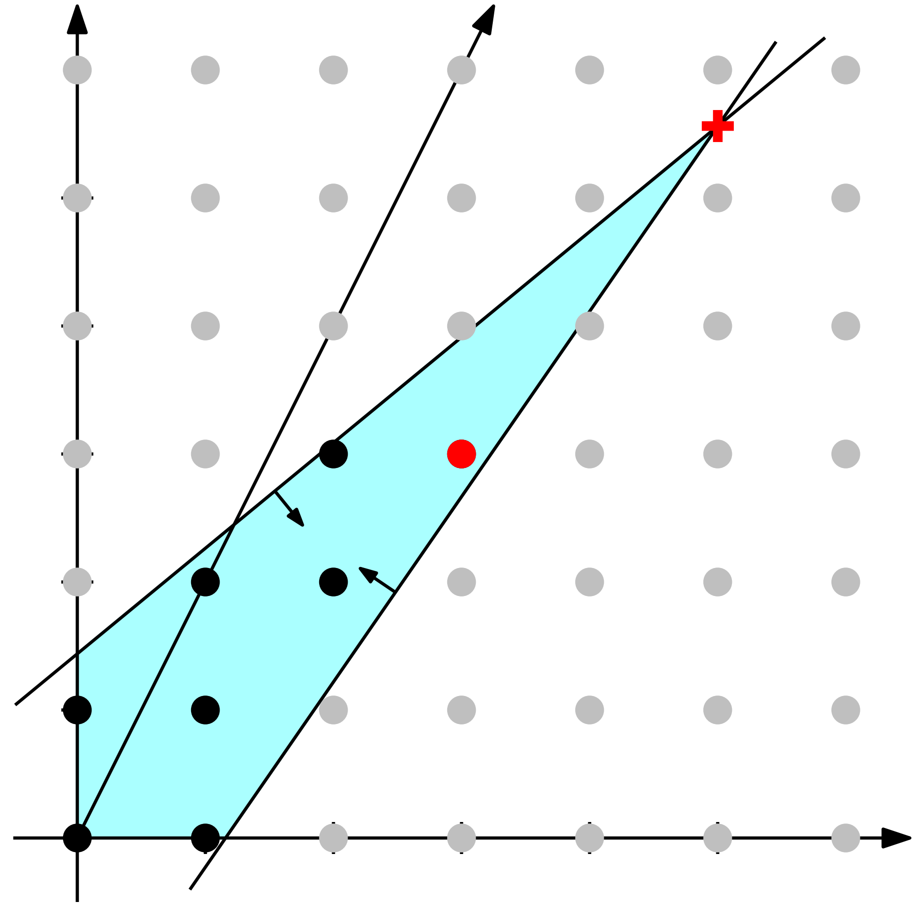
Gemischt ganzzahlige Programmierung (MIP)

- erlaubt reellwertige und ganzzahlige Variablen
- ist im Allg. NP-schwer
- liefert dennoch oft brauchbare und praktikable Lösungen in der Praxis

(Gemischt) Ganzzahlige Programmierung

Lineare Programmierung (LP) ist ein effizientes Optimierungsverfahren für

- lineare Nebenbedingungen
- lineare Zielfunktion
- reellwertige Variablen



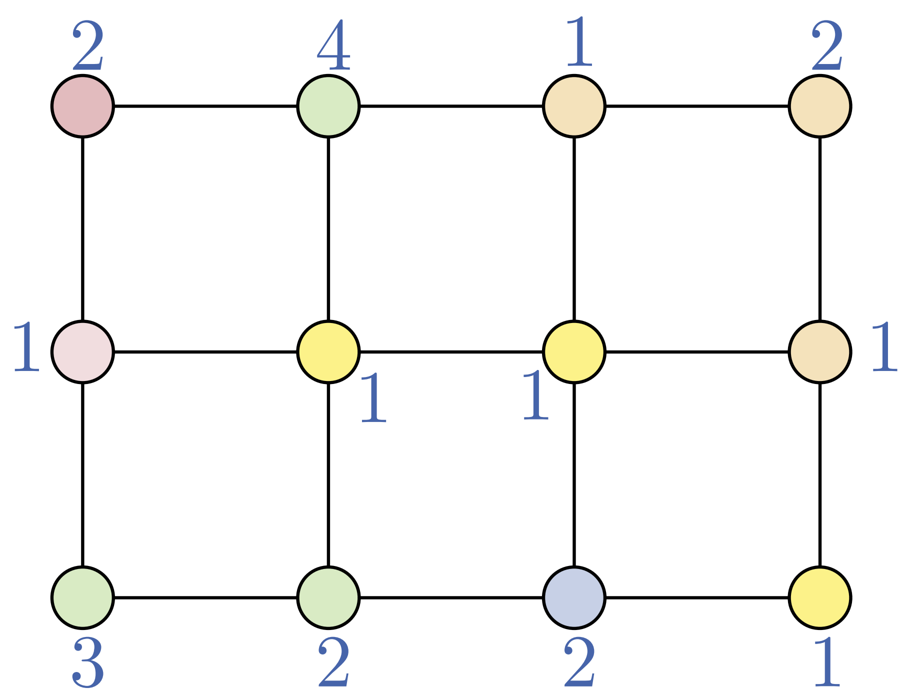
Beispiel:
maximiere $x + 2y$ sodass
$y \leq 0,9x + 1,5$
$y \geq 1,4x - 1,3$

Gemischt ganzzahlige Programmierung (MIP)

- erlaubt reellwertige und ganzzahlige Variablen
- ist im Allg. NP-schwer
- liefert dennoch oft brauchbare und praktikable Lösungen in der Praxis

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

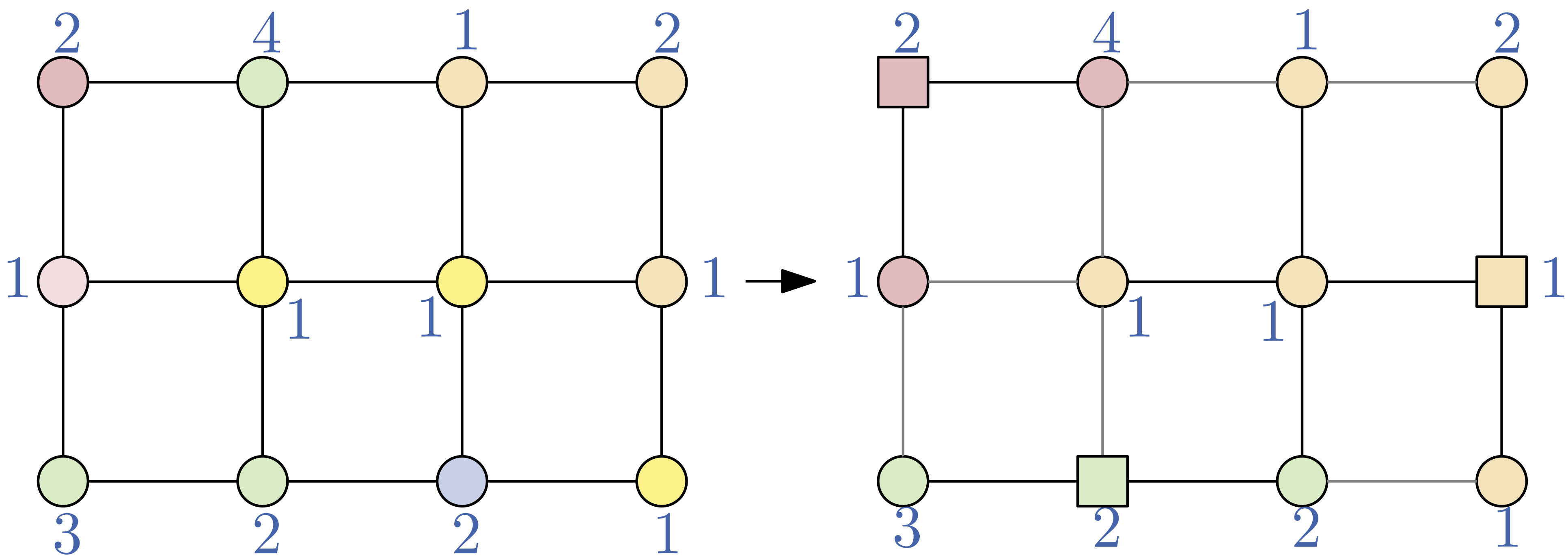
Idee: Flussnetzwerk N basierend auf G, sodass Zusammenhangskomp. aller Kanten mit Fluss > 0 die Aggregation definieren.



MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Idee: Flussnetzwerk N basierend auf G, sodass Zusammenhangskomp. aller Kanten mit Fluss > 0 die Aggregation definieren.

Betrachte Lösung:

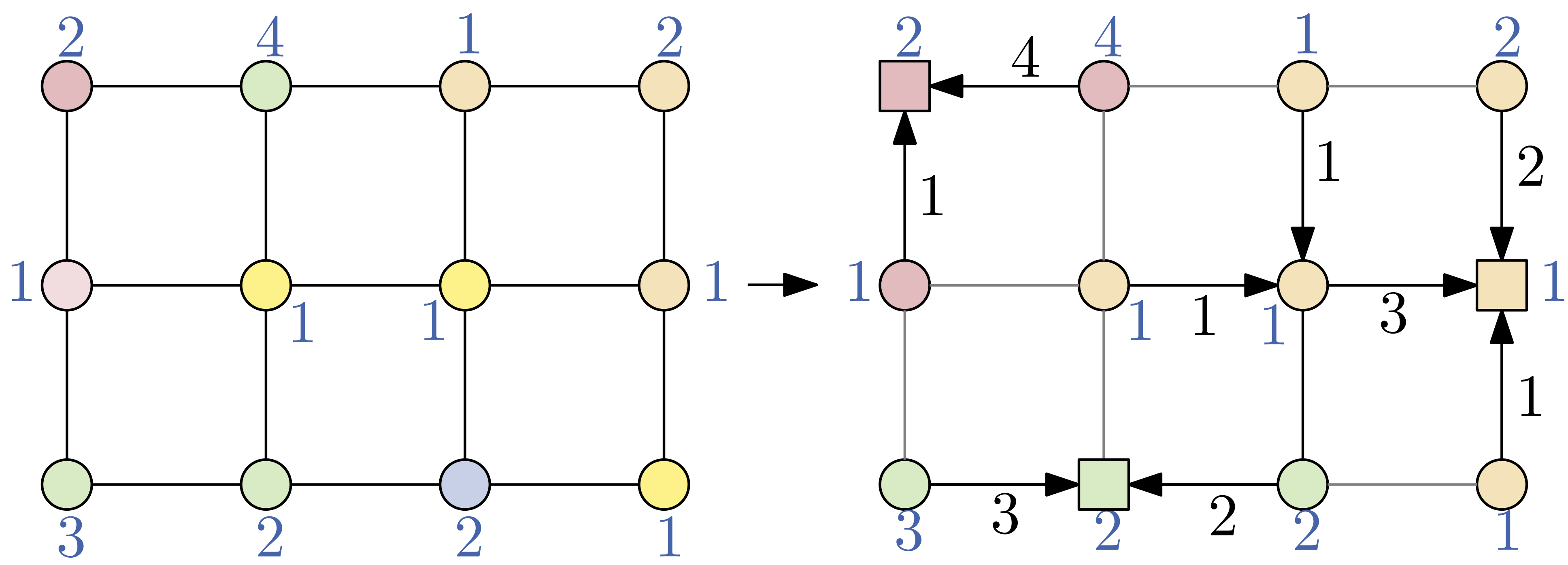


$\theta = 7$

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Idee: Flussnetzwerk N basierend auf G, sodass Zusammenhangskomp. aller Kanten mit Fluss > 0 die Aggregation definieren.

Betrachte Lösung:

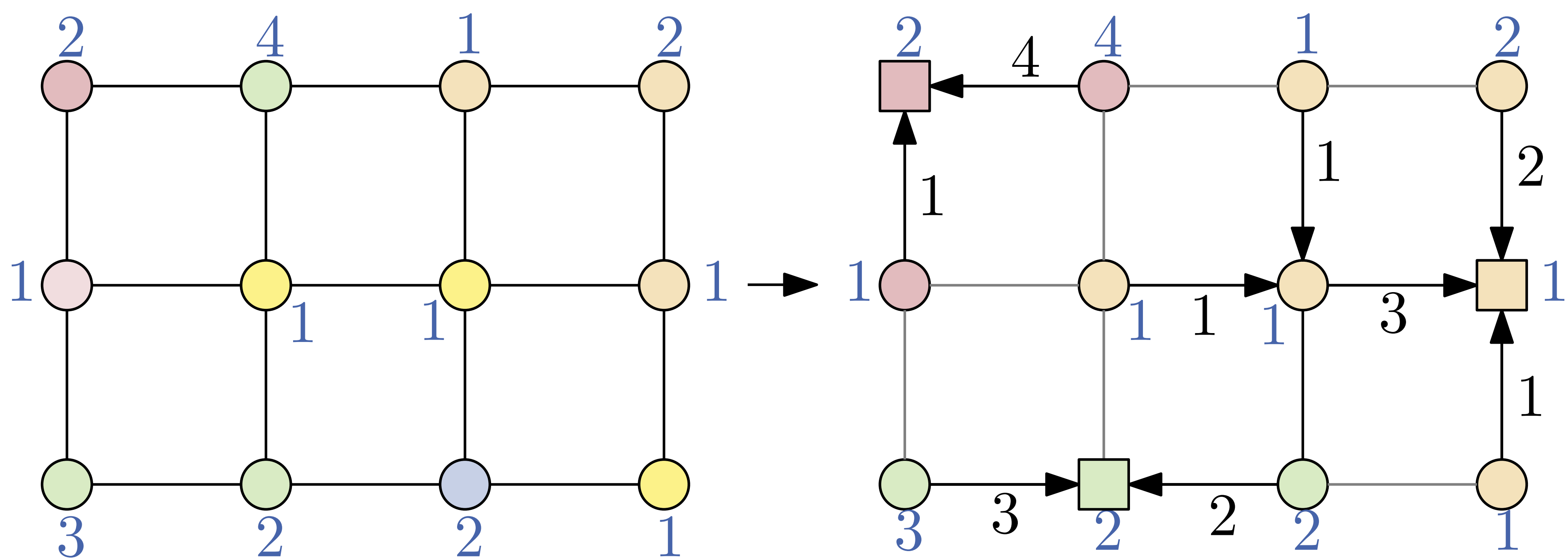


$\theta = 7$ Lösung kann als Flussnetzwerk aufgefasst werden.

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Idee: Flussnetzwerk N basierend auf G, sodass Zusammenhangskomp. aller Kanten mit Fluss > 0 die Aggregation definieren.

Betrachte Lösung:



$\theta = 7$

Lösung kann als Flussnetzwerk aufgefasst werden.

Problem: Senken sind nicht im Voraus bekannt.
Fluss darf nicht geteilt werden.

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Variablen:

- $f_a \in [0, M]$ definiert den Fluss über Kante a $(M = \sum_{v \in V} w(v))$
- $F_a \in \{0, 1\}$ mit $F_a = 1$ falls $f_a > 0$
- $b_{vk} \in \{0, 1\}$ mit $b_{vk} = 1$ falls $\gamma'(v) = k$
- $s_v \in \{0, 1\}$ mit $s_v = 1$ falls v eine Senke ist

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Variablen:

- $f_a \in [0, M]$ definiert den Fluss über Kante a $(M = \sum_{v \in V} w(v))$
- $F_a \in \{0, 1\}$ mit $F_a = 1$ falls $f_a > 0$
- $b_{vk} \in \{0, 1\}$ mit $b_{vk} = 1$ falls $\gamma'(v) = k$
- $s_v \in \{0, 1\}$ mit $s_v = 1$ falls v eine Senke ist

Zielfunktion:

minimiere

$$s \sum_{v \in V} \sum_{k \in \Gamma} w(v) \cdot d(\gamma(v), k) \cdot b_{vk} + (1 - s) \sum_{a = uv \in A} \delta(u, v) \cdot f_a$$

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Variablen:

- $f_a \in [0, M]$ definiert den Fluss über Kante a $(M = \sum_{v \in V} w(v))$
- $F_a \in \{0, 1\}$ mit $F_a = 1$ falls $f_a > 0$
- $b_{vk} \in \{0, 1\}$ mit $b_{vk} = 1$ falls $\gamma'(v) = k$
- $s_v \in \{0, 1\}$ mit $s_v = 1$ falls v eine Senke ist

Zielfunktion:

minimiere

$$s \sum_{v \in V} \sum_{k \in \Gamma} w(v) \cdot d(\gamma(v), k) \cdot b_{vk} + (1 - s) \sum_{a = uv \in A} \delta(u, v) \cdot f_a$$

Welches Kompaktheitsmaß wird hier verwendet?

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

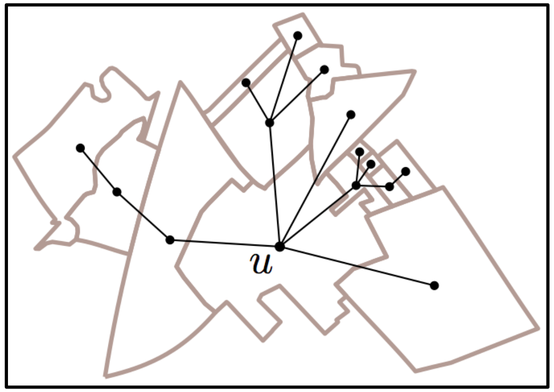
Variablen:

- $f_a \in [0, M]$ definiert den Fluss über Kante a $(M = \sum_{v \in V} w(v))$
- $F_a \in \{0, 1\}$ mit $F_a = 1$ falls $f_a > 0$
- $b_{vk} \in \{0, 1\}$ mit $b_{vk} = 1$ falls $\gamma'(v) = k$
- $s_v \in \{0, 1\}$ mit $s_v = 1$ falls v eine Senke ist

Zielfunktion:

minimiere

$$s \sum_{v \in V} \sum_{k \in \Gamma} w(v) \cdot d(\gamma(v), k) \cdot b_{vk} + (1 - s) \sum_{a = uv \in A} \delta(u, v) \cdot f_a$$



MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Nebenbedingungen:

$$M \cdot F_a \geq f_a \qquad \text{für alle } a \in A \qquad (1)$$

$$\sum_{k \in \Gamma} b_{vk} = 1 \qquad \text{für alle } v \in V \qquad (2)$$

$$b_{uk} \geq b_{vk} + (F_{uv} - 1) \qquad \text{für alle } uv \in A, k \in \Gamma \qquad (3)$$

$$b_{v\gamma(v)} \geq s_v \qquad \text{für alle } v \in V \qquad (4)$$

$$\sum_{a=vu \in A} f_a - \sum_{a=uv \in A} f_a \geq w(v) - M s_v \qquad \text{für alle } v \in V \qquad (5)$$

$$\sum_{a=vu \in A} f_a - \sum_{a=uv \in A} f_a \leq w(v) - s_v \theta(\gamma(v)) \qquad \text{für alle } v \in V \qquad (6)$$

$$\sum_{a=uv \in A} F_a \leq 1 - s_u \qquad \text{für alle } u \in V \qquad (7)$$

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Nebenbedingungen:

$$M \cdot F_a \geq f_a \qquad \text{für alle } a \in A \qquad (1)$$

$$\sum_{k \in \Gamma} b_{vk} = 1 \qquad \text{für alle } v \in V \qquad (2)$$

$$b_{uk} \geq b_{vk} + (F_{uv} - 1) \qquad \text{für alle } uv \in A, k \in \Gamma \qquad (3)$$

$$b_{v\gamma(v)} \geq s_v \qquad \text{für alle } v \in V \qquad (4)$$

$$\sum_{a=vu \in A} f_a - \sum_{a=uv \in A} f_a \geq w(v) - M s_v \qquad \text{für alle } v \in V \qquad (5)$$

$$\sum_{a=vu \in A} f_a - \sum_{a=uv \in A} f_a \leq w(v) - s_v \theta(\gamma(v)) \qquad \text{für alle } v \in V \qquad (6)$$

$$\sum_{a=uv \in A} F_a \leq 1 - s_u \qquad \text{für alle } u \in V \qquad (7)$$

Interpretation der Nebenbedingungen?

MIP A: Netzwerkfluss [Haunert, Wolff 2010]

Nebenbedingungen:

$$M \cdot F_a \geq f_a \qquad \text{für alle } a \in A \qquad (1)$$

$$\sum_{k \in \Gamma} b_{vk} = 1 \qquad \text{für alle } v \in V \qquad (2)$$

$$b_{uk} \geq b_{vk} + (F_{uv} - 1) \qquad \text{für alle } uv \in A, k \in \Gamma \qquad (3)$$

$$b_{v\gamma(v)} \geq s_v \qquad \text{für alle } v \in V \qquad (4)$$

$$\sum_{a=vu \in A} f_a - \sum_{a=uv \in A} f_a \geq w(v) - M s_v \qquad \text{für alle } v \in V \qquad (5)$$

$$\sum_{a=vu \in A} f_a - \sum_{a=uv \in A} f_a \leq w(v) - s_v \theta(\gamma(v)) \qquad \text{für alle } v \in V \qquad (6)$$

$$\sum_{a=uv \in A} F_a \leq 1 - s_u \qquad \text{für alle } u \in V \qquad (7)$$

Wie lässt sich Vollständigkeit modellieren?

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

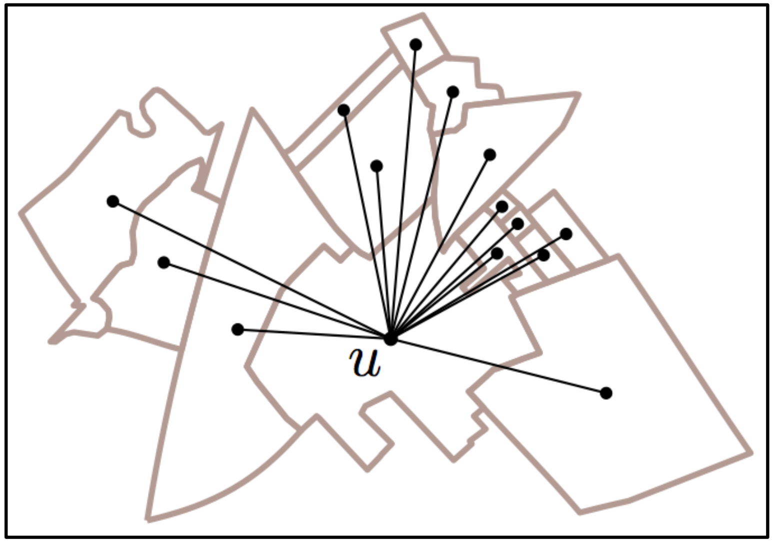
Variablen:

- $x_{uv} \in \{0, 1\}$ mit $x_{uv} = 1$ gdw. Knoten v zum Zentrum u gehört

Idee: Modelliere Zentrumszugehörigkeit

$x_{uu} = 1$ gdw. u ist Zentrum.

$x_{uv} = 1$ gdw. v gehört zu Zentrum von u.



MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Variablen:

- $x_{uv} \in \{0, 1\}$ mit $x_{uv} = 1$ gdw. Knoten v zum Zentrum u gehört

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Variablen:

- $x_{uv} \in \{0, 1\}$ mit $x_{uv} = 1$ gdw. Knoten v zum Zentrum u gehört

Zielfunktion:

minimiere

$$s \sum_{(u,v) \in V^2} w(v) \cdot d(\gamma(v), \gamma(u)) \cdot x_{uv} + (1 - s) \sum_{(u,v) \in V^2} w(v) \cdot \delta(u, v) \cdot x_{uv}$$

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Variablen:

- $x_{uv} \in \{0, 1\}$ mit $x_{uv} = 1$ gdw. Knoten v zum Zentrum u gehört

Zielfunktion:

minimiere

$$s \sum_{(u,v) \in V^2} w(v) \cdot d(\gamma(v), \gamma(u)) \cdot x_{uv} + (1 - s) \sum_{(u,v) \in V^2} w(v) \cdot \delta(u, v) \cdot x_{uv}$$

Nebenbedingungen:

$$\sum_{u \in V} x_{uv} = 1 \quad \text{für alle } v \in V \qquad (8)$$

$$\sum_{v \in V} w(v) \cdot x_{uv} \geq \theta(\gamma(u)) \cdot x_{uu} \quad \text{für alle } u \in V \qquad (9)$$

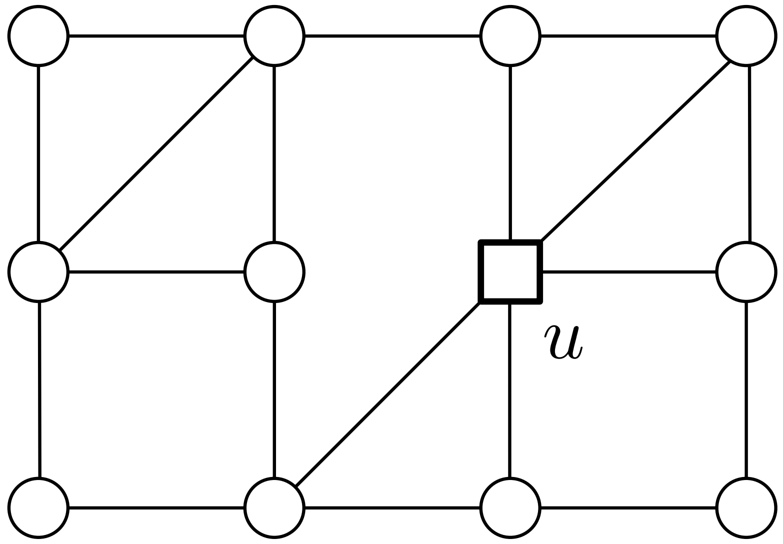
$$\sum_{w \in P_u(v)} x_{uw} \geq x_{uv} \quad \text{für alle } (u, v) \in V^2, u \neq v \qquad (10)$$

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Vorgängerrelation:

$$P_u(v) = \{w \in V \mid D(w,u) < D(v,u) \text{ und } \{v,w\} \in E\}$$

$D(w,u)$ ist kürzester w–u-Weg in $G' = (V, A)$ mit
$A = \{(u,v),(v,u) \mid \{u,v\} \in E\}$ und Kantenlänge $\alpha(u,v) = w(v)$



$w(v) = 1$ für alle $v \in V$

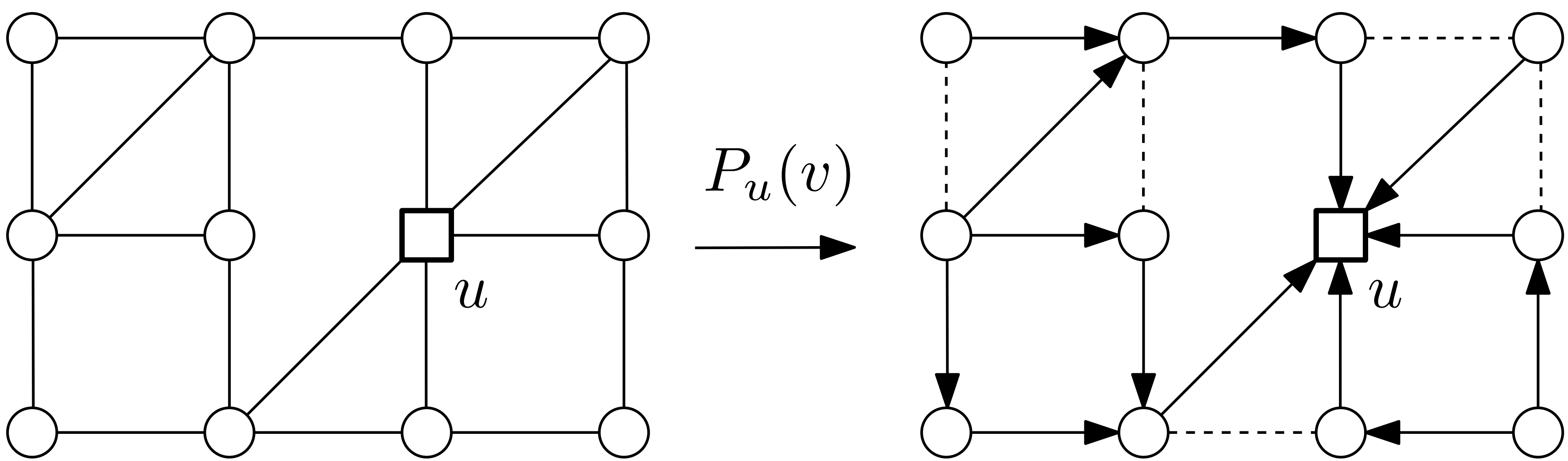
MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Karlsruhe Institute of Technology

Vorgängerrelation:

$$P_u(v) = \{w \in V \mid D(w,u) < D(v,u) \text{ und } \{v,w\} \in E\}$$

$D(w,u)$ ist kürzester w–u-Weg in $G' = (V, A)$ mit
$A = \{(u,v),(v,u) \mid \{u,v\} \in E\}$ und Kantenlänge $\alpha(u,v) = w(v)$



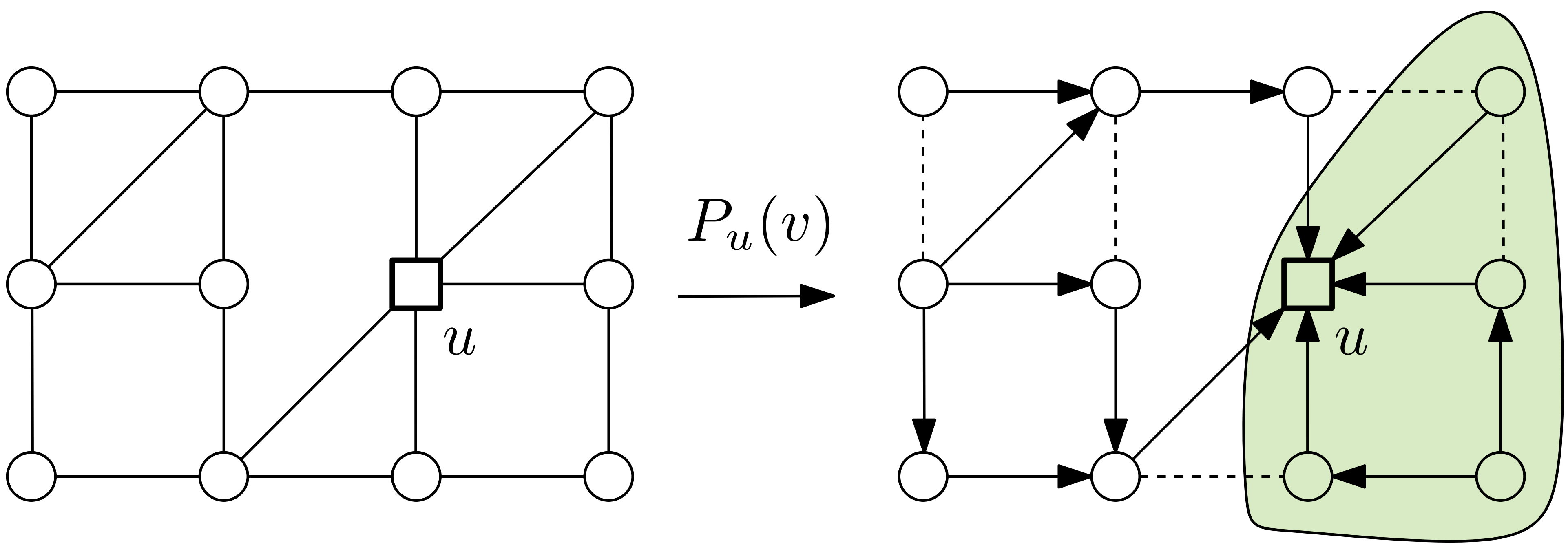
$w(v) = 1$ für alle $v \in V$

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Vorgängerrelation:

$$P_u(v) = \{w \in V \mid D(w,u) < D(v,u) \text{ und } \{v,w\} \in E\}$$

$D(w,u)$ ist kürzester w–u-Weg in $G' = (V, A)$ mit
$A = \{(u,v),(v,u) \mid \{u,v\} \in E\}$ und Kantenlänge $\alpha(u,v) = w(v)$



$w(v) = 1$ für alle $v \in V$

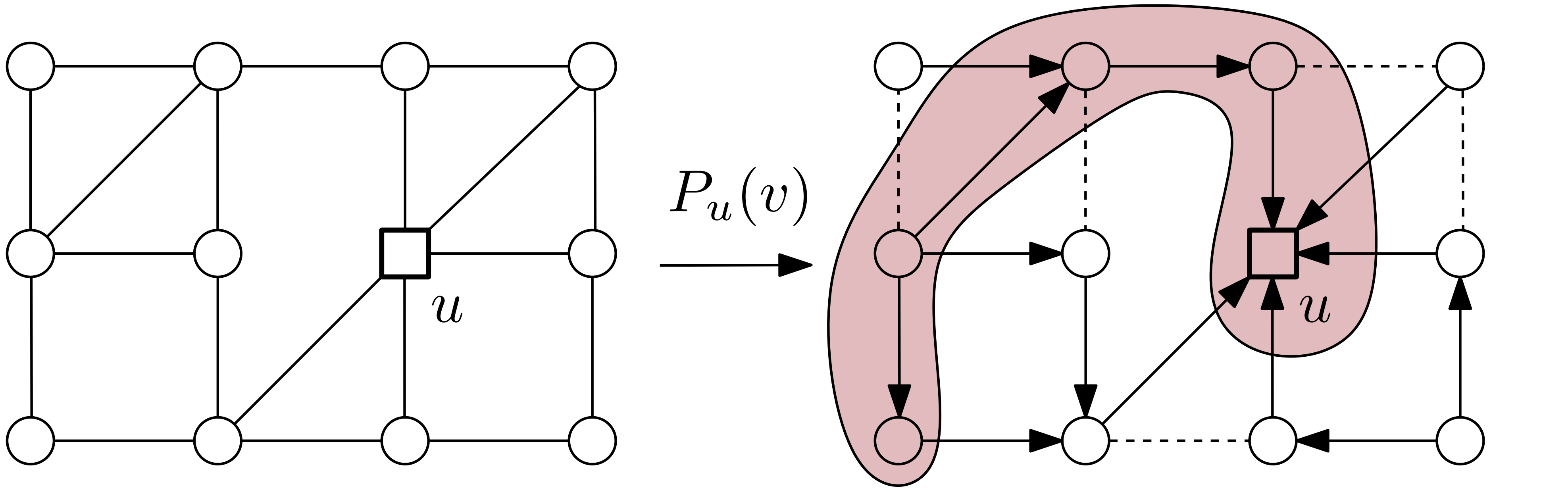
Zulässige Aggregation.

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Vorgängerrelation:

$$P_u(v) = \{w \in V \mid D(w,u) < D(v,u) \text{ und } \{v,w\} \in E\}$$

$D(w,u)$ ist kürzester w–u-Weg in $G' = (V, A)$ mit
$A = \{(u,v),(v,u) \mid \{u,v\} \in E\}$ und Kantenlänge $\alpha(u,v) = w(v)$



$w(v) = 1$ für alle $v \in V$

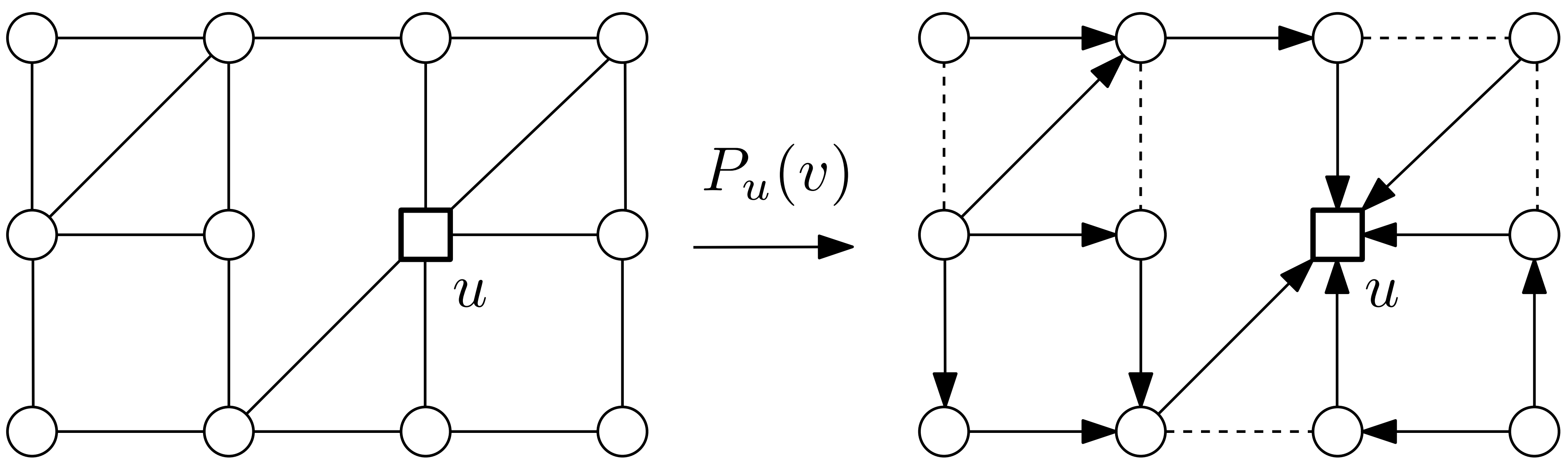
Nicht zulässige Aggregation.

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Vorgängerrelation:

$$P_u(v) = \{w \in V \mid D(w,u) < D(v,u) \text{ und } \{v,w\} \in E\}$$

$D(w,u)$ ist kürzester w–u-Weg in $G' = (V, A)$ mit
$A = \{(u,v),(v,u) \mid \{u,v\} \in E\}$ und Kantenlänge $\alpha(u,v) = w(v)$



$w(v) = 1$ für alle $v \in V$

$\rightarrow$ Optimale Lösung kann verloren gehen.

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Variablen:

- $x_{uv} \in \{0,1\}$ mit $x_{uv} = 1$ gdw. Knoten v zum Zentrum u gehört

Zielfunktion:

minimiere

$$s \sum_{(u,v)\in V^2} w(v) \cdot d(\gamma(v), \gamma(u)) \cdot x_{uv} + (1-s) \sum_{(u,v)\in V^2} w(v) \cdot \delta(u,v) \cdot x_{uv}$$

Nebenbedingungen:

$$\sum_{u\in V} x_{uv} = 1 \quad \text{für alle } v \in V \qquad (11)$$

$$\sum_{v\in V} w(v) \cdot x_{uv} \geq \theta(\gamma(u)) \cdot x_{uu} \quad \text{für alle } u \in V \qquad (12)$$

$$\sum_{w\in P_u(v)} x_{uw} \geq x_{uv} \quad \text{für alle } (u,v) \in V^2, u \neq v \qquad (13)$$

MIP B: Vorgängerrelation [Haunert, Wolff 2010]

Variablen:

- $x_{uv} \in \{0, 1\}$ mit $x_{uv} = 1$ gdw. Knoten v zum Zentrum u gehört

Zielfunkt Wie lässt sich Vollständigkeit modellieren?

minimiere

$$s \sum_{(u,v)\in V^2} w(v) \cdot d(\gamma(v), \gamma(u)) \cdot x_{uv} + (1-s) \sum_{(u,v)\in V^2} w(v) \cdot \delta(u, v) \cdot x_{uv}$$

Nebenbedingungen:

$$\sum_{u\in V} x_{uv} = 1 \qquad \text{für alle } v \in V \qquad (11)$$

$$\sum_{v\in V} w(v) \cdot x_{uv} \geq \theta(\gamma(u)) \cdot x_{uu} \qquad \text{für alle } u \in V \qquad (12)$$

$$\sum_{w\in P_u(v)} x_{uw} \geq x_{uv} \qquad \text{für alle } (u, v) \in V^2, u \neq v \qquad (13)$$

MIP C: Umfangskosten [Haunert, Wolff 2010]

wie MIP B, aber zusätzlich

Variablen:

- $y_{ue} \in [0, 1]$ mit $y_{ue} = 0$ falls mind. ein Endpunkt von $e \in E$ nicht zum Zentrum u gehört

MIP C: Umfangskosten [Haunert, Wolff 2010]

wie MIP B, aber zusätzlich

Variablen:

- $y_{ue} \in [0, 1]$ mit $y_{ue} = 0$ falls mind. ein Endpunkt von $e \in E$ nicht zum Zentrum u gehört

Nebenbedingungen:

$$y_{ue} \leq x_{uw} \qquad y_{ue} \leq x_{uv} \qquad \text{für alle } u \in V, e = \{v, w\} \in E \tag{14}$$

MIP C: Umfangskosten [Haunert, Wolff 2010]

wie MIP B, aber zusätzlich

Variablen:

- $y_{ue} \in [0, 1]$ mit $y_{ue} = 0$ falls mind. ein Endpunkt von $e \in E$ nicht zum Zentrum u gehört

Nebenbedingungen:

$$y_{ue} \leq x_{uw} \qquad y_{ue} \leq x_{uv} \qquad \text{für alle } u \in V, e = \{v, w\} \in E \tag{14}$$

Zielfunktion:

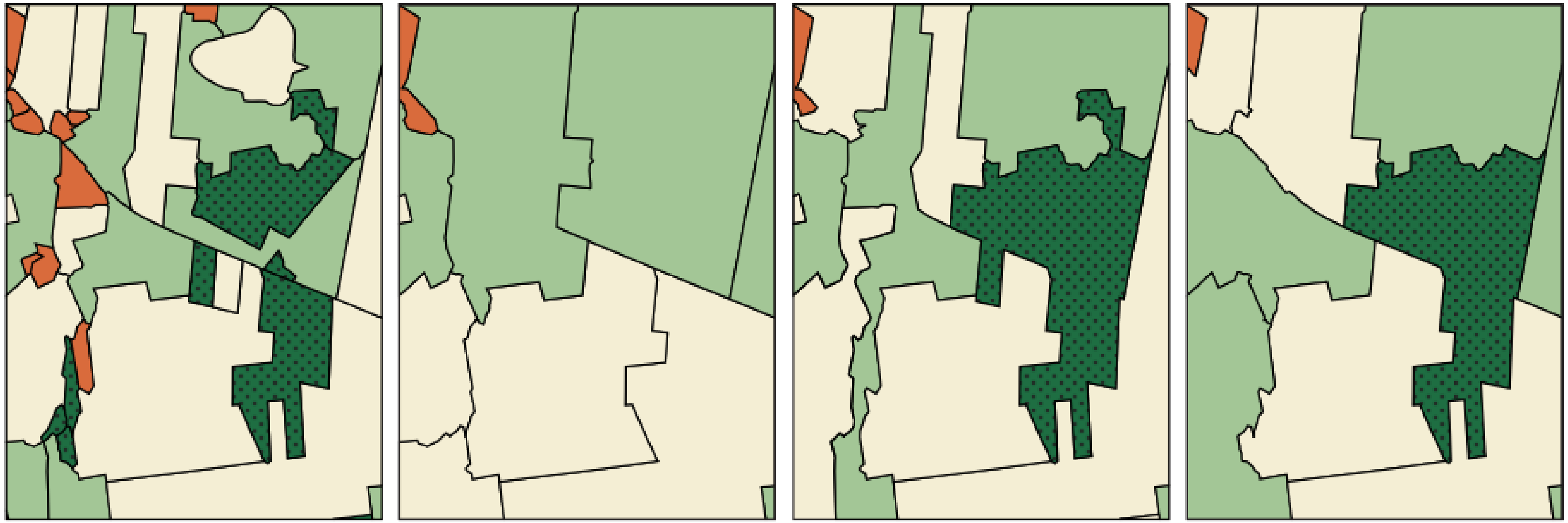
minimiere

$$s \sum_{(u,v) \in V^2} w(v) \cdot d(\gamma(v), \gamma(u)) \cdot x_{uv} - (1 - s) \cdot 2 \sum_{u \in V} \sum_{e \in E} \lambda(e) \cdot y_{ue}$$

wobei $\lambda \colon E \to \mathbb{R}^+$ die Kantenlängen angibt

Evaluation

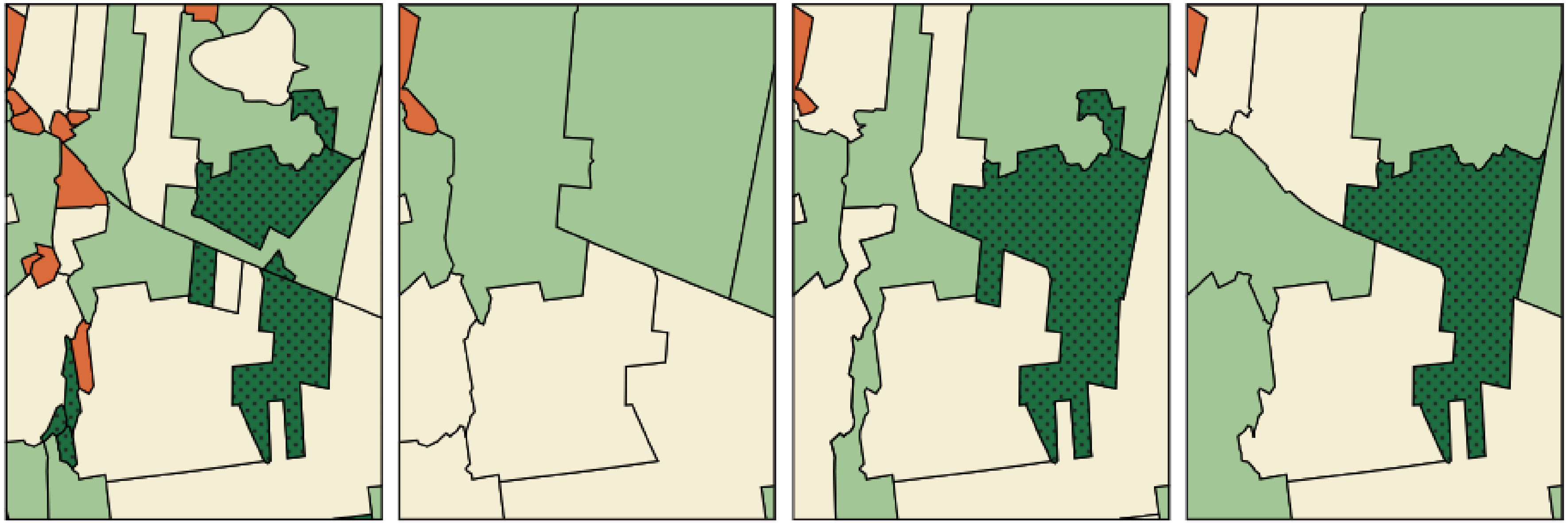
Visueller Eindruck:



Eingabe Region Growing Semantik Semantik + Kompaktheit

Evaluation

Visueller Eindruck:



Eingabe | Region Growing | Semantik | Semantik + Kompaktheit

Evaluation

Performance:

- MIP A: recht langsam, selbst für kleinere Instanzen
- MIP B/C: deutlich schneller, aber erst durch zusätzliche Heuristiken werden realistische Instanzgrößen erreicht

	Alg. 1	flow MIP				P-R MIP							
		pure		centre heuristic		pure		distance heuristic		centre heuristic		centre + dist. heuristic	
nodes	$\bar{d}$	time	$\bar{d}$	time	$\bar{d}$	time	$\bar{d}$	time	$\bar{d}$	time	$\bar{d}$	time	$\bar{d}$
30	29.30	223.2	9.20	8.8	9.20	2.8	9.20	3.0	9.20	0.04	12.81	0.01	12.81
40	22.31	3.9 **h**	6.97	41.6	6.97	31.1	7.34	16.8	7.34	0.04	7.59	0.03	7.59
50	15.30	*20.0 **h**	5.18	432.1	5.18	60.9	5.18	62.5	5.18	0.34	5.64	0.15	5.64
60	15.49					253.0	5.85	236.0	5.85	0.45	6.66	0.72	6.66
100	12.48									33.1	5.93	12.5	5.93
200	9.85									84.3	4.68	87.0	4.68
300	6.92									267.2	4.50	363.3	4.50
400	7.04									397.7	4.64	346.1	4.64

aus [Haunert, Wolff 2010]