
Algorithmen und Datenstrukturen

Zeichenerkennung, Subgraph-Isomorphie, Pruning von Suchräumen

Prof. Dr. Ralf Möller

Universität zu Lübeck

Institut für Informationssysteme

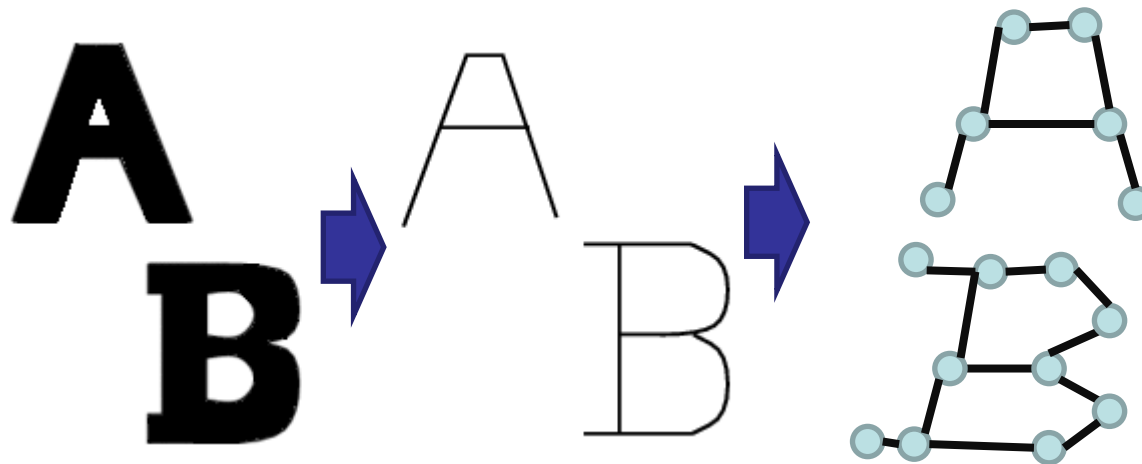
Magnus Bender (Übungen)

sowie viele Tutoren

OCR – Zeichenerkennung durch Graphabgleich

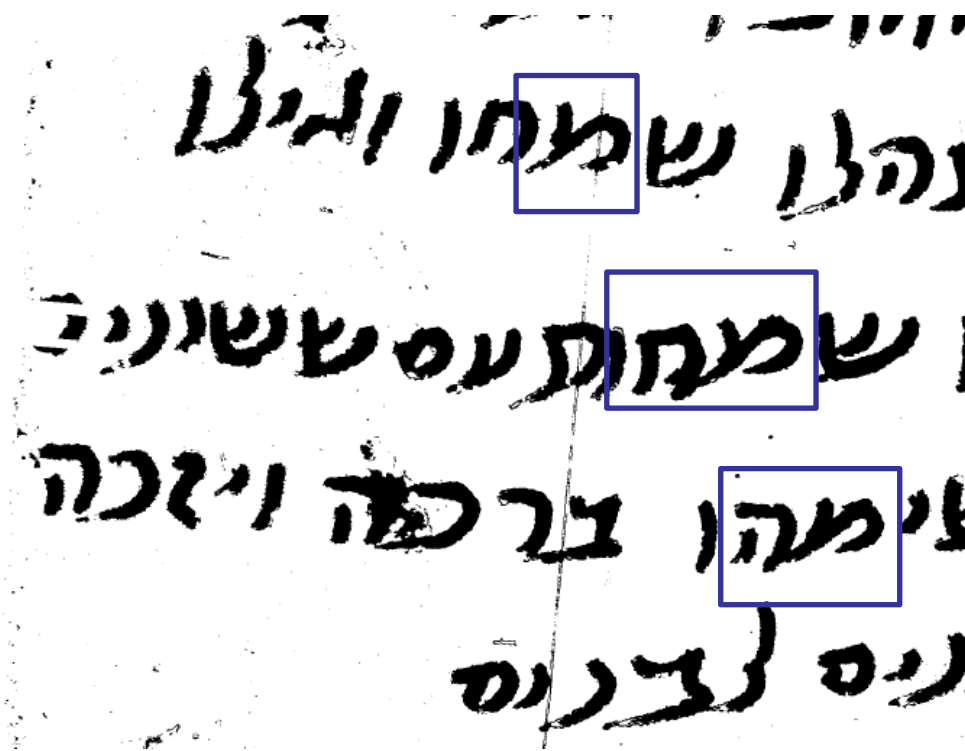
Bestimmung einer Datenstruktur:

Segmentation → Verdünnung → Graphrepräsentation

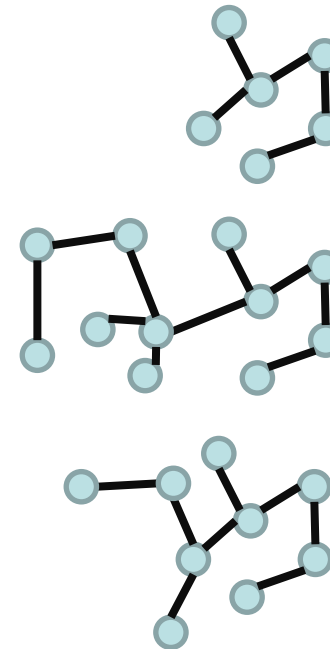


OCR durch Graphabgleich (Matching)

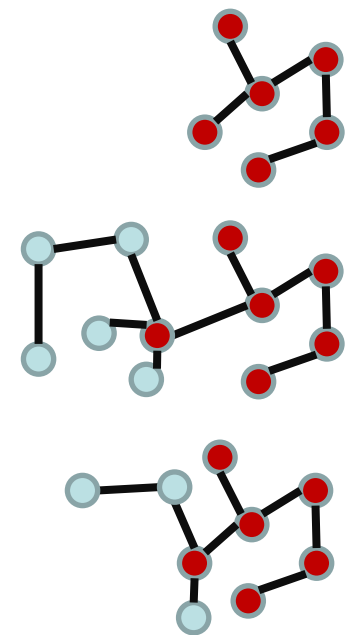
Durch Subgraphabgleich Kandidaten finden



Graphische Modellierung



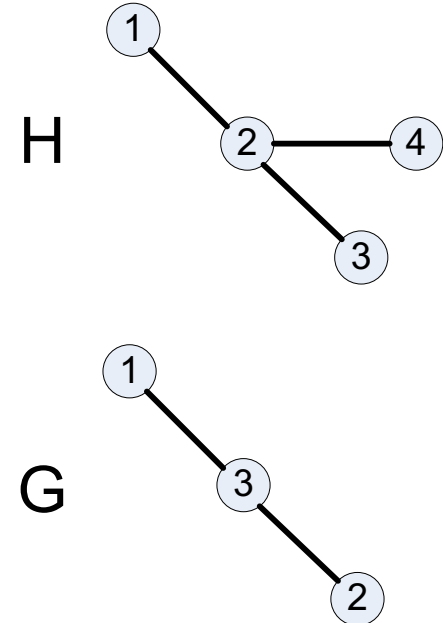
Graphischer Abgleich



Subgraph-Isomorphie

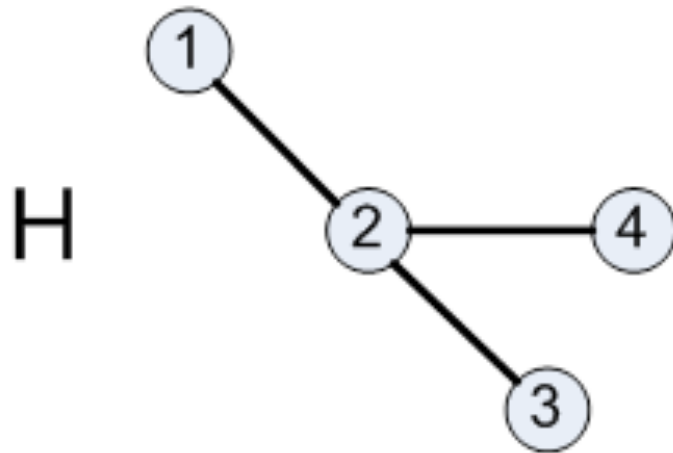
Subgraph-Isomorphismus-Problem

- Gegeben zwei Graphen **H** and **G**. Bestimme, ob **H** einen Subgraphen hat, der isomorph zu **G** ist, also bis auf Knotenumbenennung die Gestalt von **G** hat
- Zur Lösung müssen wir Korrespondenzen finden
- Beispiel: (Es gibt weitere Lösungen)
 - $1_G - 1_H$
 - $2_G - 3_H$
 - $3_G - 2_H$
 - Antwort: Ja
- Sollen wir alle möglichen Korrespondenzen generieren und testen?



Ullmanns Algorithmus

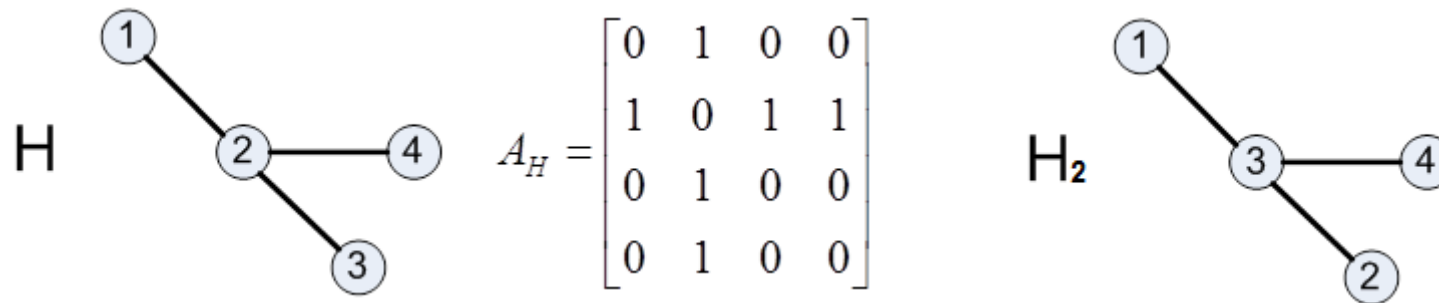
- Verwendung von algebraischen Formulierungen des Subgraph-Isomorphie-Problems
- Adjazenzmatrix A_H eines Graphen H :



$$A_H = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Aufbau des Suchraums nach Korrespondenzen

- Verwendung der sog. Permutationsmatrix
- Über Permutationsmatrix kann isomorphe Korrespondenz ausgedrückt werden



Isomorphe Korrespondenz

$$F = \begin{matrix} \mathbf{1}_{H_2} - \mathbf{1}_H \\ \mathbf{2}_{H_2} - \mathbf{3}_H \\ \mathbf{3}_{H_2} - \mathbf{2}_H \\ \mathbf{4}_{H_2} - \mathbf{4}_H \end{matrix}$$

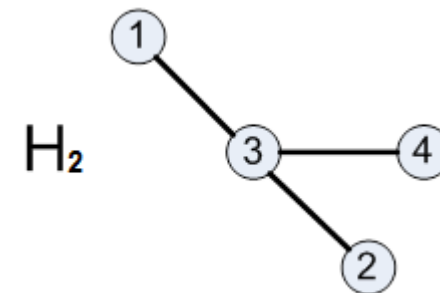
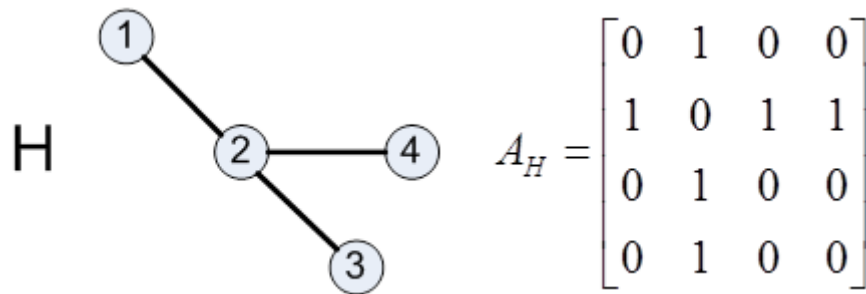
Permutationsmatrix

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$F \sim P$$

Isomorphiekriterium

Zwei Graphen H und H_2 sind isomorph mit Korrespondenz F genau dann, wenn A_H ähnlich zu A_{H_2} ist, also eine Permutationsmatrix $P \sim F$ existiert.



Isomorphe Korrespondenz \sim Permutationsmatrix

$$F = \begin{matrix} 1_{H_2} - 1_H \\ 2_{H_2} - 3_H \\ 3_{H_2} - 2_H \\ 4_{H_2} - 4_H \end{matrix} \quad P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$F \sim P$

Isomorphiekriterium:

$$A_{H_2} = P A_H P^T$$

gdw H_2 ist isomorph zu H mit Korrespondenz $F \sim P$

Skeptisch?

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad A_H = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

P "moves" the row and columns of A_H using the mapping

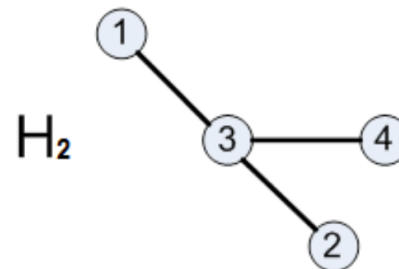
$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}$$

1: Moves only the rows

$$PA_H = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

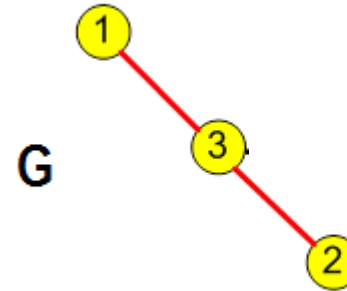
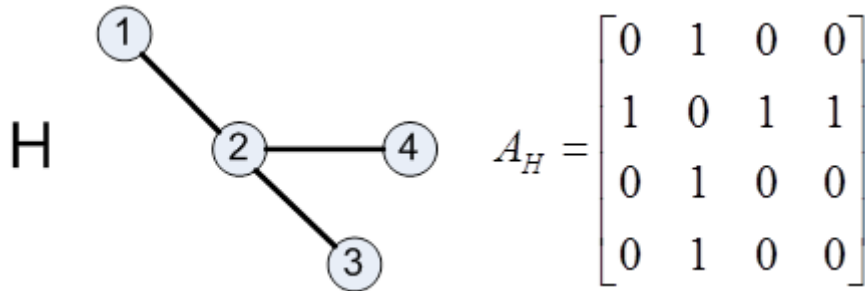
2: Moves only the columns

$$PA_H P^T = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = A_{H_2}$$



Ullmanns Algorithmus

In gleicher Weise können wir ein algebraisches Kriterium für **Subgraph**-Isomorphie bestimmen



Isomorphe Korrespondenz

~ Permutationsmatrix

$$F = \begin{matrix} \mathbf{1}_G - \mathbf{1}_H \\ \mathbf{2}_G - \mathbf{3}_H \\ \mathbf{3}_G - \mathbf{2}_H \\ \mathbf{4}_G - \varnothing \end{matrix}$$

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Eine 1 pro Zeile

Ullmanns Algorithmus

- Gegeben: Graphen G und H
- Gesucht: Permutationsmatrix P mit Dimensionen $|V_G| \times |V_H|$, so dass das Subgraph-Isomorphie-Kriterium erfüllt
- Suchraum über alle möglichen Permutationsmatrizen P mit jeweiliger Prüfung des Kriteriums aufspannen?
- Das geht besser!

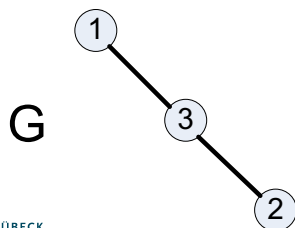
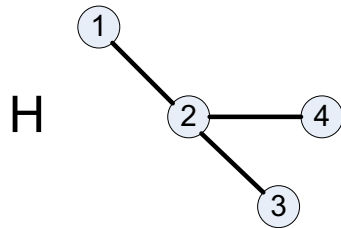
Ullmanns Algorithmus

- Konstruiere Matrix $M^{(0)}$ der Dimension der P-Matrizen:

$$m_{i,j}^{(0)} = \begin{cases} 1 & \text{if } \deg(V_{H_j}) \geq \deg(V_{G_i}) \\ 0 & \text{otherwise} \end{cases}, m_{i,j} \in \{0,1\}$$

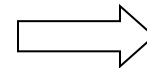
- Generiere aus $M^{(0)}$ alle P durch Wahl **einer 1 pro Zeile**
- Subgraph-Isomorphismus gefunden, wenn

$$A_G = P A_H P^T$$



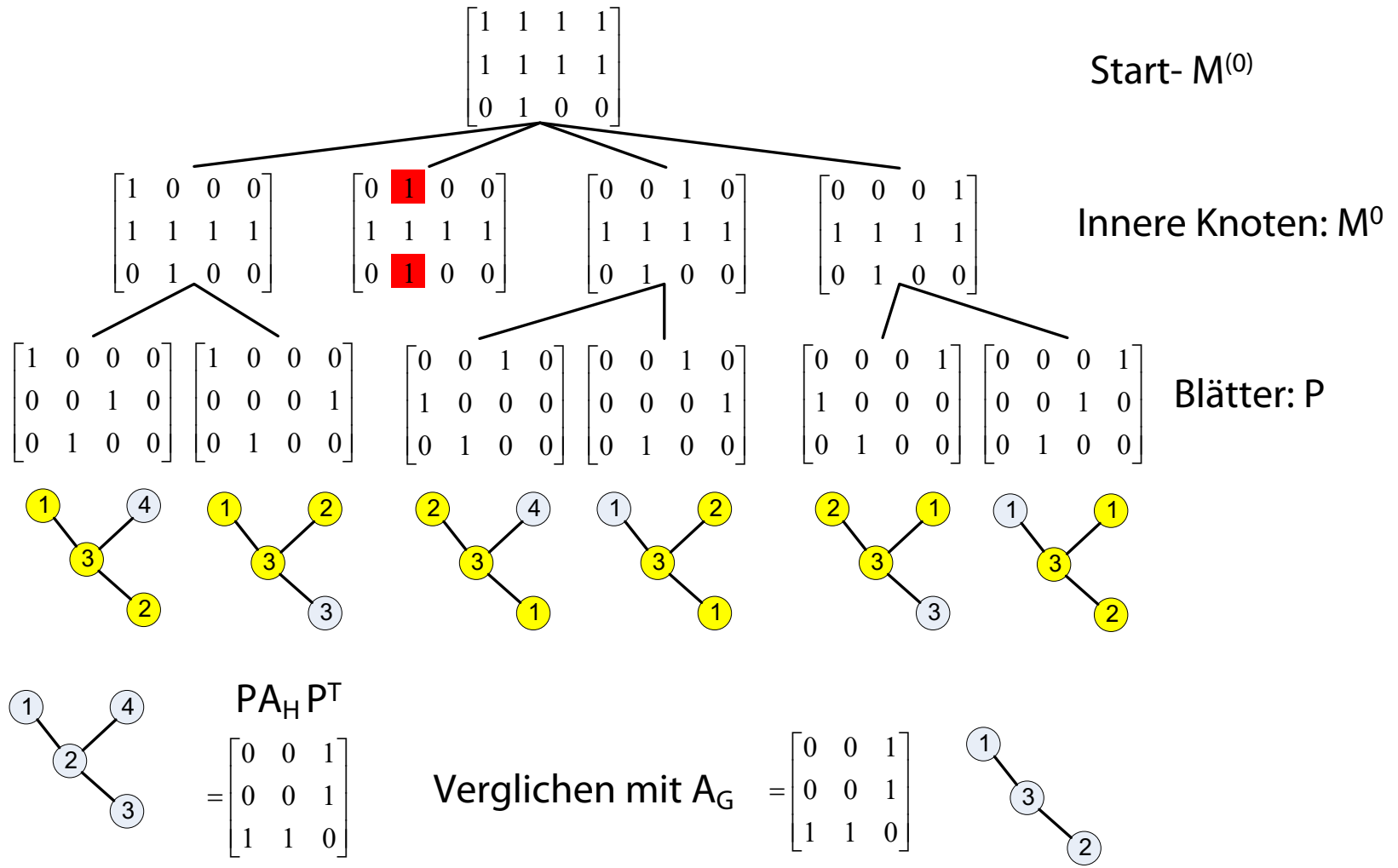
$$A_H = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$A_G = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$



$$M^0 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Ullmanns Algorithmus (mit einfachem Pruning)



Take-Home Messages

- Pruning kann anwendungsübergreifend erfolgen
 - Alpha-Beta-Pruning
 - Pure Literal (freie Wahl), Forcierte Werte, Propagierung,
 - Lernen aus Backtracking-Sackgassen
 - Nicht-chronologisches Backtracking
- Pruning kann anwendungsspezifisch erfolgen
 - Subgraph-Isomorphie,
 - Zulässige Heuristik bei A^*
- Beste Lösung wird nicht verfehlt