Algorithmen und Datenstrukturen

Prof. Dr. Ralf Möller

Universität zu Lübeck

Institut für Informationssysteme

Tanya Braun (Übungen)

sowie viele Tutoren

Danksagung

Nach einem Vortrag von

Optical Character Recognition:
Using the Ullmann Algorithm for Graphical Matching

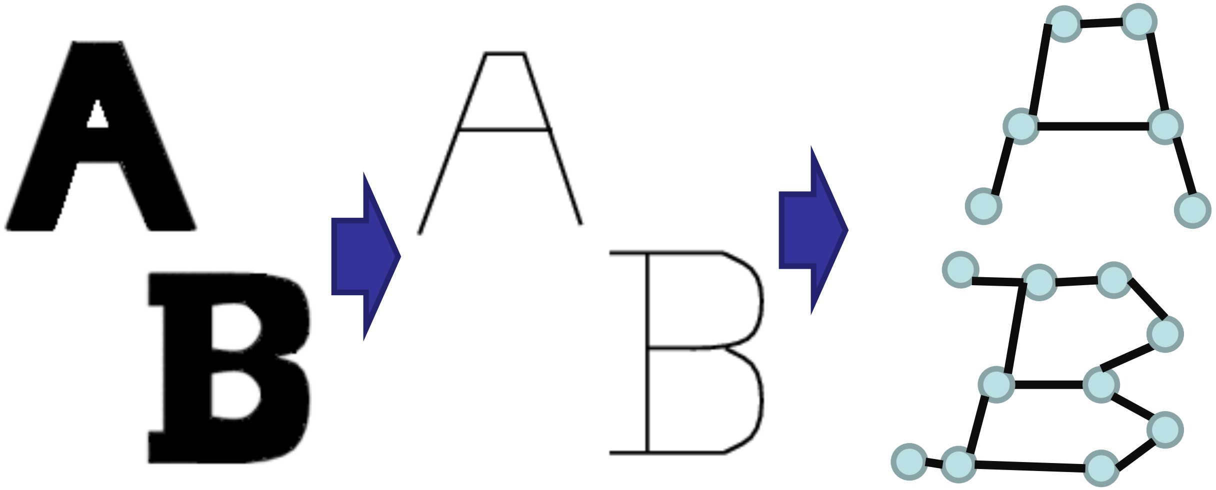
von

Iddo Aviram

OCR – Zeichenerkennung durch Graphabgleich

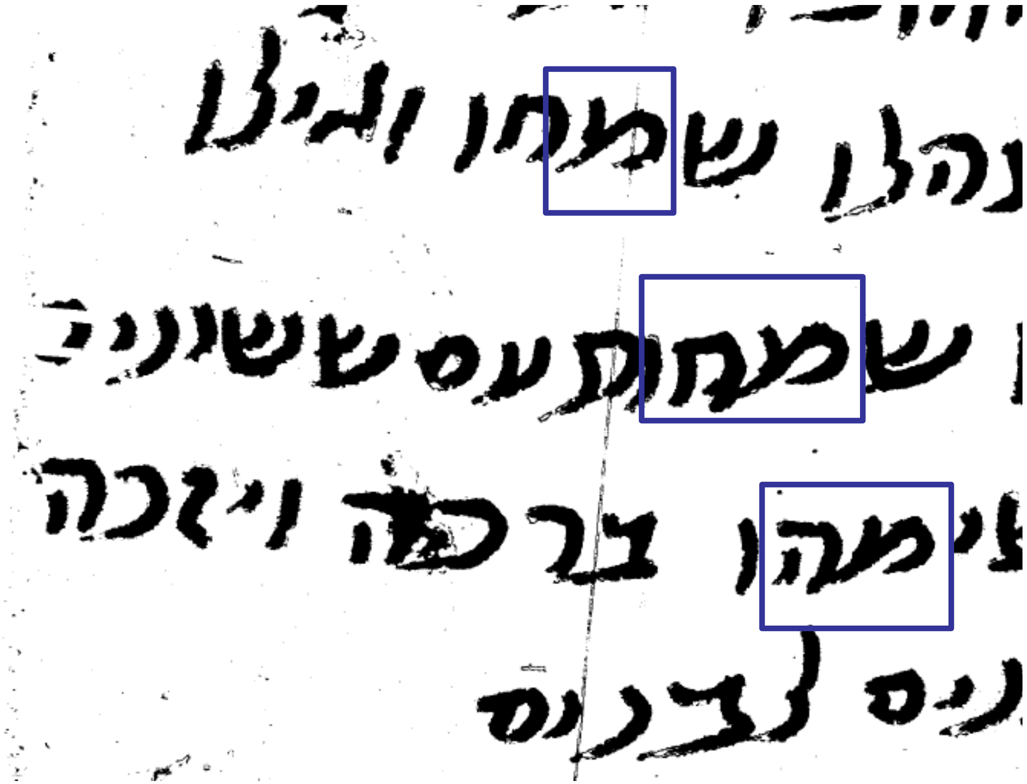
Bestimmung einer Datenstruktur:

Segmentation ⟶ Verdünnung ⟶ Graphrepräsentation



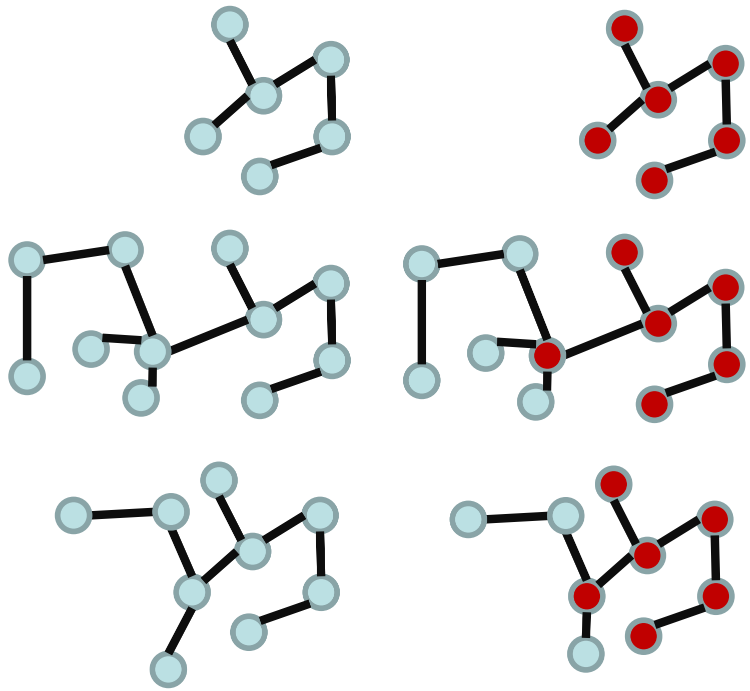
OCR durch Graphabgleich (Matching)

Durch Subgraphabgleich Kandidaten finden



Graphische Modellierung

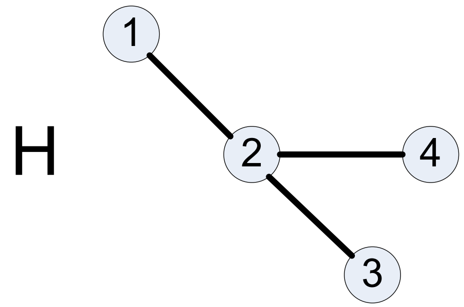
Graphischer Abgleich

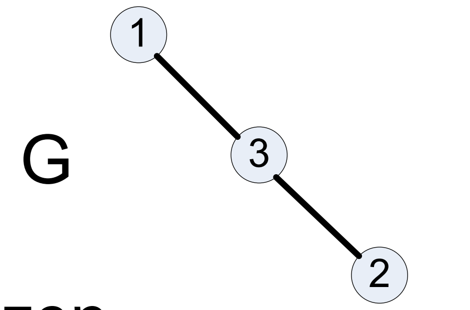


Subgraph-Isomorphie

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)
  - $\mathbf{1}_G$-$\mathbf{1}_H$
    $\mathbf{2}_G$-$\mathbf{3}_H$
    $\mathbf{3}_G$-$\mathbf{2}_H$
  - Antwort: Ja
- Sollen wir alle möglichen Korrespondenzen generieren und testen?

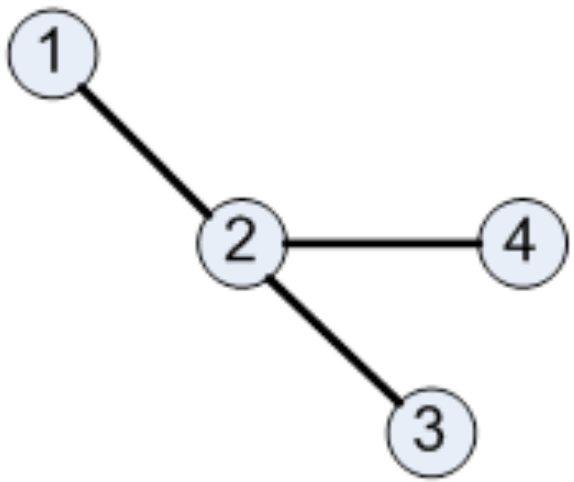




Ullmanns Algorithmus

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

H



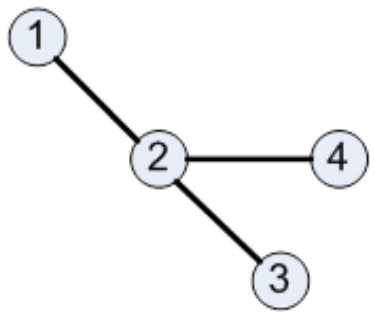
$$A_H = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

J. R. Ullmann, An Algorithm for Subgraph Isomorphism, Journal of the ACM, **1976**

Aufbau des Suchraums nach Korrespondenzen

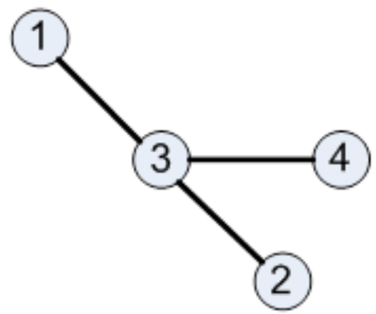
- Verwendung der sog. Permutationsmatrix
- Über Permutationsmatrix kann isomorphe Korrespondenz ausgedrückt werden

H



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

H_2



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

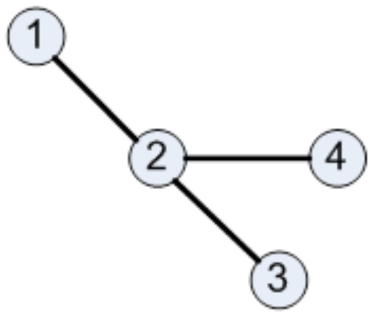
$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$F \sim P$

Isomorphiekriterium

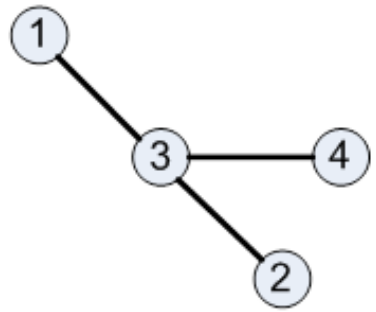
Zwei Graphen H und H_2 sind isomorph mit Korrespondenz F genau dann, wenn A_H ähnlich zu A_{H_2} ist, also eine Permutationsmatrix P~F existiert.

H



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

H_2



Isomorphe Korrespondenz ~ Permutationsmatrix

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

F~P

Isomorphiekriterium:

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

gdw H_2 ist isomorph zu H mit Korrespondenz F~P

Skeptisch?

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

P "moves" the row and columns of A_H using the mapping $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}$

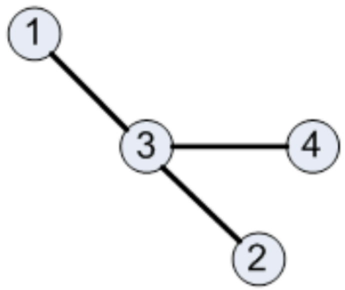
1: Moves only the rows

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

2: Moves only the columns

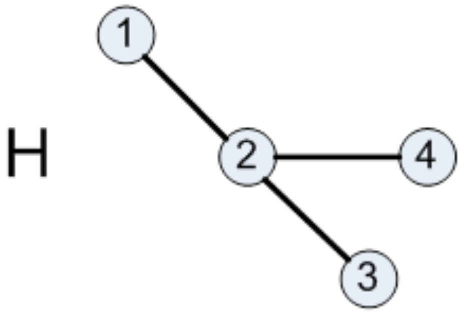
$$PA_H P^T = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix} = A_{H2}$$

H_2

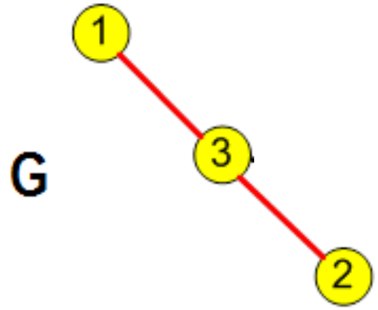


Ullmanns Algorithmus

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



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



Isomorphe Korrespondenz ~ Permutationsmatrix

F= $\mathbf{1}_G$-$\mathbf{1}_H$, $\mathbf{2}_G$-$\mathbf{3}_H$, $\mathbf{3}_G$-$\mathbf{2}_H$, $\mathbf{4}_G$-φ

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

Ullmanns Algorithmus

- Gegeben: Graphen G und H
- Gesucht: Permutationsmatrix P mit Dimensionen $|V_G|$x$|V_H|$, so dass das Subgraph-Isomorphie-Kriterium erfüllt

- Suchraum über alle möglichen Permutationsmatrizen P mit jeweiliger Prüfung des Kriteriums aufspannen?

- Das geht besser!

Ullmanns Algorithmus

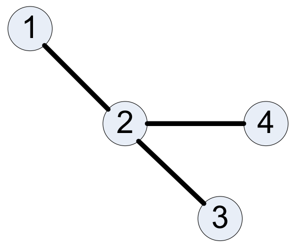
– Konstruiere Matrix $M^{(0)}$ der Dimension der P-Matrizen:

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

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

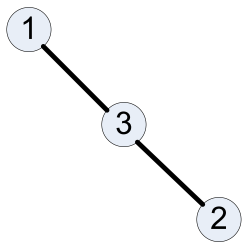
$$(a_{G\,i,j} = 1) \Rightarrow (p_{i,j} = 1) \quad \text{mit } P = M'(M'A_H)^T$$

H



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

G



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

Ullmanns Algorithmus (mit einfachem Pruning)

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

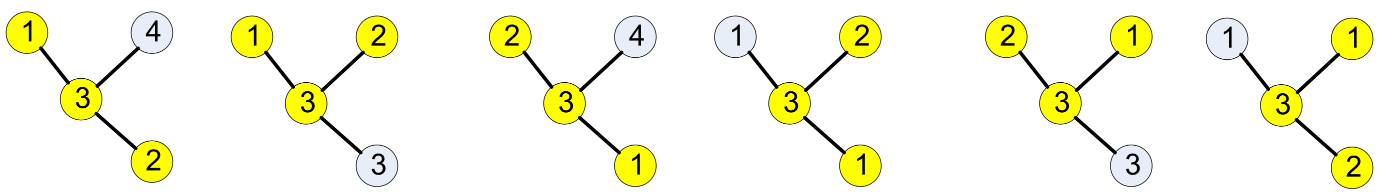
Start- $M^{(0)}$

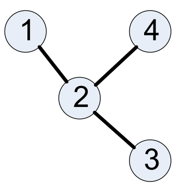
$$\begin{bmatrix}1&0&0&0\\1&1&1&1\\0&1&0&0\end{bmatrix} \quad \begin{bmatrix}0&\boxed{1}&0&0\\1&1&1&1\\0&\boxed{1}&0&0\end{bmatrix} \quad \begin{bmatrix}0&0&1&0\\1&1&1&1\\0&1&0&0\end{bmatrix} \quad \begin{bmatrix}0&0&0&1\\1&1&1&1\\0&1&0&0\end{bmatrix}$$

Innere Knoten: M

$$\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\end{bmatrix} \begin{bmatrix}1&0&0&0\\0&0&0&1\\0&1&0&0\end{bmatrix} \quad \begin{bmatrix}0&0&1&0\\1&0&0&0\\0&1&0&0\end{bmatrix} \begin{bmatrix}0&0&1&0\\0&0&0&1\\0&1&0&0\end{bmatrix} \quad \begin{bmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix} \begin{bmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\end{bmatrix}$$

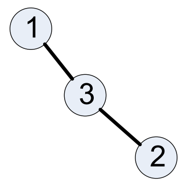
Blätter: M'





$$P = M'(M'A_H)^T = \begin{bmatrix}0&0&1\\0&0&1\\1&1&0\end{bmatrix}$$

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



Weitere Verfeinerungen des Prunings sind möglich

Take-Home Messages

- Pruning kann anwendungsübergreifend erfolgen
  - Alpha-Beta-Pruning

- Pruning kann anwendungsspezifisch erfolgen
  - Subgraph-Isomorphie,
  - Zulässige Heuristik bei A*

- Unter bestimmten Bedingungen gilt:
  Beste Lösung wird nicht verfehlt