String-Matching mit endlichen Automaten

Satz: Sei $L = \{x \in \Sigma^* | y \text{ ist Suffix von } x\}$.
Der Äquivalenzklassen DFA M_{R_L} hat die Zustände $[y_1 \cdots y_k]$, $k = 0, \ldots, m$

Beweis: Zur Erinnerung: $x_1 \; R_L \; x_2 \iff L_{x_1} = L_{x_2}$, wobei $L_x = \{z \in \Sigma^* \mid xz \in L\}$. In unserem Fall ist

$L_x = \{z \mid y \text{ ist Suffix von } xz\}$. Es ist zweierlei zu zeigen:

1. $\forall x \in \Sigma^* : \exists k : L_x = L_{y_1 \cdots y_k}$
2. $\forall j, k \leq m, j \neq k : L_{y_1 \cdots y_j} \neq L_{y_1 \cdots y_k}$

String-Matching mit endlichen Automaten

1. Es ist zu zeigen: $\forall x \in \Sigma^* : \exists k : L_x = L_{y_1 \cdots y_k}$

Sei $x \in \Sigma^*$ beliebig, $L_x = \{z \mid xz \in L\} = \{z \mid y \text{ ist Suffix von } xz\}$.
Wähle $k := \max\{j \leq m \mid y_1 \cdots y_j \text{ ist Suffix von } x\}$.
Zu zeigen $L_x = L_{y_1 \cdots y_k}$. Es gilt: $y_1 \cdots y_k$ ist ein Suffix von x.
$\supseteq$:

$$z \in L_{y_1 \cdots y_k} \Rightarrow \underbrace{y_1 \cdots y_k}_{\text{Suffix von } x} z \in L \Rightarrow xz \in L \Rightarrow z \in L_x$$

$\subseteq$: $z \in L_x \Rightarrow xz \in L \Rightarrow y$ Suffix von $xz \Rightarrow$
$|z| \geq |y|$ oder x endet mit einem Präfix von y.
Fall $|z| \geq |y|$: $z \in L = L_{y_1 \cdots y_m} = L_y$
Fall x endet mit einem Präfix von y: $xz = \cdots x_{n-k} \underbrace{y_1 \cdots y_j} z \in L$

$y_1 \cdots y_k$ das längste solche Präfix von y das zugleich Suffix von x
$\Rightarrow y$ Suffix von $y_1 \cdots y_k z \Rightarrow y_1 \cdots y_k z \in L \Rightarrow z \in L_{y_1 \cdots y_k}$.

String-Matching mit endlichen Automaten

2. Es ist zu zeigen: $\forall j, k \leq m, j \neq k : L_{y_1 \cdots y_j} \neq L_{y_1 \cdots y_k}$

Sei $j < k \leq m$. Dann ist $y_{k+1} \cdots y_m \in L_{y_1 \cdots y_k}$, kann aber nicht in $L_{y_1 \cdots y_j}$ sein, da $|y_1 \cdots y_j y_{k+1} \cdots y_m| < m$.

Satz: $\hat{\delta}(z_0, x) = z_k$ mit $k = \max\{j \leq m \mid y_1 \cdots y_j$ ist Suffix von $x\}$.

Zur Erinnerung: $\delta(z_k, a) = z_j$ mit
$j = \max\{j \leq m \mid y_1 \cdots y_j$ ist Suffix von $y_1 \cdots y_k a\}$.
Beweis des Satzes durch Induktion Ã$\frac{1}{4}$ber $|x|$.
IA. $|x| = 0$: $x = \varepsilon$: $\hat{\delta}(z_0, x) = \hat{\delta}(z_0, \varepsilon) = z_0$
IS. Sei $x = x_1 \cdots x_n$. Dann

$$\hat{\delta}(z_0, x) \quad = \quad \hat{\delta}(z_0, x_1 \cdots x_n) = \delta(\underbrace{\hat{\delta}(z_0, x_1 \cdots x_{n-1})}_{z_{k'}}, x_n)$$

String-Matching mit endlichen Automaten

$$\hat{\delta}(z_0, x) = \hat{\delta}(z_0, x_1 \cdots x_n) = \delta(\underbrace{\hat{\delta}(z_0, x_1 \cdots x_{n-1})}_{z_{k'}}, x_n)$$

$$= \begin{cases} z_{k'+1} & \text{falls } \; y_1 \cdots y_{k'+1} \text{ ist Suffix von } x_1 \cdots x_n \\ z_k & \text{falls } \; k = \max\{j \leq m \mid \underbrace{y_1 \cdots y_j} x_n \\ & \\ & \text{ist Suffix von } y_1 \cdots y_{k'} x_n\} \end{cases}$$

$$= \begin{cases} z_{k'+1} & \text{falls } \; x_n = y_{k'+1} \\ z_k & \text{falls } \; k = \max\{j \leq m \mid \quad \underbrace{y_1 \cdots y_{k'}}_{\text{Suff}(x_1 \cdots x_{n-1})} \quad x_n \\ & \\ & \quad \text{ist Suffix von } x\} \end{cases}$$

$k' = \max\{j \leq m \mid y_1 \cdots y_j \text{ ist Suffix von } x_1 \cdots x_{n-1}\}$
$k = \max\{j \leq m \mid y_1 \cdots y_j \text{ ist Suffix von } x = x_1 \cdots x_{n-1} x_n\}$
$y_1 \cdots y_j$ ist das längste Präfix von y, das auch Suffix von x.