

1. Übung Bioinformatik

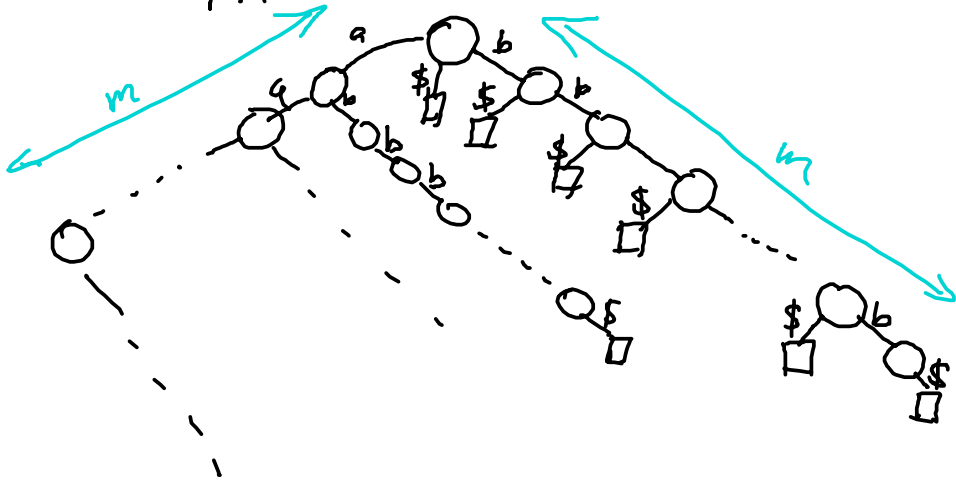
Sebastian
wild@cs...

~

- ① Einfache Suffixbäume
→ einfach Induktion über i .

② a) $a^m b^m \$ = T$

$|T| = n = 2m + 1$

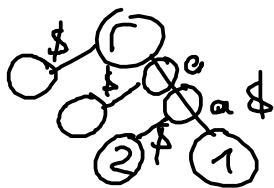


$\sim m^2$ Knoten $(m+1)^2 = \Theta(n^2)$

- b) kompakter Suffixbaum

$$T_n = a^{n-1} \$ \sim 2n-1 \text{ Knoten}$$

aaaa \$
 aaaa \$
 aa \$
 a \$



worst case:

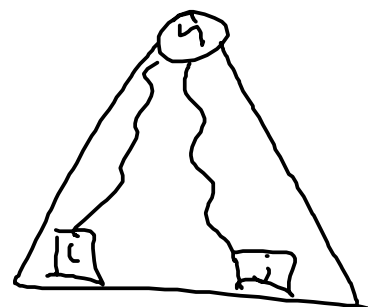
- o innerer Knoten Grad ≥ 2
- o Baum mit Grad = 2 für alle inneren Knoten
- $n-1$ Blätter
- ⇒ $2n-1$ Knoten

③ Maximal Repeats

① repeats sind Teilwörter von T

Bsp. dabaada

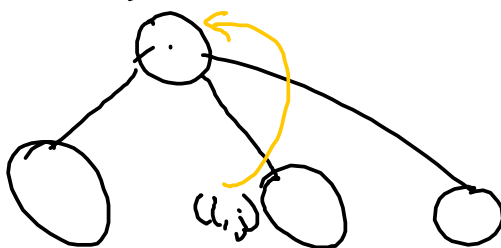
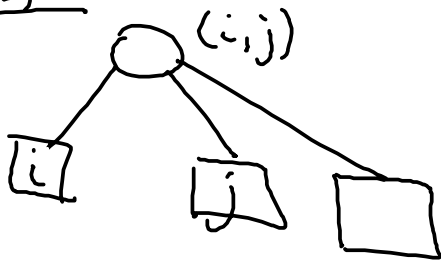
→ aa daa sind MR.



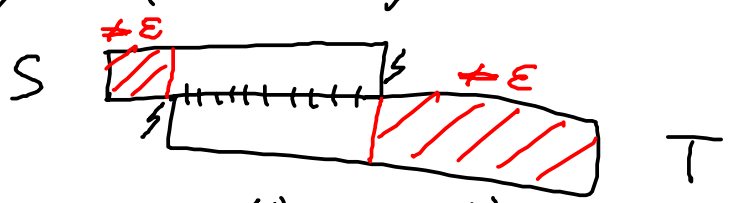
$$T_{i-1} \neq T_{j-1}$$

allgemein:

Basisfall:



④ S, T strings



$$S = xy \quad x, z \in \Sigma^+$$

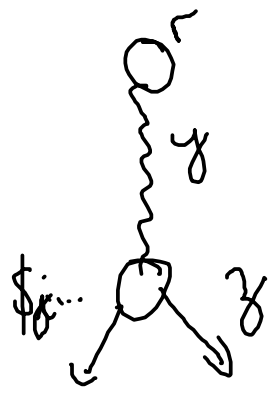
$$T = yz$$

Gay $T = (T^{(1)}, \dots, T^{(m)})$

Ges: $ov[i, j] = ov(T^{(i)}, T^{(j)})$

Suffixbaum $T^{(1)} \$_1 T^{(2)} \$_2 \dots T^{(m)} \$_m \quad \$_{1, \dots, m} \notin \Sigma$

für $i=1, \dots, m$: traversiere $T^{(i)}$. Beobachte dabei ob ein Besucher Knoten eine Kante mit Markierung $\$_x$ hat. Speichere Tupel (i, x, s) für Stringtiefe s .



- laufzeit:
- $O(n+m)$ für Suffixbaum
 - $O(n)$ für alle Traversierungen
 - für alle Knoten $\$_j$ -Kante prüfen
- #Knoten $O(n) \cdot (m + |\Sigma|)$
- $$\Sigma = O(n \cdot m)$$