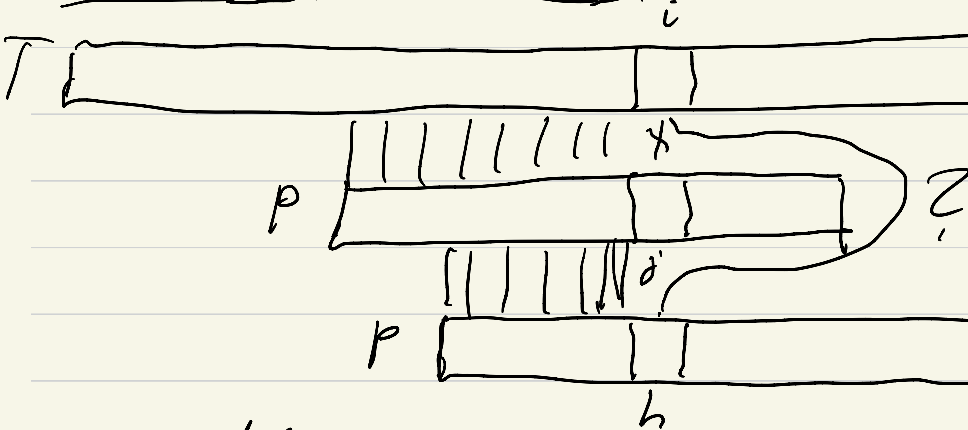
KMP algorithm

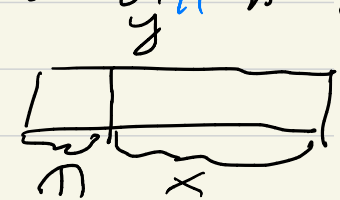
$T : \Sigma[n] \,;\, P : \Sigma[m]$



$\max h : P[0..h) \sqsupset P[0..j)$

$P_j := P[0..j) \;||\; P_h \sqsupset P_j$

$x \sqsupset y$:



Σ^*

D. $H(j) \overset{def}{\Longleftrightarrow} \{ h \in 0..j-1 \mid P_h \sqsupset P_j \}$ $(j \in 1..m)$

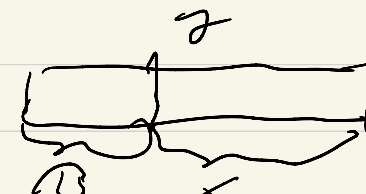
$x \sqsubset\!\!\sqsupset y \overset{def}{\Longleftrightarrow} x \sqsubset y \land x \sqsupset y$

$H(j) = \{ |x| \mid x \sqsubset\!\!\sqsupset P_j \}$

D. $next : 1..m \longrightarrow 0..m-1$

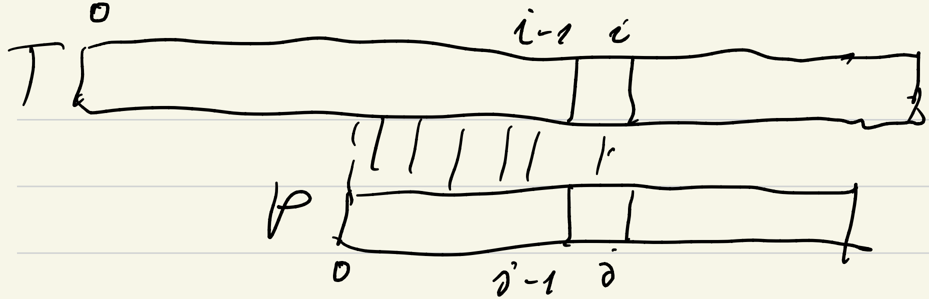
$next(j) = \max H(j)$

$x \sqsupset y$



Σ^+

$x \sqsubseteq y \;||\; x \sqsubset y$



$$P_{j+1} \sqsupset T_{i+1}$$

$$\Updownarrow$$

$$P_j \sqsupset T_i \wedge P[j] = T[i]$$

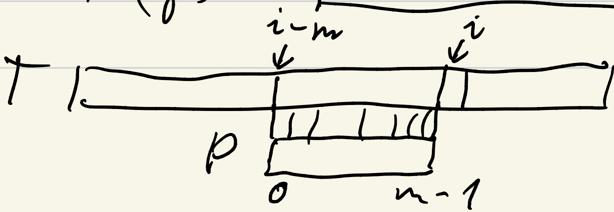
Tul

$$P_{j+1} \sqsupset T_{i+1} \iff P_j \sqsupset T_i \wedge P[j] = T[i]$$

$next(j) \in 0..j-1 \quad (j \in 1..m]$

$next(j+1) \le next(j)+1 \quad (j \in 1..m-1)$

$P[j-1] =$	B	A	B	A	B	B	A	B
$j =$	1	2	3	4	5	6	7	8
$next(j) =$	0	0	1	2	3	1	2	3



$KMP(T:\Sigma[n];\ P:\Sigma[m];\ S:\mathbb{N}\{\})$

```
next/1 : N[m]; init(next, P)
S := {}; i := j := 0
while i < n                                   Inv P_j ⊐ T_i
                                              ∧ 0 ≤ j ≤ i ≤ n
                                              ∧ j < m
    if P[j] = T[i]
        i++; j++
        if j = m
            S := S ∪ {i−m}
            j := next[m]
        else
            SKIP
    else
        if j = 0
            i++
        else
            j := next[j]
```

$t(i,j) := 2i - j \in 0..2n$

Initially: $t(i,j) = 0$

with each iteration: $t(i,j)$ strictly increases

$\Rightarrow$ KMP: max. $2n$ iteration.

KMP: min. n iteration
(initially: $i = 0$ ∧ finally: $i = n$
i increases max. by 1 in each iteration)

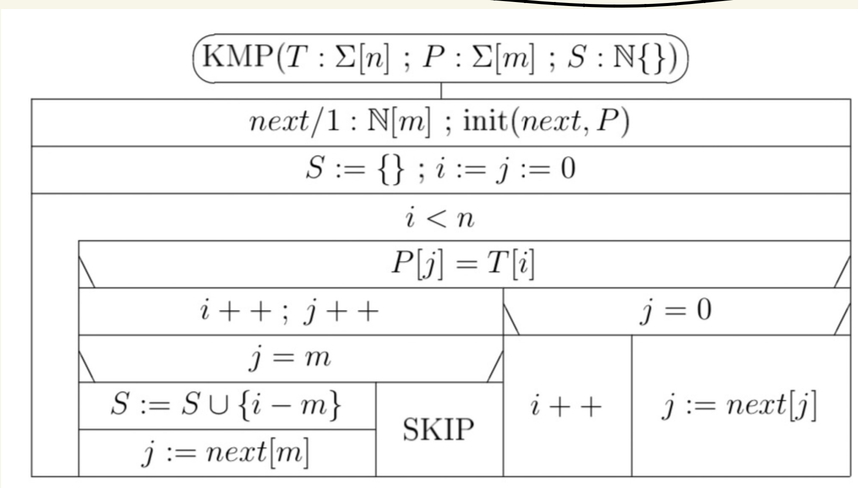
$$mT(n,m),\ MT(n,m) \in \Theta(n) + \Theta(m)$$

$$\Uparrow$$

$$mT_{init}(m),\ MT_{init}(m) \in \Theta(m)$$

$$m \le n$$

$$\boxed{mT(n,m),\ MT(n,m) \in \Theta(n)}$$



$next(1) = 0 \quad (1 \leq j < m)$

$next(j+1) = \max H(j+1) =$

$= \max \{h \in 0..j \mid P_h \sqsupset P_{j+1}\} =$ // $i := h-1$

$= \max \{i+1 \in 0..j \mid P_{i+1} \sqsupset P_{j+1}\} =$

$= \max(\{i+1 \in 1..j \mid P_{i+1} \sqsupset P_{j+1}\} \cup \{0\}) =$

$= \max(\{i+1 \in 1..j \mid \underbrace{P_i \sqsupset P_j}_{i \in H(j)} \wedge P[i] = P[j]\} \cup \{0\})$.

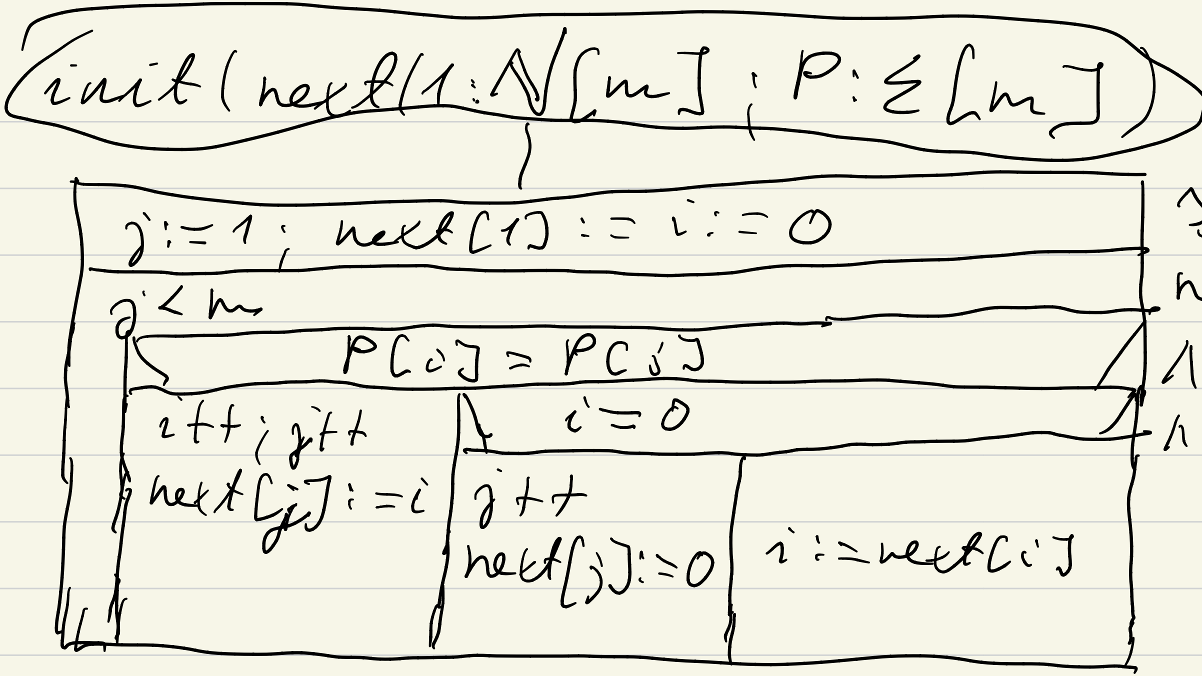
$h(j) := \{i \in 0..j-1 \mid i \in H(j) \wedge P[i] = P[j]\}$

$h(j) \subseteq H(j)$

$$next(j+1) = \begin{cases} \max h(j) + 1 & \underline{if}\ h(j) \neq \{\} \\ 0 & \underline{if}\ h(j) = \{\} \end{cases}$$

$\underline{I}_1 \ \max_{l+1} H(j) = next(\max_l H(j))$

$(\max_1 H(j) = next(j))$



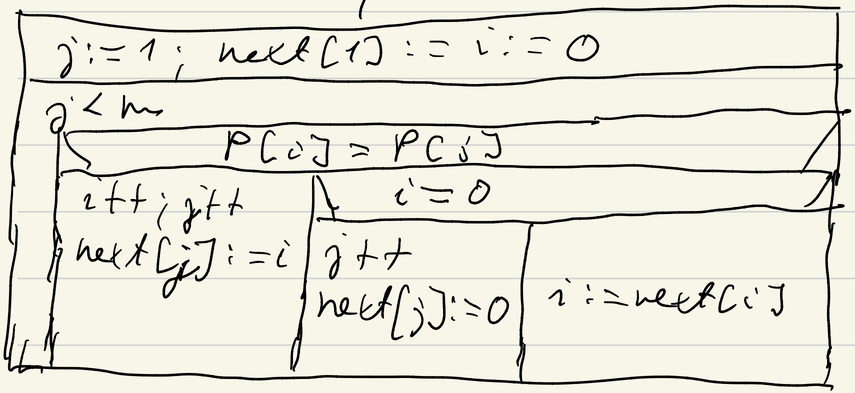
Invariant

$next[1..j] = next(1..j)$
$\wedge\ 0 \le i < j \le m$
$\wedge\ P_i \sqsupset P_j \quad (i \in H(j))$

to prove: $mT_{init}(m),\ MT_{init}(m) \in \Theta(m)$

See in the printed lecture notes !!!

init(next[1:N[m]; P:Σ[m])



P[j-1] =	B	A	B	A	B	B	A	B
j =	1	2	3	4	5	6	7	8
next[j] =	0	0	1	2	3	1	2	3

i	j	next[j]	0 1 2 3 4 5 6 7 B A B A B B A B
0	1	0	~~B~~
0	2	0	B
1	3	1	B A
2	4	2	B A B
3	5	3	B A B ~~A~~
1	5	3	B ~~A~~
0	5	3	B
1	6	1	B A
2	7	2	B A B
3	8	3	