More KMP

Review:

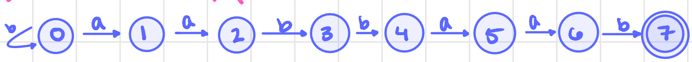
- we have skeleton machine M_p
- pattern p, text t
- looking for first occurrence of p in t

Failure function ~ computed based on p

ex) $p = aabbaab$ → $length(p) = 7$

$t = abaabaabbaab$

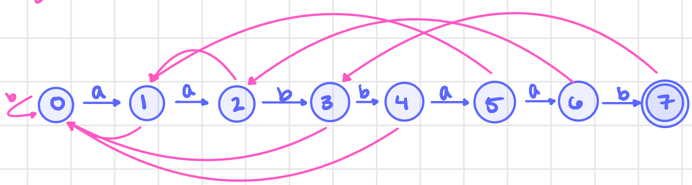
skeleton machine M_p:



failure function:

p_i	a	a	b	b	a	a	b
i	1	2	3	4	5	6	7
$f(i)$	0	1	0	0	1	2	3

add edges to Mp based on failure function:



* pink arrows represent failure

fail: so follow failure arrow without advancing Mp head

input: a b a a b a a b b a a b

State: 0 1 0 1 2 3 1 2 3 4 5 6 7

Here are the first few steps we took to determine the sequence of states:

1. Initially, Mp is in state 0
2. upon reading first letter of t, Mp enters state 1 because $t_1 = p_1 =$ "a"
3. next, the machine reads that $t_2 \neq p_2$, so Mp enters state 0 (follow the failure arrow for state 1) without moving the input head
4. since $t_2 \neq p_1$, case 2* prevails, and Mp remains in state 0 and advances its head to t_3

...

* see previous class notes

Failure Function alg:

input: pattern of length ≥ 1

output: failure function $f(p)$

function FAILURE FUNCTION($p = p_1 \dots p_l$)
L1 { $f(1) \leftarrow 0$
   $i \leftarrow 0$
L2 **for** $j \in \{2, \dots, l\}$ **do**
L3 $\quad i \leftarrow f(j-1)$
L4 $\quad$ **while** $p_j \neq p_{i+1}$ **and** $i > 0$ **do**
$\quad\quad i \leftarrow f(i)$
$\quad$ **end while**
L5 $\quad$ **if** $p_j \neq p_{i+1}$ **and** $i = 0$ **then**
$\quad\quad f(j) \leftarrow 0$
$\quad$ **else**
L6 $\quad\quad f(j) \leftarrow i + 1$
$\quad$ **end if**
end for
end function

Let's walk through the algorithm on $p = aabbaab$

$f(1) = 0$

$i = 0$

for $j = 2$

$i = f(j-1) = f(1) = 0$

$p_j = p_2 = a$
$p_{i+1} = p_1 = a$ } they're equal

while $p_j \neq p_{i+1}$ ✗ and $i > 0$ ✗

do not enter while loop

if $p_j \neq p_{i+1}$ ✗ and $i = 0$ ✓

skip to "else"

$f(j) = i + 1$ so $f(2) = 1$

for $j = 3$

$i = f(j-1) = f(2) = 1$

while $\overset{(b)}{p_3} \neq \overset{(a)}{p_2}$ ✓ and $i > 0$ ✓

$i = f(i) = f(1) = 0$

if $\overset{(b)}{p_3} \neq \overset{(a)}{p_1}$ ✓ and $i = 0$ ✓

$f(j) = f(3) = 0$

for $j = 4$:

$i = f(3) = 0$

while $\overset{(b)}{p_4} \neq \overset{(a)}{p_1}$ ✓ and $i > 0$ ✗

if $p_4 \neq p_{i+1}$ ✓ and $i = 0$ ✓

$f(4) = 0$

for $j = 5$

$i = f(4) = 0$

while $\overset{(a)}{p_5} \neq \overset{(a)}{p_1}$ ✗ and $i > 0$ ✗

if $p_5 \neq p_1$ ✗ and $i = 0$ ✓

else:

$f(5) = i + 1 = 0 + 1 = 1$

for $j = 6$

$i = f(5) = 1$

while $\overset{(a)}{p_6} \neq \overset{(a)}{p_2}$ ✗ and $i > 0$ ✓

if $p_6 \neq p_2$ ✗ and $i = 0$ ✗

else: $f(6) = i + 1 = 1 + 1 = 2$

for $j = 7$

$i = f(6) = 2$

while $\overset{(b)}{p_7} \neq \overset{(b)}{p_3}$ ✗ and $i > 0$ ✓

if $p_7 \neq p_3$ ✗ and $i = 0$ ✗

else: $f(7) = i + 1 = 2 + 1 = 3$

i	1	2	3	4	5	6	7
$f(i)$	0	1	0	0	1	2	3

✓

yippee it works!