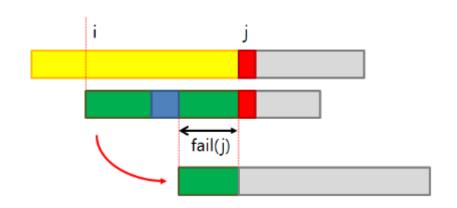
KMP알고리즘

김태희 khy@jaen.kr

KMP 알고리즘

- Knuth–Morris–Pratt Algorithm
- O(N+M)
- 패턴 정보를 활용하여 검색 시간을 단축하는 방식
- ex) S="aaaaaaaaab", W="aaaab"
- S와 W에 a가 계속 반복되므로 S의 매 위치마다 최소 5번의 비교



Х	x위치까지 접미사	fail(x)
0	a	0
1	aa	1
2	aaa	2
3	aaaa	3
4	aaaab	0

KMP 알고리즘

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
S	а	b	а	b	а	b	а	b	C	а	b	а	b	а	b	d					
W	а	b	а	b	а	b	C														

접미사

	0	1	2	3	4	5	6	7	8	9	10	11	12
S	а	b	а	b	а	С	а						
W	а	b	a	b	а	С	а						

접두사

X	x위치까지접미사	fail(x)
0	a	0
1	ab	0
2	aba	1
3	abab	2
4	ababa	
5	ababac	
6	ababaca	

W를 한 칸씩 옮겨 가면서 S(W와 같음)와 몇 칸이나 일치하는지계산.

i=0일 때는 생략

접미사

	0	1	2	3	4	5	6	7	8	9	10	11	12
S	а	b	a	b	а	С	а						
W			a	b	а	b	а	С	а				

X	x위치까지접미사	fail(x)
0	a	0
1	ab	0
2	aba	1
3	abab	
4	ababa	
5	ababac	
6	ababaca	

접미사

	0	1	2	3	4	5	6	7	8	9	10	11	12
S	а	b	a	b	а	С	а						
W			a	b	а	b	а	С	а				

X	x위치까지접미사	fail(x)
0	a	0
1	ab	0
2	aba	1
3	abab	2
4	ababa	
5	ababac	
6	ababaca	

접미사

	0	1	2	3	4	5	6	7	8	9	10	11	12
S	а	b	а	b	a	С	а						
W			а	b	a	b	а	С	а				

X	x위치까지접미사	fail(x)
0	a	0
1	ab	0
2	aba	1
3	abab	2
4	ab a ba	3
5	ababac	
6	ababaca	

접미사

	0	1	2	3	4	5	6	7	8	9	10	11	12
S	а	b	а	b	а	С	а						
W			a	b	a	b	а	С	а				

접두사

X	x위치까지접미사	fail(x)
0	а	0
1	ab	0
2	aba	1
3	abab	2
4	ab a ba	3
5	ababac	
6	ababaca	

j=3위치에서 불일치 발생

(2위치까지는 일치) fail(3-1) = 1 이용하여 → j = 1

접미사

	0	1	2	3	4	5	6	7	8	9	10	11	12
S	а	b	a	b	а	С	а						
W					a	b	а	b	а	С	а		

접두사

X	x위치까지접미사	fail(x)
0	a	0
1	ab	0
2	aba	1
3	abab	2
4	ab a ba	3
5	ababac	
6	ababaca	

j=1위치에서 불일치 발생 fail(1-1) = 0 이용하여 j = 0

접미사

	0	1	2	3	4	5	6	7	8	9	10	11	12
S	а	b	а	b	а	С	а						
W						а	b	а	b	а	С	a	

접두사

X	x위치까지접미사	fail(x)
0	a	0
1	ab	0
2	aba	1
3	abab	2
4	ab a ba	3
5	ababac	0
6	ababaca	

일치 하지 않으므로 j++, i++

접미사

	0	1	2	3	4	5	6	7	8	9	10	11	12
S	a	b	a	b	а	С	a						
W							a	b	а	b	а	С	а

X	x위치까지접미사	fail(x)
0	a	0
1	ab	0
2	aba	1
3	abab	2
4	ab a ba	3
5	ababac	0
6	<mark>a</mark> babac <mark>a</mark>	1