

KMP알고리즘

김태희

khy@jaen.kr

KMP 알고리즘

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



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

[illegible]

실패함수

접미사

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

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 일치하는지 계산.
i=0일 때는 생략

실패함수

접미사

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

접두사

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

실패함수

접미사

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

접두사

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

실패함수

접미사

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

접두사

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

실패함수

접미사

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

접두사

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

j=3위치에서 불일치 발생

(2위치까지는 일치)
 $\text{fail}(3-1) = 1$ 이용하여
 $\rightarrow j = 1$

실패함수

접미사

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

접두사

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

j=1위치에서 불일치 발생
 $\text{fail}(1-1) = 0$ 이용하여
 $j = 0$

실패함수

접미사

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

접두사

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

일치 하지 않으므로
j++, i++

실패함수

접미사

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

접두사

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