

## 0. 개요

전광판에 꾸준히 흘러가는 문자열을 통해 실제로 보여주고 싶은 문자열을 구해본다. 예를 들어 전광판에 지금 보이는 문자열이

aabaaa

인 경우 사실 보여주고 싶은 문자열은 aabaa인데 이를 반복시킨 문자열의 첫 여섯 문자만 보여진 것일 수 있다.

aabaa aabaa aabaa aabaa  
전광판에 보여지는 문자열

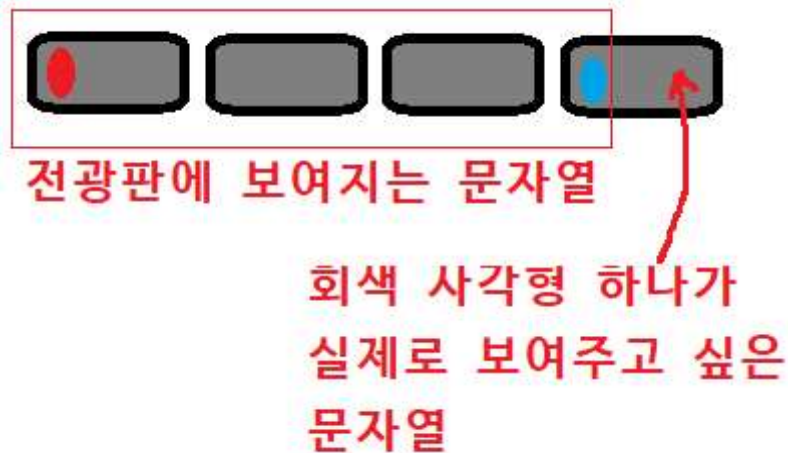
혹은 보여주고 싶은 문자열은 aabaaab인데 이 문자열이 잘려서 첫 여섯 문자만 보여진 것일 수도 있다.

aabaaab aabaaab aabaaab  
전광판에 보여지는 문자열

전광판에 보이는 문자열에 대해서 실제 보여주고 싶은 문자열은 다양할 수 있다. 이 중 가장 짧은 것이 우리의 목표이다. 이 경우 실제 보여주고 싶은 문자열 중 가장 짧은 것은 aaba 이다.

aaba aaba aaba aaba aaba  
전광판에 보여지는 문자열

전광판 문자열  $s$ 에 대해 실제로 보여주고 싶은 문자열을  $s$ 와 같게 두더라도 문제의 조건에 부합한다. 그러므로 굳이 전광판 문자열  $s$  보다 긴 문자열을 고려할 필요는 없다. 실제 보여주고 싶은 문자열이 전광판 문자열  $s$ 보다 짧은 것이었다면 다음과 같은 그림을 그려볼 수 있다.



$3 < f < 4$ 의 어떤 실수  $f$ 에 대해

(보여주고 싶은 문자열) \*  $f$  = 전광판 문자열  $s$ 가  
성립하는 하나의 예시이다.

전광판에는 보여주고 싶은 문자열이 반복하여 출력하므로  
전광판 문자열에서 붉은색 타원 영역과 푸른색 타원 영역은 서로 같다.

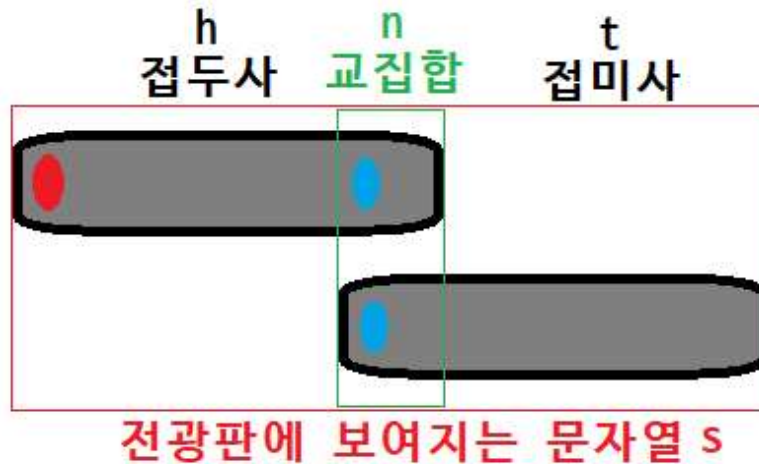
따라서 이 문제에서는 전광판 문자열에서 실패 함수의 값  
(서로 일치하는 최장 접두사·접미사의 길이)은 매우 유용하다.

위 예시에서는 전광판에 보여주고 싶은 문자열이 2 개 이상 출력되는 경우이지만, 전광판에 보여주고 싶은 문자열이 [한 개 약간 더] 출력되는 경우에도 전광판에 서로 일치하는 접두사·접미사가 나타난다. 따라서 이 문제에서는 전광판 문자열의 실패 함수의 값은 매우 중요하다고 볼 수 있다.

## 1. 두 가지 경우

전광판 문자열  $s$ 의 (서로 일치하는) 점두사  $h$ 와 점미사  $t$ 를 구할 때 두 가지 경우를 생각해 볼 수 있다. (여기서 말하는 점두사  $h$ 와 점미사  $t$ 는 최장 점두사 및 최장 점미사일 필요는 없다.)

1.a 점두사와 점미사가 서로 겹치는 경우 (점두사  $h$ 의 길이가 전광판 문자열  $s$ 의 길이의 절반을 초과하는 경우)



점두사  $h$ 와 점미사  $t$ 가 서로 겹칠 때 겹치는 영역을 교집합  $n$ 이라고 하자

이 경우 실제 보여주고 싶은 문자열은 점두사  $h$ 에 교집합  $n$ 을 뺀 부분이다. 이 문자열이 지나간 다음에는 교집합  $n$  영역에 들어가는데 이 영역의 시작은 위치 상 점미사  $t$ 의 시작이고, 점미사  $t$ 의 시작은 점두사  $h$ 의 시작과 문자가 같다. 따라서 첫 번째 보여주고 싶은 문자열이 지나간 다음 보여주고 싶은 문자열이 이전 보여주고 싶은 문자열을 반복한다고 볼 수 있다.

교집합  $n$ 의 길이는 다음 식으로 구할 수 있다.

(두 집합  $A, B$ 에 대해  $(A \cup B)$ 의 크기) = ( $A$ 의 크기) + ( $B$ 의 크기) - ( $A \cap B$ 의 크기) 이므로,

$(A \cap B)$ 의 크기) = ( $A$ 의 크기) + ( $B$ 의 크기) - ( $A \cup B$ 의 크기)가 성립한다. 이 식을 염두에 두면 다음에 나오는 식을 유도할 수 있다.)

$$n.length = 2 * h.length - s.length;$$

보여주고 싶은 문자열은 점두사  $h$ 에서 교집합  $n$ 을 뺀 부분이므로 보여주고 싶은 문자열의 길이는 다음과 같다.

$$\begin{aligned} \text{보여주고 싶은 문자열.length} &= h.length - n.length \\ &= s.length - h.length \end{aligned}$$

1.b 점두사와 점미사가 겹치지 않는 경우 (점두사  $h$ 의 길이가 전광판 문자열  $s$ 의 길이 절반 이하인 경우)



점두사  $h$ 와 점미사  $t$  사이에 점두사와 점미사에 속하지 않는 영역을 중립  $n$ 이라고 하자

점두사  $h$ 의 길이가 전광판 문자열  $s$  길이의 정확한 절반인 경우

중립  $n$ 은 빈 문자열 (길이 0짜리 문자열)이 된다.

이 경우 실제 보여주고 싶은 문자열은 점두사  $h$ 에 중립  $n$ 을 합한 영역이라고 볼 수 있다. 이 문자열이 지나간 다음에는 점미사 부분에 들어가는데 이 점미사 부분의 시작은 점두사 부분의 시작과 같다. 따라서 첫 번째 보여주고 싶은 문자열이 지나간 다음 보여주고 싶은 문자열이 이전 보여주고 싶은 문자열을 반복한다고 볼 수 있다. (또한 두 번째 보여주고 싶은 문자열은 중립  $n$ 영역이 잘려서 부분만 출력 됨을 알 수 있다.)

중립 n의 길이는 다음 식으로 구할 수 있다. (전광판 문자열 s에서 접두사 h, 접미사 t를 빼면 중립 n이 나오므로 다음 식이 유도된다.)

$$n.length = s.length - 2 * h.length;$$

보여주고 싶은 문자열은 접두사 h와 중립 n을 합한 영역이므로 반복 마디의 길이는 다음과 같다.

$$\begin{aligned} \text{보여주고 싶은 문자열.length} &= h.length + n.length \\ &= s.length - h.length \end{aligned}$$

두 경우에 대해 전부

$$\text{보여주고 싶은 문자열.length} = s.length - h.length$$

가 성립한다. 따라서 케이스 분기를 사용하지 않고 해당 식으로 반복 마디의 길이를 구할 수 있다.

## 2. 최장 접두사

전광판 문자열 s에 대해 (서로 일치하는) 접두사, 접미사는 다양하게 구할 수 있고, 그 다양한 접두사들에 대해 등식

$$\text{보여주고 싶은 문자열.length} = s.length - h.length$$

이 성립한다. (등식을 유도하는 과정에서 접두사가 가장 길어야 한다는 가정은 사용하지 않았으므로 최장 접두사에서만 이 등식이 성립하는 것은 아니다.) 하지만 이 문제의 목표는 보여주고 싶은 문자열을 최소화하는 것이다.

$\text{보여주고 싶은 문자열.length} = s.length - h.length$
<div style="display: flex; justify-content: space-around;"><div style="text-align: center;"><p>이 값이 최소가 되길 원하는데...</p></div><div style="text-align: center;"><p>이 값은 고정되어 있다.</p></div><div style="text-align: center;"><p>이 값이 최대가 되어야 좌변의 값이 최소가 된다.</p></div></div>

보여주고 싶은 문자열의 최소 길이를 구해야 하므로 접두사의 길이(h.length)의 최대치를 구해야 한다. 이게 바로 전광판 문자열 s의 최장 접두사를 구해야 하는 이유이다. (서로 일치하는) 접두사, 접미사 중 가장 긴 것을 구해야 보여주고 싶은 문자열 후보 중 가장 짧은 것을 구할 수 있다.