"최대 길이 공통 괄호 문자열" 문제 풀이

작성자: 최은수

부분문제 1

S의 부분 문자열인 괄호열을 S의 부분 괄호열이라 하자. 모든 괄호열은 여는 괄호 하나에서 시작해 닫는 괄호 하나에서 끝나기 때문에 A의 부분 괄호열의 개수는 많아야 $\frac{|A|^2}{4}$개이다. 마찬가지로, B의 부분 괄호열의 개수는 많아야 $\frac{|B|^2}{4}$개이다.

A의 부분 괄호열 중, B의 부분 괄호열인 것만을 모아, 이 중 최대 길이인 것을 구하면 $O(|A|^5 + |B|^5)$에 문제를 해결할 수 있다.

부분문제 2

A의 부분 괄호열과 B의 부분 괄호열 쌍 모두를 비교하는 것이 아니라, 각각을 정렬하여 비교하면 $O(|A|^3 + |B|^3)$에 문제를 해결할 수 있다.

부분문제 3

A의 모든 접미사에 대한 트라이를 만들고, 이 중 올바른 괄호열인 문자열에 해당하는 노드에 표시를 하자.

B의 모든 접미사를 트라이에서 검색하여, 검색하며 만나는 표시된 노드 중 깊이가 가장 큰 노드의 깊이가 답이 된다.

트라이의 구성 및 검색에는 문자열의 길이의 합만큼의 시간이 소요되므로, 시간복잡도는 $O(|A|^2 + |B|^2)$이다.

부분문제 4

어떤 문자열 S가 있을 때, S의 i번째 문자까지 있는 여는 괄호의 개수에서 닫는 괄호의 개수를 뺀 값을 문자열 S의 i번째 높이 h_i로 정의하자. 편의상 $h_0 = 0$으로 정의하자.

$A[i..j]$가 올바른 괄호열인 조건과, $h_{i-1} = h_j$이며, $i \le k < j$를 만족하는 모든 k에 대해 $h_k \ge h_j$인 조건은 동치이다.

$1 \le i \le |A|$를 만족하는 i에 대해, 다음 방법을 통해 $A[i..j]$가 올바른 괄호열인 가장 큰 j를 구할 수 있다.

먼저, 이분탐색을 통해 $i \le k$이면서, $h_{i-1} - 1 = h_k$인 가장 작은 k를 찾는다. 없다면 $k = |A|$로 둔다.

다음으로, $j \le k$이면서, $h_{i-1} - 1 = h_k$인 가장 큰 j를 찾는다. 만약 $j = i - 1$이라면 i번째 문자에서 시작하는 올바른 괄호열은 없다.

이런 $A[i..j]$들 중 가장 긴 문자열의 길이가 답이 된다. 시간복잡도는 $O(|A| \log |A|)$이다.

부분문제 5

두 문자열 A와 B를 합친 $A * B$의 접미사 배열을 구하자. 그 후, 이 접미사 배열의 LCP 배열을 구하자.

A의 접미사 $A[i..|A|]$에 대해, 이 문자열의 B의 부분 문자열인 가장 긴 접두사는 다음의 두 가지 경우이다.

- 사전순으로 $A[i..|A|]$보다 앞인 B의 접미사 중 가장 뒤에 있는 접미사와의 최장 공통 접두사

- 사전순으로 $A[i..|A|]$보다 뒤인 B의 접미사 중 가장 앞에 있는 접미사와의 최장 공통 접두사

두 경우의 길이 모두 스위핑을 통하여 모든 접미사에 대해 $O(|A|+|B|)$에 구할 수 있으며, 둘 중 최댓값의 배열을 $l_1, l_2, ..., l_{|A|}$라 하자.

답의 경우 부분문제 4의 마지막 이분탐색에서 $j < i + l_i$라는 조건만 추가되었기 때문에, k 대신 $min(k, i + l_i - 1)$를 사용하여 구할 수 있다. 시간복잡도는 $O(|A| \log |A| + |B| \log |B|)$이다.