What is z?

Given a string of length .

is the length of the longest substring starting from which is also a prefix of .

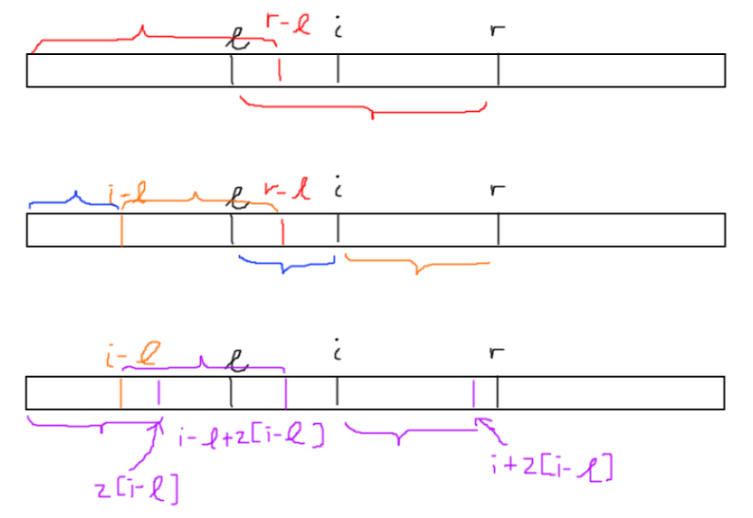
General concept

Iterate over .

Maintain an interval , which is the interval with maximum such that and is a prefix-substring.

The algorithm

Suppose we have for and .

if , there does not exist any prefix-substring such that . Calculate new value from naively.

Otherwise, let .

Since is a prefix-substring, .

We can divide into and , similarly for . Therefore, .

We want to find , so we can use as .

2 cases:

: longest prefix-substring starting from i is length , no need to update

: update to , increment until is maximum prefix-substring,

Linear because is like 2 pointers.

Problem 1: standard pattern matching

Given 2 strings and of size and , find all occurences of in as a substring.

Just run z algorithm on the string , where # is a character not appearing in or .

Problem 2: number of distinct substrings

Given a string of length , find the number of distinct substrings of .

Add to one by one, for each iteration count new substrings formed (that is not already in )

For each iteration, reverse and do z algorithm on it. Maximum is the length of maximum that already exist, so add to answer.

Time complexity

Can recalculate answer after adding or deleting a character at the end of the string

Problem 3: string compression

Given a string , find the shortest string such that concatenated times ( can be any number) is equal to .

calculate z function, answer is first such that is and

