

6.7 找到重复的子字符串

力扣题号：459.重复的子字符串。

【题目描述】

给出一个非空字符串，判断是否有重复子字符串。

【示例一】

输入："abcabc"。

输出：True。

【示例二】

输入："asdfa"。

输出：False。

【思路】

这其实是一道标准的 KMP 的题目，为什么寻找重复子字符串也涉及 KMP 算法了呢？

这里就要说一说 next 数组了，next 数组记录的就是最长相等前后缀的长度，如果 $\text{next}[\text{len}-1] \neq -1$ ，则说明字符串有相同的前后缀。

最长相等前后缀的长度为 $\text{next}[\text{len}-1]+1$ ，数组长度为 len。

如果 $\text{len} \% (\text{len} - (\text{next}[\text{len}-1]+1)) = 0$ ，则说明（数组长度-最长相等前后缀的长度）正好可以被数组的长度整除，该字符串中有重复的子字符串。

数组长度减去最长相等前后缀的长度相当于第一个重复子字符串的长度，也就是一个重复周期的长度，如果这个周期可以被整除，则说明整个数组就是这个周期的循环。

强烈建议读者把 next 数组打印出来，分析 next 数组的规律，有助于理解 KMP 算法。

以字符串 asdfasdfasdf 为例，如图 6-12 所示。

此时 $\text{next}[\text{len}-1]=7$ ， $\text{next}[\text{len}-1]+1=8$ ，8 就是字符串 asdfasdfasdf 的最长相等前后缀的长度。

$(\text{len} - (\text{next}[\text{len}-1]+1)) = 12$ （字符串的长度） $- 8$ （最长公共前后缀的长度） $= 4$ ，4 正好可以被 12（字符串的长度）整除，说明有重复的子字符串（asdf）。