

四月OA两道题目：

1. 给定长度l，找出一个字串里面所有长为l的子串。字串的字符不能重复。双指针O(n)扫一次，然后加个hash map 或者 trie 都可以。

2. 一个字串，每个字符表示一个场景。两个相同字符之间认为是一个连续的场景。例如: abcda, 可以认为这五个字符是同一个场景。或者acafghbeeb可以认为又aca和beb两个场景。场景之间有重合那么就把场景合起来，例如abcab，这里abca和bcab是重合的，那么认为这五个字符是同一个场景。给一个字符串，求最长场景。

这个其实就是求出多个区间，然后求链接起来的最长区间了，没啥好说。

已经收到安塞的邀请。求祝福。求大米。

两道亚马逊OA遇到的题

1.get all substring with k distinct character and length k

需要把所有满足条件的子串放入集合并去重， sliding window + hashmap

他题说的是长度为k有k distinct character, 所以就是你说这意思，长度k,没重复

2.寻找最短的sub sequence contains 所有的tag

input 大概是这样

tag_list = ["made","in","china"]

all_tags = ["made", "a","b","c","in", "china","made","b","c","d"]

答案则为 [4,6], 不需要保存顺序，只需返回最短的subsequence包含所有tag_list的元素

当时时间有点紧，开始也用错了方法，后来想到应该是用前向型双指针，希望可以帮到大家

详情请看：

<http://www.1point3acres.com/bbs/thread-348589-1-1.html>

第一题：给一个string，是一篇小说内容，比如"Jimmy has an apple, it is on the table" 再给一个exclude word list，比如 "an" "a" "is" "the"，让你给出这个string里出现频率最高的单词。

第二题，给你一个log file，List<String>，每个元素代表log file一行，在每个元素里面，有substring，他们用空格分开，比如"a1ws efg7 i90s"，String里面只有字母和数字，第一个substring是id，依据id后面的substring给log file排序，字母在top。

比如input是

"fhie 1df8 sdfs"

"fdsf 2def sees"

"efe2 br9o fjsd"

"asd1 awer jik9"

output是：

"asd1 awer jik9"

"efe2 br9o fjsd"

"fhie 1df8 sfd8"

"fd8f 2def sees"

第一题是词频题，给一个String，然后一个exclusiveWordList (List<String>), 找出String里去掉exclusive word 之后的最高频词汇的list。

两个Map，一个Map记录词频，什么类型的Map均可。另一个Map这样设计，TreeMap<Integer, Set<String>>, key是单词的出现次数，value是哪些单词出现了这么多次，每次新单词进来两个Map都更新。必须用TreeMap，因为TreeMap可以按照顺序输出key，可以做到logN复杂度，这样输出前k个，就是kLogN复杂度。

需要输出top N的时候，就把TreeMap的前k个输出出去。

另外，最近Alexa在狂招人，后面可能还有hiring event，想去的抓住机会啊。。

