

# HW 10

1. 试说明如何扩展 Rabin-Karp 算法用于处理以下问题: 在一个  $n \times n$  的二维字符数组中搜索一个给定的  $m \times m$  的模式。(该模式可以在水平方向和垂直方向移动, 但是不可以旋转。)

对于水平的每一行字符串, 对其求哈希值, 再对这  $m$  个哈希值再求一次哈希即可。

对于水平移动, 分别对每一行运用  $t_{s+1} = (d(t_s - T[s+1]h) + T[s+m+1]) \bmod q$ , 再对这  $m$  个结果再求一次哈希即可得到结果。对于垂直移动, 已经得到了对之前  $m$  行的哈希值:  $t_i, t_{i+1}, \dots, t_{i+m-1}$ , 故垂直向下移动一行的哈希值  $t' = (d(t[i, \dots, i+m-1] - t_i * d^{m-1}) + t_{i+m}) \bmod q = (d(t - t_i * h) + t_{i+m}) \bmod q$

2. 对字母表  $\Sigma = \{a, b\}$ , 画出与模式 **ababbabbababbababbabb** 对应的字符串匹配自动机的状态转换图

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
a	b	a	b	b	a	b	b	a	b	a	b	b	a	b	a	b	b	a	b	b

状态	a	b
0	1	0
1	1	2
2	3	0
3	1	4
4	3	5
5	6	0
6	1	7
7	3	8
8	9	0
9	1	10
10	11	0
11	1	12
12	3	13
13	14	0
14	1	15
15	16	8
16	1	17
17	3	18
18	19	0
19	1	20
20	3	21
21	9	0

3. 描述解决这一问题的算法, 并对算法效果与复杂度进行分析 (不用完成 oj 上的题目!!!) <https://202.38.86.171:1443/problem/E6-1>

直接用kmp算法：通过 $\pi[q] = \max\{k : P_k \sqsubset P_q\}$ 计算  $\pi$ ，在将字符串和模式串进行匹配，失配时使用  $\pi$  进行滑动。对于每个模式串和原串，用数组p和s存储，每次记录下其长度。

**时间复杂度：**计算前缀函数时，由于 k 最多增加 m-1次，且 k 的减少是在while循环中减少的，k永远不可能为负值，采用聚合分析，故此处 k 的减少次数也就是while循环中的次数，最多为 m-1次，故ComputePrefixFunction的时间复杂度为 $O(m_t)$ ，同理，kmp算法的时间复杂度为 $O(n_t)$ ，故总时间复杂度为

$$\sum_{t=1}^T (O(m_t) + O(n_t)) = \sum_{t=1}^T O(m_t + n_t) = O\left(\sum_{t=1}^T (m_t + n_t)\right)$$

**空间复杂度：**p、s的大小分别为最大的那个模式串、原串的长度，故总空间复杂度为 $O(\max(m_t) + \max(n_t))$

由于题中数据的模式串长度不超过10,000, 原串长度不超过1,000,000，故实际所需内存约为

$$1000000 * \text{sizeof}(\text{char}) + 10000 * \text{sizeof}(\text{char}) = 1MB + 10KB \approx 1MB$$

4. 对于在无向图中寻找最长简单回路这一问题，给出其形式化的定义并给出其相关的判定问题。另外，给出与该判定问题对应的语言。

形式化定义：无向图中从源节点到目的节点的最长的简单路径

判定问题：i=<G, u, v, k>是一个实例，如果存在 G 中从 u 到 v 存在一条简单路径的长度最少为 k 条边，则 output 为 1，否则为 0。

语言：

$PATH = \{ \langle G, u, v, k \rangle : G = (V, E) \text{ 是一个无向图}, u, v \in V, k \geq 0 \text{ 是一个整数, 即 } G \text{ 中从 } u \text{ 到 } v \text{ 存在一条长度最少为 } k \text{ 条边} \}$