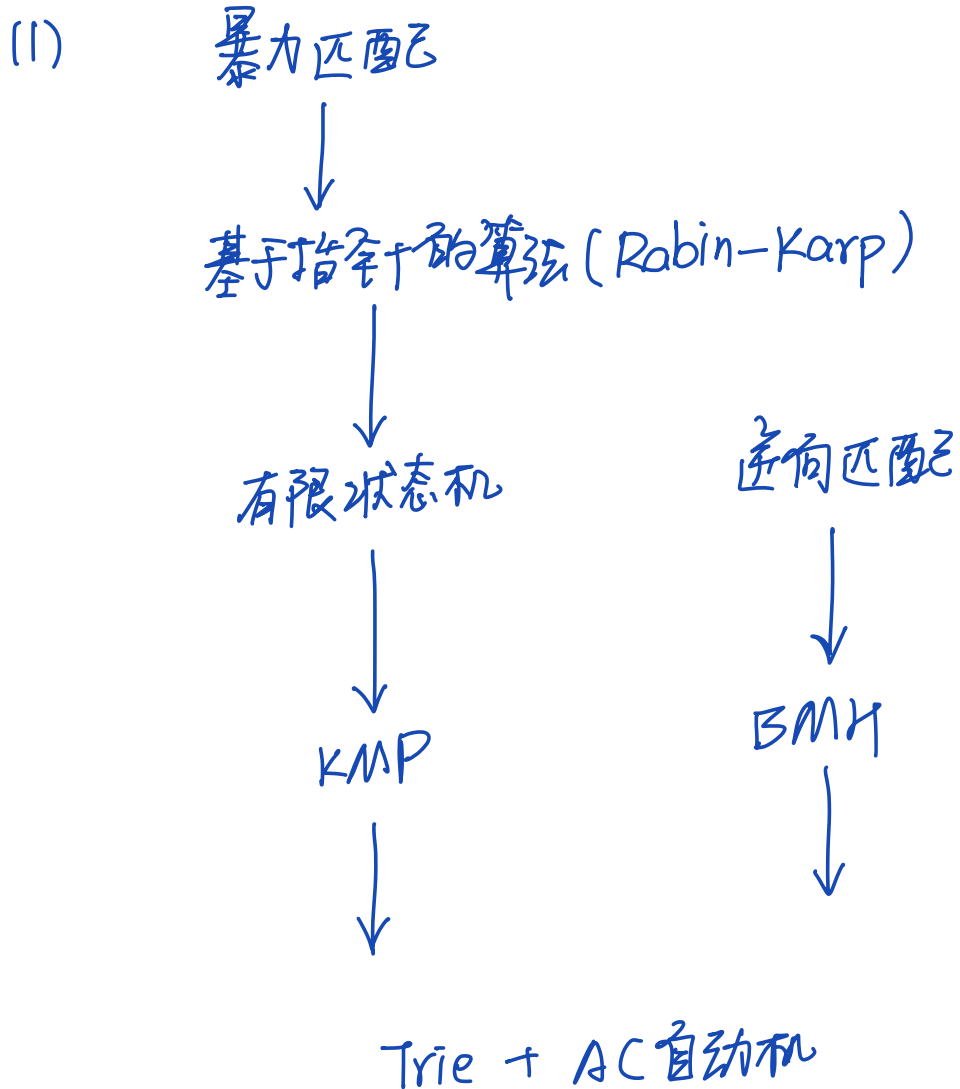


# 字符串匹配问题



主要讲: KMP, BMH

(2) KMP:

计算前缀数组

COMPUTE - PREFIX (P):

$m \leftarrow \text{len}(P)$

new  $\pi[0 \dots m]$   $m+1$  长度

$\pi[0] \leftarrow 0, \pi[1] \leftarrow 0$  相当于DP的初始化

for  $q \leftarrow 2$  to  $m$  do:

$k \leftarrow \pi[q-1]$  前缀的最后一个元素

    while  $k > 0$  and  $P[k] \neq P[q-1]$

$k \leftarrow \pi[k]$

    if  $P[k] = P[q-1]$ :

$k \leftarrow k+1$

$\pi[q] \leftarrow k$

return  $\pi$

---

主算法:

KMP-SEARCH(T, P)

$m \leftarrow \text{len}(P)$

$n \leftarrow \text{len}(T)$

    flag  $\leftarrow 1$

    ...  $\text{COMPUTE\_PREFIX}(P)$

$\pi \leftarrow \text{COMPUTE-TRIE-FACTOR}$

$q \leftarrow 0$  *是模式串的指针*

for  $i \leftarrow 0$  to  $n-1$  do:

while  $q > 0$  and  $P[q] \neq T[i]$ :

$q \leftarrow \pi[q]$

if  $P[q] = T[i]$ :

$q \leftarrow q + 1$

if  $q = m$ : *搜集结果*

    Print "ans: "  $i - m + 1$ ,

    flag  $\leftarrow 0$

$q \leftarrow \pi[q]$   *$\pi[m]$ 就是成功匹配后, 可能会重复*

if flag = 1 then:

    return -1 *没有找到*

(3) BMH:

BMH-SEARCH( $T, P$ ):

flag  $\leftarrow 1$

for  $c \leftarrow 0$  to  $\Sigma$ : *字典*

    shift[c] = m *字典*

*初始化*

for  $k \leftarrow 0$  to  $m-2$ :

*生成转移表*

```

shift[P[k]] = m - 1 - k
s ← 0
while s ≤ n - m do:
    j ← m - 1
    while T[s + j] = P[j] do:
        j ← j - 1
    if j < 0 then:
        print "ans: " s
        flag ← 0
    s ← s + shift[T[s + m - 1]]
if flag = 1 then:
    return -1

```

$shift[w] = \begin{cases} 1) m - 1 - \max\{i < m-1 \mid w \in P[0..m-2]\} \\ 2) m, \text{ otherwise} \end{cases}$

最右下标  
 不忘最后一个元素  
 对齐除尾元素以外的最右相同元素

## 会画转移表

(2) 如下表格是BMH算法的匹配过程，请写出第2到6行的内容。

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
T	a	b	d	a	c	b	a	c	d	b	c	a	c	a	b	c	a	c
1	a	b	c	a	c													
2			a	b	c	a	c											
3				a	b	c	a	c										
4						a	b	c	a	c								
5									a	b	c	a	c					
6											a	b	c	a	c			
7														a	b	c	a	c