

本节内容

KMP算法

朴素模式匹配算法
的优化

王道考研/CSKAOYAN.COM

1

KMP算法



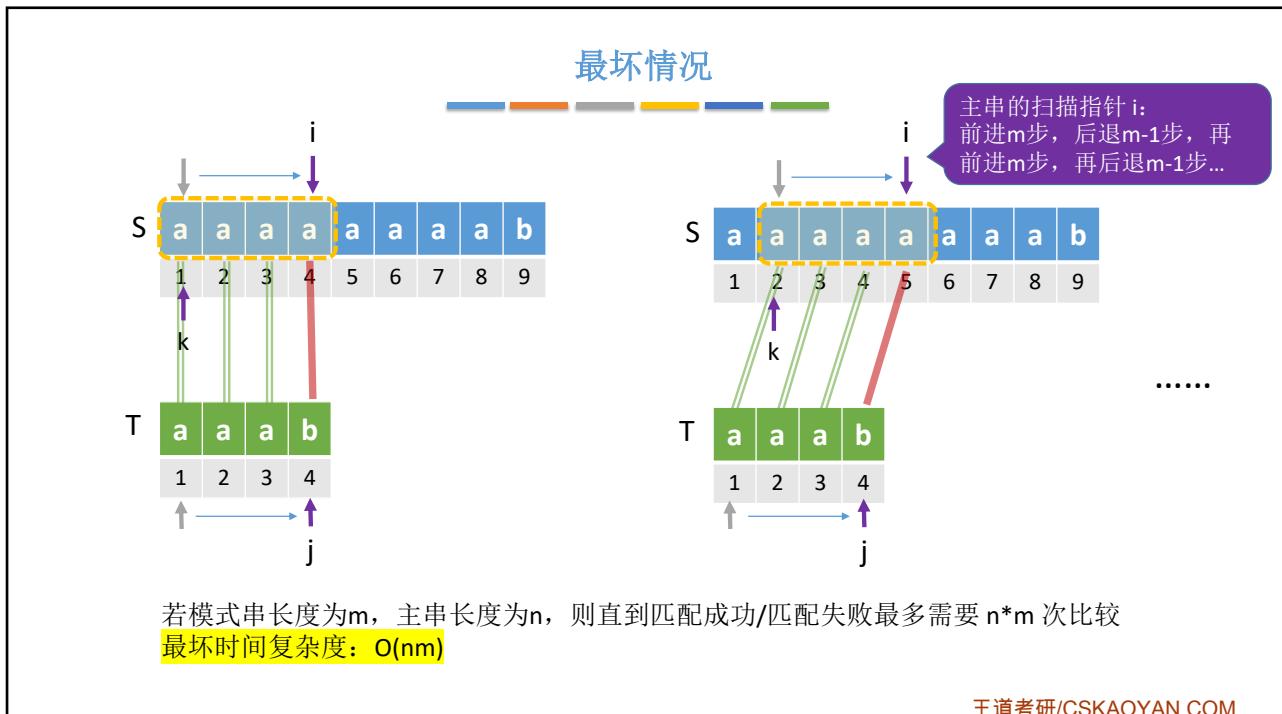
由D.E.Knuth, J.H.Morris和V.R.Pratt提出，因此称为 KMP算法



你不要凶我 我害怕。

王道考研/CSKAOYAN.COM

2

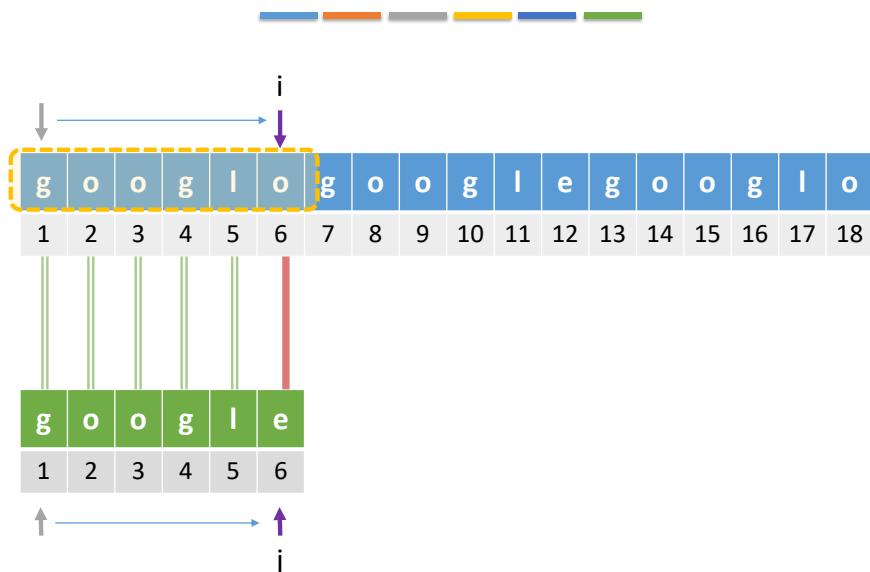


3



4

比较坏的情况



王道考研/CSKAOYAN.COM

5

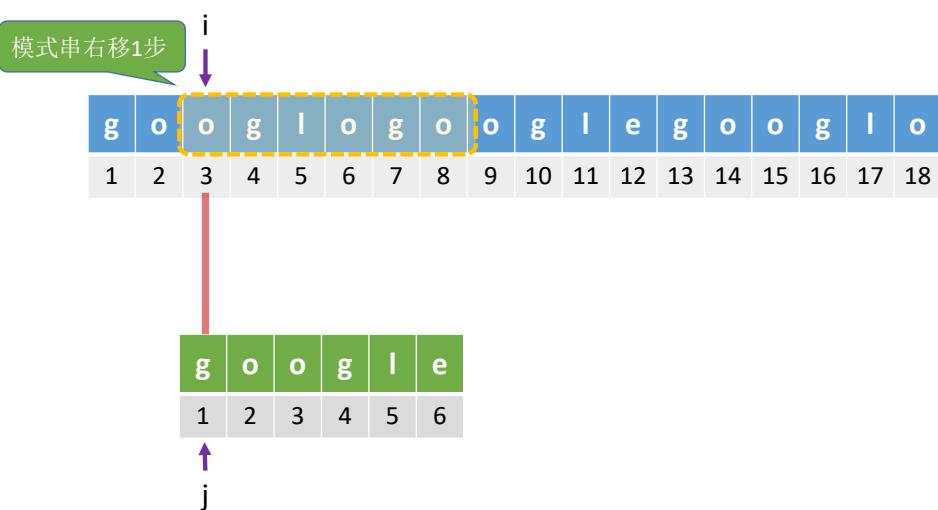
比较坏的情况



王道考研/CSKAOYAN.COM

6

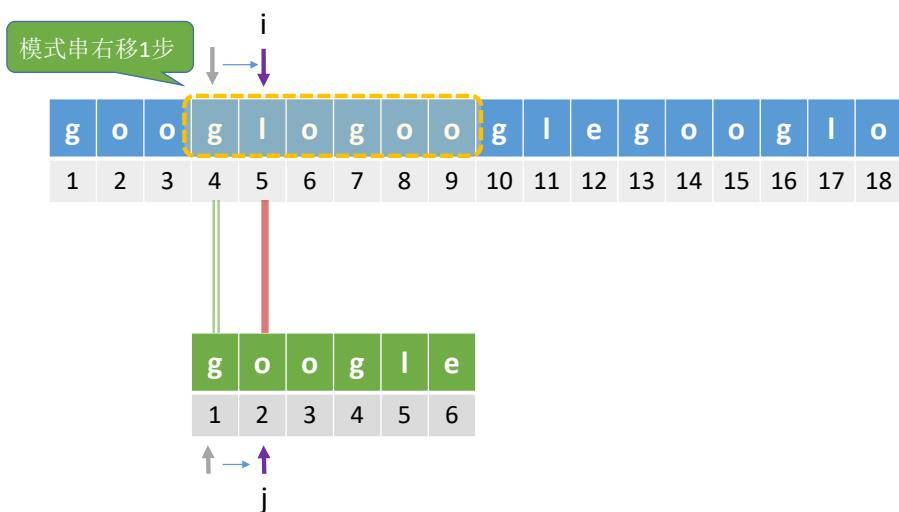
比较坏的情况



王道考研/CSKAOYAN.COM

7

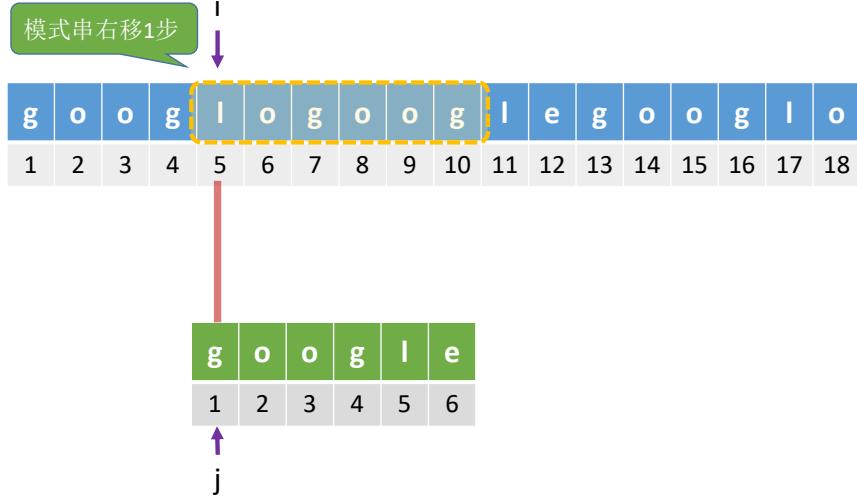
比较坏的情况



王道考研/CSKAOYAN.COM

8

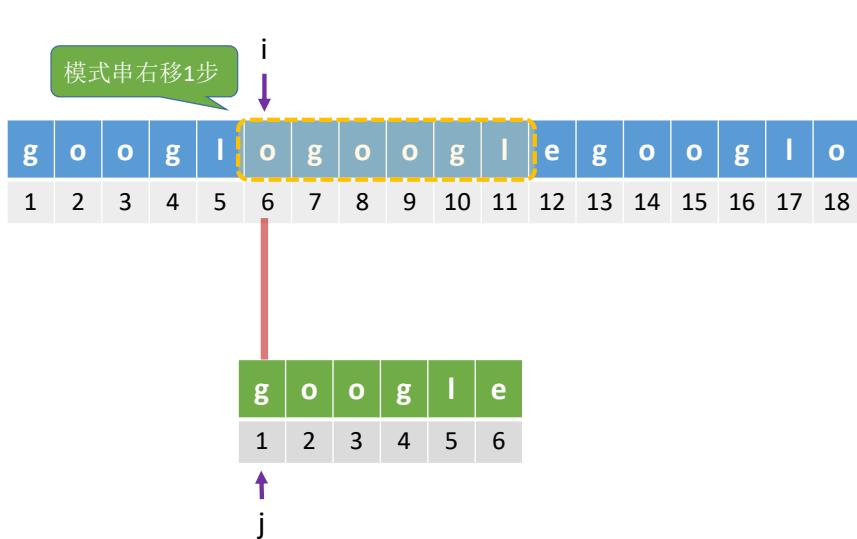
比较坏的情况



王道考研/CSKAOYAN.COM

9

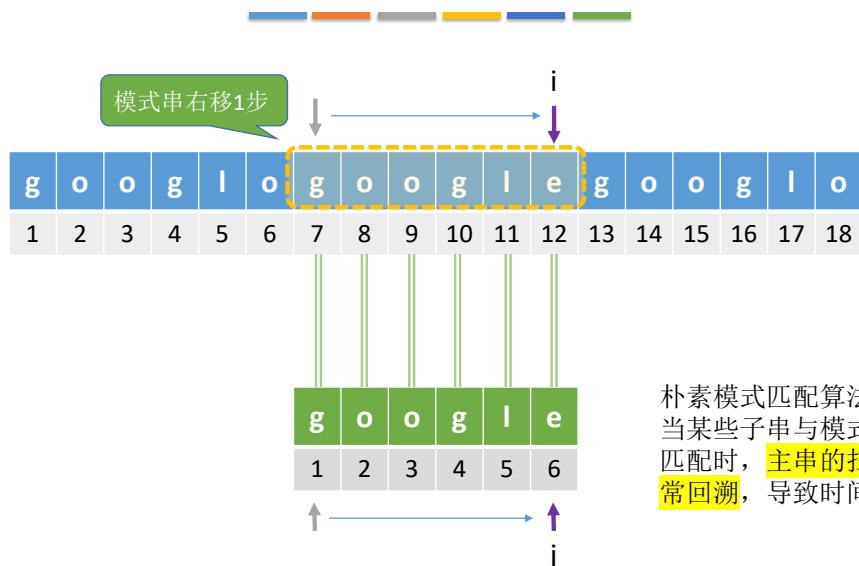
比较坏的情况



王道考研/CSKAOYAN.COM

10

比较坏的情况



王道考研/CSKAOYAN.COM

11

看看人脑的反应

马 赛 克

任务：找到'google'

要求：你的目光
不要回溯

你的目光相当于指针 i

王道考研/CSKAOYAN.COM

12

解决主串回溯问题

任务：找到'google'

王道考研/CSKAOYAN.COM

13

模式串指针应该回溯到哪儿？

王道考研/CSKAOYAN.COM

14

模式串指针应该回溯到哪儿？

主串的指针没有回溯

x	x	x	x	x	g	o	o	g	l	?	?	?	?	?	?	?	?	?
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	

少了5次对比

i 值得期待
是否是 g

若 j=6 时发生不匹配，则应让 j 回到 1

也可以理解为模式串往右移动了5步

g	o	o	g	l	e
1	2	3	4	5	6

j

王道考研/CSKAOYAN.COM

15

模式串指针应该回溯到哪儿？

这个字符是'g'，
该怎么做？

应该让 i++,
j++

x	x	x	x	x	g	o	o	g	l	g	?	?	?	?	?	?	?	?
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	

g	o	o	g	l	e
1	2	3	4	5	6

j

王道考研/CSKAOYAN.COM

16

模式串指针应该回溯到哪儿？

x	x	x	x	x	g	o	o	g	l	?	?	?	?	?	?	?	?	?
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	

这个字符不是'g',
该怎么做?

i
应该让 i 后移，检查
下一个字符是否是 g

若 j=1 时发生不匹配，则应让 i++, 而
j 依然是 1

g	o	o	g	l	e
1	2	3	4	5	6

j

王道考研/CSKAOYAN.COM

17

模式串指针应该回溯到哪儿？

x	x	x	x	g	o	o	g	?	?	?	?	x	x	x	x	x	x
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

这个字符不是'r',
该怎么做?

i
值得期待
是否是 o

如果 j=k 时才发现匹配失败，
说明 1~k-1 都匹配成功

若 j=5 时发生不匹配，则应让 j 回到 2

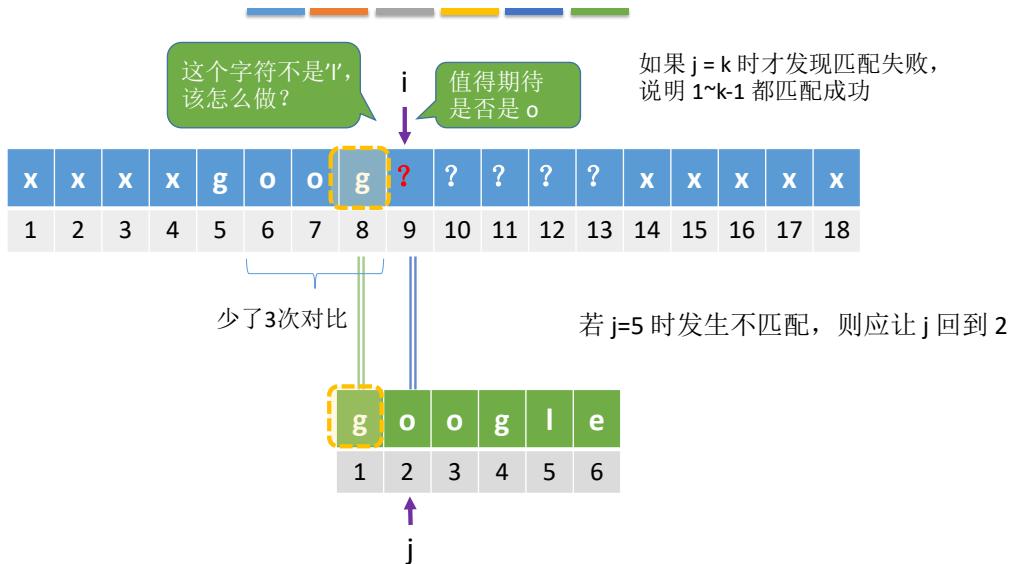
g	o	o	g	l	e
1	2	3	4	5	6

j

王道考研/CSKAOYAN.COM

18

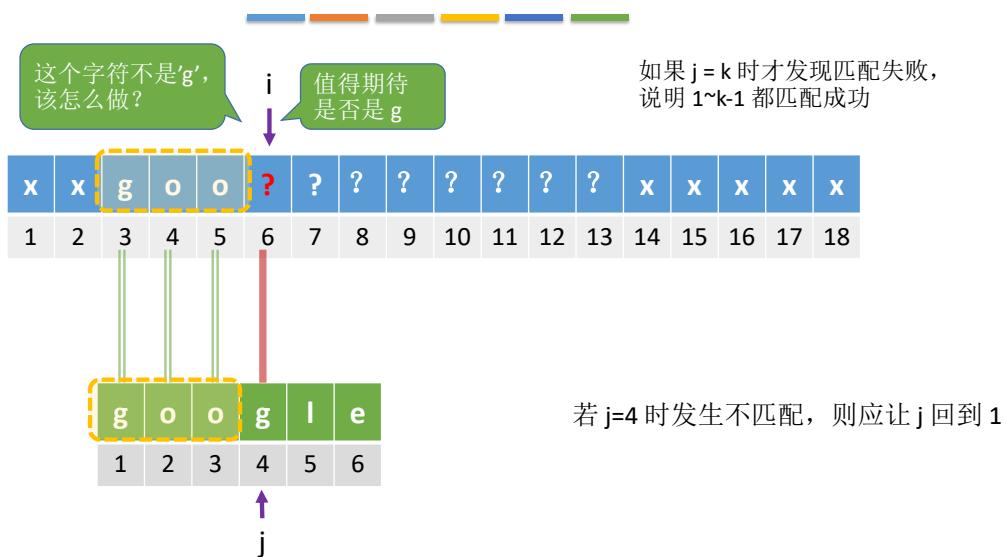
模式串指针应该回溯到哪儿？



王道考研/CSKAOYAN.COM

19

模式串指针应该回溯到哪儿？



王道考研/CSKAOYAN.COM

20

模式串指针应该回溯到哪儿？

这个字符不是'g'，
该怎么做？

i 值得期待
是否是 g

如果 $j = k$ 时才发现匹配失败，
说明 1~k-1 都匹配成功

x	x	g	o	o	?	?	?	?	?	?	?	?	x	x	x	x	x
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

少了2次对比

g	o	o	g	l	e
1	2	3	4	5	6

若 $j=4$ 时发生不匹配，则应让 j 回到 1

21

模式串指针应该回溯到哪儿？

这个字符不是'o'，
该怎么做？

i 值得期待
是否是 g

如果 $j = k$ 时才发现匹配失败，
说明 1~k-1 都匹配成功

x	x	g	o	?	?	?	?	?	?	?	?	?	x	x	x	x	x
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

g	o	o	g	l	e
1	2	3	4	5	6

若 $j=3$ 时发生不匹配，则应让 j 回到 1

22

模式串指针应该回溯到哪儿？

这个字符不是'o'，
该怎么做？

i
值得期待
是否是 g

如果 $j = k$ 时才发现匹配失败，
说明 $1 \sim k-1$ 都匹配成功

x	x	g	o	?	?	?	?	?	?	?	?	?	x	x	x	x	x
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

少了1次对比

g	o	o	g	l	e
1	2	3	4	5	6

若 $j=3$ 时发生不匹配，则应让 j 回到 1

王道考研/CSKAOYAN.COM

23

模式串指针应该回溯到哪儿？

这个字符不是'o'，
该怎么做？

i
值得期待
是否是 g

如果 $j = k$ 时才发现匹配失败，
说明 $1 \sim k-1$ 都匹配成功

x	x	g	?	?	?	?	?	?	?	?	?	?	x	x	x	x	x
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

g	o	o	g	l	e
1	2	3	4	5	6

若 $j=2$ 时发生不匹配，则应让 j 回到 1

王道考研/CSKAOYAN.COM

24

模式串指针应该回溯到哪儿？

这个字符不是'o'，
该怎么做？

i

值得期待
是否是 g

如果 $j = k$ 时才发现匹配失败，
说明 $1 \sim k-1$ 都匹配成功



若 $j=2$ 时发生不匹配，则应让 j 回到 1

王道考研/CSKAOYAN.COM

25

模式串指针应该回溯到哪儿？

如果 $j = k$ 时才发现匹配失败，说明 $1 \sim k-1$ 都匹配成功

g	o	o	g	l	e
1	2	3	4	5	6

↑
j

int next[7];

0	1	2	3	4	5	6
0	1	1	1	2	1	

当 $j=k$ 且发现字符不匹配时，令 $j=next[k]$

先令 $j=0$ ，
再进行 $j++$

若当前两个字符匹配，则 $i++$, $j++$

若 $j=1$ 时发生不匹配，则应让 $i++$, 而 j 依然是 1

若 $j=2$ 时发生不匹配，则应让 j 回到 1

若 $j=3$ 时发生不匹配，则应让 j 回到 1

若 $j=4$ 时发生不匹配，则应让 j 回到 1

若 $j=5$ 时发生不匹配，则应让 j 回到 2

若 $j=6$ 时发生不匹配，则应让 j 回到 1

王道考研/CSKAOYAN.COM

26

KMP算法代码

```

int Index_KMP(SSString S,SSString T,int next[]){
    int i=1, j=1;
    while(i<=S.length&&j<=T.length){
        if(j==0||S.ch[i]==T.ch[j]){
            ++i;
            ++j; //继续比较后继字符
        }
        else {
            j=next[j]; //模式串向右移动
        }
        if(j>T.length)
            return i-T.length; //匹配成功
        else
            return 0;
    }
}

```

g	o	o	g	l	e
1	2	3	4	5	6

↑
j

int next[7];

0	1	2	3	4	5	6
0	1	1	1	2	1	

当 j=k 且发现字符不匹配时，令j=next[k]

王道考研/CSKAOYAN.COM