

作业:

- S= GCATCGCAGAGAGTATACAGTACG
- 76543210
- T= GCAGCGAG

给出BM算法的bmBc和bmGs[i]数组? 对文本串S的BM算法匹配过程进行手动推导, 计算对齐次数和字符比较次数?

解: bmBc

	C	A	G	C	T
bmBc[c]	1	2	3	8	

bmGs[i]	i	0	1	2	3	4	5	6	7
x[i]		G	C	A	G	C	G	A	G
bmGs[i]		7	7	7	7	7	4	2	1

第1次对齐

GCATCGCAGAGAGTATACAGTACG

GCAGCGAG * 匹配一次

$$\text{shift}(Bc) = \text{bmBc}['A'] - 0 = 1$$

$$\text{shift}(Gs) = \text{bmGs}[i] = 1$$

→ 1

第2次对齐

GCATCGCAGAGAGTATACAGTACG

GCAGCGAG * 匹配三次

$$\text{shift}(Bc) = \text{bmBc}['C'] - 2 = 1$$

$$\text{shift}(Gs) = \text{bmGs}[i] = 4$$

→ 4

第3次对齐

GCATCGCAGAGAGTATACAGTACG

GCAGCGAG * 匹配四次

$$\text{shift}(Bc) = \text{bmBc}['A'] - 3 = -2 \Rightarrow \text{shift}(Bc) = 1$$

$$\text{shift}(G_s) = \text{bm}_{G_s}[i] = 7$$

$$\rightarrow 7$$

第4次对齐

GCATCGCAGAGAGTATACAGTACG

$$\begin{array}{c} \text{GCAGCGAG} \quad \text{匹配3次} \\ \text{shift}(B_c) = \text{bm}_{B_c}[c] - 2 = 1 \\ \text{shift}(G_s) = \text{bm}_{G_s}[i] = 4 \\ \rightarrow 4 \end{array}$$

第5次对齐

GCATCGCAGAGAGTATACAGTACG

$$\begin{array}{c} \text{GCAGCGAG} \quad \text{匹配2次} \\ \text{GCAGCGAG} \end{array}$$

匹配结束 共5次对齐, $1+3+4+3+2=13$ 次比较