

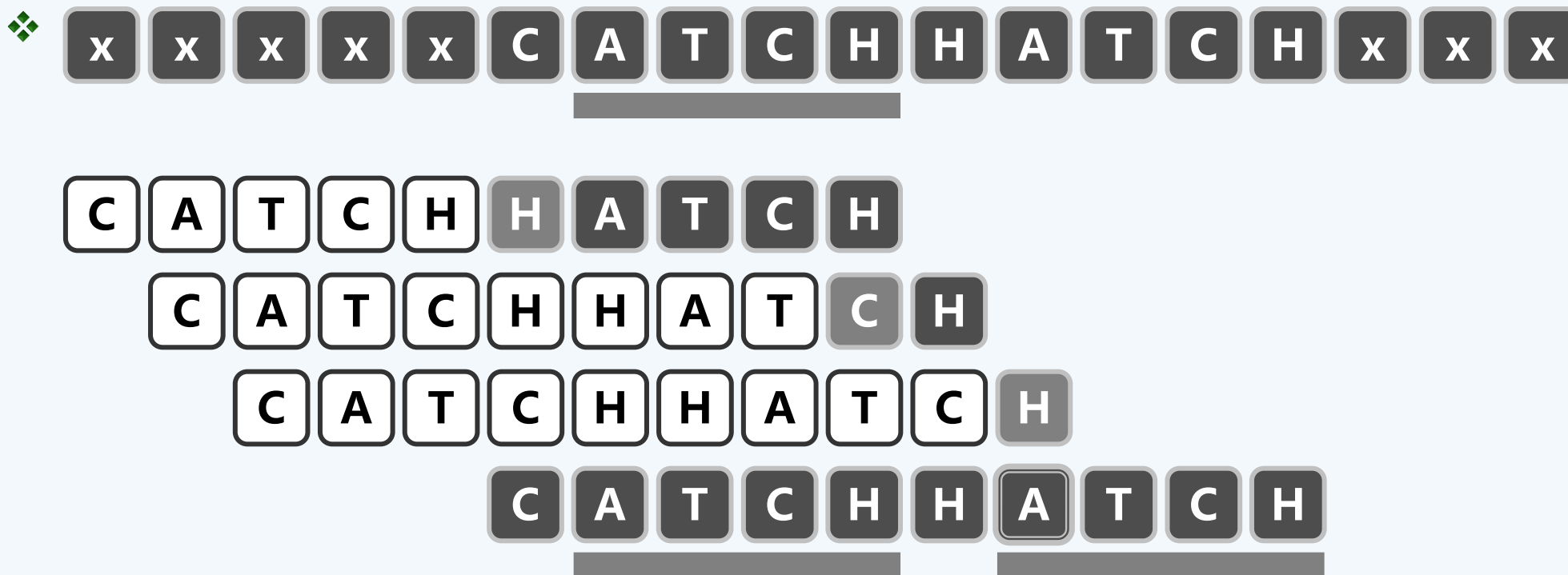
11. 串

(e1) BM_GS算法：好后缀

邓俊辉

deng@tsinghua.edu.cn

经验 = 匹配的后缀



首趟比对虽以失败结束，却积累了足够的经验（匹配的后缀 **ATCH**） // 好后缀

❖ 据此，可以省去中间的两趟，而直接转至最后一趟（P右移5个字符）

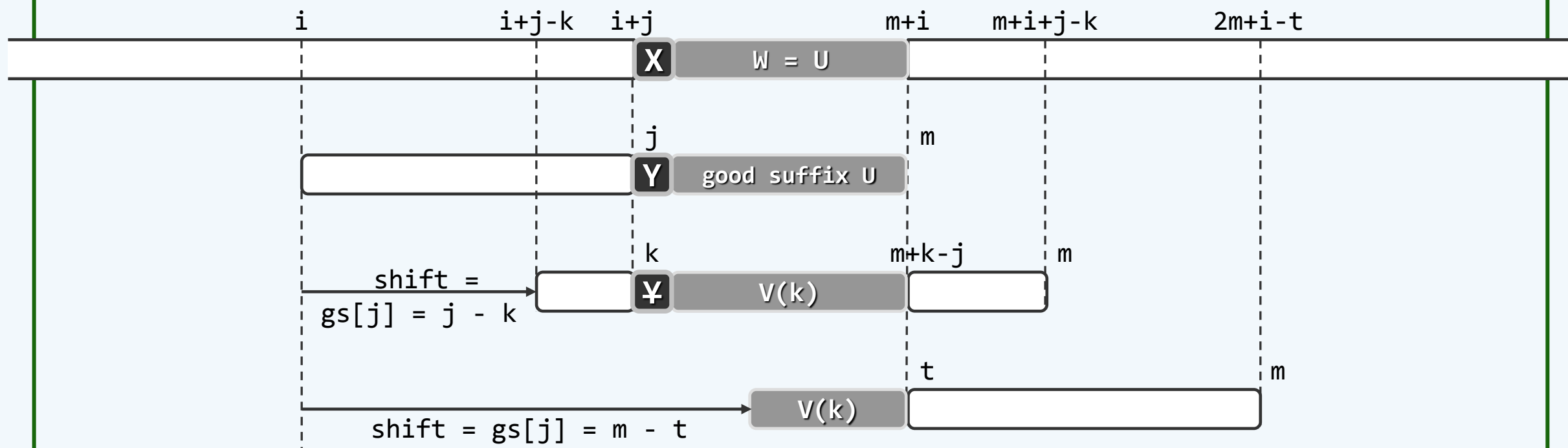
❖ 这一规律与技巧，与KMP完全一致，只不过前后颠倒而已...

Good-Suffix Shift

❖ 扫描比对中断于 $T[i + j] = \boxed{X} \neq \boxed{Y} = P[j]$ 时, $U = P(j, m)$ 必为好后缀

❖ 故下一对齐位置必须使:

- 1) U 重新与 $V(k) = P(k, m + k - j)$ 匹配, 且
- 2) $P[k] = \boxed{\text{¥}} \neq \boxed{Y} = P[j]$

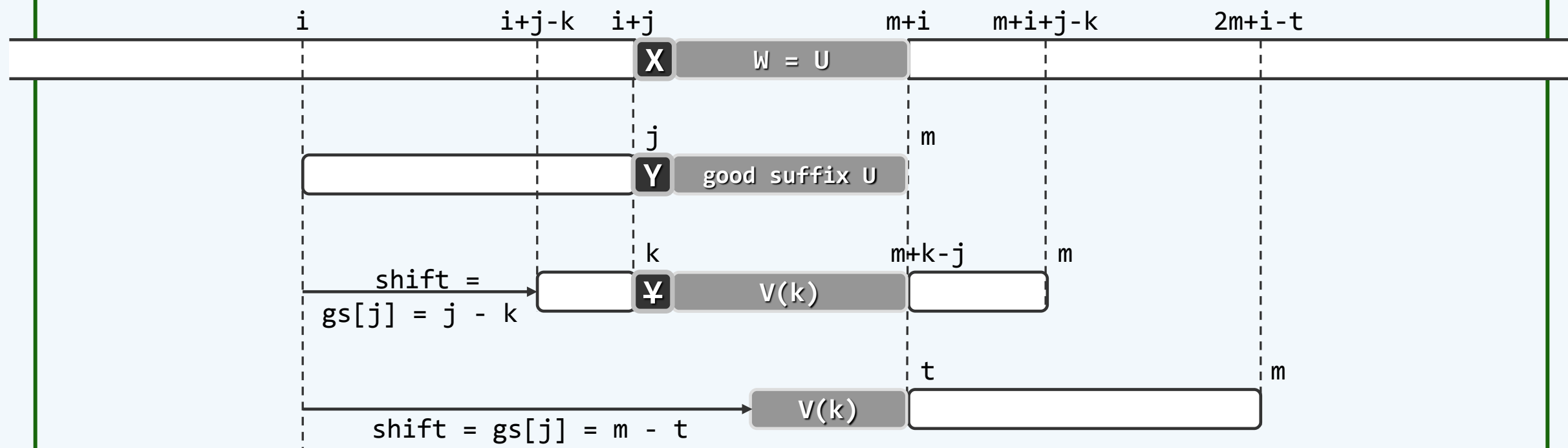


Good-Suffix Shift

❖ 若P中的确存在这样的子串 $V(k)$ ，则可

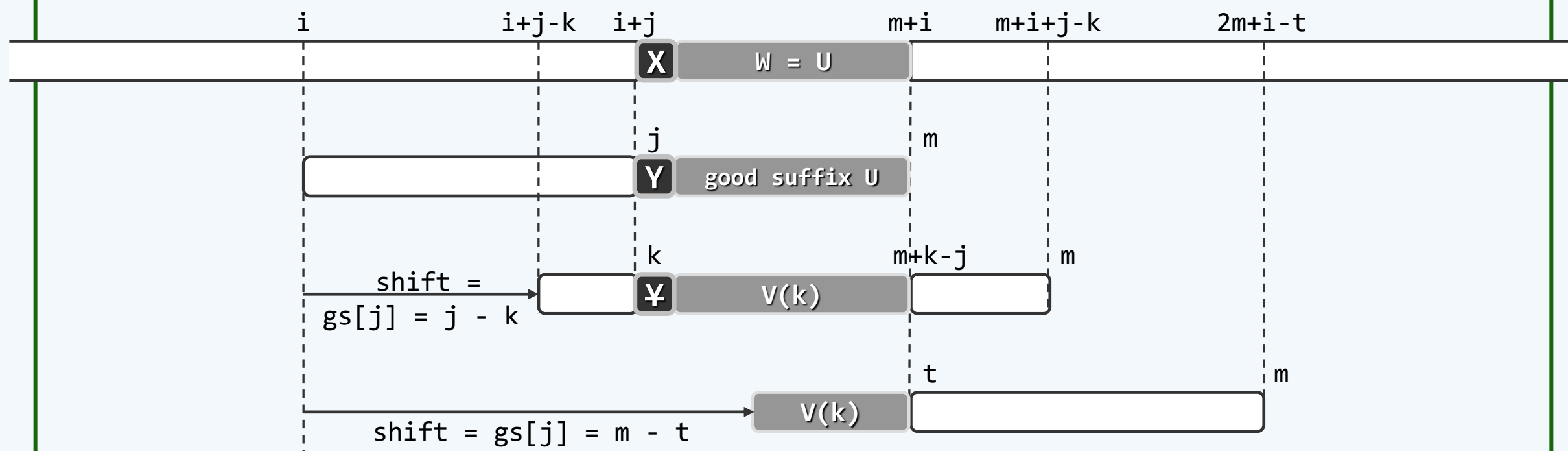
选择其中 k 最大者（尽可能靠后）

然后，通过右移使之与 U 对齐（移动距离尽可能小）



Good-Suffix Shift

- ❖ 否则，在所有前缀 $P[0, t)$ 中，取与 U 的 **后缀** 匹配的最长者 // 注意：有可能 $t = 0$
- ❖ 无论如何，位移量仅取决于 j 和 P 本身——亦可 **预先计算**，并 **制表待查**



实例

8 8 8 8 8 4 8 1
非 曰 静 也 善 故 静 也

圣 人 之 静 也 非 曰 静 也 善 故 静 也

非 曰 静 也 善 故 静 也

+1 → 非 曰 静 也 善 故 静 也

+4 → 非 曰 静 也 善 故 静 也