

11. 串

(c3) KMP算法：理解next[]表

吴用再使时迁扮作伏路小军，去曾头市寨中，探听他不出何意，所有陷坑，暗暗地记着，离寨多少路远，总有几处。时迁去了一日，都知备细，暗地使了记号，回报军师。

邓俊辉

deng@tsinghua.edu.cn

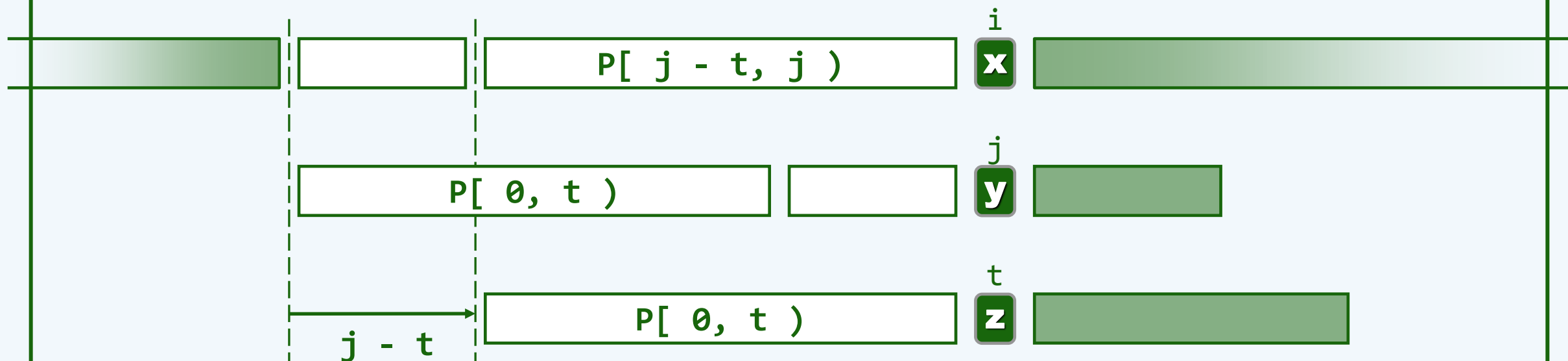
自匹配 = 快速右移

❖ 对任意 j ，考察集合：

$$N(P, j) = \{ 0 \leq t < j \mid P[0, t) == P[j - t, j) \}$$

亦即，在 $P[j]$ 的前缀 $P[0, j)$ 中，所有匹配真前缀和真后缀的长度

❖ 因此，一旦 $T[i] \neq P[j]$ ，可从 $N(P, j)$ 中取某个 t ，令 $P[t]$ 对准 $T[i]$ ，并继续比对



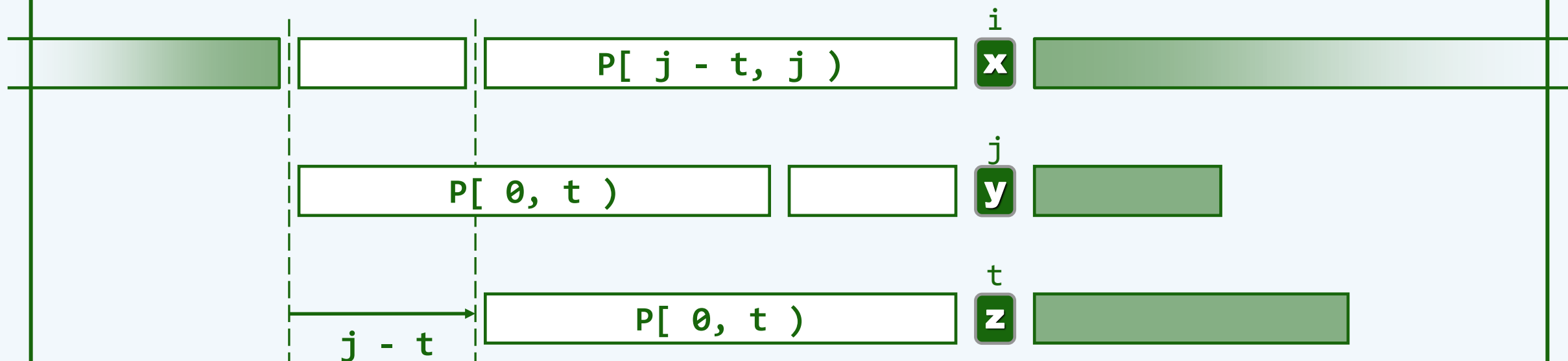
最长自匹配 = 快速右移 + 避免回退

❖ $|N(P, j)| > 1$ 时，难道需要遍历其中的每一个 t ？

❖ 不必！

❖ 观察：位移量 = $j - t$ ，与 t 成反比

❖ 因此，若选用最大的 t ，则必然最安全



next[0]

❖ 只要 $j > 0$, 必有 $[\emptyset] \in N(P, j)$ //空串是任何非空串的真子串

❖ 但若 $j = 0$, 则有 $N(P, 0) = \emptyset$ //空串没有真子串

❖ 不妨取 $\text{next}[0] = [-1] \dots$

❖ 回顾主算法：行之有效！如何理解？

