Développement : automate des occurrences

Pierron Théo

13 avril 2014

Soit T un texte, P un motif. On cherche un automate minimal qui reconnaît $\Sigma^* P$.

On note P_k le k-ème préfixe de P, T_k celui de T.

Définition 1 On note $x \sqsupset y$ si x est suffixe de y. On définit aussi

$$\sigma(x) = \max\{k, P_k \sqsupset x\}$$

le plus grand préfixe de P qui est suffixe de x.

On remarque que $\sigma(x) = |P|$ ssi $P \sqsupset x$.

Lemme 1
Si $x \sqsupset y$, $\sigma(x) \leqslant \sigma(y)$.

Démonstration. Tout suffixe de x est suffixe de y donc $P_{\sigma(x)} \sqsupset x \sqsupset y$ donc $\sigma(y) \geqslant \sigma(x)$. ■

Lemme 2
$\sigma(xa) \leqslant \sigma(x) + 1$.

Démonstration. Si $\sigma(xa) = 0$, c'est bon. Sinon, on a $P_{\sigma(xa)} \sqsupset xa$ donc $P_{\sigma(xa)-1} \sqsupset x$.

Ainsi, $\sigma(x) \geqslant \sigma(xa) - 1$. ■

Lemme 3
$\sigma(xa) = \sigma(P_{\sigma(x)}a)$.

Démonstration. $P_{\sigma(x)} \sqsupset x$ donc $P_{\sigma(x)}a \sqsupset xa$ donc $\sigma(xa) \geqslant \sigma(P_{\sigma(x)}a)$.

Réciproquement, on a $P_{\sigma(x)}a \sqsupset xa$ et $P_{\sigma(xa)} \sqsupset xa$. Ce sont des suffixes d'un même mot dont l'un est suffixe de l'autre. On regarde donc les tailles. Comme $|P_{\sigma(xa)}| = \sigma(xa) \leqslant \sigma(x) + 1 = |P_{\sigma(x)}a|$, on a $P_{\sigma(xa)} \sqsupset P_{\sigma(x)}a$ donc $\sigma(xa) \leqslant \sigma(P_{\sigma(x)}a)$. ■

On définit l'automate $A = (Q, \Sigma, \delta, I, F)$ par $Q = [\![0, |P|]\!]$, $I = \{0\}$, $F = \{|P|\}$ et

$$\delta(q, a) = \sigma(P_q a)$$

On pose alors la fonction de transition étendue ϕ définie par induction : $\phi(\varepsilon) = 0$ et $\phi(wa) = \delta(\phi(w), a)$.

<u>Théorème 1</u> $\phi(T_i) = \sigma(T_i)$.

Démonstration. Par récurrence sur i.

Si $i = 0$, $T_0 = \varepsilon$ donc $\phi(T_0) = 0 = \sigma(T_0)$. Si $\phi(T_i) = \sigma(T_i)$, on a

$$\phi(T_{i+1}) = \delta(\phi(T_i), T[i+1]) = \delta(\sigma(T_i), T[i+1]) = \sigma(P_{\sigma(T_i)}T[i+1]) = \sigma(T_iT[i+1]) = \sigma(T_{i+1})$$ ■

<u>Théorème 2</u> *A est minimal.*

Démonstration. On écrit $P = P_iQ_i$. On a $\phi_i(Q_i) = \phi(P_iQ_i) = \phi(P) = |P|$ donc $Q_i \in L_i(A)$.

De plus, si $i > j$, $\phi_j(Q_i) = \phi(P_jQ_i) < |P|$ car $|P_jQ_i| < |P|$ donc $Q_i \notin L_j(A)$.

Les états i et j sont donc séparés par la congruence de Nérode, donc A est minimal. ■