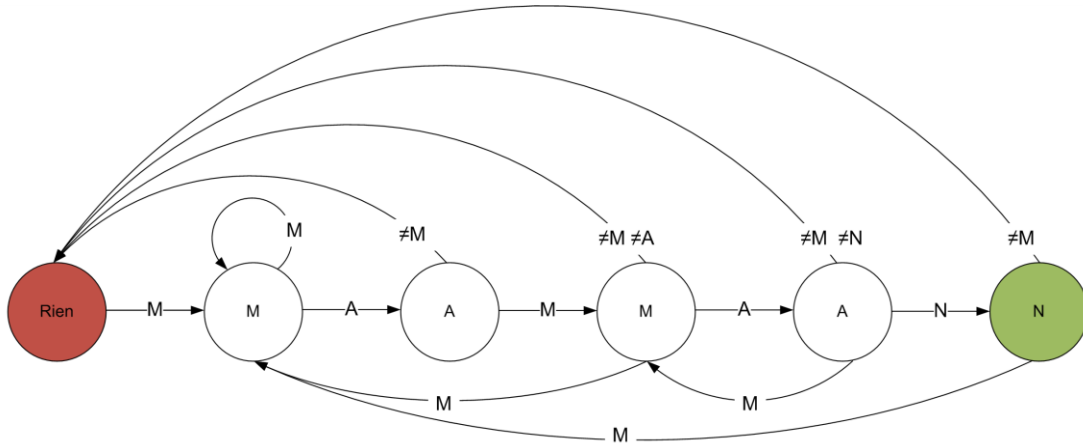


Notations pour la construction d'un automate fini pour la reconnaissance d'un patron dans un texte

Dans ce document, on considère l'exemple de la construction de l'automate permettant de reconnaître le patron « maman » :



- On note P le patron à rechercher. P a une longueur m : $P[1 \dots m]$
- On note T le texte dans lequel P est recherché. T a une longueur n : $T[1 \dots n]$
- On note Σ l'alphabet dont sont issus P et T . Généralement l'ensemble des caractères dans un jeu de caractères:

$$\Sigma = \{'A', 'B', \dots, ' ', '., \dots'\}$$

Σ est fini. Son cardinal $|\Sigma|$ est un entier.

- On note Σ^* l'ensemble des chaînes de caractères (finies) construites à partir de l'alphabet Σ .
- On note xy la concaténation de deux chaînes de caractères x et y de Σ^* .
- On note $w \subset x$: w préfixe de x . Par exemple, « télé » est préfixe de « téléphone » : « télé » \subset « téléphone »
- On note $w \supset x$: w suffixe de x . Par exemple, « age » est suffixe de « partage » : « age » \supset « partage »

- On note P_k le préfixe de longueur k du patron P ($k \leq m$)
- On note Σ_p le plus petit alphabet (sous-ensemble de Σ) dont est issu P : $\Sigma_p = \{'M', 'A', 'N'\}$
- On note Q l'ensemble des états (nombre de caractères + 1) :

$$Q = \{q_0, q_1, q_2, q_3, q_4, q_5\}$$

q_0 : l'état noté « Rien » dans la figure

q_5 : l'état noté N dans la figure

Q est fini (contrairement à infini). Son cardinal $|Q|$ est un entier.

Pour des raisons de simplicité, un état peut-être représenté par son index : $q_i = i$

- On note q_0 est l'état initial : $q_0 = 0$

- On note δ la fonction de transition de l'automate. δ est une fonction de $Q \times \Sigma \rightarrow Q$ associant à chaque paire {état, caractère} un état de Q :

$$\delta(q,a) \in Q$$

Dans notre exemple, nous aurons $\delta(q_0, 'M') = q_1$, $\delta(q_1, 'M') = q_1$, $\delta(q_5, 'A') = q_0 \dots$

Il est possible de représenter la fonction $\delta(q,a)$ par un tableau de taille $(m+1) \times (|\Sigma_p| + 1)$:

$q \setminus a$	'M'	'A'	'N'	Autre
0	1	0	0	0
1	1	2	0	0
2	3	0	0	0
3	1	4	0	0
4	3	0	5	0
5	1	0	0	0

La construction de ce tableau peut se réaliser grâce à cet algorithme :

Initialiser δ à 0

Pour $q = 0 : m$

Pour chaque caractère a dans Σ_p

$$k = \min(m+1, q+2)$$

Répéter

$$k = k - 1$$

Tant que $k > 0$ et $(P_k \supset P_q a)$ est faux

$$\delta(q, a) = k$$

Fin Pour

Fin Pour