AUTOMATE MINIMAL

Sakarovitch p120 - 125.

Soit Σ un alphabet fini.
Soit L un langage sur Σ

FAIRE UN EX

Déf L'automate des quotients de L, ou automate minimal de L est $\mathcal{A}_L = (Q_L, \Sigma, \delta_L, \{L\}, F_L)$

où $Q_L = \{u^{-1}L \mid u \in \Sigma^*\}$

$F_L = \{u^{-1}L \mid u \in L\}$

$\delta_L = \{(u^{-1}L, a, (ua)^{-1}L) \mid u \in \Sigma^*, a \in \Sigma\}$

⚠ $u \neq v \not\Rightarrow u^{-1}L \neq v^{-1}L$.

Rq il est toujours complet et déterministe car $u^{-1}L = v^{-1}L \Rightarrow (ua)^{-1}L = (va)^{-1}L$.

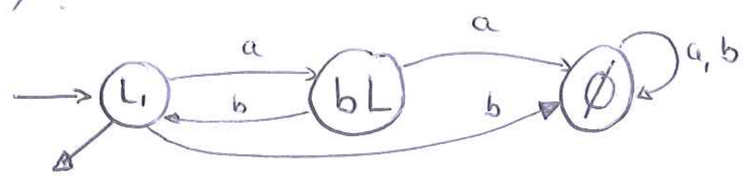
ex — $\Sigma = \{a,b\}$ — $L_1 = \mathcal{L}((ab)^*)$

Si $u \in (ab)^*$, $u^{-1}L_1 = L_1$. Donc $F_{L_1} = \{L_1\}$

Si $u \in (ab)^*a$, $u^{-1}L_1 = bL_1$

Sinon, $u^{-1}L_1 = \emptyset$.

D'où $\mathcal{A}_{L_1}$ =



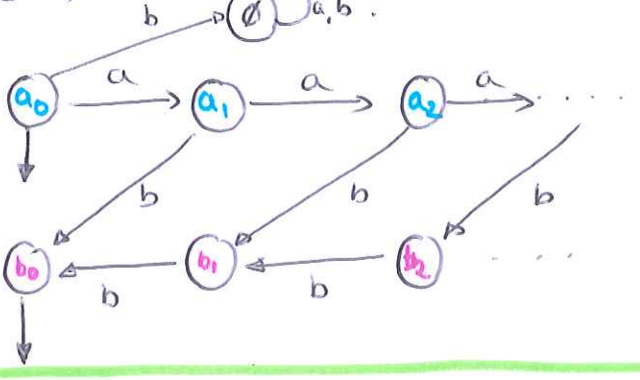
ex — $\Sigma = \{a,b\}$ — $L_2 = \{a^n b^n \mid n \in \mathbb{N}\}$

Si $u = a^k$ $u^{-1}L_2 = \{a^m b^{m+k} \mid m \in \mathbb{N}\}$ qu'on notera a_k

Si $u = a^{k+j} b^j$, $u^{-1}L_2 = \{b^k\}$ qu'on notera b_k

Sinon $u^{-1}L_2 = \emptyset$

$\mathcal{A}_{L_2}$:



Si $u \in L_2$
$u = a^k$ avec $k = 0$
ou $u = a^{k+j} b^j$ avec $k = 0$
donc $F_L = \{a_0, b_0\}$

$L_2 = \varepsilon^{-1} L_2 = (a^0)^{-1} L_2$
donc l'état initial est a_0

Pté $\underline{\mathcal{L}(\mathcal{A}_L) = L}$ "L'automate des quotients d'un langage reconnaît ce langage."

<u>Preuve</u> Montrons d'abord par récurrence sur la longueur de $u \in \Sigma^*$, que pour tt état $X \in Q_L$, $\delta_L(X,u) = u^{-1}X$

(où δ_L désigne l'extension de la f° de transi° des lettres aux mots, bien définie car $\mathcal{A}_L$ est complet dét.)

- Si $|u| = 1$, u s'écrit a pour $a \in \Sigma$.

  Pour $X \in Q_L$, il existe $v \in \Sigma^*$ tel que $X = v^{-1}L$.

  or par déf de Q_L, $(v^{-1}L, a, (va)^{-1}L)$ est une transi° de $\mathcal{A}_L$

  qu'on peut réécrire $(X, u, a^{-1}(v^{-1}L)) = (X, u, u^{-1}X) \in \delta_L$

  soit $\delta(X,u) = u^{-1}X$.

  D'où la propriété au rang 1.

- Si $|u| = n+1$ et si la propriété est vraie au rg n.

  On écrit u comme au' où $a \in \Sigma$, $u' \in \Sigma^*$, et $|u'| = n$.

  Pour $X \in Q_L$ on a
$$\begin{aligned}
\delta(X,u) &= \delta(X, au') \\
&= \delta(\delta(X,a), u') \\
&= \delta(a^{-1}X, u') && \text{pté au rg 1} \\
&= u'^{-1}(a^{-1}X) && \text{HR, pté au rg } n. \\
&= (au')^{-1}X \\
&= u^{-1}X.
\end{aligned}$$

  D'où la propriété au rang $n+1$.

On peut maintenant conclure : pour $w \in \Sigma^*$

$w \in \mathcal{L}(\mathcal{A}_L) \iff \exists X \in I_L = \{L\},\ \delta(X,w) \in F_L$

$\iff \delta(L,w) \in F_L$

$\iff w^{-1}L \in F_L \iff w \in L$

car il existe $\tilde{w} \in L$ tel que $w^{-1}L = \tilde{w}^{-1}L$

Puisque $\tilde{w} \in L$, $\tilde{w}\varepsilon \in L$ et donc $\varepsilon \in \tilde{w}^{-1}L$

donc $\varepsilon \in w^{-1}L$ donc $w \in L$.

$\Leftarrow$ par déf de F_L.

Minimalité et csq

Pté [Si L est le langage reconnu par un automate complet. dét. accessible A non néc. fini, alors il existe une surjec° de $\underset{\text{états de } A}{Q}$ dans $\underset{\text{états de } \mathcal{A}_L}{Q_L}$

↳ **Cor** [Si L est langage reconnaissable, alors $\mathcal{A}_L$ est l'automate complet dét. ayant le moins d'états reconnaissant L. (minimalité)

↳ **Cor.** $\boxed{L \text{ rec} \Leftrightarrow \{f^{-1}L \mid f \in \Sigma^*\} \text{ est fini}}$

[Rq : et tout autre automate minimal en ce sens lui est isomorphe (par φ bijective).

<u>Preuve pté</u> On va en fait montrer plus précisém' qu'il existe un morphisme d'automate surj de A sur A_L.

Cela signifie qu'il existe $\varphi \in \mathcal{F}(Q, Q_L)$ telle que

$\rightarrow \forall q \in Q, \forall a \in A, \ \varphi(\delta(q,a)) = \delta_L(\varphi(q), a)$

$\rightarrow \varphi(i) = L$

$\rightarrow \varphi(F) \subset F_L$

Puisqu'on a supposé A accessible, Q s'écrit $\{\delta(i,f) \mid f \in \Sigma^*\}$.

On définit alors $\varphi = \left(\begin{array}{ccc} Q & \longrightarrow & Q_L \\ \delta(i,f) & \longmapsto & f^{-1}L \end{array} \right)$

On s'assure que c'est une bonne déf : si $\underline{\delta(i,f) = \delta(i,g)}$

$w \in f^{-1}L \Leftrightarrow fw \in L \Leftrightarrow \delta(i, fw) \in F \Leftrightarrow \delta(\delta(i,f), w) \in F$ dc $\underline{f^{-1}L = g^{-1}L}$.

$\Updownarrow$

$w \in g^{-1}L \Leftrightarrow gw \in L \Leftrightarrow \delta(i, gw) \in F \Leftrightarrow \delta(\delta(i,g), w) \in F$

• $\varphi(i) = \varphi(\delta(i, \varepsilon)) = (\varepsilon^{-1})L = L$.

• Si $q \in F$ il s'écrit $\delta(i,f)$ où néc $f \in L$ donc $\varphi(q) = f^{-1}L \in F_L$ car $f \in L$.

• $\varphi(\delta(q,a)) = \varphi(\delta(\delta(i,f),a))$ où $f \in \Sigma^*$ tq $\delta(i,f) = q$.

$= \varphi(\delta(i, fa)) = (fa)^{-1}L = a^{-1}(f^{-1}L) = a^{-1}(\varphi(\delta(i,f)))$

$= a^{-1}\varphi(q) = \delta_L(\varphi(q), a)$.

↳ Surjectivité $\forall q \in Q_L, \exists f \in \Sigma^*, q = f^{-1}L = \varphi(\delta(i,f))$ existe car A est complet

<u>Preuve Cor 1</u> Si $L = \mathcal{L}(A)$ où A est fini, $Q_L = \varphi(Q)$ est fini car Q l'est, et néc $\#Q \leq \#Q_L$.

<u>Preuve Cor 2</u> L rec $\Rightarrow Q_L$ fini or $Q_L = \{f^{-1}L \mid f \in \Sigma^*\}$. Réc si Q_L est fini $L = \mathcal{L}(A_L)$ est rec car A_L fini.

<u>Rq</u> Si L n'est pas rec. on peut quand m construire $\mathcal{A}_L$, il est comp. et dét mais plus fini $\boxed{\text{ex } L_2}$

Congruences de Néřode

On veut pouvoir calculer efficacement $\mathcal{A}_L$ dans le cas où L nous est donné comme $\mathcal{L}(A)$ où A est 1 automate fini et dét

On note $A = (Q, \Sigma, \delta, I, F)$, et pour $q \in Q$, $L_q = \{w \in \Sigma^* \mid \delta(p, w) \in F\}$

Déf | L'équivalence de Néřode, ou congruence de Néřode est la relation d'équivalence sur Q définie par

$$\boxed{q_1 \sim q_2 \text{ ssi } L_{q_1} = L_{q_2}}$$

Pté

* $p \sim q \Leftrightarrow \forall f \in \Sigma^*,\ \delta(p, f) \sim \delta(q, f)$
* $p \sim q \Rightarrow (p \in F \text{ ssi } q \in F)$

Preuve * $\Leftarrow$ clair avec $w = \varepsilon$.

$$\begin{aligned}
\Rightarrow L_{\delta(p,f)} &= \{w \in \Sigma^* \mid \delta(\delta(p,f), w) \in F\} \\
&= \{w \in \Sigma^* \mid \delta(p, fw) \in F\} \\
&= \{w \in \Sigma^* \mid fw \in L_p\} \\
&= \{ \text{——} \mid fw \in L_q\} \\
&= \{ \text{——} \mid \delta(q, fw) \in F\} \\
&= \{ \text{——} \mid \delta(\delta(q,f), w) \in F\} \\
&= L_{\delta(q,f)}
\end{aligned}$$

D'où $\delta(p,f) \sim \delta(q,f)$.

* $p \in F \Leftrightarrow \varepsilon \in L_p \Longleftrightarrow \varepsilon \in L_q \Leftrightarrow q \in F$.

sous l'hyp $L_p = L_q$ ie $p \sim q$

Déf On définit l'automate <u>quotient par congruence de Néirode</u> de A comme étant $A/_{\mathcal{N}} = (Q_{\mathcal{N}}, \Sigma, \delta_{\mathcal{N}}, I_{\mathcal{N}}, F_{\mathcal{N}})$

où $Q_{\mathcal{N}} = Q/_{\mathcal{N}}$ = l'ens. des classes modulo la congruence de $\mathcal{N}$.

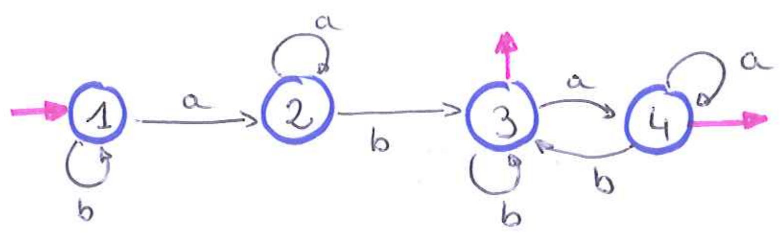
$\delta_{\mathcal{N}}(\bar{p}, a) = \bar{q}$ ssi $\delta(p,a) = q$, pr $\forall a \in \Sigma, (p,q) \in Q^2$.

$I_{\mathcal{N}} = \overline{q_i}$ où $I = \{q_i\}$

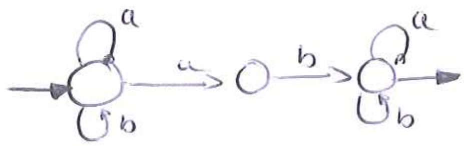
$F_{\mathcal{N}} = F/_{\mathcal{N}}$.

Rq → C'est bien défini grâce aux pté préc.
→ On applique ça à A <u>déterministe</u>.

ex $\mathcal{A}$



Rq $\mathcal{A}$ est le déterminisé de



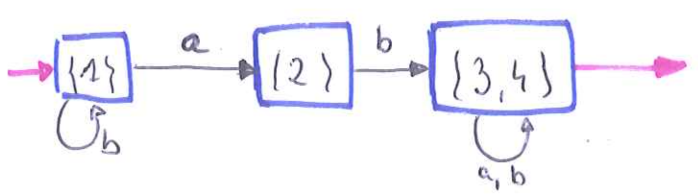
donc $L(\mathcal{A}) = (a+b)^* ab (a+b)^*$.

- $L_1 = L$ • $L_2 \neq L_1$ car $b \in L_2$ mais $b \notin L_1$.
- $L_3 \neq L_1$ car $\varepsilon \in L_3$ mais $\varepsilon \notin L_1$. • idem $L_3 \neq L_2$.
- $L_4 \neq L_1$ et $L_4 \neq L_2$ pour les m̂ raisons.
- Soit $u \in L_3$. Si $u = \varepsilon$, $u \in L_4$.
  Si $u = av$, $v \in L_4$ or ④↺^{a} donc $av \in L_4$ soit $u \in L_4$
  Si $u = bv$, comme ③↺^{b}, $v \in L_3$, et puisque $④ \xrightarrow{b} ③$, $bv \in L_4$ soit $u \in L_4$.

Donc $L_3 \subset L_4$. De m̂ on MQ $L_4 \subset L_3$. D'où $L_4 = L_3$.

→ Finalemt $Q/_{\mathcal{N}} = \{\{1\}, \{2\}, \{3,4\}\}$

et $\mathcal{A}_{\mathcal{N}}$:



Pté Dans l'automate minimal $\mathcal{A}_L$ d'un langage L, la congruence de Nérode est l'égalité puisque pour tt état $X \in Q_L$, $L_X = X$

BBC p 315 pté 5.5.

Preuve X s'écrit $u^{-1}L$ pour un certain $u \in \Sigma^*$.

Par déf $L_X = \{ w \in \Sigma^* \mid \delta_L(X, w) \in F_L \}$

$= \{ \text{———} \mid w^{-1}X \in F_L \}$

$= \{ \text{———} \mid w^{-1}u^{-1}L \in F_L \}$

$= \{ \text{———} \mid (uw)^{-1}L \in F_L \}$

$= \{ \text{———} \mid uw \in L \}$ cf fin de la preuve de $\mathcal{L}(\mathcal{A}_L) = L$.

$= u^{-1}L$

$= X$.

Pté L'automate de Nérode $\mathcal{A}/_N$ est isomorphe à l'automate minimal $\mathcal{A}_L$ où $L = \mathcal{L}(\mathcal{A})$.

BBC p 317. pté 5.7.

Preuve On considère à nouveau $\varphi \begin{pmatrix} Q \to Q_L \\ q \mapsto L_q \end{pmatrix}$ et $\bar{\varphi} = \begin{pmatrix} Q/\!\!\sim \to Q_L \\ [q] \mapsto L_q \end{pmatrix}$

(Qu'on écrivait avant $\delta(i, f) \to f^{-1}L$.

C'est bien la m̂ chose car si $q = \delta(i, f)$
$L_q = \{ w \mid \delta(q, w) \in F \}$
$= \{ w \mid \delta(\delta(i, f), w) \in F \}$
$= \{ - \mid \delta(i, fw) \in F \}$
$= \{ w \mid fw \in L \}$ car $\{i\} = I$ et $\mathcal{L}(A) = L$
$= f^{-1}L$.

On sait déjà que φ est surjectif. donc $\bar{\varphi}$ l'est aussi

La déf des congruences de Nérode assure l'injectivité de $\bar{\varphi}$

en effet $\varphi(q) = \varphi(q') \Leftrightarrow L_q = L_{q'} \Leftrightarrow q \sim q'$

donc $\bar{\varphi}([q]) = \bar{\varphi}([q']) \Leftrightarrow q \sim q' \Leftrightarrow [q] = [q']$.

Calcul de l'automate de Nérode

L'idée est d'arriver à la congruence de Nérode en affinant successivement des relations d'équivalences - et les partitions associées - sur les sommets.

Deux états sont congrus selon Nérode si les mots qui leur permettent d'atteindre un état final sont les mêmes.

On introduit pour $k \in \mathbb{N}$, une relation moins forte qui assure que les mots de taille $\leq k$ qui permettent d'arriver à l'état final sont les mêmes.

$$p \sim_k q \quad \text{ssi} \quad L_p^{(k)} = L_q^{(k)} \quad \text{où} \quad L_p^{(k)} = \{w \in \Sigma^* / |w| \leq k, w \in L_p\}$$

Évidemment ces propriétés sont de plus en plus fortes, on dit que les relations sont de plus en plus fines, et la congruence de Nérode est la plus fine car $p \sim_{\mathcal{N}} q$ ssi $\forall k \in \mathbb{N}\ p \sim_k q$ ☆1

puisque $L_p = \sup_{k \in \mathbb{N}} L_p^{(k)}$.

Pté $\boxed{p \sim_k q \Leftrightarrow p \sim_{k-1} q \text{ et } \forall a \in \Sigma,\ \delta(p,a) \sim_{k-1} \delta(q,a)}$ ☆2

Preuve $p \sim_k q \Leftrightarrow L_p^{(k)} = L_q^{(k)}$

$\Leftrightarrow \forall w \in \Sigma^*, |w| \leq k \Rightarrow w \in L_p$ ssi $w \in L_q$

$\Leftrightarrow \forall w \in \Sigma^*, |w| \leq k-1 \Rightarrow$ ——

et $\forall a \in \Sigma\ \forall w \in \Sigma^*, |w| \leq k-1 \Rightarrow aw \in L_p$ ssi $aw \in L_q$

$\Leftrightarrow L_p^{(k-1)} = L_q^{(k-1)}$ et $\forall a \in \Sigma$, $\forall w \in \Sigma^* |w| \leq k-1$, $w \in L_{\delta(p,a)}$ ssi $w \in L_{\delta(q,a)}$

$\Leftrightarrow L_p^{(k-1)} = L_q^{(k-1)}$ et $\forall a \in \Sigma$, $L_{\delta(p,a)}^{(k-1)} = L_{\delta(q,a)}^{(k-1)}$

$\Leftrightarrow p \sim_{k-1} q$ et —— $\delta(p,a) \sim_{k-1} \delta(q,a)$.

Notons Q_k la partition de Q selon $\sim_k$. (ie $Q_k = Q/\sim_k$)

Puisque $p \sim_0 q \Leftrightarrow \varepsilon \in L_p$ ssi $\varepsilon \in L_q \Leftrightarrow p \in F$ ssi $q \in F$

on a $Q_0 = \{Q \setminus F, F\}$. C'est là notre initialisation.

A chaque étape on calcule la partition suivante en utilisant la pté suivante.

Dès que $Q_k = Q_{k+1}$, on a $\forall l \geq k \; Q_l = Q_k$.

(En effet si $Q_k = Q_{k+1}$ $p \sim_k q$ ssi $p \sim_{k+1} q$
et alors $p \sim_{k+2} q$ ssi $p \sim_{k+1} q$ et $\forall a \in \Sigma$, $\delta(p,a) \sim_{k+1} \delta(q,a)$
ssi — k ——— k —
ssi $p \sim_{k+1} q$
soit $Q_{k+2} = Q_{k+1}$ et on itère..

De plus on est assuré que cette stabilisation a lieu au + tard au rang $n-2 = \# Q - 2$ puisque $\# Q_0 = 2$, tant que $Q_{k+1} \neq Q_k$, c-à-d $\sim_{k+1}$ est strictement plus fine que Q_k, on a $\# Q_{k+1} \geq Q_k + 1 \underset{\text{itérant}}{\geq} k+2$, en particulier $\# Q_{n-2}$ serait $\geq n-2+2 = \# Q$, or il ne peut y avoir + de n classes dans 1 partition d'1 ens à n élém^t.

$$Q_{\omega} = \limsup_{k \to +\infty} Q_k = Q_{n-2}.$$

D'où l'algorithme suivant : (très haut niveau non optimisé en espace (n^2m au lieu de $2n^2$))

Nérode (A)

Soit M une matrice indicée par $Q \times Q \times [0..n-2]$ // $M(p,q,i)$ représentera si $p \sim_i q$ ou non.

Pour tout $(p,q) \in Q^2$
- Si $p \in F$ et $q \in F$, $M(p,q,0) \leftarrow$ vrai
- Si $p \notin F$ et $q \notin F$ $M(p,q,0) \leftarrow$ vrai
- Sinon $M(p,q) \leftarrow$ faux

en $O(n^2)$

Pour $i = 1$ à $n-2$
- Pour $(p,q) \in Q^2$
  - Pour $a \in \Sigma$
    - $M(p,q,i) \leftarrow M(p,q,i-1) \wedge M(\delta(p,a), \delta(q,a), i-1)$

en $O(n^2 n\, m)$

Retourner $M(p,q,(n-2))$

D'où un algorithme en $O(mn^3)$