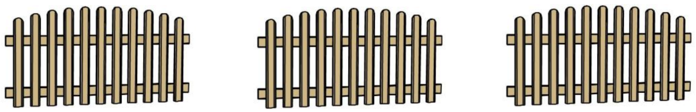
5 – Clôture des langages rationnels



Récapitulatif

- L'ensemble des langages rationnels est clos par :
  - concaténation
  - union
  - étoile
- Cette classe est également close par :
  - complémentation
  - intersection
  - quotient à droite ou à gauche
- Chaque langage rationnel est :
  - caractérisé (et reconnu) par son unique automate minimal
  - décrit par une infinité d'E.R.
  - reconnu par une infinité de A.F.D. ou de A.F.N.
  - engendré par une infinité de grammaires ...

Clôture par complémentation

L langage rationnel sur Σ

$\Rightarrow$

$\overline{L} = \Sigma^*\backslash L$ langage rationnel sur Σ

Déterministe + complet suffiraient ici

Construction :

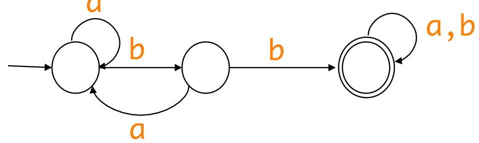
- L un L.R. sur $\Sigma \Rightarrow$ il existe un A.F.D. minimal $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ tel que $L(\mathcal{A}) = L$
- on construit $\mathcal{C}$ à partir de $\mathcal{A}$ pour qu'il reconnaisse $\overline{L}$:

$$\mathcal{C} = (Q, \Sigma, \delta, q_0, Q\backslash F)$$

➔ *tout état non final est devenu final et vice versa.*

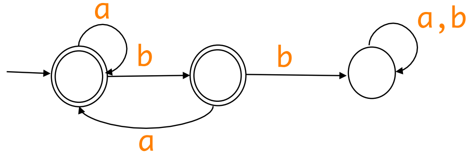
Exemple

$\mathcal{A}$



$L(\mathcal{A}) = (a+b)^*bb(a+b)^*$

$\mathcal{C}$



$L(\mathcal{C}) = \overline{L(\mathcal{A})}$

i.e. les mots exempts du facteur bb sur {a,b}

Cette construction n'est correcte que parce que $\mathcal{A}$ est complet et déterministe !

Clôture par intersection

L et K langages rationnels sur Σ

⇒

L∩K langage rationnel sur Σ

Démonstration :

- l'ensemble des L.R. clos par complémentation donc $\overline{L}$ et $\overline{K}$ sont rationnels
- l'ensemble des L.R. clos par union donc $\overline{L} \cup \overline{K}$ est rationnelle
- de par la clôture par complémentation, on obtient que

$$\overline{\overline{L} \cup \overline{K}}$$

  est un langage rationnel.
- par identité ensembliste on déduit : $L \cap K$ est rationnel.

Construction de l'intersection*

- Soient L et K deux langages rationnels et les A.F.D. :
  - $A = (Q, \Sigma, \delta, q_0, F)$ tel que $L(A) = L$
  - $B = (Q', \Sigma, \delta', q_0', F')$ tel que $L(B) = K$

  alors

  $C = (Q\mathrm{x}Q', \Sigma, \delta_C, [q_0,q_0'], F\mathrm{x}F')$

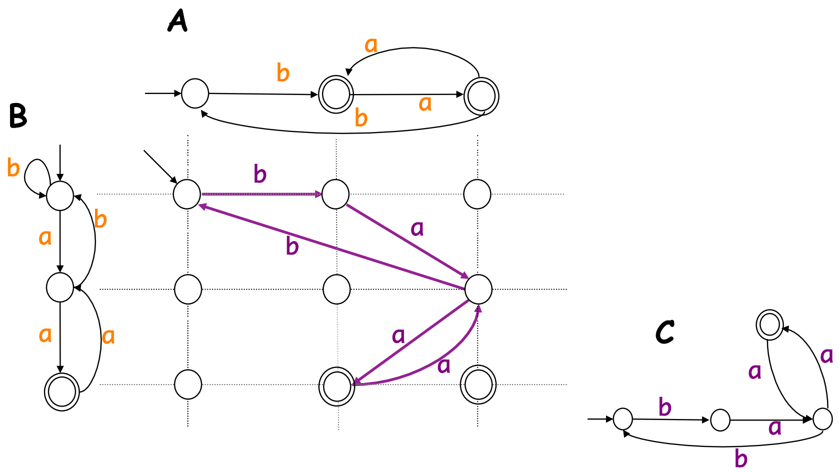
  avec pour tout couple p,q de QxQ',
  pour tout σ de Σ,

  $\delta_C([p,q] , \sigma) = [\delta(p, \sigma), \delta'(q,\sigma)]$

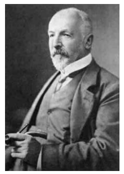
- l'automate fini C reconnaît le langage rationnel $L(C)$ égal à L∩K.

* On pourrait construire l'union de quasi-presque similaire : *cf. TD !*

Exemple



Des langages non rationnels ?



Georg Ferdinand Ludwig Philip Cantor
(1845-1918)

On raisonne sur un alphabet donné.

- Par un argument de diagonalisation, on a obtenu (cours 1) que l'ensemble des langages *n'est pas* dénombrable.

Démontré en TD !
(cf. ex.6 TD 1)

- L'ensemble des expressions régulières est dénombrable.
- Donc l'ensemble des langages rationnels est dénombrable.

➔ on en déduit qu'il existe des *langages non rationnels*.

D'un raisonnement similaire sur les automates, on aurait aussi pu déduire l'existence de langages non rationnels.

Problème

- Donnée un langage L sur l'alphabet Σ
- Problème le langage L est-il rationnel ?

Approche du problème :

- le théorème de l'étoile va servir à montrer qu'un langage <u>n'est pas</u> rationnel.
- dans le cas où le théorème s'avère inutilisable*, il faut une démonstration *ad hoc*.

* ce théorème ne permet pas de montrer la non-rationalité de tous les langages non-rationnels : ce n'est pas une caractérisation des langages rationnels.

Théorème de l'étoile

- Lors de la lecture d'un mot de longueur m ou plus accepté par un automate fini A comportant n états avec n≤m, l'automate A passe forcément deux fois par le même état.
- Plus précisément, le théorème énonce :
  « quand un mot w a un facteur v de longueur plus grande que le nombre d'états de l'automate, la lecture de v boucle forcément quelquepart ... »

Théorème

soit L un langage rationnel, il existe un entier n, n ≥ 1 tel que :

- pour tout mot w de L de longueur ≥ n,
- pour toute factorisation de w en xvy avec v facteur de longueur ≥ n
- il existe des mots r, s et t tels que :

i) $0 < |s| \leq n$

ii) $\forall\, i \geq 0,\ x\, r\, s^i\, t\, y \in L$

Démonstration

- $A = (Q, \Sigma, \delta, q_0, F)$ automate fini à n états tel que $L = L(A)$
- prenons un mot $w = xvy$ de L avec $|v| \geq n$: $w = \sigma_1\sigma_2 \ldots \sigma_{|w|}$
  on sait que $\delta^*(q_0, \sigma_1\sigma_2\ldots\sigma_{|w|}) \in F$
- pendant la lecture du facteur v de w,
  A passe forcément deux fois dans le même état q de Q.
- il existe donc deux indices j et k, avec $j<k$,
  $|x|<j \leq |xv|$ et $|x|<k \leq |xv|$ tels que :

$$q = \delta^*(q_0, \sigma_1 \ldots \sigma_j) = \delta^*(q_0, \sigma_1 \ldots \sigma_k)$$

- si on pose :

$$x\,r = \sigma_1 \ldots \sigma_j$$
$$s = \sigma_{j+1} \ldots \sigma_k$$
$$t\,y = \sigma_{k+1} \ldots \sigma_{|w|}$$

  pour tout $i \geq 0$, les mots $xrs^ity \in L$
- $j<k$ donc $|s| \geq 1$ et $|s| \leq n$ en prenant j et k minimaux.

Exemple

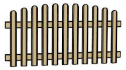
- soit le langage $L = \{ 0^n 1^n,\ n \geq 0 \}$
- supposons par l'absurde qu'il est rationnel
- quel que soit l'entier $n \geq 1$ fourni par le théorème :
- soit le mot $w = 0^n1^n$ de L, on a bien sa longueur $2n \geq n$
- et soit la factorisation de w en xvy avec

$$x = \varepsilon$$
$$v = 0^n$$
$$y = 1^n$$

- pour toute factorisation de v en r, s et t tels que
  $0 < |s| \leq n$
  on cherche un i tel que
  $x\,r\,s^i\,t\,y \notin L$
- par exemple prenons $i=0$: $z = x\ r\ t\ y = 0^m1^n \notin L$ car $m<n$
  $\Rightarrow$ contradiction avec le théorème
- Conclusion $L = \{ 0^n1^n,\ n \geq 0 \}$ n'est pas rationnel.

Autre idée ...

Utiliser les propriétés de clôture !



Elles peuvent suffire à démontrer qu'un langage n'est pas rationnel, dès lors que l'on connaît déjà des langages non rationnels ...

Exemple :

- Considérons le langage $L = \{w \in \{0,1\}^*, |w|_0 = |w|_1\}$.
- Est-il rationnel ? ... supposons que oui.
- On sait que le langage K décrit par 0^*1^* est rationnel.
- $K \cap L = \{ 0^n1^n,\ n \geq 0 \}$ donc $K \cap L$ n'est pas rationnel.
- Or on sait que l'intersection de deux langages rationnels est un langage rationnel $\Longrightarrow$ contradiction

Conclusion : $L = \{w \in \{0,1\}^*, |w|_0 = |w|_1\}$ n'est pas rationnel.