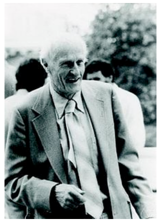
# 2 - Théorème de Kleene



Stephen C. Kleene
Mathématicien et logicien américain
1909-1994

## Théorème de Kleene

Théorème : Un langage sur un alphabet $\Sigma$ est rationnel si et seulement si il est reconnu par un automate fini.

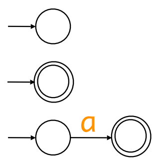
Premier sens : preuve par induction

- On cherche des A.F.N pour chaque langage régulier dénoté par les expressions régulières de base

**Expressions régulières**

Définition inductive

(B)
$\varnothing \in$ ER
$\varepsilon \in$ ER
$a \in$ ER, pour tout $a \in \Sigma$
(I) ...



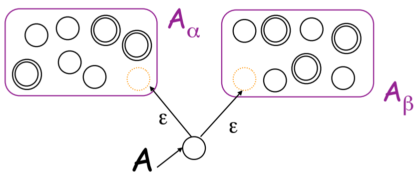

## Suite preuve (opération +)

- par hypothèse d'induction, on suppose qu'il existe deux A.F.N. $A_\alpha$ et $A_\beta$ tels que $L(A_\alpha)$ est le langage décrit par l'exp.rég. $\alpha$ et $L(A_\beta)$ celui décrit par l'exp.rég. $\beta$.

**Expressions régulières**

Définition inductive

(B) ...
(I) si $\alpha, \beta \in$ ER alors
$(\alpha + \beta) \in$ ER
...




- $A$ reconnaît le langage décrit par $(\alpha + \beta)$

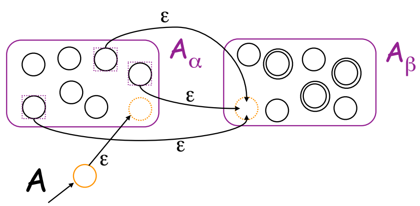
## Suite preuve (opération .)

- $A$ reconnaît le langage décrit par $(\alpha.\beta)$

**Expressions régulières**

Définition inductive

(B) ...
(I) si $\alpha, \beta \in$ ER alors
...
$(\alpha . \beta) \in$ ER
...



## Suite preuve (opération *)

- $\mathcal{A}$ reconnaît le langage décrit par $\alpha^*$

Expressions régulières

Définition inductive

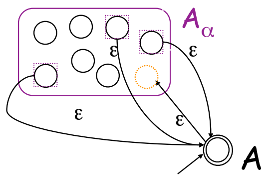
(B) ...

(I) si $\alpha, \beta \in$ ER alors

...

...

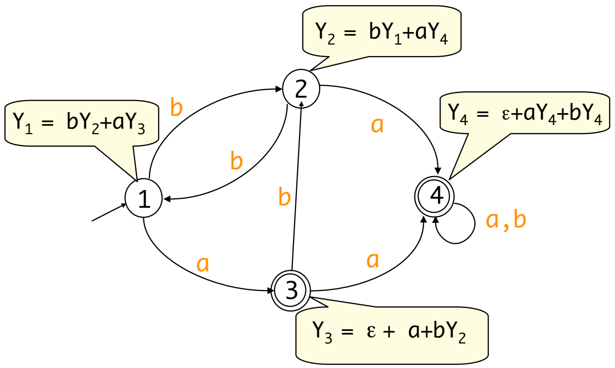
$\alpha^* \in$ ER




## Passage d'un A.F.D. à une E.R.

- Réciproquement, à partir d'un A.F.D, peut-on construire une expression régulière qui décrit le langage reconnu par $\mathcal{A}$ ?
- Les systèmes d'équations linéaire à droite permettent de ramener le calcul d'une exp.rég. à la résolution d'un système d'équations.
- A chaque état $q$ on associe une expression régulière $Y_q$ dénotant le langage $L_q(\mathcal{A})$ associé à cet état.
- On obtient un système d'équations dont les inconnues sont des exp.rég. dénotant des langages.
- Si $q$ est l'état initial, $Y_q$ décrit le langage reconnu par $\mathcal{A}$.

## Exemple




$Y_q$ décrit le langage des mots w tels que $\delta^*(q,w) \in F$

$\delta^*$ : prolongement de $\delta$ aux mots

## Exemple

Attention !!!

Ne pas oublier les $\varepsilon$ ...

$$\begin{cases} Y_1 = bY_2 + aY_3 \\ Y_2 = bY_1 + aY_4 \\ Y_3 = \varepsilon + aY_4 + bY_2 \\ Y_4 = \varepsilon + (a+b)\, Y_4 \end{cases}$$

Résoudre un tel système revient à calculer $Y_1$, car il est associé à l'état initial 1.

On procède par substitutions et on va avoir besoin de résoudre l'équation :

$$Y = A\ Y + B$$

## Equation $Y = A Y + B$

- $A^*B$ est bien solution :

  $A^*B = A A^*B + B = A^+ B + B = A^*B$

- $A^*B$ est une solution minimale :

  si Y est une solution, alors $A^*B \subseteq Y$

  Récurrence sur la hauteur d'étoile :
  - si i=0, $A^0B = B \subseteq Y$ car $Y = AY + B$
  - hyp. réc. pour i=n : $A^nB \subseteq Y$
  - pour n+1 : $A^{n+1}B = A\, A^nB \subseteq A\, Y \subseteq A\, Y + B = Y$, cqfd.

## Equation $Y = A Y + B$ (suite)

- si $\varepsilon \notin A$ alors $A^*B$ est l'unique solution.

  On suppose la non-unicité de la solution $A^*B$.
  Soit X une autre solution
  et soit un mot w de longueur minimale tel que $w \in X \setminus A^*B$.

  $w \in X = A\,X + B$ et $w \notin B$ donc $w = uv$ avec $u \in A$ et $v \in X$.
  Or $v \notin A^*B$ (sinon w aussi) donc $v \in X \setminus A^*B$.
  Contradiction : la longueur de w n'était donc pas minimale !

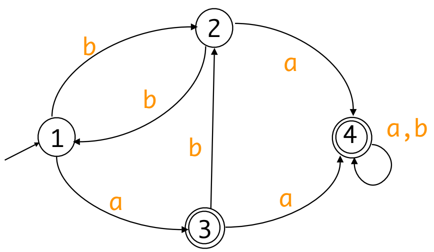
- si $\varepsilon \in A$ alors pour tout $C \subseteq \Sigma^*$, $Y = A^*B + A^*C$ est aussi solution :

  $A^*B + A^*C = A\,(A^*B + A^*C) + B = A^+B + A^+C + B = A^*B + A^+C = A^*B + A^*C$

- si l'automate de départ est déterministe, on a toujours : $\varepsilon \notin A$

## Exemple

On garde toujours à l'esprit que si $\varepsilon \notin A$ alors $A^*B$ est l'unique solution de $Y=AY+B$.



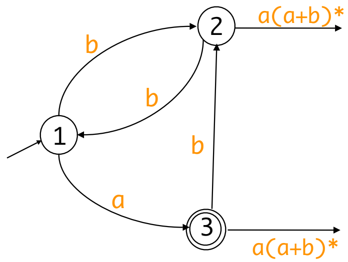

$$\begin{cases} Y_1 = bY_2 + aY_3 \\ Y_2 = bY_1 + aY_4 \\ Y_3 = \varepsilon + aY_4 + bY_2 \\ Y_4 = \varepsilon + (a+b)Y_4 \end{cases}$$

$\Rightarrow$

$Y_4 = (a+b)^*$

($Y_4$ est alors éliminé)

## Exemple




$$\begin{cases} Y_1 = bY_2 + aY_3 \\ Y_2 = bY_1 + a(a+b)^* \\ Y_3 = \varepsilon + a(a+b)^* + bY_2 \end{cases}$$

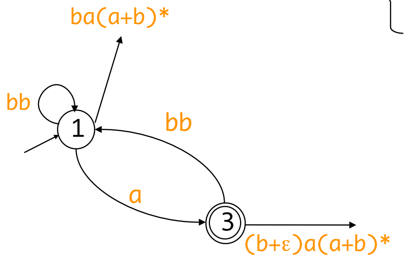
$\Rightarrow$

$$\begin{cases} Y_1 = bbY_1 + ba(a+b)^* + aY_3 \\ Y_3 = \varepsilon + a(a+b)^* + bbY_1 + ba(a+b)^* \end{cases}$$

($Y_2$ est alors éliminé)

## Exemple

On garde toujours à l'esprit que si $\varepsilon \notin A$ alors $A^*B$ est l'unique solution de $Y=AY+B$.




$$\begin{cases} Y_1 = bbY_1 + ba(a+b)^* + aY_3 \\ Y_3 = \varepsilon + bbY_1 + (b+\varepsilon)a(a+b)^* \end{cases}$$

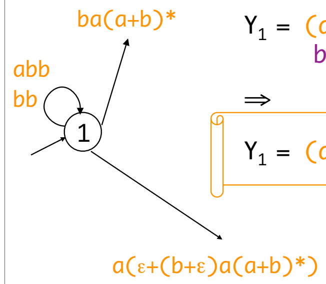
$\Longrightarrow$

$$Y_1 = bbY_1 + ba(a+b)^* + a\varepsilon + abbY_1 + a(b+\varepsilon)a(a+b)^*$$

($Y_3$ est alors éliminé)

## Exemple

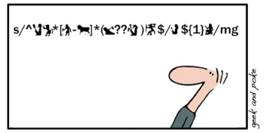
On garde toujours à l'esprit que si $\varepsilon \notin A$ alors $A^*B$ est l'unique solution de $Y=AY+B$.




$$Y_1 = (abb+bb)\, Y_1 + ba(a+b)^* + a(\varepsilon+(b+\varepsilon)a(a+b)^*)$$

$\Longrightarrow$

$$Y_1 = (abb+bb)^*(ba(a+b)^*+a(\varepsilon+(b+\varepsilon)a(a+b)^*))$$



ANCIENT EGYPTIAN REGEXP