Langages formels

1) Mots, langages

2) Automates déterministes

3) Grammaires et langages algébriques

Définition 4:

Les langages rationnels sont les langages reconnaissables par des automates finis.

(LP n'est donc pas rationnel)

3) Grammaires et langages algébriques

- **3a) définitions**
- Définition 5:

Une grammaire algébrique est un triplet G=(X,Y,*P*) où:

X est un ensemble fini (alphabet terminal).

Y est un alphabet disjoint de X (alphabet auxiliaire)

P est un sous-ensemble de Y x (X U Y)*
(ensemble des productions ou règles de réécriture)

Exemple 3:

$$G_1 : X = \{a,b\}; Y = \{S\}; P = \{(S,\varepsilon),(S,aSbS)\}$$

Notation abrégée de P:

$$P : S \rightarrow \varepsilon \mid aSbS$$

Exemple 4

$$G_2 : X = \{a, b, c, +, -, *, /\} \; Y = \{S\}$$

$$P : S \rightarrow S+S \mid S-S \mid S*S \mid S/S \mid a \mid b \mid c$$

- **<u>Définitions 6</u>**: Soit G=(X,Y,*P*) une grammaire algébrique et f et g deux mots de (XUY)*.

On dit que ***g dérive immédiatement de f***, on note **f** $\vdash$**g**, lorsque:

$\exists$ u, v, α $\in$(XUY)* et S$\in$Y

1) f = u•S•v .

2) g= u•α•v

3) S$\rightarrow$α est une règle de réécriture.

Dérivations immédiates :

$$P : S \rightarrow \varepsilon \mid aSbS$$

Exemple :

S ⊢aSbS ⊢aaSbSbS ⊢aaSbSb ⊢ aaSbb⊢ aabb

Dérivations immédiates :

G2 : S → S+S | S*S | S-S | S/S | a | b | c

Comment obtenir par une suite de dérivations immédiates (commençant par S) le mot **a+b*c** ?

Dérivations immédiates :

G2 : P $\rightarrow$ S+S | S*S | S-S | S/S | a | b | c .

Comment obtenir par une suite de dérivations immédiates (commençant par S) le mot **a+b*c** ?

S ⊢S+S ⊢a+S ⊢a+S*S ⊢a+b*S ⊢a+b*c

ou

S ⊢S*S ⊢S+S*S ⊢a+S*S ⊢a+b*S ⊢a+b*c

Définitions 6 (suites)

On dit que **g dérive à l'ordre k de f** ($k \geq 0$), et on

écrit **f** $\vdash^{k}$ **g** lorsqu'il existe $f_0, f_1, \ldots, f_k \in (X \cup Y)^*$ tels que:

1) $f = f_0$ et $g = f_k$.

2) $f_{i-1} \vdash f_i$ pour $i \in [1, k]$

Définitions 6 (suites)

On dit que ***g dérive de f***,
on note $\mathbf{f} \vdash^{*} \mathbf{g}$, lorsqu'il existe k $\in$IN tel que
$\mathbf{f} \vdash_{k} \mathbf{g}$

Dérivations:

$$P : S \rightarrow \varepsilon \mid aSbS$$

S $\vdash$aSbS $\vdash$aaSbSbS $\vdash$aaSbSb $\vdash$aaSbb $\vdash$ aabb

On peut écrire aussi:

S $\vdash^5$aabb

Ou plus simplement :

S $\vdash^*$aabb

On dira que “aabb dérive de S” (ou “S engendre aabb”)

Définitions 7

- Etant donné une grammaire algébrique $G=(X,Y,P)$ et un mot $f \in (XUY)^*$, on appelle **langage engendré par f** dans la grammaire G, que l'on note L(G, f) l'ensemble des mots ***terminaux*** qui dérivent de f.

$$L(G, f) = \{ m \in X^* / f \vdash^* m\}$$

Définitions 7 (suite)

- On appelle **langage algébrique** tout langage L tel qu'il existe une grammaire algébrique G et un auxiliaire S tel que $L = L(G, S)$.

3b) Propriétés fondamentales

- Propriétés 3:
  Soit G=(X, Y, *P*) une grammaire algébrique
  Soient f, f ', g, g', h, h_1, h_2 $\in$(XUY)* ,
  et **u, v $\in$X*.**

  - 1) Si f $\vdash^*$g alors $h_1 \bullet f \bullet h_2 \vdash^* h_1 \bullet g \bullet h_2$
  - 2) Si f $\vdash^*$g et g $\vdash^*$h alors f $\vdash^*$h
  - 3) Si f $\vdash^*$g et f ' $\vdash^*$g' alors f • f ' $\vdash^*$ g • g'
  - 4) Si f = u•f ' • v et si f $\vdash^*$g alors g se factorise en
    g = u• g'• v et f ' $\vdash^*$g'

Propriétés 3 (suite)

- Pour tout $\mathbf{f_1, f_2, g \in (XUY)^*}$ et $k \in IN$:

$$f_1 \bullet f_2 \vdash^k g \iff$$

$\exists\, g_1, g_2 \in (XUY)^*$ et k_1 et $k_2 \in IN$ tels que:

$$\boxed{\begin{array}{l} 1)\ f_1 \vdash^{k_1} g_1 \\ 2)\ f_2 \vdash^{k_2} g_2 \\ 3)\ k_1 + k_2 = k \\ 4)\ g = g_1 \bullet g_2 \end{array}}$$

Exemple 5

$$G_3 = \left(X = \{a, b\}, Y = \{S\}, P_3 = \{S \rightarrow aSb / \varepsilon\} \right)$$

Comment démontrer que,
$L(G_3,S)=\{a^n b^n$,avec n un entier positif ou nul$\}$?

$$G_3 = \left(X = \{a, b\}, Y = \{S\}, P_3 = \{S \rightarrow aSb / \varepsilon\} \right)$$

$L(G_3, S) = \{a^n b^n$, avec n un entier positif ou nul$\}$?

1) Pour tout entier n, $S \vdash^* a^n b^n$?
(par récurrence)

2) Si $S \vdash^* m$ alors
m est de la forme $a^n b^n$?
(Par récurrence sur le nombre de dérivations)

$$G_3 = \left(X = \{a,b\}, Y = \{S\}, P_3 = \{S \rightarrow aSb / \varepsilon\} \right)$$

$L(G_3,S)=\{a^n b^n$,avec n un entier positif ou nul} ?

1)Pour tout entier n, $S \vdash^* a^n b^n$?
(par récurrence)

Par convention $a^0 b^0 = \varepsilon$

P(0) ? $S \vdash^* a^0 b^0 = \varepsilon$ car $S \rightarrow \varepsilon$ est une règle de réécriture.

P(n)=>P(n+1) ?

par **HR**, $S \vdash^* a^n b^n$ donc (propriété 1)

$aSb \vdash^* a a^n b^n b = a^{n+1} b^{n+1}$

et puisque $S \vdash^* aSb$ (c'est une règle) on a :

$S \vdash^* a^{n+1} b^{n+1}$ (P(n+1)) (propriété 2)

$$G_3 = \left(X = \{a,b\}, Y = \{S\}, P_3 = \{S \rightarrow aSb / \varepsilon\} \right)$$

L(G_3,S)={$a^n b^n$,avec n un entier positif ou nul} ?

2) Si S⊢*m alors
m est de la forme $a^n b^n$?
(Par récurrence sur le
nombre de dérivations)

P(k) :Si S⊢^{k} m alors
m est de la forme $a^n b^n$
(m est terminal)

P(1) ? Ε est le seul mot terminal qui dérive de S en une dérivation.
P(1) est vrai.

P(n)=>P(n+1) ? Soit m un mot terminal tel que S⊢$^{k+1}$m avec k>0.
Donc : S⊢aSb⊢^{n}m
Par hypothèse de récurrence et d'après la propriété 4,
m=a•m'•b avec m'= $a^n b^n$.
Donc m= $a^{n+1} b^{n+1}$.

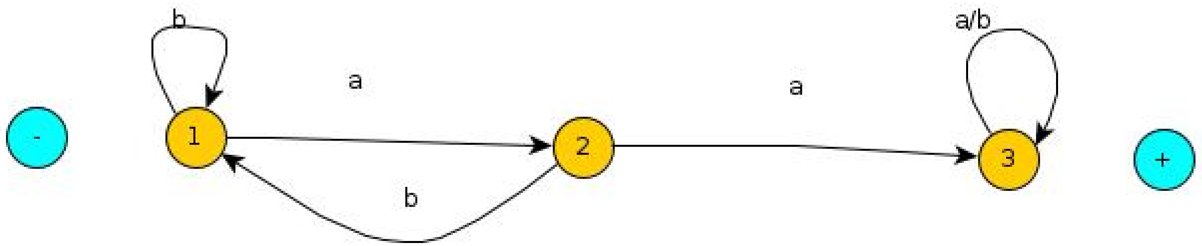
3c) Langages rationnels et langages algébriques

- Théorème:

Tous les langages rationnels sont algébriques.

Remarque: cette inclusion est stricte.

Passage de l'automate à la grammaire



L(A)=L(G,1) avec G définie par :

Car 1 est
l'état initial

1 → b1/a2
2 → a3/b1
3 → a3/b3/ε

Car 3 est un
état final