NGUYEN Thi Minh Tuyen

Théorie des automates et langages formels

Formes canoniques des Grammaires

Théorème

Théorème 1: Tout langage régulier est un langage non contextuel.

Démonstration: On démontre que pour tout automate fini, il existe une grammaire non contextuelle tel que le langage engendré par la grammaire est le même que le langage reconnu par l'automate.

Par algorithme constructif.

Entrée : automate fini.

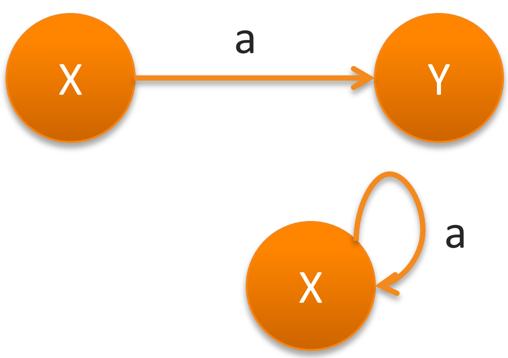
Sortie : grammaire.

Automate fini et grammaire non contextuelle

- L'alphabet de l'automate fini sont les terminaux de la grammaire.
- Chaque état correspond a une variable de la grammaire.

(L'état de départ corresponds à la variable du départ S.)

- Pour chaque transition, on crée une production correspondante



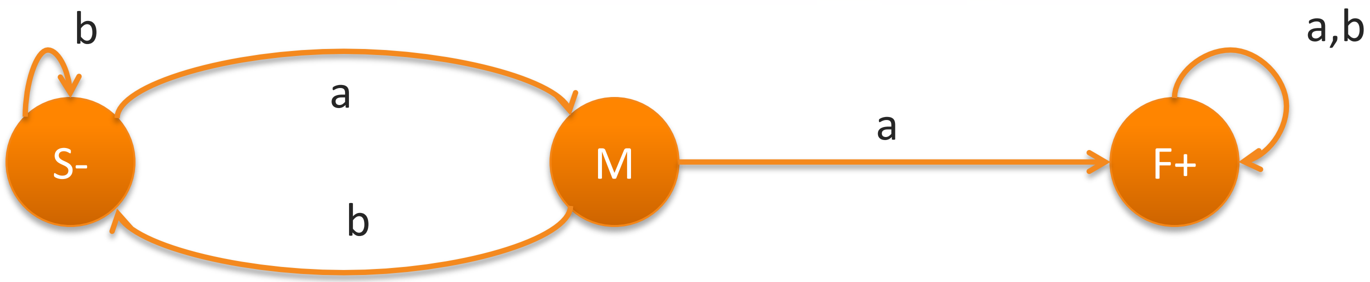
$X \rightarrow aY$

$X \rightarrow aX$

- Pour chaque état final X, on crée une production de la forme:

$X \rightarrow \Lambda$

Exemple



S → aM | bS

M → aF| bS

F → aF| bF| Λ

babbaaba

$$S \xrightarrow{b} S \xrightarrow{a} M \xrightarrow{b} S \xrightarrow{b} S \xrightarrow{a} M \xrightarrow{a} F \xrightarrow{b} F \xrightarrow{a} F$$

$S \Rightarrow bS \Rightarrow baM \Rightarrow babS \Rightarrow babbS \Rightarrow babbaM \Rightarrow babbaaF \Rightarrow babbaabF \Rightarrow babbaabaF \Rightarrow babbaaba$

Définition

- Un mot partiel est une suite de terminaux (qui peut être vide) suivie d'une seule variable.

  (terminal)(terminal)…(terminal)(Variable)

- Une grammaire non contextuelle est une grammaire régulière si toutes les productions sont de la forme:

  Variable →mot partiel

ou

  Variable →mot (suite de terminaux ou Λ)

Théorème 2

Tout langage engendré par une grammaire régulière est un langage régulier.

Démonstration. On construit un graphe de transition pour le langage.

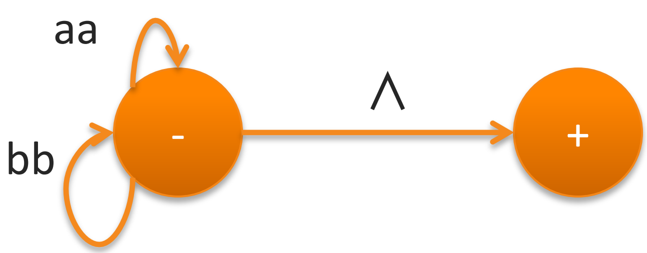
- L'alphabet Σ du graphe de transition est l'alphabet terminal.
- Un état pour chaque variable. L'état nommé S est l'état du départ. On rajoute un état final +.
- Les transitions:

$$N_x \rightarrow w_y N_z \qquad\qquad N_p \rightarrow w_q$$

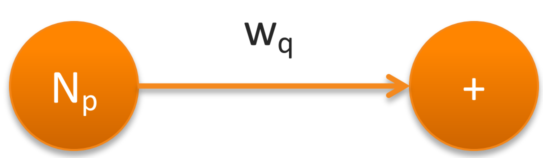
$$N_x \xrightarrow{w_y} N_z \qquad\qquad N_p \xrightarrow{w_q} +$$

Exemple

$S \rightarrow aaS$ $S \rightarrow bbS$ $S \rightarrow \Lambda$



(aa+bb)*



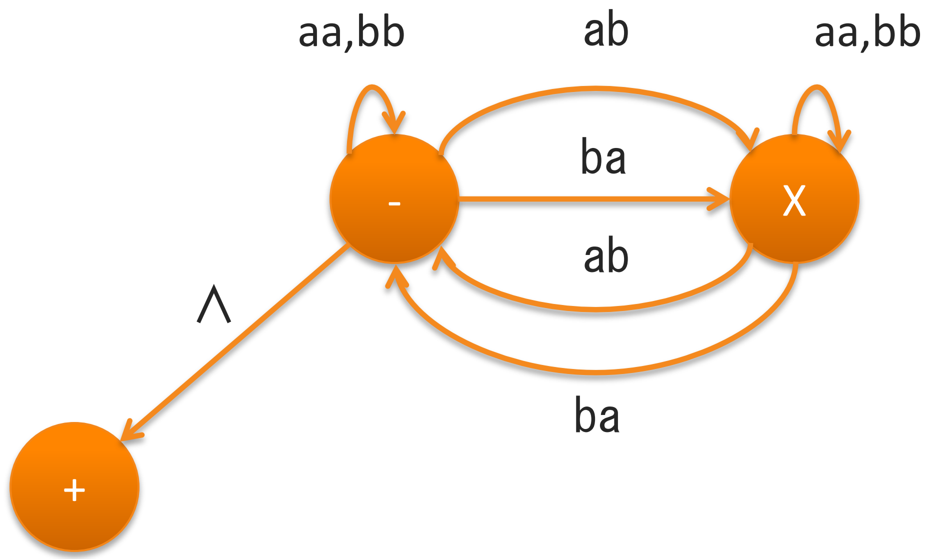
<u>Algorithme alternatif</u>:

1. $N_p \rightarrow w_q$ tant que w_q n'est pas Λ
2. Pour chaque transition de la forme $N \rightarrow \Lambda$, on met un + dans l'état pour N.

Exemple

S →aaS | bbS |abX | baX | Λ

X →aaX | bbX | abS | baS



PAIR-PAIR

Production $X \rightarrow \Lambda$

- Si Λ est dans le langage, alors on a nécessairement une production de la forme $N \rightarrow \Lambda$.
- Par contre, une production de la forme $N \rightarrow \Lambda$ ne garantie pas que Λ est dans le langage
- Exemple: $S \rightarrow aX$ $S \rightarrow a$ $X \rightarrow \Lambda$

Théorème 3. Soit L un langage engendré par une grammaire non contextuelle G. Il existe une grammaire non contextuelle sans productions de la forme $X \rightarrow \Lambda$ tel que:

$L - \{\Lambda\}$ est engendré par la nouvelle grammaire.

Théorème

Théorème 4. Soit L un langage engendré par une grammaire non contextuelle G sans productions de la forme N → Λ. Il existe une grammaire non contextuelle G' sans productions de la forme N → **Λ** (production nulle) et N → M (production unitaire) telle que L est engendré par G'.

Théorème 5. Soit L un langage engendré par une grammaire non contextuelle G. Il existe une grammaire G' qui génère tous les mots de L - {Λ} telle que toutes les productions de G' sont de la forme:

Variable →suite de Variables

Variable → un seul terminal

Forme normale de Chomsky

Définition. Une grammaire non contextuelle est dite sous la forme normale de Chomsky si toutes les productions sont de la forme:

Variable → (Variable)(Variable)

Variable → terminal

Théorème 6. Soit L un langage engendré par une grammaire non contextuelle G. Il existe une grammaire G' sous la forme normale de Chomsky et qui génère tous les mots de L - {Λ}.

Dérivation gauche

<u>Définition</u>. Dans une suite de variables et terminaux, la première variable (si elle existe) est appelée la variable gauche.

<u>Définition</u>. Une dérivation où, à chaque étape, une production est appliqué à la variable gauche est dite une dérivation gauche.

<u>Exemple</u>. $S \rightarrow aSX \mid b$

$X \rightarrow Xb \mid a$

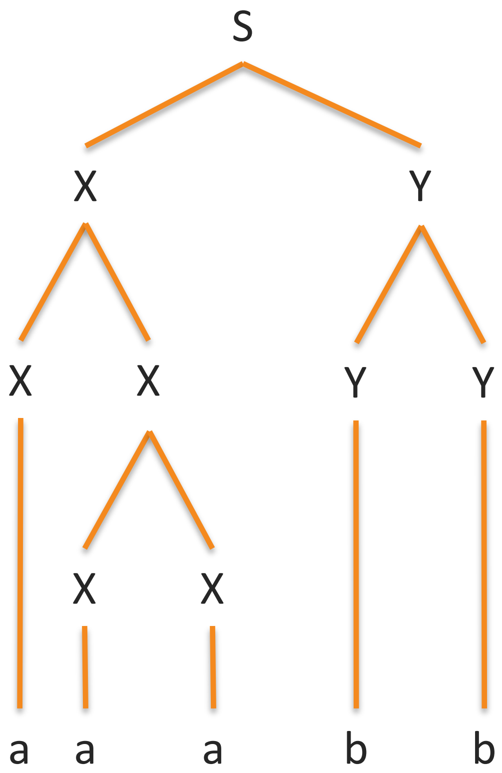
$S \Rightarrow a\underline{S}X \Rightarrow aa\underline{S}XX \Rightarrow aab\underline{X}X \Rightarrow aab\underline{X}bX \Rightarrow aabab\underline{X} \Rightarrow aababa$

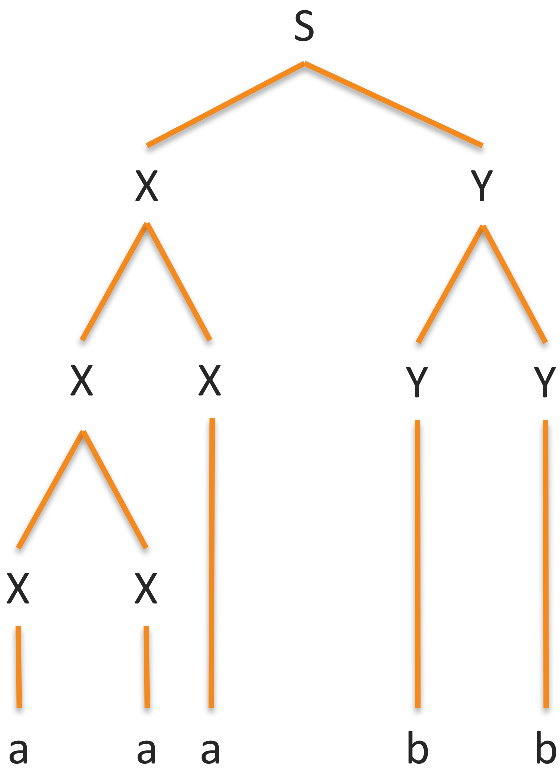
Exemple

$S \rightarrow XY$

$X \rightarrow XX \mid a$

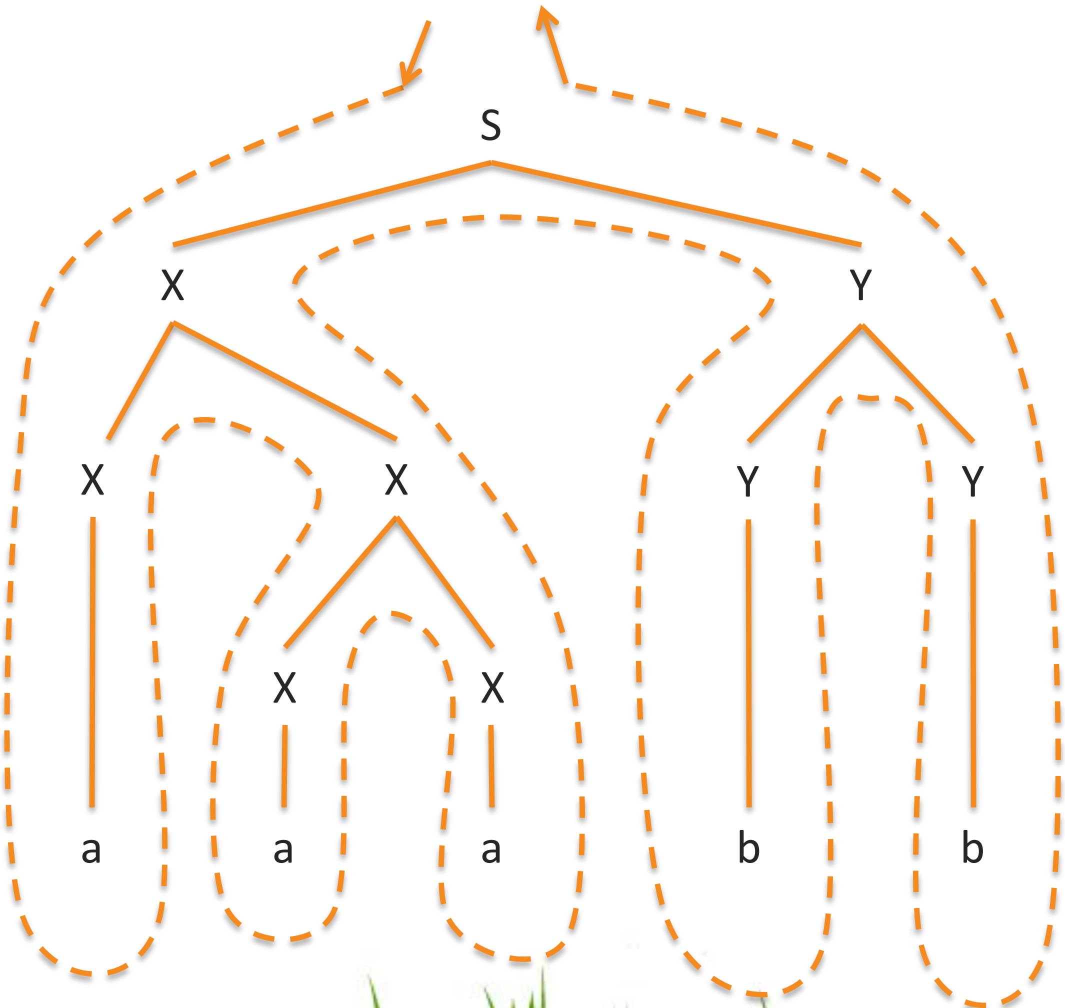
$Y \rightarrow YY \mid b$





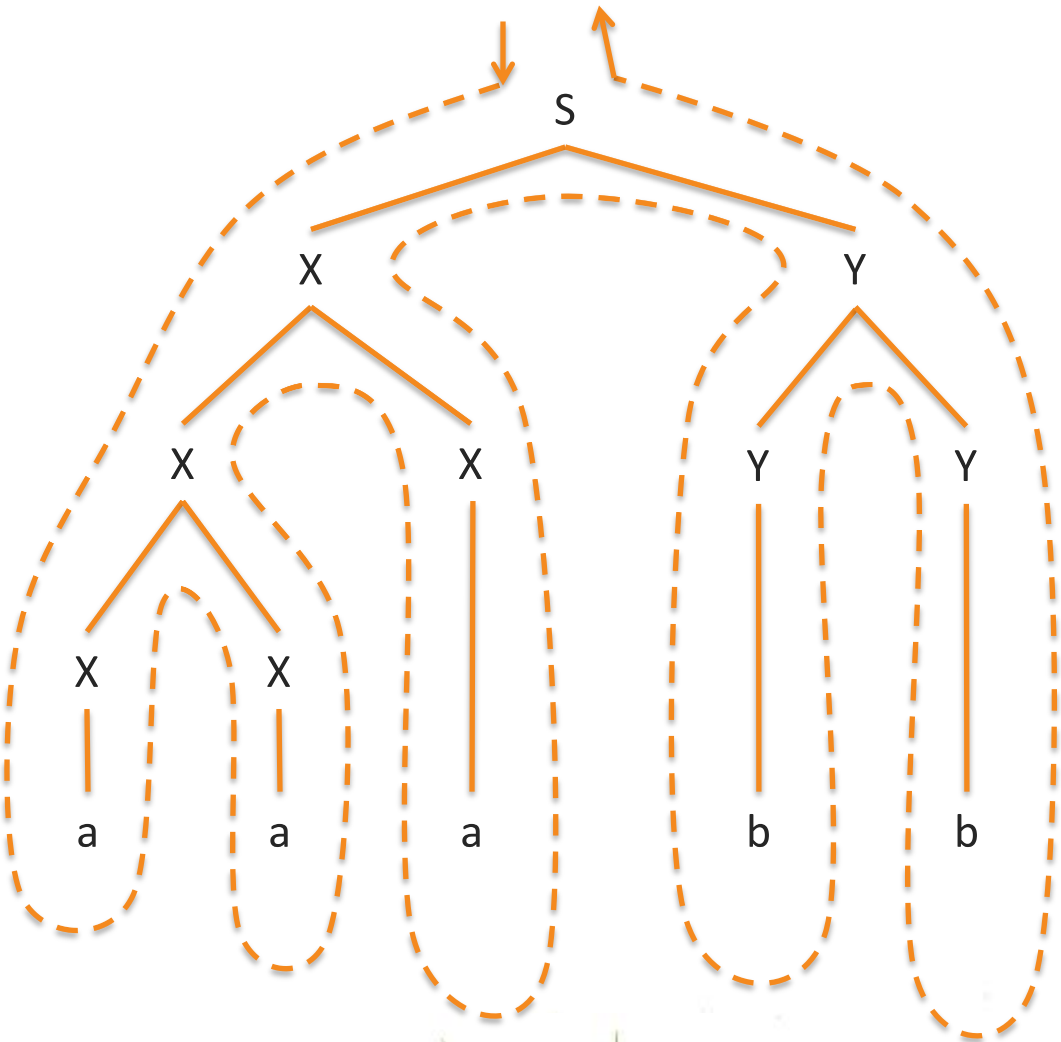
Exemple: Dérivation gauche pour aaabb

$$
\begin{aligned}
S &\Rightarrow XY \\
&\Rightarrow XXY \\
&\Rightarrow aXY \\
&\Rightarrow aXXY \\
&\Rightarrow aaXY \\
&\Rightarrow aaaY \\
&\Rightarrow aaaYY \\
&\Rightarrow aaabY \\
&\Rightarrow aaabb
\end{aligned}
$$



Exemple: Dérivation gauche pour aaabb

$$
\begin{aligned}
S &\Rightarrow XY \\
&\Rightarrow XXY \\
&\Rightarrow XXXY \\
&\Rightarrow aXXY \\
&\Rightarrow aaXY \\
&\Rightarrow aaaY \\
&\Rightarrow aaaYY \\
&\Rightarrow aaabY \\
&\Rightarrow aaabb
\end{aligned}
$$



Théorème 7

Tout mot dans un langage engendré par une grammaire non contextuelle a une dérivation gauche.

Exemple. $S \rightarrow (S) \mid S \supset S \mid {\sim}S \mid p \mid q$

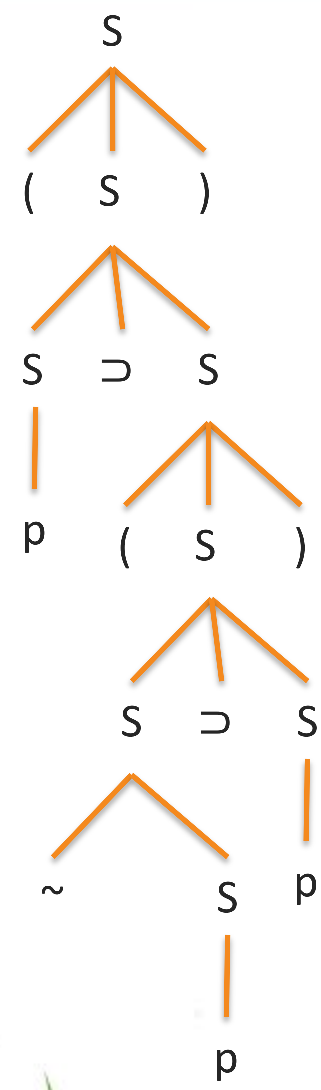
$S \Rightarrow (S) \Rightarrow (S \supset S) \Rightarrow (p \supset S)$

$\Rightarrow (p \supset (S)) \Rightarrow (p \supset (S \supset S))$

$\Rightarrow (p \supset ({\sim}S \supset S))$

$\Rightarrow (p \supset ({\sim}p \supset S))$

$\Rightarrow (p \supset ({\sim}p \supset q))$



Question?