

4- Grammaire

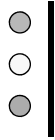
- Comment définir rigoureusement un langage ?
 - En donnant des règles (syntaxiques) de « grammaire »
 - Qui explicitent comment construire les mots sur l'alphabet
 - Définition « constructive » du langage
 - Intuitivement, il faut pour cela 4 composants
 - Un alphabet
 - Les règles de construction, dites de « production »
 - Pour exprimer ces règles, un ensemble de symboles spécifiques
 - Parmi les symboles spécifiques, un symbole particulier à partir duquel toute construction commence
- NB : On verra qu'il existe d'autres moyens pour définir les langages

17

Grammaires : définition

- Une grammaire G est un quadruplet $\langle X, N, P, S \rangle$ où :
 - X est un ensemble fini de terminaux (l'alphabet)
 - N est un ensemble fini de non terminaux (avec $N \cap X = \emptyset$)
 - P est un ensemble fini de règles de production (règles de réécriture)
 - $P = \{ \alpha \rightarrow \beta \text{ où } \begin{array}{l} \alpha \text{ est une séquence non vide de symboles de l'ensemble } N \cup X \\ \beta \text{ est une séquence non vide de symboles de l'ensemble } N \cup X \\ \text{ou } \beta = \lambda \end{array} \}$
 - La signification intuitive de $\alpha \rightarrow \beta$ est que la partie gauche α de la règle peut être remplacée par la partie droite β
 - En pratique, nous manipulerons des grammaires t.q. $\alpha \in N$
 - C.-à-d. uniquement 1 non terminal en partie gauche des règles
 - Les règles peuvent être récursives
 - $S \in N$ appelé axiome

18



Grammaires : définition

- Dérivation

- Soient

- $G = \langle N, X, P, S \rangle$
 - $w_1 \in (N \cup X)^* \setminus \{\lambda\}$
 - $w_2 \in (N \cup X)^*$

- On dit que w_1 se dérive en une étape en w_2 (ou que w_2 dérive directement de w_1), noté $w_1 \Rightarrow w_2$,
si $\exists x, y \in (N \cup X)^*$ tels que $w_1 = x u y$ et $w_2 = x v y$
et si la règle $u \rightarrow v \in P$

- On dit que w_1 se dérive (en un nombre quelconque d'étapes) en w_2 ,
noté $w_1 \Rightarrow^* w_2$,
si $w_1 = w_2$
ou si $\exists w_3 \in (N \cup X)^*$ t.q. $w_1 \Rightarrow w_3$ et $w_3 \Rightarrow^* w_2$

19



Grammaires : définition

- Langage « engendré » par une grammaire

- Soit $G = \langle N, X, P, S \rangle$
 - Le langage engendré par la grammaire G , noté $L(G)$, est l'ensemble des mots $\{w \in X^* / S \Rightarrow^* w\}$

20

● | Grammaires : définition

- Arbres de dérivation
 - Soit $G = \langle N, X, P, S \rangle$ et $w \in L(G)$
 - Un arbre de dérivation du mot w est une représentation (structurelle, non séquentielle) des dérivations de w à partir de S
 - Aux feuilles d'un arbre de dérivation, il n'y a que des terminaux ! (on dit que l'arbre est clos)
 - Rq : $L(G)$ est l'ensemble des mots de X^* qui admettent un arbre de dérivation

21

● | Grammaires : définition

- Ambiguïté
 - Une grammaire est ambiguë si $\exists w \in L(G)$ qui admet au moins deux arbres de dérivations différents
 - Rq : En pratique, les grammaires ambiguës posent problème puisqu'un mot peut être « compris » de plusieurs façons différentes
 - $2 \times 3 + 4$

22

Égalité sur les langages

- Soient
 - Une grammaire G
 - G engendre un langage noté $L(G)$
 - Un langage L
- Comment montrer que $L(G) = L$?
 - Par « double inclusion » c.-à-d.
 - $L \subseteq L(G)$
 - Par ex., par récurrence sur la longueur (ou ...) des mots
 - $L(G) \subseteq L$
 - Par ex., par récurrence sur le nombre d'utilisation d'une ou des règles

23

Hiérarchie de Chomsky

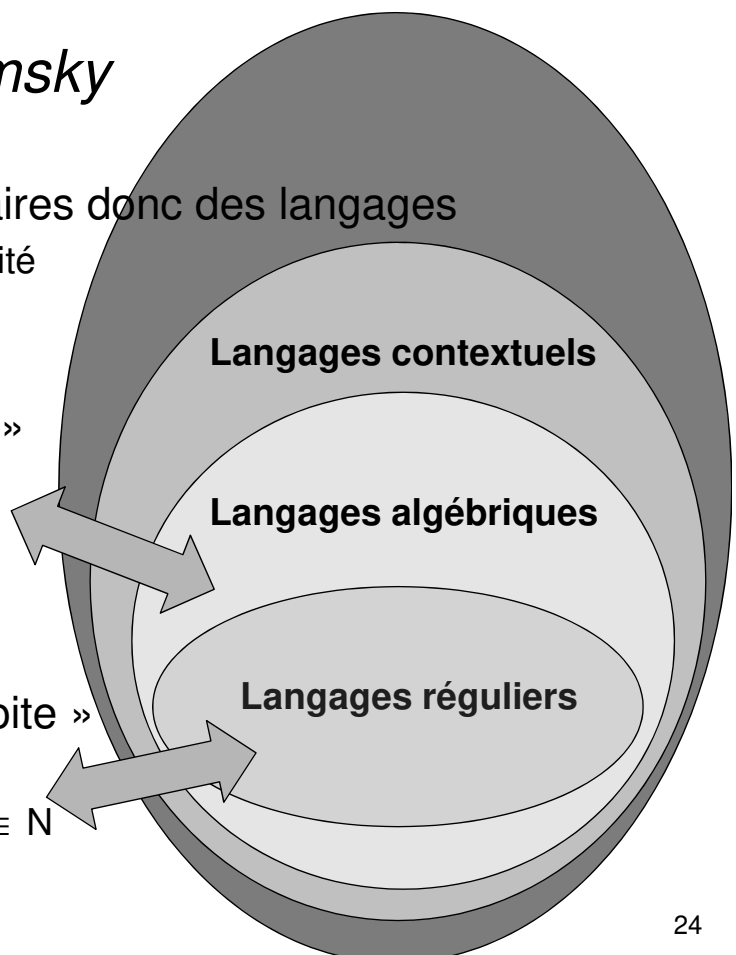
- Classification des grammaires donc des langages
 - En fonction de leur complexité

- Grammaire « context-free »

- Les règles sont de la forme
$$A \rightarrow \beta \quad \text{avec } A \in N$$
$$(\text{et } \beta \in (N \cup X)^*)$$

- Grammaire « linéaire à droite »

- Les règles sont de la forme
$$A \rightarrow a B \mid \lambda \quad \text{avec } A, B \in N$$
$$\text{et } a \in X$$



24