Langages formels

Xavier Urbain

2024-2025

Il y aura une page web...

Surveillez vos emails (rappels, informations. . .)
Vous m'écrivez ? ⇝ "LIFLF" + groupe dans le sujet politesse de base
depuis `@etu.univ-lyon1.fr`

Modalités de contrôle des connaissances : essentiellement CC

- Plus ou moins une interro "surprise" par TD
- TP noté
- ECA

SILENCE DANS L'AMPHI

Dessins Mots Phrases Sens Bin./Exéc.

Analyse lexicale

Dessins Mots Phrases Sens Bin./Exéc.

Analyse syntaxique

Dessins Mots Phrases Sens Bin./Exéc.

Analyse sémantique

Notion de langage

Langage : ensemble de mots sur alphabet fini (vocabulaire)

Definition.

Soit V vocabulaire fini (éléments : symboles ou lettres)

mot : suite finie $u = \{\alpha_i\}_{i\in[1..n]}$ (longueur $|u| = n$)

Notation plus lisible $u = \alpha_1 \cdots \alpha_n$

Pour $n = 0$: ε

Notion de langage

Composition interne : produit de concaténation $\bullet$

$\{\alpha_i\}_{i\in[1..m]} \cdot \{\beta_i\}_{i\in[1..n]} = \{\gamma_i\}_{i\in[1..m+n]}$

tel que $\begin{array}{ll} \gamma_i = \alpha_i & i \in [1..m] \\ \gamma_{m+i} = \beta_i & i \in [1..n] \end{array}$

Exemple.

Sur $V = \{\alpha, \beta\}$ $\quad \alpha\alpha\beta \cdot \beta\alpha\beta = \alpha\alpha\beta\beta\alpha\beta$

Associatif $(u \cdot v) \cdot m = u \cdot (v \cdot m)$

Neutre $u \cdot \varepsilon = \varepsilon \cdot u = u$

Puissances par récurrence : $u^0 = ?\varepsilon \qquad u^{p+1} = u \cdot u^p$

$$(\alpha_1 \cdots \alpha_n)^p = \underbrace{(\alpha_1 \cdots \alpha_n) \cdot \ldots \cdot (\alpha_1 \cdots \alpha_n)}_{p \text{ fois}}$$

Notion de langage : ensemble de mots

Opérations :

Produit de concaténation : $L_1 \circ L_2 = \{u \cdot v \mid u \in L_1, v \in L_2\}$

Puissance : $L^0 = \{\varepsilon\} \qquad L^{p+1} = L \circ L^p$

Union : $L_1 \cup L_2 = \{u \mid u \in L_1$ ou $u \in L_2\}$

Intersection : $L_1 \cap L_2 = \{u \mid u \in L_1$ et $u \in L_2\}$

Itéré : $L^\star = \cup_{p\geq 0} L^p$

Itéré strict : $L^+ = \cup_{p\geq 1} L^p$

$\leadsto V^\star$: ensemble des mots sur V

Complémentaire : $\overline{L} = \{u \mid u \in V^\star$ et $u \notin L\}$

Notion de langage :

Problème : comment décrire ?

3 écoles :

- Génération ⇝ un moyen de le générer — Grammaires
- Reconnaissance ⇝ un moyen de décider l'appartenance — Automates
- Caractérisation algébrique ⇝ notations et équations — Systèmes

Grammaires

Une phrase simple est constituée d'un sujet suivi d'un verbe suivi d'une préposition suivie d'un complément. Le sujet est un prénom ou un pronom. Les prénoms sont « Alain » ou « Béatrice » ou « Clotilde ». Les pronoms sont « Il » ou « Elle ». Les verbes sont « va » ou « vient »...

- Différencier descripteur et décrit
- Ambiguïté
- Concepts descripteurs peu nombreux

⇝ Notation spécialisée

Grammaires

Phrase	→	Sujet Verbe Préposition Complément
Sujet	→	Prénom | Pronom
Prénom	→	alain | béatrice | clotilde
	⋮	

Dans la suite : majuscule = non terminal

Grammaire : formellement

Grammaire (hors contexte) = (V_t, V_n, S, R) où $(V = V_t \cup V_n)$

- V_t ensemble fini de symboles terminaux, vocabulaire terminal
- V_n ensemble fini de symboles non terminaux $(V_n \cap V_t = \varnothing)$ vocabulaire non terminal
- $S \in V_n$ source (ou axiome) — souligné
- R ensemble fini de règles $N \to u$ où $N \in V_n, u \in V^\star$ — $(V_t \cup V_n)^\star$

$$\left\{\begin{array}{lllllll} \underline{N} & \to & 0 & & M & \to & \varepsilon \\ \underline{N} & \to & 1M & & M & \to & 0M \\ & & & & M & \to & 1M \end{array}\right\}$$

$$\begin{array}{lll} \underline{N} & \to & 0 \mid 1M \\ M & \to & \varepsilon \mid 0M \mid 1M \end{array}$$

Grammaire : formellement

Récriture par G

$G = (V_t, V_n, S, R), u \in V^\star, v \in V^\star, \quad u \to v$ si :

- $u = w_1 N w_2$ où w_1 et $w_2 \in V^\star, N \in V_n$
- $v = w_1 w w_2$ où $w \in V^\star$
- $N \to w \in R$

Remplacement de non terminaux par membres à droite de règles

$u \in V^\star, v \in V^\star, u \to^\star v$ si $\exists w_0, \cdots, w_n$ mots de $V^\star$ tels que

• $w_0 = u$ • $w_i \to w_{i+1}$ pour tout $i \in [0, n-1]$ • $w_n = v$

Langage de G, $\mathcal{L}(G) = \{$mots $\in V_t^\star$ dérivant par G de $S\}$

Grammaire : formellement, exemple

$$\left\{\begin{array}{lllllll} \underline{N} & \to & 0 & & M & \to & \varepsilon \\ N & \to & 1M & & M & \to & 0M \\ & & & & M & \to & 1M \end{array}\right\} \qquad \begin{array}{lll} \underline{N} & \to & 0 \mid 1M \\ M & \to & \varepsilon \mid 0M \mid 1M \end{array}$$

Source N

$$N \to 1M \to 11M \to 110M \to 1101M \to 1101$$

Grammaire : formellement, exemple

$$\left\{\begin{array}{lllllll} \underline{N} & \to & 11 & & N & \to & 1001 \\ N & \to & N0 & & N & \to & NN \end{array}\right\}$$

1100 ? 1101 ? 10101 ?

Grammaire : preuves

$$\left\{\begin{array}{lllllll} \underline{N} & \to & 11 & & N & \to & 1001 \\ N & \to & N0 & & N & \to & NN \end{array}\right\}$$

$[\![\_]\!] : V^\star \to \mathbb{N}$

w : écriture en binaire de $[\![w]\!]$

Pour tout $w \in \mathcal{L}(G)$, $[\![w]\!]$ est-il divisible par 3 ?