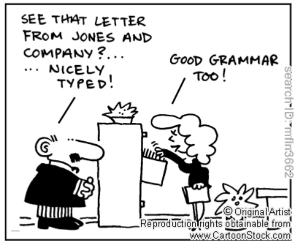
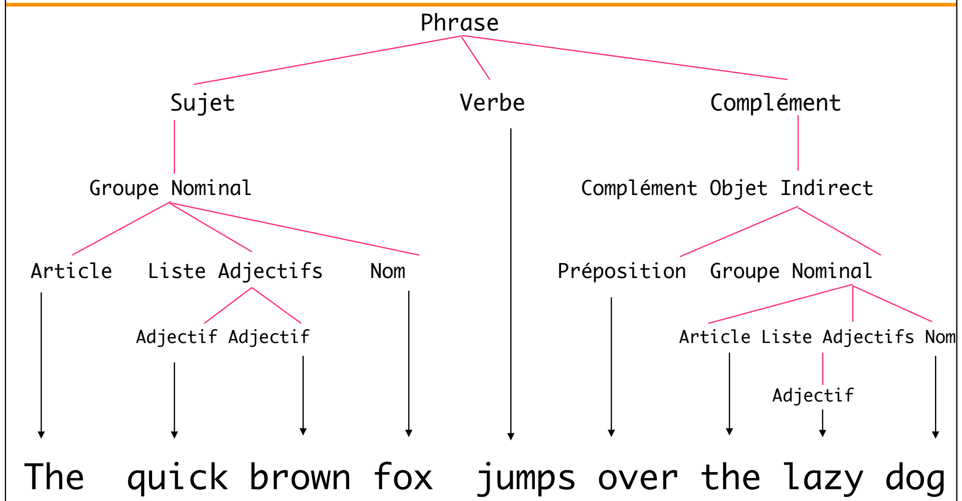
4 – Grammaires régulières



Exemple



Réécriture

- mécanisme qui permet d'engendrer des familles d'objets (figures géométriques, étiquettes de graphes, polyominos, mots, fractales ...)
- on dispose d'un objet initial : l'axiome A
- ainsi que de règles de réécriture i.e. un ensemble d'éléments du type
  $\alpha \rightarrow \beta$ (lire : *l'objet α se réécrit en celui β.*)
- elle s'appuie sur une relation de dérivation binaire transitive notée $\Rightarrow^*$ et définie par :
  - si $\alpha \rightarrow \beta$ alors $\alpha \Rightarrow \beta$
  - si $\alpha \rightarrow \beta$ et si en remplaçant dans γ l'objet α par celui β on obtient δ alors $\gamma \Rightarrow \delta$
  - $\Rightarrow^*$ est la fermeture transitive de $\Rightarrow$
- l'ensemble engendré est : $\{\varphi, A \Rightarrow^* \varphi\}$

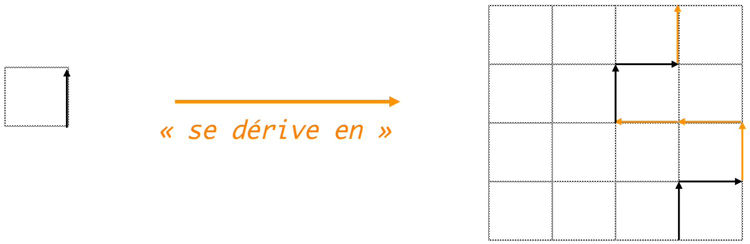
Application : image de synthèse

Exemple

Une seule règle de réécriture :
(F signifie « flèche »)

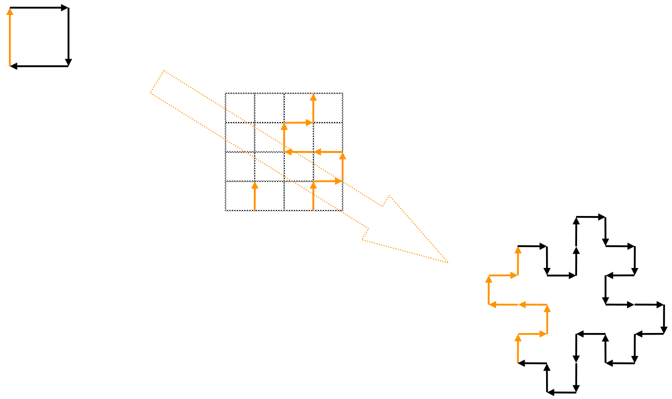
F → F + F - F - FF + F + F - F

(+ signifie que l'on tourne à droite
 - signifie que l'on tourne à gauche)

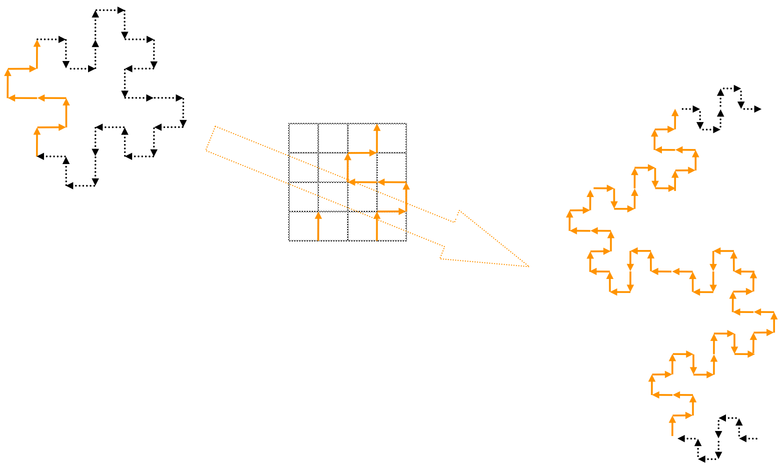


Une étape de réécriture

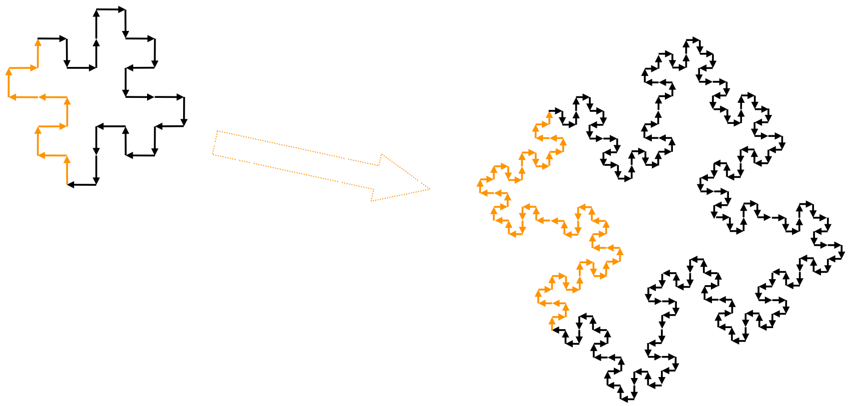
F + F + F + F



Début de l'étape 2



2e étape entière



Grammaire

- Une grammaire G est un quadruplet (N,T,P,A) où :
  - N est un ensemble fini de symboles non-terminaux
  - T est l'alphabet de symboles terminaux
  - P est un ensemble fini de productions (ou de règles)
    $\{ \varphi \rightarrow \psi \ , \ \varphi \in (N \cup T)^+, \ \psi \in (N \cup T)^* \}$
  - A est l'axiome
- Propriétés
  - le vocabulaire est $V = N \cup T$
  - $N \cap T = \varnothing$
  - (P,A) est un système de réécriture.
- Le langage engendré par la grammaire G est
  - $L(G) = \{ \varphi \in T^*, \ A \Rightarrow^* \varphi \}$
  - L(G) : ensemble des mots sur T dérivant de l'axiome A.

Exemples

- G = (N,T,P,A) avec :
  - N = {A, B} ensemble des non-terminaux
  - T = {a, b} ensemble des terminaux
  - P = { A → a A
    A → b B
    B → b B
    B → ε }
  - A est l'axiome

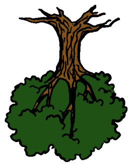
  ➥ L(G) est décrit par a^*bb^*

- G' = (N',T', P',A) avec :
  - N' = { A } , axiome A
  - T' = { 0, 1 }
  - P' = { A → 0 A 1
    A → ε }

ATTENTION
à ne pas inventer une expression régulière pour un langage peut-être pas rationnel ...

➥ $L(G') = \{0^n 1^n, n \geq 0\}$ (non rationnel).

Arbre de dérivation



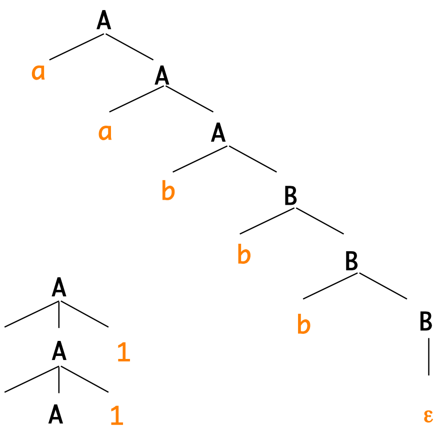
Soit G = (N,T,P,A) une grammaire avec un seul non-terminal par partie gauche des productions.

Un arbre de dérivation pour un mot w engendré par G est un arbre dont :

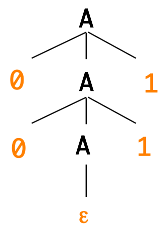
- la racine est étiquetée par l'axiome A
- les feuilles sont étiquetées par des éléments de
  $T \cup \{\varepsilon\}$
- les nœuds internes le sont par des éléments de N
- un nœud interne étiqueté B a des fils, de gauche à droite, étiquetés $\alpha_1, \alpha_2, \ldots \alpha_n$, s'il existe dans P une production :
  $B \rightarrow \alpha_1 \alpha_2 \ \ldots \ \alpha_n$
- w est formé de la concaténation des feuilles lues dans un parcours (ici indifféremment infixe, préfixe ou postfixe) de l'arbre.

Exemples

- G = (N,T,P,A) avec :
  - N = { A, B }
  - T = { a, b }
  - P ={ A → a A
    A → b B
    B → b B
    B → ε }



- G' = (N',T',P',A) avec :
  - N' = { A }
  - T' = { 0, 1 }
  - P' = { A → 0 A 1
    A → ε }



Grammaires régulières*

- Une grammaire (N,T,P,A) est régulière *à droite* si les éléments de P sont de la forme :

$$B \to a\ C \quad \text{ou} \quad B \to a \quad \text{ou} \quad B \to \varepsilon$$

avec $B \in N$, $C \in N$ et $a \in T$

- Une grammaire (N,T,P,A) est régulière *à gauche* si les éléments de P sont de la forme :

$$B \to C\ a \quad \text{ou} \quad B \to a \quad \text{ou} \quad B \to \varepsilon$$

avec $B \in N$, $C \in N$ et $a \in T$

Théorème un langage L est rationnel si et seulement si il existe une grammaire régulière qui engendre les mots de L.

* aussi dites grammaires linéaires.

Exemple & contre-exemple

- G = (N,T,P,A) est régulière à droite :
  - P = { $A \to a\ A$, $A \to b\ B$, $B \to b\ B$, $B \to \varepsilon$ }
  - L(G) est décrit par l'exp° régulière : a^*bb^*
- G' = (N',T',P',A) n'est pas une grammaire régulière :
  - P' = { $A \to 0\ A\ 1$, $A \to \varepsilon$ }
  - L(G') = $\{ 0^n1^n,\ n \geq 0 \}$

Un langage non-rationnel ne peut pas être engendré par une grammaire régulière.

A.F.D. → grammaire régulière*

Le théorème précédent se prouve par double construction.

Premier sens :

- Soit un automate fini déterministe donné
  $\mathcal{A} = (\Sigma, Q, \delta, q_0, F)$
- Construisons G = (N,T,P,A), une gram. rég. pour L($\mathcal{A}$) :
  - l'ensemble des non-terminaux N est celui des états Q de $\mathcal{A}$ **
  - l'ensemble des terminaux T égale Σ
  - l'axiome A de G correspond à l'état q_0
  - pour tout (q,σ,q') de δ il existe $q \to \sigma\ q'$ dans P **
  - pour tout f de F, il existe $f \to \varepsilon$ dans P **

* on peut aussi partir d'un A.F.N.
** à un renommage près.

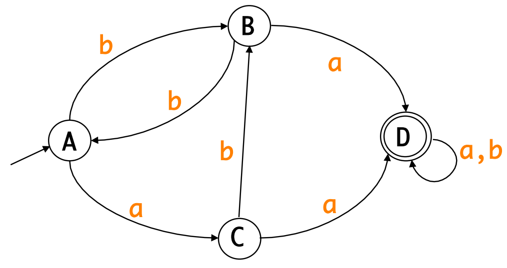
Exemple

$\mathcal{A} = (\Sigma, Q, \delta, A, F)$

Σ = {a,b}, Q = {A,B,C,D}, F = {D}

G = (N,T,P,A)

N = {A,B,C,D}
T = {a,b}
P = { $A \to a\ C$, $A \to b\ B$, $B \to a\ D$, $B \to b\ A$, $C \to a\ D$, $C \to b\ B$, $D \to a\ D$, $D \to b\ D$, $D \to \varepsilon$ }



L(G) = L($\mathcal{A}$)

Grammaire régulière ⟶ A.F.N.

Construction réciproque :

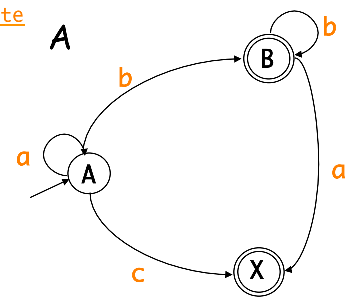
- soit $G = (N,T,P,A)$ une grammaire régulière à droite
  construisons un automate fini $\mathcal{A}$ qui accepte L(G) :
  $$\mathcal{A} = (\Sigma,\ Q,\ \delta,\ q_0,\ F)$$
- on ajoute un nouveau non-terminal X pour les productions $B \to \alpha$
  et on les remplace par $B \to \alpha\ X$ et $X \to \varepsilon$.
- les états de Q deviennent les non-terminaux $N \cup \{X\}$.*
- l'alphabet Σ devient l'ensemble T des terminaux.
- l'état q_0 correspond à l'axiome A de la grammaire
- pour toute production $B \to \alpha\ C$ de P, il existe (B,α,C) dans δ.*
- pour toute production $B \to \varepsilon$, l'état correspondant B est final dans $\mathcal{A}$.

* à un renommage près.

Exemple

$G = (N,T,P,A)$ une grammaire régulière à droite

$N = \{A,B\}$
$T = \{a,b,c\}$
$P = \{\ A \to a\ A$
$A \to b\ B$
$A \to c$ (devient $A \to c\ X$ et $X \to \varepsilon$)
$B \to b\ B$
$B \to a$ (devient $B \to a\ X$)
$B \to \varepsilon\ \}$



$\mathcal{A} = (\Sigma,\ Q,\ \delta,\ A,\ F)$

$\Sigma = \{a,b,c\}$
$Q = \{A,B,X\}$
$F = \{B,X\}$

$L(G) = L(\mathcal{A})$

$E\ :\ a^*\ b\ b^*\ (a+\varepsilon)\ +\ a^*c$