Grammaires régulières

Montaigne 2023-2024

– mpi23@arrtes.net –

Prolégomènes

Les chapitres précédents ont introduits les concepts de :

- **DFA/NFA** qui sont des modèles de machines qui acceptent les mots d'un langage régulier ;
- **expressions régulières** qui décrivent formellement des langages réguliers.

Prolégomènes

Existe-t-il un autre moyen d'engendrer tous les mots d'un langage régulier ?

L'idée est de **voir** mot se construire.

La réponse à la question est **une grammaire**.

Grammaires

Comme nous l'avons vu, une **grammaire** $\mathcal{G}$ est un quadruplet $(\mathcal{V}_N, \mathcal{V}_T, \mathcal{R}, \mathcal{S})$ où :

- $\mathcal{V}_N$ est l'ensemble des **symboles non terminaux** ou **variables** ;
- $\mathcal{V}_T$ est l'ensemble des **symboles terminaux** ;
- $\mathcal{R}$ est l'ensemble des **règles de production** ;
- $\mathcal{S}$ est **le symbole initial**.

Une grammaire permet de construire tous les **mots** d'un langages, c'est-à-dire des **suites de symboles terminaux**.

Grammaires

De manière très générale, en posant $\mathcal{V} = \mathcal{V}_T \cup \mathcal{V}_N$, une règle définit comment un élément α de $\mathcal{V}^*$ peut être transformé en un autre élément β de $\mathcal{V}^*$.

$$\alpha \to \beta$$

Se donner l'ensemble des règles suffit pour caractériser une grammaire, comme l'illustre l'exemple suivant.

$$\begin{cases} S \to a \mid Sb \mid A \\ A \to b \mid SA \end{cases}$$

Ici, $\mathcal{V}_N = \{S, A\}$ et $\mathcal{V}_T = \{a, b\}$. S est le symbole initial.

Partant de S, il est possible de dériver le mot *abb*.

$$S \to Sb \to Ab \to SAb \to Sbb \to abb$$

On note : $S \Rightarrow^* abb$.

Grammaires régulières

Une **grammaire régulière** $\mathcal{G}_{reg}$ est une grammaire dans laquelle les règles ne peuvent être de de quatre types.

$$\begin{cases} A \to \varepsilon & A \in \mathcal{V}_N \\ A \to a & A \in \mathcal{V}_N, a \in \mathcal{V}_T \\ A \to B & A, B \in \mathcal{V}_N \\ A \to aB & A, B \in \mathcal{V}_N, a \in \mathcal{V}_T \end{cases}$$

Cette grammaire régulière est dite **à droite**. Dans une **grammaire régulière à gauche**, la dernière règle est de la forme $A \to Ba$.

Grammaires régulières

Les grammaires régulières reconnaissent les langages réguliers.

Pour établir ce résultat, il convient de montrer qu'à partir d'une grammaire régulière, il est possible de construire un objet qui représente un langage régulier, à savoir un DFA/NFA ou une expression régulière.

D'une grammaire régulière vers un NFA

Soit $\mathcal{G}_{reg} = (\mathcal{V}_N, \mathcal{V}_T, \mathcal{R}, S)$ une grammaire régulière. Construisons un NFA $\mathcal{A} = (Q, \Sigma, q_0, F, \delta)$.

À une règle de la forme $A \rightarrow aB$, **on associe** deux états A et B et une transition étiquetée par a de A vers B.

À une règle de la forme $A \rightarrow B$, **on associe** une ε-transition entre deux états A et B.

À une règle $A \rightarrow \varepsilon$, **on associe** une transition d'un état A vers un état acceptant q_f étiquetée par ε.

À une règle $A \rightarrow a$, **on associe** une transition d'un état A vers le même état acceptant q_f étiquetée par a.

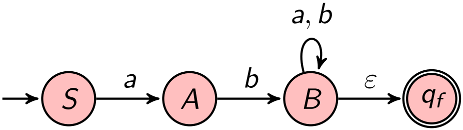
Ainsi, on obtient :

$$q_0 = S \qquad F = \{q_f\} \qquad \Sigma = \mathcal{V}_T \qquad Q = \mathcal{V}_N \cup \{q_f\}$$

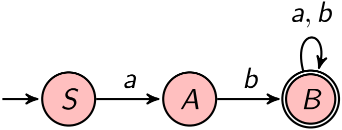
D'une grammaire régulière vers un NFA

Construction d'un NFA à partir de la grammaire suivante.

$$S \to aA \qquad A \to bB \qquad B \to aB \mid bB \mid \varepsilon$$

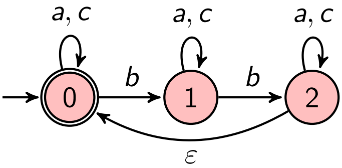


Il est possible de construire directement un DFA.



D'un NFA/DFA vers une grammaire régulière

Considérons l'automate suivant.



On peut lui associer la grammaire régulière suivante.

- Terminaux $\mathcal{V}_T = \{a, b, c, \}$
- Non terminaux $\mathcal{V}_N = \{S, A, B\}$ avec $S = 0$, $A = 1$, $B = 2$.
- Règles

$$\begin{cases} S \rightarrow aS \mid cS \mid bA \mid \varepsilon \\ A \rightarrow aA \mid cA \mid bB \\ B \rightarrow aB \mid cB \mid S \end{cases}$$

S'entraîner

Construire un NFA associé à la grammaire régulière suivante. En déduire une expression régulière qui dénote le langage reconnu par le NFA.

$$\begin{cases} S \to aS \\ A \to \varepsilon \mid bS \mid bA \mid aB \\ B \to B \mid aS \end{cases}$$