LYCÉE LECONTE DE LISLE

# Langages et monoïdes

Vincent Picard

# Monoïde

Un **monoïde** est un ensemble muni d'une loi de composition interne $(M, \#)$ qui vérifie :

1. # est **associative** :

$$\forall x, y, z \in M, x \# e = e \# x = x$$

2. Il existe un **élément neutre** $e \in M$ vérifiant :

$$\forall x \in M, x \# e = e \# x = x$$

- Un groupe $(G, \cdot)$ est un monoïde, en particulier : $(\mathbb{Z}, +)$, $(\mathbb{K}, +)$, $(\mathbb{K}^*, \times)$, $(\mathscr{L}_n(\mathbb{K}), \circ)$, $(\mathfrak{S}_n, \circ)$ ...sont des monoïdes.
- $(\mathbb{N}, +)$ est un monoïde mais n'est pas un groupe...
- $(\Sigma^*, .)$ est le monoïde des mots sur $\Sigma$ mais n'est pas un groupe...

# Morphisme de monoïdes

Soit $(M, \#)$ et $(N, \star)$ deux monoïdes d'éléments neutre $e_M$ et $e_N$ respectivement. Un **morphisme de monoïdes** est une application $\varphi : (M, \#) \to (N, \star)$ qui vérifie :

1. $\forall x, y \in M, \quad \varphi(x \# y) = \varphi(x) \star \varphi(y)$
2. $\varphi(e_M) = e_n$

- L'application module :

$$|\cdot| : (\mathbb{C}, \times) \to (\mathbb{R}^+, \times)$$

  est un morphisme de monoïdes.

- Si $E$ est un $\mathbb{K}$-espace vectoriel de dimension finie alors :

$$\det : (\mathcal{L}(E), \circ) \to (\mathbb{K}, \times)$$

  est un morphisme de monoïdes.

# Exercices

1. Justifier que l'application conjuguée $z \mapsto \bar{z}$ sur les nombres complexes est un morphisme de monoïdes.

2. Montrer que l'application qui à tout mot $u$ associe $|u|$ est un morphisme de monoïdes.

3. Soit $a \in \Sigma$ une lettre. Montrer que l'application qui à tout mot $u$ sur $\Sigma$ associe $|u|_a$ est un morphisme de monoïdes.

4. Montrer que la composée de deux morphismes de monoïdes est un morphisme de monoïdes.

5. Soit $\varphi$ un morphisme de monoïdes défini sur $(\Sigma^*, .)$, montrer que $\varphi$ est uniquement déterminée par sa restriction à $\Sigma$.

Ainsi, lorsque nous aurons à définir un morphisme sur $(\Sigma^*, .)$, il suffira de donner sa valeur $\varphi(x)$ pour tout $x \in \Sigma$

# Reconnaissance par monoïde

Un langage $L$ sur l'alphabet $\Sigma$ est **reconnu par monoïde** s'il existe un monoïde $(M, \#)$, une partie $P \subset M$ et un morphisme de monoïdes $\varphi : (\Sigma^*, .) \to (M, \#)$ telle que :

$$L = \varphi^{-1}(P)$$

- Sur $\Sigma = \{a, b\}$, le langage

$$L = \{u \in \Sigma^* \ / \ |u|_a = |u|_b\}$$

des mots contenant autant de $a$ que de $b$ est reconnu par :

$$\varphi : u \mapsto |u|_a - |u|_b$$

avec $(M, \#) = (\mathbb{Z}, +)$ et $P = \{0\}$.

# Exercices

Donner des exemples de morphismes de monoïdes reconnaissant les langages suivants :

1. Mots contenant plus de $a$ que de $b$.
2. Mots contenant deux fois plus de $a$ que de $b$.
3. Mots de longueur impaire.
4. Mots $u$ sur $\Sigma = \{a, b, c\}$ tels que $|u|_c = |u|_a + |u|_b$
5. Mots de longueur paire contenant autant de $a$ que de $b$.

Petite remarque : tout langage est reconnu par le morphisme identité bien sur...

# Langages réguliers

On a la caractérisation algébrique suivante des langages réguliers :

Un langage $L$ est **régulier** si et seulement s'il existe un monoïde **fini** $(M, \#)$ qui le reconnaît, c'est-à-dire tel qu'il existe un morphisme $\varphi$ et une partie $P \subset M$ telle que $\varphi^{-1}(P) = L$.

- Le langage des mots de longueur impaire :

$$\begin{aligned} \varphi : (\Sigma^*, .) &\to (\mathbb{Z}/2\mathbb{Z}, +) \\ u &\mapsto \mathrm{Cl}(|u|) \end{aligned}$$

  avec $P = \{1\}$.

- Le langage sur $\Sigma = \{a, b, c\}$ des mots qui ne contiennent pas de $c$.

$$\begin{aligned} \varphi : (\Sigma^*, .) &\to (\{0, 1\}, \times) \\ u &\mapsto 0 \text{ si } c \in u, \quad 1 \text{ sinon.} \end{aligned}$$

  avec $P = \{1\}$.

# Si un langage est reconnu par monoïde fini alors il est régulier

- Soit un langage reconnu par monoïde fini $M$ : $L = \varphi^{-1}(P)$
- **Idée :** concevoir un automate qui calcule $\varphi(u)$.

Soit $L = \varphi^{-1}(P)$ un langage reconnu par monoïde fini $(M, \#)$, on lui associe l'automate fini déterministe complet $A = (Q, q_0, F, \delta)$ suivant :

- $Q = M$
- $q_0 = e_M$
- $F = P$
- Pour tout état $q \in Q$ et toute lettre $a \in \Sigma$, $\delta(q, a) = q \# \varphi(a)$

# Du monoïde vers l'automate : exemple

| # | $x$ | $y$ | $z$ |
|---|---|---|---|
| $x$ | $x$ | $y$ | $z$ |
| $y$ | $y$ | $z$ | $x$ |
| $z$ | $z$ | $x$ | $y$ |

$$\varphi : (\Sigma^*, .) \to (M, \#)$$

$$\varepsilon \mapsto x$$

$$a \mapsto y$$

$$b \mapsto x$$

- Construire l'automate associé à ce monoïde fini et au morphisme $\varphi$ avec $P = \{y, z\}$. Quel langage reconnaît-il ?

# Du monoïde vers l'automate : exemple

| # | $x$ | $y$ | $z$ |
|---|---|---|---|
| $x$ | $x$ | $y$ | $z$ |
| $y$ | $y$ | $z$ | $x$ |
| $z$ | $z$ | $x$ | $y$ |

$$\varphi : (\Sigma^*, .) \to (M, \#)$$

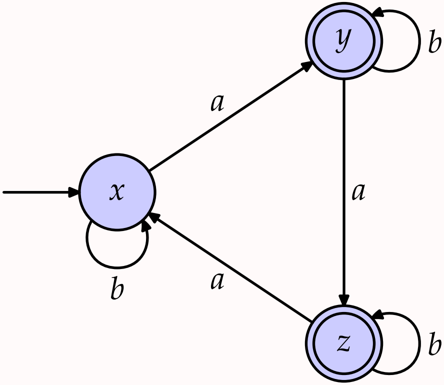
$$\varepsilon \mapsto x$$

$$a \mapsto y$$

$$b \mapsto x$$

- Construire l'automate associé à ce monoïde fini et au morphisme $\varphi$ avec $P = \{y, z\}$. Quel langage reconnaît-il ?

- Solution :




- $L = \{u \in \{a, b\}^* \ / \ |u|_a \not\equiv 0 \ [3]\}$

# Du monoïde vers l'automate : preuve

- Montrons que pour tout mot $u \in \Sigma^*$, $\delta^*(q_0, u) = \varphi(u)$. On procède par récurrence sur la longueur du mot $u$.
- **Initialisation :** Si $u = \varepsilon$, on a d'une part $\delta^*(q_0, \varepsilon) = q_0 = e_M = \varphi(\varepsilon)$ car $\varphi$ est un morphisme de monoïdes.
- **Hérédité :** Soit $u$ un mot de longueur $n+1$ que l'on peut décomposer en $u = v.a$ avec $v$ un mot de longueur $n$ et $a$ une lettre. Alors

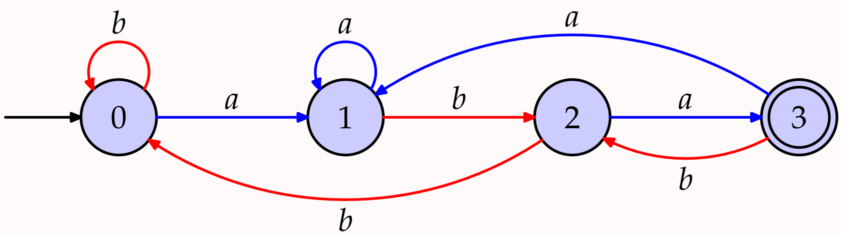
$$\begin{aligned}
\delta^*(q_0, u) &= \delta^*(q_0, v.a) \\
&= \delta(\delta^*(q_0, v), a) \\
&= \delta(\varphi(v), a) \\
&= \varphi(v) \# \varphi(a) \\
&= \varphi(v.a) \\
&= \varphi(u)
\end{aligned}$$

- **Conclusion :** $u \in \mathcal{L}(A) \Leftrightarrow \delta^*(q_0, u) \in F \Leftrightarrow \varphi(u) \in P \Leftrightarrow u \in \varphi^{-1}(P) \Leftrightarrow u \in L$

# De l'automate vers le monoïde

- On s'attaque maintenant à la réciproque : soit $L$ un langage reconnaissable par automate fini $A = (Q, q_0, F, \delta)$, montrons qu'il est reconnaissable par monoïde fini.

- Exemple : mots finissant par *aba*




- Soit $u \in \Sigma^*$ on définit l'application :

$$\begin{aligned} \tau_u : Q &\to Q \\ q &\mapsto \delta^*(q, u) \end{aligned}$$

C'est l'application qui à tout état calcule son état d'arrivée par la lecture de $u$.

- Exemples :

$$\begin{aligned} \tau_a(0) &= 1 & \qquad \tau_b(0) &= 0 \\ \tau_a(1) &= 1 & \qquad \tau_b(1) &= 2 \\ \tau_a(2) &= 3 & \qquad \tau_b(1) &= 0 \\ \tau_a(3) &= 1 & \qquad \tau_b(1) &= 2 \end{aligned}$$

$\tau_\varepsilon = \mathrm{Id}_Q$

# Monoïde des transitions

- On remarque que $(Q^Q, \#)$, l'ensemble des applications $Q \to Q$ muni de la loi :

$$f \# g = g \circ f$$

  est un **monoïde fini**.

- De plus, on remarque que :

$$\varphi : (\Sigma^*, .) \to (Q^Q, \#)$$

$$u \mapsto \tau_u$$

  est un morphisme de monoïdes.

- En effet, remarquons d'abord que pour tout couple de mots $u, v$ on a

$$\forall q \in Q, \quad (\tau_u \# \tau_v)(q) = (\tau_v \circ \tau_u)(q) = \tau_v(\tau_u(q)) =$$

$$\tau_v(\delta^*(q, u)) = \delta^*(\delta^*(q, u), v) = \delta^*(q, uv) = \tau_{uv}(q)$$

- et donc $\varphi(u.v) = \tau_{uv} = \tau_u \# \tau_v = \varphi(u) \# \varphi(v)$.

- De plus, on a bien : $\varphi(\varepsilon) = \tau_\varepsilon = \mathrm{Id}_Q$

# Bien choisir la partie $P$

- Nous avons maintenant :
  - un monoïde fini $(Q^Q, \#)$,
  - Un morphisme de monoïde $\varphi : (\Sigma^*, .) \to (Q^Q, \#)$,
- Reste à choisir correctement la partie $P$, afin d'avoir $\mathcal{L}(A) = \varphi^{-1}(P)$.
- On pose $P = \{f \in Q^Q \,/\, f(q_0) \in F\}$.
- Alors on a bien, pour tout mot $u$:

$$u \in \mathcal{L}(A) \Leftrightarrow \delta^*(q_0, u) \in F \Leftrightarrow \tau_u(q_0) \in F \Leftrightarrow \varphi(u) \in P \Leftrightarrow u \in \varphi^{-1}(P)$$

# Application : propriétés de fermeture de $\mathrm{RAT}(\Sigma)$

- En utilisant la reconnaissance par monoïde, montrer que :

  1. Le complémentaire d'un langage régulier est un langage régulier.
  2. L'intersection de deux langages régulier est un langage régulier.
  3. L'union de deux langages réguliers est un langage régulier.
  4. Le miroir $\tilde{L}$ d'un langage $L$ régulier est un langage régulier.

- Remarque : on peut aussi montrer la stabilité pour la concaténation, mais c'est bien plus difficile...

# Application : racine carrée d'un langage régulier

- Soit $L$ un langage sur $\Sigma$, on définit :

$$\sqrt{L} = \{u \in \Sigma^* \ / \ u^2 \in L\}$$

1. Soit $L_1$ le langage des mots sur $\{a, b\}$ contenant le facteur $bb$, déterminer $\sqrt{L_1}$
2. Comparer $L$ et $\sqrt{L^2}$ pour un langage $L$ quelconque
3. Comparer $L$ et $\sqrt{L}^2$ pour un langage $L$ quelconque
4. Montrer que si $L$ est régulier alors $\sqrt{L}$ est régulier

- Remarque : ce résultat se généralise à la racine $n$-ième d'un langage.