Langages, Compilation, Automates. Partie 8: Langages non hors-contexte, Automates à pile déterministes.

Florian Bridoux

Polytech Nice Sophia

2022-2023

Table des matières

Table des matières

Lemme de l'étoile

Lemme de l'étoile

Si L est régulier, alors il existe un nombre n tel que pour tout mot w de L, si $|w| \geq n$, alors w peut être factorisé en $w = xyz$ de telle sorte que

- $1 \leq |y| \leq |xy| \leq n$.
- $\forall t \geq 0,\ xy^t z \in L$.

Idée de preuve (très simplifiée):

- Soit A l'AFD minimum qui reconnaît L et n son nombre d'états.
- Prenons $w \in L$. On a: $q_0 \xrightarrow{w} q_f$.
- Si $|w| \geq n$, alors $q_0 \xrightarrow{x} q_i \xrightarrow{y} q_i \xrightarrow{x} q_f$.
- Donc pour tout $t \geq 0$, $q_0 \xrightarrow{x} q_i \xrightarrow{y^t} q_i \xrightarrow{z} q_f$.

Langages non réguliers: Exemple d'utilisation

Exemple d'utilisation: montrons que $L = \{a^t b^t \mid t \in \mathbb{N}\}$ n'est pas régulier.
Par l'absurde, supposons que L est régulier. Donc il existe un n tel que tout mot $w \in L$, w puisse être factorisé en $w = xyz$ de telle sorte que

- $1 \leq |y| \leq |xy| \leq n$.
- $\forall t \geq 0$, $xy^t z \in L$.

Prenons $w = a^n b^n$.

- Donc, $xy \in \{a\}^*$.
- Donc, $y = a^\alpha$ avec $\alpha \geq 1$.
- Donc, $xy^0 z = xz = a^{n-\alpha} b^n \notin L$.
- Absurde, donc L n'est pas régulier.

Table des matières

Lemme d'itération pour les langages hors-contexte

Lemme d'itération pour les langages hors-contexte

Soit L un langage hors-contexte. Il existe $n \geq 1$ tel que tout mot w de L de longueur $|w| \geq n$ possède une factorisation $w = xuyvz$ telle que :

- $1 \leq |uv|$,
- $|uyv| \leq n$,
- $xu^t yv^t z \in L$ pour tout entier $n \geq 0$.

Il permet de prouver que certains langages ne sont pas hors-contexte.

$L = \{a^t b^t c^t \mid t \in \mathbb{N}\}$

Corollaire

$L = \{a^t b^t c^t \mid t \in \mathbb{N}\}$ n'est pas hors-contexte.

- Supposons que $L = \{a^t b^t c^t \mid t \in \mathbb{N}\}$ soit hors-contexte.
- Donc, il existe $n \geq 1$ tel que tout $w \in L$ avec $|w| \geq n$ possède une factorisation $w = xuyvz$ avec $1 \leq |uv|$, $|uyv| \leq n$ et $xu^t yv^t z \in L$ pour tout entier $t \geq 0$.
- Prenons $w = a^n b^n c^n$ et sa factorisation $w = xuyvz$.
- Comme $xu^2yv^2z \in L$, u et v ne peuvent qu'avoir un seul type de lettre chacun (car les a doivent précéder les b et les b les c).
- Donc, uv ne contient qu'une ou deux lettres différentes.
- Donc $xu^2yv^2z \in L$ n'a pas le même nombre de a, b et c. Absurde, donc L est non hors-contexte.

$L = \{ss \mid s \in \{a, b\}^*\}$

Corollaire

$L = \{ss \mid s \in \{a, b\}^*\}$ n'est pas hors-contexte.

- On prend $s = a^n ba^n b$ et $w = ss = a^n ba^n ba^n ba^n b \in L$.
- Soit $xuyvz = w$ avec $1 < |uv|, |uyv| \leq n$, $xu^t yv^t z \in L$.
- Comme $|uyv| < n$, uv ne contient pas plus d'un b.
- Si u (resp. v) contient un b, alors xu^2yv^2z en a 5: absurde.
- Donc $u = a^\alpha$, $v = a^\beta$.
- Comme $|uyv| < n$, u et v sont dans deux groupes de a adjacent (ou le même groupe de a).
- (on ne traite qu'un des 7 cas) $xu^2yv^2z = a^n ba^{n+\alpha} b$ $a^{n+\beta} ba^n b \notin L$

Exercice

Montrer que $L = \{a^{t^2} \mid t \in \mathbb{N}\}$ n'est pas hors-contexte.

Intersection de deux langages hors-contexte

Théorème Intersection

L'intersection de deux langages hors-contexte n'est pas nécessairement hors-contexte.

Exemple:

- $L_1 = \{a^n b^n c^m \mid n \in \mathbb{N}, m \in \mathbb{N}\}$.
- $L_2 = \{a^m b^n c^n \mid n \in \mathbb{N}, m \in \mathbb{N}\}$.
- $L_1 \cap L_2 = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ qui n'est pas hors-contexte.
- Mais L_1 et L_2 peuvent être engendré par la grammaire hors-contexte:

$$\begin{array}{lll} S & \to & EF \\ E & \to & aEb \mid \epsilon \\ D & \to & cD \mid \epsilon \end{array} \quad \text{et} \quad \begin{array}{lll} S & \to & EF \\ E & \to & aE \mid \epsilon \\ D & \to & bDc \mid \epsilon \end{array}$$

Union et complément de deux langages hors-contexte

Exercice

Est-ce que l'union de deux langages hors-contexte est hors-contexte? Prouvez-le.

Exercice

Même question pour le complément d'un langage hors-contexte.

Intersection de deux langages hors-contexte

Théorème Intersection

L'intersection d'un langage hors-contexte et d'un langage régulier est hors-contexte.

Exercice

Montrer que $L = \{w \in \{a, b, c\}^* \mid |w|_a = |w|_b = |w|_c\}$ n'est pas hors-contexte.

Intersection de deux langages hors-contexte

Théorème Intersection

L'intersection d'un langage hors-contexte et d'un langage régulier est hors-contexte.

Exercice

Montrer que $L = \{w \in \{a, b, c\}^* \mid |w|_a = |w|_b = |w|_c\}$ n'est pas hors-contexte.

- $N = a^*b^*c^*$. N est régulier.
- Soit $M = N \cap L$.
- Si L est hors-contexte, M aussi.
- Mais $M = \{a^n b^n c^n \mid n\mathbb{N}\}$ qui n'est pas hors-contexte.
- Absurde, donc L n'est pas hors-contexte.

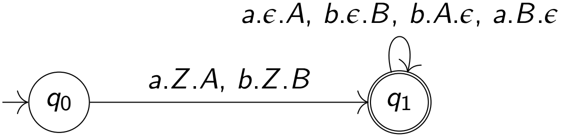
$L = \{w \in \{a, b, c\}^* \mid |w|_a = |w|_b = |w|_c\}$

Exemple de grammaire pour engendrer
$L = \{w \in \{a, b, c\}^* \mid |w|_a = |w|_b = |w|_c\}$:

$$
\begin{array}{lcl}
S & \rightarrow & ABCS \mid \epsilon \\
AB & \rightarrow & BA \\
AC & \rightarrow & CA \\
BA & \rightarrow & AB \\
BC & \rightarrow & CB \\
CA & \rightarrow & AC \\
CB & \rightarrow & BC \\
A & \rightarrow & a \\
B & \rightarrow & b \\
C & \rightarrow & c
\end{array}
$$

Table des matières

Non-déterminisme



Arbre des transitions possibles pour *aabb*:

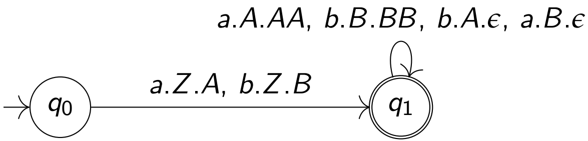
$$(q_0, Z) \xrightarrow{a.Z.A} (q_1, A) \xrightarrow{a.\epsilon.A} (q_1, AA)$$

$$(q_1, AA) \xrightarrow{b.\epsilon.B} (q_1, BAA) \xrightarrow{b.\epsilon.B} (q_1, BBAA)$$

$$(q_1, AA) \xrightarrow{b.A.\epsilon} (q_1, A) \xrightarrow{b.\epsilon.B} (q_1, BA)$$

$$(q_1, A) \xrightarrow{b.A.\epsilon} (q_1, \epsilon)$$

Déterminisme



Arbre des transitions possibles pour *aabb*:

$$(q_0, Z) \xrightarrow{a.Z.A} (q_1, A) \xrightarrow{a.\epsilon.A} (q_1, AA) \xrightarrow{b.A.\epsilon} (q_1, A) \xrightarrow{\quad b.A.\epsilon \quad} (q_1, \epsilon)$$

On dit que le langage $L = \{a^n b^n\}$ peut être reconnu par un automate à pile **déterministe**.

Automates à pile déterministe

Informellement, un automate à pile est déterministe dans chaque configuration de l'automate, il n'y a au plus une transition possible.

Définition (Automates à pile déterministe)

Un automate à pile est **déterministe** si pour tout $q_i \in Q$, $a \in \Sigma$, $\gamma \in \Gamma$, $|\delta(q_i, a, \gamma) \cup \delta(q_i, \epsilon, \gamma) \cup \delta(q_i, a, \epsilon) \cup \delta(q_i, \epsilon, \epsilon)| \leq 1$.

Automates à pile déterministe

Contrairement aux AFI qui sont aussi puissant que les AFD (ils reconnaissent les langages réguliers), les AP déterministes sont strictement moins puissant que les AP indéterministes.

Théorème (Palindromes et déterminisme)

Le langage hors-contexte L des palindromes sur l'alphabet $\{a, b\}$ n'est pas reconnu par un AP déterministe.

Idée de preuve:

- Disons qu'un AP déterministe A, reconnaît L.
- Donc, quand il lit un mot palindrome il doit finir dans un état acceptant avec une pile vide.
- A a un nombre fini d'état acceptant. On peut trouver deux palindromes différents u et v que A accepte en finissant dans le même état acceptant.
- Donc A va accepter uu et vu ou refuser uu et vu. Absurde.

Automates à pile déterministe

On peut montrer que palindrome est un langage hors-contexte grâce à cette grammaire:

$$S \quad \rightarrow \quad aSa \mid bSb \mid b \mid a \mid \epsilon$$

On remarque que cette grammaire n'est pas ambigüe. Donc, grammaire non ambigüe $\neq$ automate à pile déterministe.

Automates à pile déterministe