

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année



## INF 302 : LANGAGES & AUTOMATES

### Chapitre 10 : Langages Non Réguliers et Lemmes de l'itération

Yliès Falcone

yliès.falcone@univ-grenoble-alpes.fr — www.ylies.fr

Univ. Grenoble-Alpes, Inria  
Laboratoire d'Informatique de Grenoble - [www.liglab.fr](http://www.liglab.fr)  
Équipe de recherche LIG-Inria, CORSE - [team.inria.fr/corse/](http://team.inria.fr/corse/)

Année Académique 2021 - 2022

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    1 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- 1 Motivations
- 2 Lemme de l'itération
- 3 Constante d'itération
- 4 Une condition suffisante pour la non-régularité
- 5 Utilisation des propriétés de fermeture
- 6 Résumé

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    2 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- 1 Motivations
- 2 Lemme de l'itération
- 3 Constante d'itération
- 4 Une condition suffisante pour la non-régularité
- 5 Utilisation des propriétés de fermeture
- 6 Résumé

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    3 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Deux exemples

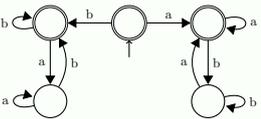
**Exemple (Le langage  $\{a^n b^n \mid n \in \mathbb{N}\}$  n'est pas d'état finis)**

- Si nous essayons de trouver un AEFD pour ce langage, nous observons que cet automate doit se souvenir du nombre de  $a$  qui ont été lus en entrée.
- Comme le nombre de  $a$  n'est pas limité, il y a un nombre infini de possibilités.
- Cela ne peut être fait avec un nombre fini d'états.

Le "nombre infini de possibilités" n'empêche pas un langage d'être d'états finis.

**Exemple (Langage d'états finis)**

L'ensemble des mots sur  $\{a, b\}$  qui contiennent autant de facteurs  $ab$  que de facteurs  $ba$ .



Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    4 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Motivations

Montrer la non-régularité

**Comment montrer qu'un langage  $L$  n'est pas régulier / à états ?**

**Raisonnement intuitif**

On va utiliser une propriété des langages réguliers.

- Tout langage régulier  $L$  satisfait la propriété de l'itération  $P$ .
- Un langage qui ne satisfait pas  $P$  n'est pas régulier.

**Utilisation : démontrer qu'un langage est non régulier - Preuve par l'absurde**

- Supposer que  $L$  est régulier.
- Utilisation de  $P$  pour déduire des informations sur les mots du langage étudié.
- Trouver une contradiction.

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    5 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Motivations

Comprendre la régularité

**Certains langages ne peuvent être reconnus / décrits par des automates**

- par exemple :  $\{a^n \cdot b^n \mid n \in \mathbb{N}\}$ .
- Les automates ne peuvent pas compter (de manière arbitraire)

**La plupart des langages sont non réguliers**

- Les langages à états / automates sont dénombrables.
- Les langages sont indénombrables.

**Comment caractériser précisément les langages qui peuvent être reconnus par des automates ?**

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    6 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- 1 Motivations
- 2 Lemme de l'itération
- 3 Constante d'itération
- 4 Une condition suffisante pour la non-régularité
- 5 Utilisation des propriétés de fermeture
- 6 Résumé

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    7 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Lemme de l'itération

Soit  $L$  un langage régulier.

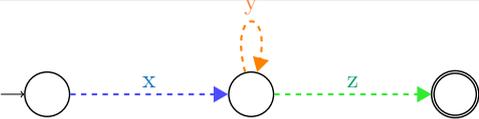
**Lemme de l'itération, Pumping lemma**

Il existe  $N \in \mathbb{N}$  tel que pour tout mot  $w \in L$  avec  $|w| \geq N$ , on peut trouver  $x, y, z \in \Sigma^*$  tels que

$$w = xyz$$

et :

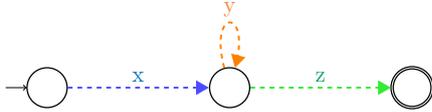
- 1  $y \neq \epsilon$ ,
- 2  $|xy| \leq N$ ,
- 3 pour tout  $k \in \mathbb{N}$ ,  $xy^kz \in L$ .



Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    8 / 46

### Démonstration du Lemme de l'itération

- Soit  $L$  un langage régulier.
- Alors, il existe un ADEF  $A = (Q, \Sigma, q_0, \delta, F)$  tel que  $L(A) = L$ .
- Soit  $n = |Q|$ .
- Soit  $w = a_1 \cdots a_m \in L$  tel que  $|w| = m \geq n$ .
- Soit  $p_i = \delta^*(q_0, a_1, \dots, a_i)$  pour  $i \leq m$ .
- Alors, ils existent  $i$  et  $j$  avec  $0 \leq i < j \leq n$  tels que  $p_i = p_j$ .
- On pose  $x = a_1 \cdots a_i$ ,  $y = a_{i+1} \cdots a_j$  et  $z = a_{j+1} \cdots a_m$ .
- Alors, on a :
  - 1  $w = xyz$ .
  - 2  $|xy| \leq n$ .
  - 3  $y \neq \epsilon$  car  $i < j$ .
  - 4  $\delta^*(p_i, y) = p_j = p_i$  et donc  $\delta^*(p_i, y^k) = p_i$ , pour tout  $k \geq 0$ .
- Donc,  $xy^kz \in L$  pour tout  $k \geq 0$ .



### Application du Lemme de l'itération

#### Exemple (Application du Lemme de l'itération sur $L = \{a^i b^j \mid i \geq 0\}$ )

##### Preuve par l'absurde :

- Supposons que  $L$  soit régulier.
- Alors, on sait par le Lemme de l'itération, qu'il existe  $n \geq 0$  tel que pour tout  $w \in L$  avec  $|w| \geq n$  ils existent  $x, y, z \in \Sigma^*$  avec :
  - 1  $w = xyz$ .
  - 2  $y \neq \epsilon$ .
  - 3  $|xy| \leq n$ .
  - 4  $xy^kz \in L$ , pour tout  $k \geq 0$ .
- Soit  $w = a^n b^n$  (où  $n$  est le  $n$  du Lemme de l'itération). On a  $w \in L$ .
- Soient  $x, y, z \in \Sigma^*$  comme ci-dessus.
- Alors, comme  $|xy| \leq n$  et  $y \neq \epsilon$ , on a  $x = a^i$  et  $y = a^j$  avec  $j \geq 0$  et  $i > 0$ .
- Soit  $w' = xy^2z = a^i a^j a^j a^{n-j-i} b^n = a^{n+i} b^n$ .
- Alors, d'une part on devrait avoir (d'après le lemme)  $w' \in L$  et d'autre part  $w' \notin L$  car  $n+i > n$ .
- Ce qui est une contradiction. Donc  $L$  n'est pas régulier.

### Application du Lemme de l'itération

#### Exemple (Application du Lemme de l'itération sur $L = \{w \mid |w|_a = |w|_b\}$ )

##### Preuve par l'absurde :

- Supposons que  $L$  soit régulier.
- Alors, on sait par le Lemme de l'itération, qu'il existe  $n \geq 0$  tel que pour tout  $w \in L$  avec  $|w| \geq n$  ils existent  $x, y, z \in \Sigma^*$  avec :
  - 1  $w = xyz$ .
  - 2  $y \neq \epsilon$ .
  - 3  $|xy| \leq n$ .
  - 4  $xy^kz \in L$ , pour tout  $k \geq 0$ .
- Soit  $w = a^n b^n$ . On a  $w \in L$  car  $|w|_a = |w|_b$ .
- Soient  $x, y, z \in \Sigma^*$  comme ci-dessus.
- Alors, comme  $|xy| \leq n$  et  $y \neq \epsilon$ , on a  $x = a^i$  et  $y = a^j$  avec  $j \geq 0$  et  $i > 0$ .
- Soit  $w' = xy^2z = a^i a^j a^j a^{n-j-i} b^n = a^{n+i} b^n$ .
- Alors, d'un côté on a  $w' \in L$  mais aussi  $w' \notin L$  car  $n+i > n$ .
- Ceci est une contradiction. Donc  $L$  n'est pas régulier.

### Application du Lemme de l'itération

#### Exemple (Application du Lemme de l'itération sur $L = \{ww \mid w \in \{a, b\}^*\}$ )

##### Preuve par l'absurde :

- Supposons que  $L$  soit régulier.
- Alors, on sait par le Lemme de l'itération, qu'il existe  $n \geq 0$  tel que pour tout  $w \in L$  avec  $|w| \geq n$  ils existent  $x, y, z \in \Sigma^*$  avec :
  - 1  $w = xyz$ .
  - 2  $y \neq \epsilon$ .
  - 3  $|xy| \leq n$ .
  - 4  $xy^kz \in L$ , pour tout  $k \geq 0$ .
- Soit  $w = a^n b a^n b$ . On a  $w \in L$ .
- Soient  $x, y, z \in \Sigma^*$  comme ci-dessus.
- Alors, comme  $|xy| \leq n$  et  $y \neq \epsilon$ , on a  $x = a^i$  et  $y = a^j$  avec  $j \geq 0$  et  $i > 0$ .
- Soit  $w' = xy^2z = a^i a^j a^j a^{n-j-i} b a^n b = a^{n+i} b a^n b$ .
- Alors, d'une part on devrait avoir (d'après le lemme)  $w' \in L$  et d'autre part  $w' \notin L$  car il n'existe pas de  $w''$  t.q.  $w' = w''w''$ .
- Ceci est une contradiction. Donc  $L$  n'est pas régulier.

### Application du Lemme de l'itération

#### Exemple (Application du Lemme de l'itération sur $L = \{a^i b^j \mid i > j\}$ )

##### Preuve par l'absurde :

- Supposons que  $L$  soit régulier.
- Alors, on sait par le Lemme de l'itération, qu'il existe  $n \geq 0$  tel que pour tout  $w \in L$  avec  $|w| \geq n$  ils existent  $x, y, z \in \Sigma^*$  avec :
  - 1  $w = xyz$ .
  - 2  $y \neq \epsilon$ .
  - 3  $|xy| \leq n$ .
  - 4  $xy^kz \in L$ , pour tout  $k \geq 0$ .
- Soit  $w = a^{n+1} b^n$ . On a  $w \in L$  car  $n+1 > n$ .
- Soient  $x, y, z \in \Sigma^*$  comme ci-dessus.
- Alors, comme  $|xy| \leq n$  et  $y \neq \epsilon$ , on a  $x = a^i$  et  $y = a^j$  avec  $j \geq 0$  et  $i > 0$ .
- On a  $w = a^i a^j a^{n+1-j-i} b^n$  avec  $j \geq 0$ . Soit  $w' = xy^0z = a^{n+1-i} b^n$ .
- Alors, d'une part on devrait avoir  $w' \in L$  et d'autre part  $w' \notin L$  car  $|w'|_a \leq |w'|_b$  ( $i \geq 1$ ).
- Ceci est une contradiction. Donc  $L$  n'est pas régulier.

### Application du Lemme de l'itération

#### Exercice : Application du Lemme de l'itération sur $L = \{a^i b^j \mid i < j\}$

Montrer que  $L$  n'est pas régulier.

#### Exercice : Application du Lemme de l'itération sur $L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$

Montrer que  $L$  n'est pas régulier.

#### Exercice : Application du Lemme de l'itération sur $L = \{a^n b^m b a^{n+m} \mid n, m \in \mathbb{N}\}$

Montrer que  $L$  n'est pas régulier.

### Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- 1 Motivations
- 2 Lemme de l'itération
- 3 Constante d'itération
- 4 Une condition suffisante pour la non-régularité
- 5 Utilisation des propriétés de fermeture
- 6 Résumé

### Constante d'itération

Selon le lemme de l'itération, il existe (au moins) un  $n \in \mathbb{N}$  tel que chaque mot  $w$  d'un langage régulier t.q.  $|w| \geq n$  pour un certain  $n \in \mathbb{N}$  possède un facteur *non vide* ( $y$ ) qui peut être itéré.  
Un tel entier  $n$  est appelé une **constante d'itération** et le facteur est dit **facteur itérant**.

- Si  $n$  est une constante d'itération, alors chaque  $n' \geq n$  est une constante d'itération.
- Si  $n$  n'est pas une constante d'itération, alors chaque  $n' \leq n$  n'est pas une constante d'itération.

#### Définition (Constante d'itération minimale)

La constante d'itération minimale d'un langage est sa plus petite constante d'itération.

#### Exemple (La constante d'itération minimale de $01^*$ est 2)

- Le mot  $w = 0$  est tel que  $w \in 01^*$ ,  $|w| = 1$  et  $w$  ne contient pas de facteur qui peut être itéré.
- Un mot  $w \in 01^*$  tel que  $|w| \geq 2$  possède un facteur qui peut être itéré à travers la décomposition  $w = xyz$  avec  $x = 0, y = 1, z = \text{reste}$

## Constante d'itération - Exercices

## Constante d'itération minimale

Trouver la constante d'itération minimale pour les langages suivants :

- $0001^*$
- $0^*$
- $0^*1^*$
- $0^*1^+0^+1^* + 10^*1$

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- Motivations
- Lemme de l'itération
- Constante d'itération
- Une condition suffisante pour la non-régularité
- Utilisation des propriétés de fermeture
- Résumé

## Une condition suffisante à la non-régularité

Soit  $L$  un langage sur un alphabet  $\Sigma$ .

## Une condition suffisante pour la non-régularité

S'il existe deux suites de mots  $(\alpha_n)_{n \in \mathbb{N}}$  et  $(\beta_n)_{n \in \mathbb{N}}$  telles que

$$\forall i, j \in \mathbb{N} : \alpha_i \cdot \beta_j \in L \text{ si et seulement si } i = j$$

alors  $L$  n'est pas régulier.

## Exemple (Application de la condition suffisante)

Pour le langage  $\{a^k b^k \mid k \in \mathbb{N}\}$ , nous pouvons prendre  $\alpha_n = a^n$  et  $\beta_n = b^n$ .

## Exercice : application de la condition suffisante

- $\{w \in \Sigma^* \mid w \text{ est un palindrome}\}$
- $\{w \cdot w \mid w \in \Sigma^*\}$

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- Motivations
- Lemme de l'itération
- Constante d'itération
- Une condition suffisante pour la non-régularité
- Utilisation des propriétés de fermeture
- Résumé

## Utilisation des propriétés de fermeture

- $L_?$  dénote un langage dont on cherche à montrer la non-régularité,
- $L_{\text{reg}}$  dénote un langage que l'on sait régulier,
- $L_{\text{nonreg}}$  dénote un langage que l'on sait non régulier.

## Utilisation de la fermeture par union et intersection

Si  $L_? \cup L_{\text{reg}} = L_{\text{nonreg}}$  ou  $L_? \cap L_{\text{reg}} = L_{\text{nonreg}}$ , alors  $L_?$  est non régulier (sinon  $L_{\text{nonreg}}$  serait régulier).

## Utilisation de la fermeture par complémentation

- Si  $\overline{L_?} = L_{\text{nonreg}}$  alors  $L_?$  est non régulier.
- Si  $\overline{L_?} = L_{\text{reg}}$  alors  $L_?$  est régulier.

## Utilisation des propriétés de fermeture

## Utilisation de la fermeture par morphisme

Soient  $L_?$  et  $L_{\text{nonreg}}$  deux langages sur  $\Sigma$  et  $\Sigma'$  respectivement et  $f$  un morphisme de  $\Sigma$  vers  $\Sigma'^*$  tels que  $L_{\text{nonreg}} = f(L_?)$ .

- Alors  $L_?$  est non régulier. En effet, supposons (par l'absurde) que  $L_?$  soit régulier. Comme les langages réguliers sont fermés par morphisme et  $L_{\text{nonreg}} = f(L_?)$ , alors  $L_{\text{nonreg}}$  serait régulier. Contradiction.

## Utilisation de la fermeture par morphisme inverse

Soient  $L_?$  et  $L_{\text{nonreg}}$  deux langages sur  $\Sigma$  et  $\Sigma'$  respectivement et  $f$  un morphisme de  $\Sigma$  vers  $\Sigma'^*$  tels que  $L_{\text{nonreg}} = f^{-1}(L_?)$ .

- Alors  $L_?$  est non régulier. En effet, supposons (par l'absurde) que  $L_?$  soit régulier. Comme les langages réguliers sont fermés par morphisme inverse et  $L_{\text{nonreg}} = f^{-1}(L_?)$ , alors  $L_{\text{nonreg}}$  serait régulier. Contradiction.

## Utilisation des propriétés de fermeture

## Exemple (Utilisation de la fermeture par intersection)

Utilisation pour démontrer que  $\{w \mid |w|_a = |w|_b\}$  est non régulier.

- Supposons que nous avons démontré que  $\{a^k \cdot b^k \mid k \in \mathbb{N}\}$  soit non régulier.
- Nous savons que  $a^* \cdot b^*$  est régulier.

Alors, comme  $\{w \in \{a, b\}^* \mid |w|_a = |w|_b\} \cap a^* \cdot b^* = \{a^k \cdot b^k \mid k \in \mathbb{N}\}$ , nous en déduisons que  $\{w \mid |w|_a = |w|_b\}$  est non régulier.

## Exemple (Utilisation de la fermeture par complémentation)

- Comme  $\{a^p \mid p \text{ premier}\}$  est non régulier, alors  $\{a^p \mid p \text{ non premier}\}$  est non régulier.
- Comme  $\{w \mid |w|_a = |w|_b\}$  est non régulier, alors  $\{w \mid |w|_a \neq |w|_b\}$  est non régulier.

## Utilisation des propriétés de fermeture

## Exemple (Utilisation de la fermeture par morphisme)

- $L_1 = \{0^k 1^k 2^k \mid k \in \mathbb{N}\}$  est non régulier car  $L_2 = \{a^k b^k \mid k \in \mathbb{N}\}$  est non régulier. En effet, considérons, le morphisme induit par

$$h = \{0 \mapsto a, 1 \mapsto b, 2 \mapsto \epsilon\},$$

nous avons  $h(L_1) = L_2$ .

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- 1 Motivations
- 2 Lemme de l'itération
- 3 Constante d'itération
- 4 Une condition suffisante pour la non-régularité
- 5 Utilisation des propriétés de fermeture
- 6 **Résumé**

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    26 / 48

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Résumé

### Langages Non Réguliers et Lemme de l'itération

- Lemme de l'itération (et ses variantes) : propriétés des langages réguliers
  - énoncé
  - preuve
- Utilisation : prouver que des langages sont non-réguliers.
  - preuves par contradiction
  - utilisation des propriétés de fermeture : opérations booléennes, morphisme (inverse)
- Exemples de démonstrations de la non-régularité de certains langages.
- Constante d'itération

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    27 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- 1 Dénombrement des langages réguliers et non réguliers
- 2 Le lemme de l'itération est une condition nécessaire à la régularité
- 3 Variantes du lemme de l'itération
- 4 **Théorème de Ehrenfeucht, Parikh et Rozenberg**

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    28 / 48

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Ensembles dénombrables et indénombrables

### Définitions

Intuitivement, un ensemble est dénombrable lorsque ses éléments peuvent être listés sans omission ni répétition.

En informatique, on accède uniquement aux ensembles dénombrables.

### Définition (Ensemble dénombrable)

Un ensemble est dénombrable lorsqu'il peut être mis en bijection avec une partie de  $\mathbb{N}$ .

### Exemple (Ensemble dénombrable)

- $\mathbb{N}$  et ses parties (entiers non nuls, entiers pairs, entiers impairs, etc)
- l'ensemble des entiers relatifs
- l'ensemble des rationnels
- l'ensemble des nombres premiers
- $\mathbb{N} \times \mathbb{N}$

### Exemple (Ensemble non dénombrable)

- $\mathbb{R}$
- $\mathcal{P}(\mathbb{N})$
- $\mathbb{R} \rightarrow \mathbb{B}$

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    29 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Propriétés des ensembles dénombrables

Cantor, 1891

### Sous ensembles d'un ensemble dénombrable

Tout sous-ensemble d'un ensemble dénombrable est dénombrable.

### Union dénombrable d'ensembles finis

Une union dénombrable d'ensembles finis est dénombrable.

### Démonstration.

Lister chaque ensemble dénombrable en ligne :

$\rightarrow$	$a_{11}$	$a_{12}$	$a_{13}$	$\dots$	$A_1$
$\rightarrow$	$a_{21}$	$a_{22}$	$a_{23}$	$\dots$	$A_2$
$\rightarrow$	$a_{31}$	$a_{32}$	$a_{33}$	$\dots$	$A_3$
$\rightarrow$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$

Énumération des éléments en diagonale :  $a_{11}, a_{21}, a_{12}, a_{31}, \dots$  □

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    30 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Propriétés des ensembles indénombrables

Cantor, 1891

### Sur ensembles d'un ensemble indénombrable

Tout sur-ensemble d'un ensemble indénombrable est indénombrable.

### Le cardinal d'un ensemble est toujours strictement inférieur à celui de ses parties

Si  $E$  est un ensemble dénombrable infini, alors  $\mathcal{P}(E)$  est indénombrable.

### Démonstration.

Soit  $E$  un ensemble dénombrable infini. Supposons qu'il existe une bijection  $f$  de  $E$  vers  $\mathcal{P}(E)$ .  
Considérons l'ensemble  $D$  des éléments qui n'appartiennent pas à leur image :

$$\{x \in E \mid x \notin f(x)\}$$

Comme  $D \subseteq E$  et que  $f$  est surjective, on peut trouver  $x_0 \in E$  tq  $f(x_0) = D$  et cet élément est unique ( $f$  est injective). Il y a deux cas :

- si  $x_0 \in D$ , par définition de  $D$ ,  $x_0 \notin f(x_0) = D$ , contradiction.
- si  $x_0 \notin D$ , par définition,  $x_0 \in f(x_0) = D$ , contradiction.

Ainsi,  $\mathcal{P}(E)$  n'est jamais en bijection avec  $E$ . En conséquence, si  $E$  est en bijection avec  $\mathbb{N}$  (dénombrable infini), alors  $\mathcal{P}(E)$  est indénombrable. □

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    31 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Argument diagonal pour démontrer la non dénombrabilité de l'ensemble des réels

Cantor, 1891

### Exemple

Argument diagonal pour montrer la non-dénombrabilité des réels

- Supposons que les réels soient dénombrables.
- Listons les dans un ordre (arbitraire).
- Définissons le réel  $X$  tel que le  $j^{\text{ème}}$  chiffre de  $X$  soit différent du  $j^{\text{ème}}$  chiffre du  $i^{\text{ème}}$  réel.
- $X$  n'est pas dans la liste.

	entier	1	2	3	4	...
1 <sup>er</sup>	0,	0	0	0	0	...
2 <sup>ème</sup>	42,	1	0	0	0	...
3 <sup>ème</sup>	12,	8	9	0	0	...
4 <sup>ème</sup>	81,	4	8	3	0	...
5 <sup>ème</sup>	0,	3	6	5	8	...
$\vdots$						
$\vdots$						
<b>X</b>	<b>1,</b>	<b>2</b>	<b>0</b>	<b>4</b>	<b>9</b>	<b>...</b>

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    32 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Dénombrement des langages réguliers

D'après le théorème de Kleene, pour dénombrer les langages réguliers, nous pouvons dénombrer l'ensemble des automates à finis.

Comme tout automate à états fini peut être représenté par un AEFD, nous allons dénombrer les AEFDs (dénombrer les automates à états finis se fait similairement).

### Dénombrement des AEFDs à $|Q|$ états

- $|Q|$  états initiaux possibles
- $2^{|Q|}$  ensemble d'états finaux possibles
- $|Q|^{|Q| \times |Q|}$  fonctions de transition possibles

$\Rightarrow |Q| \times 2^{|Q|} \times |Q|^{|Q| \times |Q|}$  AEFDs possibles

- Chacun des ensembles des AEFDs à  $|Q|$  est un ensemble fini.
- L'ensemble de tous les AEFDs est l'union de chacun des ensembles à  $|Q|$  états où  $|Q|$  prend les valeurs de  $\mathbb{N}$  (union dénombrable d'ensembles finis).

Y. Falcone (UGA - Inria)    INF 302 : Langages & Automates    Année Académique 2021 - 2022    33 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Dénombrement de l'ensemble de tous les mots

Alphabet ( $\Sigma$ ) : fini

Ensemble des mots  $\Sigma^*$

L'ensemble de tous les mots sur  $\Sigma$  est dénombrable.

Démonstration.

- Un alphabet  $\Sigma$  est fini et peut s'écrire  $\Sigma = \{a_1, \dots, a_n\}$ , pour un certain  $n \in \mathbb{N}$ .
- On ordonne les éléments de  $\Sigma^*$  avec un ordre lexicographique induit par un ordre (arbitraire) sur  $\Sigma : \{\epsilon, a_1, \dots, a_n, a_1 \cdot a_1, \dots, a_1 \cdot a_n, \dots, a_n \cdot a_n, \dots\}$   $\square$

Exemple (Ensemble de mots)

- Ensemble des phrases en français.
- Ensemble des programmes dans un langage de programmation.

Démonstration (alternative) de  $\Sigma^*$  dénombrable pour  $\Sigma = \{0, \dots, 9\}$

Utiliser la correspondance :

$$\mathbb{N} \rightarrow \Sigma^* \text{ telle que } x \mapsto 1x$$

Y. Falcone (UGA - Inria) INF 302 : Langages & Automates Année Académique 2021 - 2022 34 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Dénombrement des langages sur un alphabet

- $\Sigma$  : alphabet
- $\Sigma^*$  : ensemble de tous les mots
- $\mathcal{P}(\Sigma^*)$  : ensemble de tous les langages

Argumentation 1

- $\Sigma^*$  est dénombrable.
- L'ensemble des sous-ensembles d'un ensemble infini dénombrable est indénombrable.

Argumentation 2

- Supposons que  $\mathcal{P}(\Sigma^*)$  soit dénombrable.
- Soient  $L_1, L_2, \dots$  ces langages.
- Construisons  $L$  défini par :

$$w_i \in L \text{ ssi } w_i \notin L_i$$

- $L$  n'est pas dans la liste des langages.

	$w_1$	$w_2$	$w_3$	$w_4$	$w_5$	...
$L_1$	0	0	0	0	0	...
$L_2$	1	1	0	0	0	...
$L_3$	1	0	0	0	0	...
$L_4$	1	0	0	0	0	...
$L_5$	0	1	0	0	1	...
...						
$L$	1	0	1	1	0	...

Y. Falcone (UGA - Inria) INF 302 : Langages & Automates Année Académique 2021 - 2022 35 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- Dénombrement des langages réguliers et non réguliers
- Le lemme de l'itération est une condition nécessaire à la régularité
- Variantes du lemme de l'itération
- Théorème de Ehrenfeucht, Parikh et Rozenberg

Y. Falcone (UGA - Inria) INF 302 : Langages & Automates Année Académique 2021 - 2022 36 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Le lemme de l'itération est une condition nécessaire à la régularité

Le lemme de l'itération est une condition nécessaire à la régularité d'un langage. Ce n'est pas une condition suffisante.

C'est à dire :

$$\forall L \subseteq \Sigma^* : L \text{ régulier} \implies L \text{ satisfait le lemme de l'itération}$$

Ceci n'empêche pas :

$$L \text{ non régulier} \implies L \text{ satisfait le lemme de l'itération}$$

pour certains langages  $L \subseteq \Sigma^*$ .

En effet, nous donnons des exemples de langages non réguliers qui satisfait le lemme de l'itération.

Y. Falcone (UGA - Inria) INF 302 : Langages & Automates Année Académique 2021 - 2022 37 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Langage non régulier satisfaisant le lemme de l'itération

Exemple 1

Exemple (Un langage non-régulier satisfaisant le lemme de l'itération)

$$\{\#a^n b^n \mid n \geq 1\} \cup \{\#^k \cdot w \mid k \neq 1, w \in \{a, b\}^*\}$$

Exemple (Un langage non-régulier satisfaisant le lemme de l'itération)

- Soit  $L$  un langage non régulier.
- Nous construisons le langage  $L_{\#} = (\#)^+ \cdot L \cup \Sigma^*$ , avec  $\# \notin \Sigma$ .
- $L_{\#}$  n'est pas régulier : considérons le morphisme induit par  $h$  t.q.  $h(a) = a$  pour  $a \in \Sigma$  et  $h(\#) = \epsilon$ .
- Nous montrons que  $L_{\#}$  satisfait le lemme de l'itération. Soit  $w \in L_{\#}$  t.q.  $|w| \geq 1$ . On a  $w \in (\#)^+ \cdot L$  ou  $w \in \Sigma^*$ .
  - Si  $w \in (\#)^+ \cdot L$ , alors  $w$  peut s'écrire  $w_{\#} \cdot w'$  avec  $w_{\#} \in (\#)^+$  et  $w' \in L$ . Nous pouvons trouver la décomposition  $x = \epsilon, y = \#, z = \text{reste}$  qui satisfait les conditions du lemme.
  - Si  $w \in \Sigma^*$ , alors  $w$  peut se décomposer avec  $x = \epsilon, y = \text{première lettre de } w, z = \text{reste}$  qui satisfait les conditions du lemme.

Y. Falcone (UGA - Inria) INF 302 : Langages & Automates Année Académique 2021 - 2022 38 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Langage non régulier satisfaisant le lemme de l'itération

Exemple 2

Exemple (Un langage non-régulier satisfaisant le lemme de l'itération)

- Considérons  $L = \{w \cdot w^R \cdot u \mid w, u \in \{a, b\}^+\}$ .
- $L$  n'est pas régulier (en utilisant la deuxième version du lemme).
- $L$  satisfait les conditions du lemme de l'itération. Soit  $m \in L$  t.q.  $|m| \geq 4$ . Le mot  $m$  s'écrit  $w \cdot w^R \cdot u$  avec  $w, u \in \{a, b\}^+$ . Nous distinguons deux cas.
  - Si  $w$  est une lettre, alors  $|u| \geq 2$ . Nous prenons  $x = w \cdot w^R, y = \text{la première lettre de } u, z = \text{reste}$ .
  - Si  $|w| \geq 2$ , alors  $w = w' \cdot s$  avec  $s \in \{a, b\}$  et  $|w'| \geq 1$ . Alors  $w^R = s \cdot w'^R$ . Nous prenons  $x = w', y = s \cdot s, z = w'^R \cdot u$ .

Dans les deux cas, la décomposition de  $m = x \cdot y \cdot z$  satisfait les conditions du lemme de l'itération.

Y. Falcone (UGA - Inria) INF 302 : Langages & Automates Année Académique 2021 - 2022 39 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Langage non régulier satisfaisant le lemme de l'itération

Exemple 3

Soit  $\Sigma = \{a, b, c\}$ .

Exemple (Un langage non-régulier satisfaisant le lemme de l'itération)

- Considérons  $L = \bigcup_{n \geq 0} (a^+ \cdot b)^n \cdot (b^+ \cdot c)^n + \Sigma^* \cdot c \cdot \Sigma^*$
- $L$  n'est pas régulier.
- $L$  satisfait les conditions du lemme de l'itération.

Y. Falcone (UGA - Inria) INF 302 : Langages & Automates Année Académique 2021 - 2022 40 / 46

Univ. Grenoble Alpes, Département Licence Sciences et Technologies, Licence deuxième année

## Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération

- Dénombrement des langages réguliers et non réguliers
- Le lemme de l'itération est une condition nécessaire à la régularité
- Variantes du lemme de l'itération
- Théorème de Ehrenfeucht, Parikh et Rozenberg

Y. Falcone (UGA - Inria) INF 302 : Langages & Automates Année Académique 2021 - 2022 41 / 46

**Lemme de l'itération**

Choix du facteur itérant

Soit  $L$  un langage régulier.**Lemme de l'itération, Pumping lemma**

Il existe  $n \in \mathbb{N}$  tel que  
 pour tout mot  $w \in L$ ,  
 pour tous mots  $u, w', v$  tels que  $w = uw'v$  avec  $|w'| \geq n$ ,  
 on peut trouver  $x, y, z \in \Sigma^*$  tels que

$$w' = xyz$$

et :

- 1  $y \neq \epsilon$ ,
- 2  $|xy| \leq n$ ,
- 3 pour tout  $k \in \mathbb{N}$ ,  $uxy^kzv \in L$ .

**Lemme de l'itération par bloc**Soit  $L$  un langage régulier.**Lemme de l'itération, Pumping lemma**

Il existe  $n \in \mathbb{N}$  tel que  
 pour tout mot  $w \in L$ ,  
 pour tous mots  $u, w_1, \dots, w_n, v$  tels que  $w = uw_1 \dots w_n v$  avec  $\forall i \in [1, n] : w_i \neq \epsilon$ ,  
 on peut trouver  $k, l \in \mathbb{N}$  tels que

$$1 \quad 0 \leq k < l \leq n,$$

$$2 \quad \forall i \in \mathbb{N} : uw_1 \dots w_k (w_{k+1} \dots w_l)^i w_{l+1} \dots w_n v \in L.$$

où  $w_1 \dots w_k = \epsilon$  si  $k = 0$  et  $w_{l+1} \dots w_n = \epsilon$  si  $k = l$ .

**Plan Chap. 10 - Langages non-réguliers et lemmes de l'itération**

- 1 Dénombrement des langages réguliers et non réguliers
- 2 Le lemme de l'itération est une condition nécessaire à la régularité
- 3 Variantes du lemme de l'itération
- 4 Théorème de Ehrenfeucht, Parikh et Rozenberg

**Conditions  $(E_n)$  et  $(E'_n)$** Soit  $L$  un langage sur l'alphabet  $\Sigma$ **Définition (Condition  $(E_n)$ )**

$L$  vérifie la condition  $(E_n)$  (pour un entier  $n \in \mathbb{N}$ ) ssi pour tout mot  $w \in \Sigma^*$ , pour tous mots  $u, v, w_1, \dots, w_n$  tels que  $w = u \cdot w_1 \dots w_n \cdot v$  et  $w_i \neq \epsilon$  pour  $i \in [1, n]$ , ils existent  $k, l \in \mathbb{N}$  tels que  $0 < k \leq l \leq n$  et pour tout  $k \in \mathbb{N}$

$$w \in L \quad \text{si et seulement si} \quad u \cdot w_1 \dots w_k (w_{k+1} \dots w_l)^k w_{l+1} \dots w_n \cdot v \in L$$

**Définition (Condition  $(E'_n)$ )**

$L$  vérifie la condition  $(E'_n)$  (pour un entier  $n \in \mathbb{N}$ ) ssi pour tout mot  $w \in \Sigma^*$ , pour tous mots  $u, v, w_1, \dots, w_n$  tels que  $w = u \cdot w_1 \dots w_n \cdot v$  et  $w_i \neq \epsilon$  pour  $i \in [1, n]$ , ils existent  $k, l \in \mathbb{N}$  tels que  $0 < k \leq l \leq n$  et pour tout  $k \in \mathbb{N}$

$$w \in L \quad \text{si et seulement si} \quad u \cdot w_1 \dots w_k \cdot w_{l+1} \dots w_n \cdot v \in L$$

**Théorème de Ehrenfeucht, Parikh et Rozenberg****Théorème de Ehrenfeucht, Parikh et Rozenberg**

Les conditions suivantes sont équivalentes.

- $L$  est régulier.
- Il existe un entier  $n$  tel que  $L$  vérifie la condition  $(E_n)$ .
- Il existe un entier  $n$  tel que  $L$  vérifie la condition  $(E'_n)$ .