

# Langage: notions de base et propriétés

# Concepts de base: Alphabet

- ▶ Un alphabet  $X$  est un ensemble fini, non vide, de symboles comme par exemple les lettres et les chiffres.
- ▶ Exemple:



# Concepts de base: Mot

- ▶ Un mot  $w$  est défini sur l'alphabet  $X$  est une suite finie et ordonnée, éventuellement vide, d'éléments de l'alphabet.
- ▶ C'est une concaténation de chiffres, de lettres ou de symboles appartenant à  $X$ .
- ▶ Le mot vide, noté  $\varepsilon$  correspond à la suite de symboles vide.



# Concepts de base: Mot

- ▶ Exemple:



# Concepts de base: langage

- ▶ Un langage  $L$  défini sur l'alphabet  $X$  est un ensemble de mots défini sur l'alphabet en question. Cet ensemble de mots peut être fini comme il peut être infini.
- ▶ Exemple:



# Concepts de base: langage

- ▶ Formellement, si  $X$  est un alphabet alors un langage sur  $X$  est par définition un sous-ensemble du monoïde  $X^*$
- ▶ Exemple: soit  $x = \{a, b, c\}$
- ▶ Problème: on cherche l'ensemble des mots qui commencent par  $a$  suivi de zéro ou plusieurs  $c$  et se terminent par  $bb$ .
- ▶ Solution:



# Concepts de base: Monoïde

- ▶ Pour un alphabet fini  $X$ , le monoïde  $X^*$  librement engendré par  $X$  est l'ensemble des suites finies  $u = u_1 u_2 \dots u_n$  où  $u_i$  est une lettre de  $X$  et  $n$  désigne la longueur de  $u$  avec  $n \geq 0$ .
- ▶ La loi de composition interne du monoïde  $X^*$  est la concaténation.
- ▶ Si  $X = \{a, b\}$ , alors l'ensemble  $X^* = \{\varepsilon, a, b, aa, ab, bb, aab, bba, aba, aaa, bbb, \dots\}$
- ▶ L'élément neutre  $\varepsilon$  appelé aussi le mot vide est l'unique mot de longueur 0.

# Concepts de base: Monoïde

- ▶ Définissons tout d'abord les éléments  $X^*$  et  $X^+$ :
- ▶ Soit un alphabet  $X$ . L'ensemble  $X^+$  des mots définis sur  $X$  de longueur supérieur ou égale à un est défini par:
  1. Tout élément de  $X^+$  appartient à  $X$
  2. Si  $a \in X$  et  $w \in X^+$ , alors  $aw \in X^+$
  3.  $X^+$  ne contient que des objets obtenus en appliquant un nombre fini de fois les règles 1 et 2.
  4.  $X^* = X^+ \cup \{\varepsilon\}$

# Concepts de base: Monoïde

- ▶ Exemple:



# Opérations sur les mots: Egalité

- ▶ Deux mots sont égaux si et seulement si ils sont formés de la même suite de lettre.
- ▶ Soit  $X$  un alphabet et soient  $u \in X^*$  et  $v \in X^*$  /  
 $u = u_1.u_2\dots.u_n$  et
- ▶  $v = v_1.v_2\dots.v_m$  /  $u_i, v_i \in X$ .
  
- ▶  $u = v$  si et seulement si  $n = m$  et  $\forall i \in \{1, 2, \dots, n\}$  /  $u_i = v_i$

# Opérations sur les mots: Egalité

- ▶ Exemple:



# Opérations sur les mots: Longueur

- ▶ Soit  $X$  un alphabet et soit  $w$  un mot de  $X^*$ . La longueur de  $w$  est définie par le nombre d'occurrences de lettres de  $X$  dans  $w$ .
- ▶ Nous la notons par  $|w|$
  
- ▶ Exemple :

# Opérations sur les mots: Longueur

## Remarque:

- ▶ Le mot vide  $\varepsilon$  n'est composé d'aucun caractère, sa longueur est donc nulle ( $|\varepsilon| = 0$ ). Pour tout mot  $w$  appartenant à un langage  $L$ , on a :  $w.\varepsilon = \varepsilon.w = w$
- ▶ Nous notons  $|w|_a$  le nombre d'occurrences de la lettre  $a$  dans  $w$ .
- ▶ Exemple :

# Opérations sur les mots: Concaténation

- ▶ La concaténation de deux mots  $u \in X^*$  et  $v \in X^*$  tels que  $u = u_1.u_2 \dots u_n$  avec  $u_i \in X, \forall i \in \{1, 2, \dots, n\}$  et  $v = v_1.v_2 \dots v_m$  avec  $v_i \in X, \forall i \in \{1, 2, \dots, m\}$  est notée par  $u.v = u_1.u_2 \dots u_n.v_1.v_2 \dots v_m$ .
- ▶ Remarque: l'opération de concaténation est associative mais non commutative.
- ▶ Soit  $w = u.v$  alors  $|w| = |u| + |v|$
- ▶ Soit  $w \in X^*$  alors  $(u.v).w = u.(v.w)$  (associativité)
- ▶ Soit  $u, v \in X^*$  alors  $u.v \neq v.u$
- ▶ Soit l'élément neutre  $\varepsilon$  alors  $u.\varepsilon = \varepsilon.u = u$  (élément neutre)

# Opérations sur les mots: Concaténation

- ▶ Soient  $u$  et  $v$  deux mots définies comme suit:
- ▶  $u = u_1 u_2 u_3 \dots u_n$  et  $v = v_1 v_2 v_3 \dots v_n$
- ▶ La concaténation de  $u$  et  $v$  est le mot:

$$w = u_1 u_2 u_3 \dots u_n \cdot v_1 v_2 v_3 \dots v_n$$

**Exemple:**



# Opérations sur les mots

Exercice:

Soit  $X = \{a, b\}$ ;  $\varepsilon$  est il un langage?

.....

Soit  $X = \{a, b\}$ ;  $\emptyset$  est il un langage?

.....

Soit  $A = \{x, y\}$  et  $B = \{y, x\}$   $A^* - B^* = A^*$  ?

.....

Soit  $X = \{a, b\}$  ;  $\{\varepsilon\}$  est il un langage?

.....

# Opérations sur les mots: Facteur

- ▶ Soit  $X$  un alphabet et  $w \in X^*$ . Le mot  $u \in X^*$  est un facteur de  $w$  si et seulement si  $w = u' . u . u''$  avec  $u', u'' \in X^*$ .
- ▶ Si  $u' . u'' \neq \varepsilon$  alors dans ce cas nous désignons par  $u$  un facteur propre de  $w$ .
- ▶ Si  $u' = \varepsilon$  alors  $w = u . u''$  et dans ce cas nous désignons par  $u$  un facteur gauche de  $w$  ou préfixe de  $w$ .
- ▶ Si  $u'' = \varepsilon$  alors  $w = u' . u$  et dans ce cas nous désignons par  $u$  un facteur droit de  $w$  ou suffixe de  $w$ .

# Opérations sur les mots: Facteur

- ▶ Exemple1:



# Opérations sur les mots: Facteur

- ▶ Exemple 2:



# Opérations sur les mots: Miroir

- ▶ L'opération miroir associe à un mot  $w = u_1 u_2 \dots u_n$  le même mot mais renversé. On le note  $w^R = u_n \dots u_2 u_1$ .
- ▶ Pour les mots  $w, w_1, w_2$  définis sur l'alphabet  $X$  ont les propriétés suivantes:
  - ▶ 1 – si  $w = w^R$ , alors le mot  $w$  est appelé un **palindrome**
  - ▶ 2 –  $(w^R)^R = w$
  - ▶ 3 –  $(w_1 w_2)^R = w_2^R w_1^R$

# Opérations sur les mots: Miroir

- ▶ Exemple 1:



# Opérations sur les mots: Puissance

- ▶ La puissance d'un mot  $w$ , noté  $w_n$  où  $n \geq 0$  est définie par:
  - ▶ 1-  $w^0 = \varepsilon$
  - ▶ 2-  $w^{n+1} = w^n w$
- ▶ Soit  $X$  un alphabet et soit  $w \in X^*$
- ▶ Les deux écritures suivantes sont équivalentes
- ▶  $W = abbbba = ab^4a$

# Opérations sur les mots: Puissance

- ▶ Pour le mot  $w = abcc$  défini sur l'alphabet  $X = \{a, b, c\}$ , on a :
- ▶  $w^1 = abcc$
- ▶  $w^3 = abcc abcc abcc$
  
- ▶ **Exemple:**

# Représentation des langages:

Définition par propriétés (qualitative)

- ▶ **Définition textuelle:**  $L1$  est un langage définie sur l'alphabet  $X$  composé des mots formés d'une suite de  $a$  suivi d'une suite de  $b$  avec le nombre de  $a$  égale au nombre de  $b$ .
- ▶ **Définition qualitative:**
- ▶  $L1 = \{ w \in X^* / w = a^n b^n ; n \geq 1 \}$

# Représentation des langages:

## Définition par propriétés (qualitative)

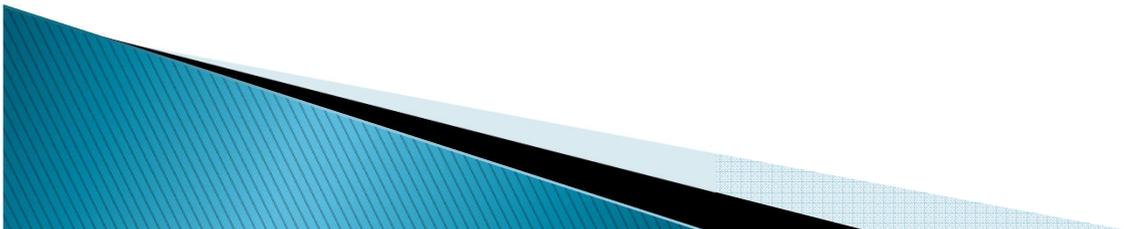
- ▶ **Exemple:**



# Représentation des langages:

## Définition récursive

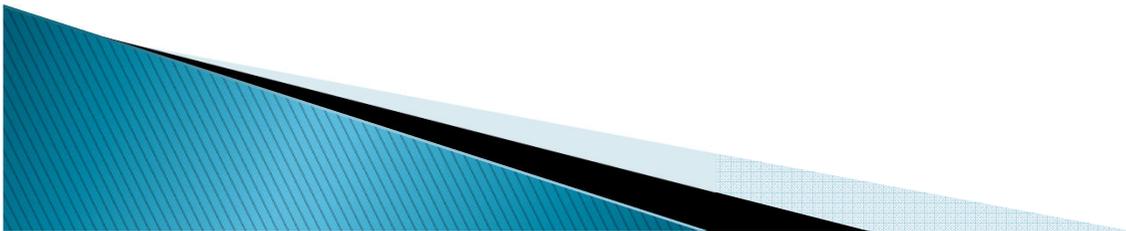
- ▶ La définition des éléments d'un langage fait référence aux éléments de ce même langage. On dit alors que cette définition est récursive.



# Exercice:

- ▶ 1– Donner la définition qualitative (par propriété) du langage dont les mots commencent par le chiffre 0, suivit de 0 ou plusieurs occurrences du chiffre 1 et se termine par 0.
- ▶ 2– Donner une définition par propriété du langage dont les mots sont formés par 0 ou plusieurs occurrences du facteur  $ab$ .
- ▶ 3– Soit  $X = \{a, b\}$ . Donner la définition qualitative du langage dont la taille des mots est impair.

# Solution



# Opération sur les langages

- ▶ Les langages sont des ensembles de mots.
- ▶ Les opérations habituelles concernant les ensembles telles que l'*union*, l'*intersection*, et la *complémentarité* sont alors tout à fait applicables aux langages.



# Opération sur les langages: Egalité

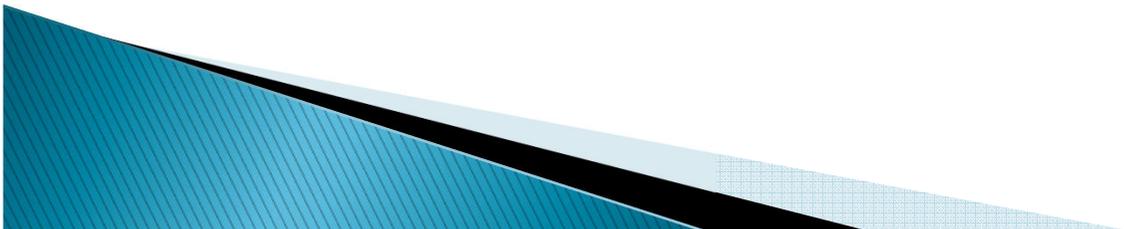
- ▶ Soit  $L_1$  et  $L_2$  deux langages engendré sur  $X^*$ . On dit que les deux langages sont égaux ( $L_1 = L_2$ ) si et seulement si  $\forall w \in L_1$  alors  $w \in L_2$  et  $\forall w \in L_2$  alors  $w \in L_1$ .

# Opération sur les langages: Union, Intersection

- ▶ L'union de deux langages  $L_1$  et  $L_2$  définis sur l'alphabet  $X$  ( $L_1 \cup L_2$ ) et leur intersection ( $L_1 \cap L_2$ ) correspondent exactement à l'union et à l'intersection d'ensembles.
- ▶ Soit  $X$  un alphabet et soit deux langages  $L_1$  et  $L_2$  engendrés sur  $X^*$ .
- ▶  $L_1 \cup L_2 = \{w \in X^* / w \in L_1 \text{ ou } w \in L_2\}$
- ▶  $L_1 \cap L_2 = \{w \in X^* / w \in L_1 \text{ et } w \in L_2\}$

# Opération sur les langages: Union, Intersection

- ▶ **Exemple:**



# Opération sur les langages:

## Complémentarité

- ▶ Soit un langage  $L$  défini sur l'alphabet  $X$ . La complémentarité de  $L$  noté  $\bar{L}$ , est définie par:

- ▶  $\bar{L} = X^* \setminus L = \{w \in X^* / w \notin L\}$

# Opération sur les langages:

## Concaténation

- ▶ Soient  $L_1$  et  $L_2$  deux langages définis sur l'alphabet  $X$ , l'opération de concaténation est définie comme suit:
- ▶  $L_1.L_2 = \{w \in X^* / w = u.v \text{ avec } u \in L_1 \text{ et } v \in L_2\}$
- ▶ L'élément neutre de l'opération de concaténation est donc l'ensemble  $L_\varepsilon = \{\varepsilon\}$  ( $L.L_\varepsilon = L_\varepsilon.L = L$ ) et l'élément absorbant le langage  $\emptyset$  ( $L.\emptyset = \emptyset.L = \emptyset$ )

# Opération sur les langages:

## Concaténation

- ▶ **Exemple:**



# Opération sur les langages:

## Puissance

- ▶ La puissance d'un langage  $L$ , noté  $L^n$  où  $n \geq 0$  est définie par:
  - ▶ 1-  $L^0 = L_\epsilon = \{\epsilon\}$
  - ▶ 2-  $L^{n+1} = L^n.L$
  
- ▶ **Exemple:**

# Opération sur les langages:

## Fermeture itérative

- ▶ La fermeture itérative d'un langage  $L$ , notée  $L^*$  est l'ensemble des mots résultant d'une concaténation d'un nombre fini de mots de  $L$ .
- ▶ Formellement on a:
  - ▶  $L^* = L^0 \cup L^1 \cup L^2 \cup \dots = \bigcup_{i \geq 0} L^i$
  - ▶ Le mot vide appartient donc à  $L^*$  mais pas forcément à  $L^+$  qui est définie comme suit:  
 $L^+ = L^1 \cup L^2 \cup L^3 \cup \dots = \bigcup_{i \geq 1} L^i = L^*.L$

# Opération sur les langages:

## Fermeture itérative

- ▶ La concaténation de langage est distributive par rapport à l'union.
- ▶ Soit  $X$  un alphabet et  $L_1$ ,  $L_2$  et  $L_3$  engendrés sur  $X^*$
- ▶  $L_1.(L_2 \cup L_3) = (L_1.L_2) \cup (L_1.L_3)$
- ▶  $(L_1 \cup L_2).L_3 = (L_1.L_3) \cup (L_2.L_3)$
- ▶ Si  $L_1 \subseteq L_2$  alors  $L_1.L_3 \subseteq L_2.L_3$