Théorie des Langages et Automates

2025-2026

Plan du cours

- Automates finis et langages réguliers
  - Notion de langage
  - Automates finis déterministes
  - Automates finis non déterministes + Déterminisation
  - Lemme de Pompage o Grammaires régulières o Expressions régulières
  - Equivalence entre automates finis, grammaires régulières et expressions régulières
  - Limites des langages réguliers
- Automates à pile, langages non contextuels
  - Automates à pile
  - Grammaires non contextuelles
  - Equivalence automates à pile et grammaires non contextuelles
  - Lemme de pompage
- Machines de Turing
  - Définitions
  - Langages Turing acceptables
  - Problème de l'arrêt

Introduction générale

- En linguistique et en informatique, on parle de théorie des langages pour décrire les langages formels.

- Plus généralement, la théorie des langages concerne tout langage fini ou infini qui peut être spécifié par une méthode ou un mécanisme fini, et explicite, qui permet de le produire ou de l'analyser.

- La théorie des langages ne se préoccupe pas du côte <u>sémantique</u> (le sens), mais plutôt syntaxique (la forme) d'un langage.

Introduction générale

- On définit un langage notamment grâce aux grammaires et on analyse les mots d'un langage grâce aux automates.

- Ce cours va s'articuler autour de ces trois concepts que sont les langages, les grammaires et les automates.

- La théorie des langages est une base pour un certain nombre de disciplines informatique dont principalement la compilation.

Chapitre 1: Automates finis et langages réguliers

Plan

- Rappel sur la théorie des ensembles
- Notion de langage
- Automates finis déterministes
- Automates finis non déterministes + Déterminisation
- Lemme de Pompage o Grammaires régulières o Expressions régulières
- Equivalence entre automates finis, grammaires régulières et expressions régulières
- Limites des langages réguliers

Rappel sur la théorie des ensembles

Définitions :

- Un ensemble est une collection d’objets sans répétition.

- Si un objet appartient à un ensemble A, on dit qu’il est élément de cet ensemble et l’on note $x \in A$.

- On distingue un ensemble particulier noté $\emptyset$ qui ne contient aucun élément.

- Un ensemble est noté par $\{...\}$ où les pointillés indiquent les éléments de l’ensemble ou une caractéristique de ses éléments

Rappel sur la théorie des ensembles

- Il existe principalement trois moyens pour définir un ensemble :
  - **Définition par extension:** consiste à donner tous les éléments d'un ensemble. Exemple: $\{0,1,2,3,4\}$.

  - **Définition par compréhension:** définit les éléments d'un ensemble par les propriétés qui les définissent. on écrira $\{x \ / \ x$ vérifie une propriété $P(x)\}$. Exemple : $\{n \in N \ / \ n \ \% \ 2 = 0\}$ définit les nombres entiers pairs.

  - **Définition par induction :** définit un ensemble par certains éléments triviaux et des règles d'induction permettant de retrouver d'autres éléments en fonction de ceux déjà connus. On écrit: $x \in A \Rightarrow f(x) \in A$ tel que f est un moyen permettant de construire d'autres éléments en fonction de l'argument. Exemple : l'ensemble des entiers peut être représenté comme suit : $N = \{0; x \in N \Rightarrow (x + 1) \in N\}$.

Opérations sur les ensembles

- **Comparaison des ensembles:**
  - **Inclusion:** $A \subseteq B$, si $\forall x \in A, x \in B$. On dira alors que A est un sous ensemble de B.
    - Exemple : $\{a, b, c\} \subseteq \{a, b, c, d\}$.
  - **Egalité:** Deux ensembles A et B sont égaux si $A \subseteq B$ et $B \subseteq A$.

Opérations sur les ensembles

- **Construction d'autres ensembles:**

  Soit A et B deux sous ensembles de de Ω:

  - L'union $A \cup B$: *comporte tout élément appartenant à A ou B*
  - L'intersection $A \cap B$: *comporte tout élément appartenant à A et B*
  - La différence $A - B$: *comporte tout élément appartenant à A et qui n'appartient pas à B*
  - Le complément, noté $\overline{A} = \Omega - A$
  - Le produit cartésien $A \times B$: *est l'ensemble des paires (a, b) telles que $a \in A$ et $b \in B$.*
  - L'ensemble des parties de A, noté 2^A : *est l'ensemble de tous les sous-ensembles de A*

Opérations sur les ensembles

- **Exemple : $A = \{a, b\}$, $B = \{a, c\}$, $\Omega = \{a, b, c\}$:**
  - $A \cap B = \{a\}$;
  - $A \cup B = \{a, b, c\}$;
  - $A - B = \{b\}$;
  - $\overline{A} = \{c\}$;
  - $A \times B = \{(a, a), (a, c), (b, a), (b, c)\}$;
  - $2^A = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$

Opérations sur les ensembles

- **Mesures sur les ensembles:**
  - La mesure la plus utilisée et celle de la cardinalité. Elle est notée card(A) et représente le nombre d'éléments de A.
  - Si l'ensemble est infini alors sa cardinalité est ∞.
  - $\mathrm{card}(A \times B) = \mathrm{card}(A).\mathrm{card}(B)$
  - $\mathrm{card}(2A) = 2\mathrm{card}(A)$.

Notion de langage

- La théorie des langages utilise un certain nombre de concepts ainsi qu'une certaine terminologie.

- Nous allons d'abord définir certaines notions capitales qui sont les bases de cette théorie:

  - Alphabet

  - Mot

  - Langage

Alphabet

- **Définitions:**
  - Un **alphabet** est un ensemble fini de symboles.
  - Un **symbole** est une entité abstraite qui ne peut pas être définie formellement comme les lettres, les chiffres, ..ect

- **Exemples :**
  - A = {0, 1}, l'alphabet binaire
  - B= {a, b, c}
  - C= {if , then, else, a, b}, un alphabet du langage C
  - F = {→, ←, ↑, ↓} un alphabet de fleches

Mot

- **Définition**:
  - Un mot sur l'alphabet *X est une séquence finie et ordonnée* d'éléments de l'alphabet.
  - C'est une concaténation de lettres.
- **Exemples:**
  - abbac et ba sont deux mots de l'alphabet {a,b,c}.
  - 00,01,10010 sont des mots de l'alphabet {0,1}
- **Notations:**
  - Le mot vide, noté ε, correspond à la suite de symboles vides.
  - L'ensemble des mots construits sur un alphabet X est un ensemble infini noté X^*
  - L'ensemble des mots construits sur un alphabet X et qui ne contient pas le mot vide est un ensemble infini noté X^+

Mot

- **Longueur d'un mot:**

  Soient X est un alphabet, $w \in X^*$, et x est un des symboles constituant le mot w:

  - $|w|$ est la longueur du mot **w**
  - $|w|_x$ est le nombre de l'occurrence de x dans le mot w

- **Exemples**:
  - $|abbac| = 5$
  - $|ba| = 2$
  - $|\varepsilon| = 0.$
  - $|ababac|_a = 3$

Mot: Opérations sur les mots

- **Concaténation**
  Soient deux mots w,w' ∈ X*, la concaténation de w et w' est définie comme la juxtaposition de w et w'. Elle est notée par w.w' ou bien ww'.
- Exemple:
  - X={0,1}, w=10, w'=00
    w.w'=1000
    w'.w=0010

Mot: Opérations sur les mots

- **Puissance d'un mot**

  Soit un alphabet X, w ∈ X*et n ∈ N , la puissance de w est donnée comme suit:

$$w^n = \begin{cases} \varepsilon \text{ si } n = 0 \\ w \text{ si } n = 1 \\ ww^{n-1} = w^{n-1}w \text{ si } n > 1 \end{cases}$$

- Exemple:

  Soit $X = \{a, b\}$ et $w = aba$

  - $w^0 = \varepsilon$
  - $w^1 = w.\varepsilon = w = aba$
  - $w^2 = w.w = ww = abaaba$

Mot: Opérations sur les mots

- **Factorisation d'un mot**

  Soit un alphabet X, w ∈ X*

- u est un facteur gauche (préfixe) de $w \Leftrightarrow \exists\, v \in X^*$ tel que $w = uv$
- u est un facteur droit (suffixe) de $w \Leftrightarrow \exists\, v \in X^*$ tel que $w = vu$
- u est un préfixe propre de $w \Leftrightarrow \exists\, v \in X^+$ tel que $w = uv$
- u est un suffixe propre de $w \Leftrightarrow \exists\, v \in X^+$ tel que $w = vu$

- **Exemple:**

  Soit $X = \{a, b\}$ et $w = babb$:

  - les préfixes de w sont : b, ba, bab, babb
  - les suffixes de w sont : b, bb, abb, babb
  - les préfixes propres de w sont : b, ba, bab
  - les suffixes propres de w sont : b, bb, abb

Mot: Opérations sur les mots

- **Inverse d'un mot ou Miroir**
  le miroir d'un mot $w=a_1a_2\ldots\ldots a_n$ est le mot noté $w^R = a_n\ldots\ldots a_2a_1$ obtenu en inversant les symboles de w.

- **Exemple:**
  - Soit w=abbc un mot de l'alphabet x={a,b,c}:
    w^R =cbba

- **Remarques:**
  - Le miroir d'un mot composé d'un seul symbole est le mot lui-même
  - $\varepsilon^R = \varepsilon$
  - Un mot est un palindrome si $w^R = w$, exple: $(aba)^R = aba$