X5I0050 - Langages et automates
Introduction

D. Béchet & T. Sadiki

Université de Nantes & Université Internationale de Rabat

10 septembre 2014

Généralités

Objectif : acquérir les bases de la théorie des langages et des automates

Notions principales : bases théoriques, principaux algorithmes, outils

Vocabulaire, langages, grammaires, classification de Chomsky, langages rationnels (ou réguliers), expressions rationnelles (ou régulières), langages algébriques (ou non contextuels ou hors-contextes), machine de Turing, automates finis déterministes et non-déterministes

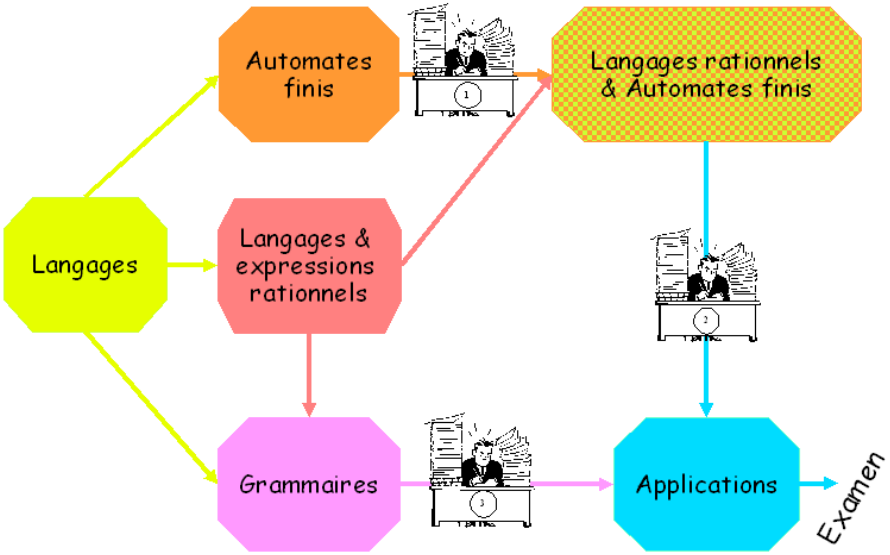
Prérequis

- Mathématiques : théorie des ensembles, démonstration par l'absurde et démonstration par récurrence
- Informatique : algorithmique, programmation impérative

Contenu

1. **Cours 1** généralités sur les langages : monoïdes, mots et monoïdes, relations entre mots, langages et opérations
2. **Cours 2** langages rationnels (ou réguliers) : expressions rationnelles, définitions rationnelles, systèmes d'équations rationnelles, lemme de l'étoile
3. **Cours 3** grammaires : hiérarchie de Chomsky
4. **Cours 4** automates finis : Turing, AFN, AFN et langages, propriétés des états, propriétés des AFN, rendre un automate déterministe, minimisation d'un AFD, combinaison d'automates
5. **Cours 5** langages rationnels et reconnaissables : équivalence entre AFN et expressions rationnelles, etc.
6. **Cours 6** applications : extraction d'informations, compilation, TALN

Dépendances entre les chapitres



Bibliographie

Livres :

- A. Aho, R. Sethi & J. Ullman, "Compilateurs : principes, techniques et outils", InterEditions, 1991.
- J.-M. Autebert, "Théorie des langages et des automates", Masson, 1994.
- P. Linz, "Formal Languages and Automata", Jones and Barnett Publishers, 2006.
- Patrice Séébold, "Théorie des automates : méthodes et exercices corrigés", série Passeport pour l'informatique, Ed. Vuibert, Paris, 1999, ISBN 2-7117-8630-7, www.vuibert.fr
- P. Verdret, "De Perl à Java : programmation des expressions régulières", Hermès, 2005

Nombreux cours en ligne...

Attention aux particularités de ce cours pour les examens

Outil pédagogique :

- JFLAP, Susan H. Rodger, http ://www.jflap.org/

Ensembles

Définition - ensemble

Un **ensemble** est une collection d'objets uniques d'un univers U. Chaque objet d'un ensemble est appelé un **élément** de cet ensemble.
Soit E un ensemble, a un élément de E, on note $\mathbf{a} \in \mathbf{E}$

Définition par extension :
$J = \{lundi, mardi, mercredi, jeudi, vendredi, samedi, dimanche\}$
Définition par intension :
$J = \{j \mid j \text{ est un jour de la semaine}\}$
L'**ensemble vide** ne contient aucun élément. Il est noté $\{\}$ ou $\emptyset$

Définition - sous-ensemble

Soit E un ensemble, A est un **sous-ensemble** de E noté $\mathbf{A} \subseteq \mathbf{E}$ si chaque élément de A est un élément de E.

Ensembles - exemples

- Dans l'univers des nombres :
  - L'ensemble des entiers positifs ou nuls $\{0, 1, 2, 3, ...\}$ est noté $\mathbb{N}$
  - $3 \in \mathbb{N}$ est vrai
  - $1,5 \in \mathbb{N}$ est faux. On note $1,5 \notin \mathbb{N}$
- Dans l'univers des êtres vivants :
  - Les chiens forment un ensemble noté $\mathcal{C}$
  - Les mammifères forment un ensemble noté $\mathcal{M}$
  - Nous avons $\mathcal{C} \subseteq \mathcal{M}$ car tous les chiens sont des mammifères.

Opération sur les ensembles

Définition - Complément, union, intersection, différence

Soient A et B des ensembles dans un univers U :
Complément de A : $\overline{\mathbf{A}}$ ou $\mathbf{Comp(A)} = \{x \mid x \in U \text{ et } x \notin A\}$;
Union : $\mathbf{A} \cup \mathbf{B} = \{x \mid x \in A \text{ ou } x \in B\}$;
Intersection : $\mathbf{A} \cap \mathbf{B} = \{x \mid x \in A \text{ et } x \in B\}$;
Différence (ou exclusion) : $\mathbf{A} \backslash \mathbf{B} = \{x \mid x \in A \text{ et } x \notin B\}$;

Attention : la définition du complément de A dépend de l'univers U

Opération sur les ensembles – exemples

- Soit P l'ensemble des entiers pairs $\{0, 2, 4, ...\}$
- $\overline{P}$ est l'ensemble des entiers qui ne sont pas pairs
- $\overline{P}$ est l'ensemble des entiers impairs
- $\overline{P} = \{1, 3, 5, ...\}$
- $\overline{P}$ est l'ensemble des entiers qui ne sont pas divisibles par 2
- $\overline{P}$ est l'ensemble des entiers dont le reste de la division par 2 est 1
- $\overline{P}$ est l'ensemble des entiers dont le modulo 2 est 1
- $\{1, 3\} \cup \{1, 2\} = \{1, 2, 3\}$
- $\{1, 3\} \cap \{1, 2\} = \{1\}$
- $\{1, 3\} \backslash \{1, 2\} = \{3\}$

Propriétés des opérations sur les ensembles

Théorème - Propriétés des opérations sur les ensembles

Soient A, B et C des ensembles de l'univers U :

Associativité : $(A \cup B) \cup C = A \cup (B \cup C)$ et $(A \cap B) \cap C = A \cap (B \cap C)$

Commutativité : $A \cup B = B \cup A$ et $A \cap B = B \cap A$

Complémentation : $A \cup \overline{A} = U$ et $A \cap \overline{A} = \emptyset$

Idempotence : $A \cup A = A$ et $A \cap A = A$

Identité : $A \cup \emptyset = A$ et $A \cap U = A$

Extrémité : $A \cup U = U$ et $A \cap \emptyset = \emptyset$

Involution : $\overline{(\overline{A})} = A$

Lois de De Morgan : $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$ et $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$

Distributivité : $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ et $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Récurrence simple

Principe de la récurrence simple

Soit un ensemble infini de propositions $\{T_n \mid n \in \mathbb{N}\} = \{T_0, T_1, T_2, ...\}$
$\forall n \in \mathbb{N}$, T_n est vraie ssi

1. T_0 est vraie
2. $\forall k \in \mathbb{N}$, si T_k est vraie alors T_{k+1} est vraie.

Démonstration par induction simple

La proposition doit être vraie pour $n = 0$:

Démontrer que T_0 est vraie

Supposer que la proposition est vraie pour $k \in \mathbb{N}$ et passer à $k + 1$:

Énoncer l'hypothèse d'induction : T_k est vraie puis

Démontrer que T_{k+1} est vraie en utilisant l'hypothèse d'induction

Conclusion : la proposition est vraie $\forall n \in \mathbb{N}$ par induction simple

Récurrence simple - Exemple

Soit la propriété T_n : $n(n+1)/2 = \sum_{1 \leq i \leq n} i = 1 + 2 + 3 + \cdots + n$
Montrons par une récurrence simple que cette propriété est vraie pour tous les entiers positifs ou nuls

1. **Cas de base** pour $n = 0$: nous montrons que T_0 est vraie
   Pour T_0 : $0(0+1)/2 = 0$ et $\sum_{1 \leq i \leq 0} i =$ somme vide $= 0$
   $\Rightarrow$ T_0 est vraie
2. **Cas général** pour $k \in \mathbb{N}$: Nous supposons que T_k est vraie
   Nous allons montrer sous cette hypothèse que T_{k+1} est vraie
   - La partie gauche de T_{k+1} :
     $(k+1)((k+1)+1)/2 = (k(k+2)+k+2)/2 =$
     $k(k+1)/2 + (k+k+2)/2 = k(k+1)/2 + k + 1$
   - Or T_k est vraie : $k(k+1)/2 = \sum_{1 \leq i \leq k} i$
   - Donc $(k+1)((k+1)+1)/2 = \sum_{1 \leq i \leq k} i + k + 1 = \sum_{1 \leq i \leq k+1} i$
     $\Rightarrow$ T_{k+1} est vraie
3. Conclusion : $\forall n \in \mathbb{N}$, T_n est vraie

Récurrence généralisée

Principe de la récurrence généralisée

Soit un ensemble de propositions $\{T_n \mid n \in \mathbb{N}\}$
$\forall n \in \mathbb{N}$, T_n est vraie ssi

1. $\forall k \in \mathbb{N}$, si T_i sont vraies $\forall i < k$ alors T_k est vraie.

Démonstration par induction généralisée

Supposer que la proposition est vraie $\forall i, 0 \leq i < k$ et passer à k

Énoncer l'hypothèse d'induction : $\forall i, 0 \leq i < k$, T_i sont vraies puis

Montrer que T_k est vraie en utilisant l'hypothèse d'induction

Conclusion : la proposition est vraie $\forall n \in \mathbb{N}$ par induction généralisée

Il n'y a pas de cas de base mais il est souvent utile de traiter à part T_0

Récurrence généralisée - Exemple

Sur l'exemple $T_n : n(n+1)/2 = \sum_{1 \leq i \leq n} i = 1 + 2 + 3 + \cdots + n$
Montrons par une récurrence généralisée que cette propriété est vraie pour tous les entiers positifs ou nuls

1. Soit $k \in \mathbb{N}$. Nous supposons que $\forall i, 0 \leq i < k \;\; T_i$ est vraie
   Nous allons montrer sous cette hypothèse que T_k est vraie
   - Si $k = 0$: dans ce cas l'hypothèse d'induction est vide mais
     $0(0+1)/2 = 0$ et $\sum_{1 \leq i \leq 0} i = \text{somme vide} = 0$
     $\Rightarrow T_k$ est vraie (dans le cas où $k = 0$)
   - Si $k > 0$,
     $T_k : k(k+1)/2 = k(k-1+2)/2 = (k(k-1)+2k)/2 = (k-1)k/2 + k$
     Or $0 \leq k-1 < k$. Donc T_{k-1} est vraie et $(k-1)k/2 = \sum_{1 \leq i \leq k-1} i$
     $\Rightarrow k(k+1)/2 = \sum_{1 \leq i \leq k-1} i + k = \sum_{1 \leq i \leq k} i$
     $\Rightarrow T_k$ est vraie (dans le cas où $k > 0$)
   - $\Rightarrow T_k$ est vraie (dans tous les cas sous l'hypothèse d'induction)
2. Conclusion : $\forall n \in \mathbb{N}, T_n$ est vraie