Langages formels et automates – cours 1

Introduction et rappels de maths discrètes

Catalin Dima

Objectifs du cours

- Introduction en théorie d'automates finis.
- Modélisation des systèmes par des automates finis
  - Parallélisme, communications synchrones et asynchrones, implémentation et abstraction.
- Algorithmes d'automates finis.
- Dualité automates/logique.
- Grammaires et hiérarchie de Chomsky.
- Automates à pile et grammaires hors contexte.

Ressources et évaluation

- J.M. Autebert, Théorie des langages et des automates, Masson 1994 (bibliothèque P12).
- J. Hopcroft, R. Motwani, J. Ullman, Automata Theory, Languages, and Computation, Pearson/Addison Wesley, 2006.
- Page web : `http://lacl.univ-paris12.fr/dima/.....`

Ressources et évaluation

- J.M. Autebert, Théorie des langages et des automates, Masson 1994 (bibliothèque P12).
- J. Hopcroft, R. Motwani, J. Ullman, Automata Theory, Languages, and Computation, Pearson/Addison Wesley, 2006.
- Page web : `http://lacl.univ-paris12.fr/dima/.....`
- Evaluation contrôle continu :
  - 2 contrôles en TD.
- Exos “devoir maison” :
  - 2 exos rendus (dans les délais) vaut 1 contrôle.
- Examen :
  - Exos de type “prouver que”
  - Exos de type “donner un algorithme”.
  - Exos “construire un automate/une grammaire/modéliser le système”.
  - Exos “prouver qu’on ne peut pas construire un automate/une grammaire”.

Pourquoi théorie des automates ?

- Automate fini = modèle élémentaire de système informatique.
  - État de l'automate = état du système.
  - Transition dans l'automate = transition d'état dans le système.
  - Trace de l'automate = comportement du systême.
  - .....
- Modélisation des notions fondamentales en informatique :
  - Modularité = intersection d'automates.
  - Implémentation = inclusion de langages.
  - Synchronisation, indépendence, comportements défaillants,...
- Bases de la compilation :
  - Arbres d'analyse syntaxique.
  - Traducteurs de langages.
  - Recherche de motifs.

Techniques et outils en théorie des automates

- Algorithmes de graphes.
- Preuves par induction.
- Preuves par réduction à l'absurde.
- Éléments d'algèbre des groupes/anneaux.
- Analyse combinatoire.
- Éléments de complexité algorithmique.

Preuves

- Preuves déductives :

$$p, p \rightarrow q \vdash q \qquad \text{(Modus ponens)}$$

- Exemple ....

Preuves

- Preuves déductives :

$$p, p \to q \vdash q \qquad \text{(Modus ponens)}$$

- Exemple ....
- Preuves inductives :

$P(n)$: formule en logique de 1er ordre, avec une variable libre : n

$$P(0), \forall n \big(P(n) \to P(n+1)\big) \vdash \forall n P(n)$$

- Exemple :

$$\sum_{i=1}^{n} i = ?$$

Preuves

- Preuves déductives :

$$p, p \rightarrow q \vdash q \qquad \text{(Modus ponens)}$$

- Exemple ....
- Preuves inductives :

$P(n)$: formule en logique de 1er ordre, avec une variable libre : n

$$P(0), \forall n\big(P(n) \rightarrow P(n+1)\big) \vdash \forall n P(n)$$

- Exemple :

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

Ensembles

- Formules de définition :

$$A = \{x \mid \phi(x)\}$$

  où $\phi(x)$ est une formule de premier ordre ayant la variable x libre (non-quantifiée).

- Égalité d'ensembles :

$$\{x \mid \phi(x)\} = \{x \mid \psi(x)\}$$

  - Preuve d'une équivalence : $\forall x. \phi(x) \Leftrightarrow \psi(x)$.
  - Donc deux preuves : $\forall x. \phi(x) \rightarrow \psi(x)$ et $\forall x. \psi(x) \rightarrow \phi(x)$.

$$\left\{n \in \mathbb{N} \mid n = 4a + 6b, a, b \in \mathbb{N}\right\} = \left\{2k \mid k \in \mathbb{N}, k \neq 1\right\}$$

Relations

- Relation : $\rho \subseteq A \times B$.
  - Deux fonctions associées $\rho : \mathcal{P}(A) \longrightarrow \mathcal{P}(B)$ et $\rho^{-1} : \mathcal{P}(B) \longrightarrow \mathcal{P}(A)$:

$$\rho(X) = \{y \in B \mid \exists x \in X, x \rho y\}$$
$$\rho^{-1}(Y) = \{x \in A \mid \exists y \in Y, x \rho y\}$$

Relations

- Relation : $\rho \subseteq A \times B$.
  - Deux fonctions associées $\rho : \mathcal{P}(A) \longrightarrow \mathcal{P}(B)$ et $\rho^{-1} : \mathcal{P}(B) \longrightarrow \mathcal{P}(A)$:

$$\rho(X) = \{y \in B \mid \exists x \in X, x\rho y\}$$
$$\rho^{-1}(Y) = \{x \in A \mid \exists y \in Y, x\rho y\}$$

- Relation d'équivalence : $\rho \subseteq A \times A$ ayant les propriétés :
  - Reflexive : $\forall x \in A, x\rho x$.
  - Symmétrique : $\forall x, y \in A, x\rho y \Rightarrow y\rho x$.
  - Transitive : $\forall x, y, z \in A, x\rho y \wedge y\rho z \Rightarrow x\rho z$.
- Quotient d'un ensemble par rapport à une relation d'équivalence :

$$A/\rho = \{X \subseteq A \mid \forall x \in X, \forall y \in A, x\rho y \Rightarrow y \in X\}$$

  - Un élément $X \in A/\rho$ s'appelle classe d'équivalence.
  - Théorème Si $X, Y \in A/\rho$ avec $X \neq Y$, alors $X \cap Y = \emptyset$.
  - Notation : pour $x \in A$, la classe d'équivalence (unique !) à laquelle x appartient est notée $[x]_\rho$.

Relations (2)

- Relation d'ordre :
  - Reflexive, antisymmétrique et transitive.

$$\forall x, y \in A, x\rho y \wedge y\rho x \Rightarrow x = y$$

  - Représentation par diagramme de Hasse – on ne représente que les relations entre éléments consecutifs.
    - $x, y \in A$ sont consécutifs par rapport à ρ si :

$$\forall z \in A, z\rho x \wedge z\rho y \Rightarrow z = x \vee z = y$$

- Ordre total : $\forall x, y \in A, x\rho y \vee y\rho x$.
  - Exemples : $(\mathbb{N}, \leq)$, $(\Sigma^*, \leq_{lex})$.
- Produit d'ordres : étant donnés (A, ρ_A) et (B, ρ_B), on crée un nouvel ordre $(A \times B, \rho)$ où ρ est défini par

$$(x, x')\rho(y, y') \text{ si et seulement si } x\rho_A x' \text{ et } y\rho_B y'$$

Exercice

Si (A, ρ_A) et (B, ρ_B) sont totaux, est-ce que $(A \times B, \rho)$ sera un ordre total ?

Opérations binaires et structures algébriques

- Opération binaire sur un ensemble A : $\oplus : A \times A \longrightarrow A$.
- $(A, \oplus)$ est un semigroupe si :
  - $\oplus$ est associative : $\forall x, y, z \in A, x \oplus (y \oplus z) = (x \oplus y) \oplus z$.
- $(A, \oplus, e)$ est un monoïde si
  - $\oplus$ est associative et
  - e est élément neutre pour $\oplus$: $\forall x \in A, x \oplus e = e \oplus x = x$.
- $\oplus$ est commutative si $\forall x, y \in A, x \oplus y = y \oplus x$.

Opérations binaires et structures algébriques

- Opération binaire sur un ensemble A : $\oplus : A \times A \longrightarrow A$.
- $(A, \oplus)$ est un semigroupe si :
  - $\oplus$ est associative : $\forall x, y, z \in A, x \oplus (y \oplus z) = (x \oplus y) \oplus z$.
- $(A, \oplus, e)$ est un monoïde si
  - $\oplus$ est associative et
  - e est élément neutre pour $\oplus$: $\forall x \in A, x \oplus e = e \oplus x = x$.
- $\oplus$ est commutative si $\forall x, y \in A, x \oplus y = y \oplus x$.
- $(A, \oplus, \otimes, \mathbf{0}, \mathbf{1})$ est un semianneau si
  - $(A, \oplus, \mathbf{0})$ est un monoïde *commutatif*.
  - $(A, \otimes, \mathbf{1})$ est un monoïde.
  - $\mathbf{0}$ est absorbant pour $\otimes$: $\forall x \in A, x \otimes \mathbf{0} = \mathbf{0} \otimes x = \mathbf{0}$.
  - $\otimes$ est distributive par rapport à $\oplus$:
    $\forall x, y, z \in A, x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$.

Alphabets, mots, langages

- Alphabet = ensemble fini de lettres/symboles.

$$\Sigma_1 = \{a, b, c, d, e\} \qquad\qquad \Sigma_2 = \{0, 1\}$$

Alphabets, mots, langages

- Alphabet = ensemble fini de lettres/symboles.

$$\Sigma_1 = \{a, b, c, d, e\} \qquad\qquad \Sigma_2 = \{0, 1\}$$

- Mot sur un alphabet Σ = séquence finie de lettres.

$$abbabcddacacacdebb \qquad 010001110101110$$

Alphabets, mots, langages

- Alphabet = ensemble fini de lettres/symboles.

$$\Sigma_1 = \{a, b, c, d, e\} \qquad\qquad \Sigma_2 = \{0, 1\}$$

- Mot sur un alphabet Σ = séquence finie de lettres.

$$abbabcddacacacdebb \qquad 010001110101110$$

- Langage sur un alphabet Σ = ensemble de mots

$L_1 = \{a, ab, abc, abcde\}$ langage sur Σ_1
$L_2 = \{a^n \mid n \in \mathbb{N}\}$ langage sur Σ_1
$L_3 = \{1w0 \mid w \text{ nombre binaire quelconque}\}$ langage sur Σ_2

Alphabets, mots, langages

- Alphabet = ensemble fini de lettres/symboles.

$$\Sigma_1 = \{a, b, c, d, e\} \qquad\qquad \Sigma_2 = \{0, 1\}$$

- Mot sur un alphabet Σ = séquence finie de lettres.

$$abbabcddacacacdebb \qquad 010001110101110$$

- Langage sur un alphabet Σ = ensemble de mots

$$L_1 = \{a, ab, abc, abcde\} \text{ langage sur } \Sigma_1$$
$$L_2 = \{a^n \mid n \in \mathbb{N}\} \text{ langage sur } \Sigma_1$$
$$L_3 = \{1w0 \mid w \text{ nombre binaire quelconque}\} \text{ langage sur } \Sigma_2$$

- Ensemble (langage) de tous les mots sur Σ noté Σ^*.

Opérations et relations sur les mots

- Concaténation : $abcc \cdot cdaba = abcccdaba$.
- Mot vide : ε.
  - $(\Sigma^*, \cdot, \varepsilon)$ est un monoïde

Opérations et relations sur les mots

- Concaténation : $abcc \cdot cdaba = abcccdaba$.
- Mot vide : ε.
  - $(\Sigma^*, \cdot, \varepsilon)$ est un monoïde (commutatif ?...)

Opérations et relations sur les mots

- Concaténation : $abcc \cdot cdaba = abcccdaba$.
- Mot vide : ε.
  - $(\Sigma^*, \cdot, \varepsilon)$ est un monoïde
- Longueur : $|abccbabcba| = 10$
  - $|\cdot| : \Sigma^* \longrightarrow \mathbb{N}$ est un morphisme de monoïdes.
- Nombre de lettres : $\#_a(abccbabcba) = 3$.
  - $\#_a : \Sigma^* \longrightarrow \mathbb{N}$ est lui aussi un morphisme de monoïdes.
- Préfixe : $abcde \preceq abcdeeabece$ mais $abcde \not\preceq ababcde$
  - $(\Sigma^*, \preceq)$ est une relation d'ordre totale.
  - Similaire : suffixe, infixe.
- Sous-mot :

$$abcd \sqsubseteq abbabacbd$$

Opérations sur les langages

- Concaténation étendue aux langages.

$$L_1 \cdot L_2 = \left\{ w_1 \cdot w_2 \mid w_1 \in L_1, w_2 \in L_2 \right\}$$

  - $\left(\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\}\right)$ est un semianneau.
  - Qu'en est-il de $\left(\mathcal{P}(\Sigma^*), \cap, \cdot, ?, \{\varepsilon\}\right)$?
- Quotient gauche de langages :

$$w \backslash L = \left\{ z \in \Sigma^* \mid wz \in L \right\}$$

  - Peut s'étendre à deux langages $L_1 \backslash L_2$.
  - Propriété de liaison avec l'union ?
- Quotient droit de langages :

$$L \backslash w = \left\{ z \in \Sigma^* \mid zw \in L \right\}$$

Opérations sur les langages (2)

- Étoile d'un langage :

$$L^* = \{w_1 w_2 \dots w_k \mid k \in \mathbb{N}, w_1, w_2, \dots, w_k \in L\}$$
$$L^+ = \{w_1 w_2 \dots w_k \mid k \geq 1, w_1, w_2, \dots, w_k \in L\}$$

  - $\varepsilon \in L^*$ pour tout L, même quand $L = \emptyset$!
  - Exemple...
- Shuffle de mots :

$$w_1 \sqcup\!\sqcup w_2 = \{a_1 b_1 a_2 \dots a_k b_k \mid k \in \mathbb{N}, w_1 = a_1 a_2 \dots a_k, w_2 = b_1 b_2 \dots b_k\}$$

  - Exemple...
  - Peut s'étendre à deux langages $L_1 \sqcup\!\sqcup L_2$.
  - $(\mathcal{P}(\Sigma^*), \cup, \sqcup\!\sqcup, \emptyset, \{\varepsilon\})$ est un semianneau aussi !