NGUYEN Thi Minh Tuyen

Théorie des automates et langages formels

Mots et Langages

Définitions

- alphabet Σ – un ensemble fini de caractères
  - Exemple:
    - ∑ = {a,b,c},
    - ∑ = { , , ♤, ♧},
    - Δ = {0,1},
    - Δ={🡨,🡪,🡩,🡫}
- lettre – ou caractère, est un élément de l'alphabet Σ
  - Exemple:
    - a, b, c,
    - , , ♤, ♧,
    - 0,1,
    - 🡨,🡪,🡩,🡫

Définitions

- Mot – est une suite finie de caractères dans Σ
- Exemple:
  - abbac, ba sont deux mots sur l'alphabet {a,b,c}
- Λ (mot vide) – le mot qui ne contient aucun caractères
- Σ* – l'ensemble de tous les mots (y compris Λ) sur l'alphabet Σ.
- Exemple:
  - {a, b, c}* = {Λ, a, b, c, aa, ab, ac, ba, bb, bc, ca, cb, cc, aaa, aab,...}
- langage sur un alphabet Σ est un ensemble de mots sur Σ. (sous ensemble de Σ*)

Exemples

	Mots-Français	Phrases-Français
alphabet		
caractère		
mot		
langage		

Langage

- Il est difficile de décrire le langage Français avec un nombre fini de règles.
- On ne peut pas décrire le langage en écrivant la liste de toutes les phrases acceptables
- Les règles grammaticales seules, ne permettent pas de décrire les phrases acceptables:

“j'ai mangé deux Lundi”

Langage

- Langages intéressants sont ceux avec les propriétés suivantes:
  - Le contenu doit être défini sans ambiguïté
  - A l'aide d'une formule, propriété ou un nombre fini de règles (grammaire) on peut reconnaître si un mot est dans le langage ou pas en un temps fini. (Syntaxe, non pas la sémantique)
- L'alphabet est un ensemble fini, mais les langages intéressants sont la plupart du temps infinis.

Exemple

- Σ={x}
- L_1={x, xx, xxx, xxxx,…} ou L_1 = {x^n | n=1, 2, 3,…}
- Σ^* = {Λ, x, xx, xxx, xxxx,…} = {x^n | n=0, 1, 2, 3,…}

  On dénote par $x^0 = \Lambda$
- L_2 = {w dans Σ^*: w a un nombre impair de caractères}

  = {x, xxx, xxxxx, … }

  = {x^n | n=1, 3, 5, … }

Opérations sur les mots

- longueur – est le nombre de caractères du mot
  - longueur (xxxxx) = 5
  - longueur (1025)=4
  - longueur (Λ)=0
- $\Sigma = \{0, 1\}$
- L_1 = ensemble des mots dans Σ^* qui commencent par 1 et qui ont une longueur au plus 3}

  = {1, 10, 11, 101, 100, 110, 111}

Inverse

- L'inverse d'un mot $u_1u_2 \ldots u_{m-1}u_m$ est le mot $u_mu_{m-1} \ldots u_2u_1$
- Exemples:
  - inverse(abc) = cba
  - inverse (xxx)=xxx
  - inverse (157)=751
- PALINDROME
  - PALINDROME={Λ et tout mot w dans Σ*:inverse(w)= w}
  - Si Σ= {a, b} alors

  PALINDROME = {Λ , a, b, aa, bb, aaa, aba, bbb, bab, ... }

Enchaînement de deux mots

- Enchaînement (ou concaténation) de deux mots est l'enchaînement du mot $u=u_1u_2\ldots u_m$ avec le mot $v=v_1v_2 \ldots v_n$ est le mot $u_v = u_1u_2\ldots u_mv_1v_2 \ldots v_n$.
- Les mots u et v sont les facteurs de cet enchaînement.
  - u=xx v=xxx u_v=xxxxx
  - u = abb v=aa u_v=abbaa
- Propriété:
  - longueur(u_v) = longueur (u) + longueur (v)

Enchaînement de deux langages

- Si L_1 et L_2 sont deux langages sur un alphabet Σ, alors l'enchaînement de L_1 avec L_2 est le langage

  $$L_1L_2 = \{uv : u \in L_1 \text{ et } v \in L_2\}$$

- Exemple:
  - $\Sigma = \{0, 1\}$
  - L_1 = {u dans Σ^*: le nombre de 0s dans u est pair}
  - L_2 = {u dans Σ^*: u commence par 0 et tous les autres caractères « s'il y en a » sont des 1}
  - L_1L_2 = {u dans Σ^*: le nombre de 0s dans u est impair}
  - L_2L_1 = {u dans Σ^*: u commence par zéro et le nombre de zéros dans u est impair}

Fermeture

- On appelle Σ^* la fermeture, ou Etoile de Kleene, de l'alphabet Σ.
- Exemples:
  - $\Sigma = \{x\}$

    $\Sigma^* = \{\Lambda, x, xx, xxx, ...\}$
  - $\Sigma = \{0, 1\}$

    $\Sigma^* = \{\Lambda, 0, 1, 00, 01, 10, 11, 000, 001, ...\}$
  - $\Sigma = \{a, b, c\}$

    $\Sigma^* = ?$

Fermeture d'un langage

- C'est un cas particulier de la Fermeture, ou « Étoile Kleene » d'un langage:
  - Soient un alphabet Σ et un ensemble de mots S dans Σ^*. La fermeture de S, dénoté S^*, est l'ensemble des mots

$$S^* = \{u \text{ dans } \Sigma^* : u = u_1 u_2 \ldots u_m \text{ , où } u_1, u_2, \ldots, u_m$$

sont tous dans $S\}$

- S^* est l'ensemble des mots qui peuvent se mettre sous la forme d'enchaînement de mots de S.

Exemples [1]

- S = {a, ab}
- S* = {Λ, a, aa, ab, aaa, aab, aba, aaaa, ...}
- abaaababa ∈ S* (ab|a|a|ab|ab|a facteurs)
- S* = {Λ ainsi que tous les mots qui commencent par a et qui ne contiennent pas deux b consécutif}
- Est-ce que la factorisation des mots de S* est unique?

Exemples [2]

- L = {xx, xxx}

  xxxxxxx ∈ L*

  xx|xx|xxx   xx|xxx|xx   xxx|xx|xx

- L* = {Λ ainsi que tous les mots qui ont au moins deux **x**}

  = {x^n : n≠1}

  Si L = Ø   Alors   L* = {Λ }

- La fermeture d'un langage L est toujours infinie sauf si L = Ø ou L ={Λ }.

Exemple

- S = {a, b, ab} T = {a, b, bb}
- Est-ce que S* = T*?
- S* = T* bien que S ≠ T
  - ab|a|a|ab|ab|a
  - a|b|a|a|a|b|a|b|a

Définition: L+

- Le langage qui contient tous les enchaînements avec au moins 1 mot de L.
- Si Λ est dans L alors, L* = L+. Sinon L+= L* - {Λ }.

  Σ+ = Σ* -{Λ }
- Exemples:
  - Σ = {x}

    Σ+ = {x, xx, xxx, …}
  - S = {aa, bbb, Λ }

    S+ = {aa, bbb, Λ , aaaa, aabbb, …} = S*

    (aaΛ = aa)

Théorème 1

- Si S est un ensemble de mots alors $S^{**} = S^{*}$
- Définitions:
  - Égalité entre deux ensembles

    $S = T$: $S \subset T$ et $T \subset S$
  - Sous-ensemble:

    $S \subset T$: pour tout x dans S, x est aussi dans T.
- Exemple: $S = \{a,b\}$
  - aaba, baaa, aaba $\in S^{*}$
  - aaba|baaa|aaba $\in S^{**}$
  - a|a|b|a|b|a|a|a|a|a|b|a $\in S^{*}$

Preuve du Théorème 1: S*=S**

- **$S^{**} \subset S^{*}$**

Chaque mot de S^{**} est l'enchaînement d'un nombre fini de mots de S^{*}. Puisque chaque mot de S^{*} est l'enchaînement d'un nombre fini de mots de S, alors chaque mot de S^{**} est l'enchaînement d'un nombre fini de mots de S. Donc $S^{**} \subseteq S^{*}$.

- **$S^{*} \subset S^{**}$**

Il est clair que pour tout ensemble de mot S, $S \subseteq S^{*}$ (un mot est l'enchaînement de lui même) et donc $S^{*} \subseteq S^{**}$.

- Conclusion : $S^{*} = S^{**}$.

Définitions récursives

- Pour définir un langage récursivement:

3 étapes:

1. Spécifier les mots de base.
2. Règles pour construire des nouveaux mots à partir des mots déjà connus.
3. Déclarer que tout mot qui n'est pas construit en suivant les règles 1 et 2 n'appartient pas au langage.

- La même méthode permet de définir les ensembles en général.

Exemple

- PAIR est l'ensemble de nombres entiers divisible par 2.
- PAIR = {2n | n = 1, 2, 3, 4, …}
- PAIR est construit par les règles
  1. 2 appartient à PAIR.
  2. Si x appartient à PAIR, x+2 appartient aussi à PAIR.
  3. Tous les éléments de PAIR sont construits en suivant 1 et 2.

- **Théorème: 14 appartient à PAIR**
  - Divisible par 2? Oui, 14/2 = 7.
  - 14 = 2n? Oui, n = 7.
- **Règle 1: 2 ∈ PAIR**
  - Règle 2: x=2, 2+2 = 4 ∈ PAIR
  - Règle 2: x=4, 4+2 = 6 ∈ PAIR
  - Règle 2: x=6, 6+2 = 8 ∈ PAIR
  - Règle 2: x=8, 8+2 = 10 ∈ PAIR
  - Règle 2: x=10, 10+2 = 12 ∈ PAIR
  - Règle 2: x=12, 12+2 = 14 ∈ PAIR

Définition alternative de l'ensemble PAIR

1. 2 appartient à PAIR.
2. Si x et y appartiennent à PAIR, x+y appartient aussi à PAIR.
   - Règle 1: 2 ∈ PAIR
   - Règle 2: x=2, y=2, 2+2 = 4 ∈ PAIR
   - Règle 2: x=4, y=2, 4+2 = 6 ∈ PAIR
   - Règle 2: x=4, y=4. 4+4 = 8 ∈ PAIR
   - Règle 2: x=6, y=8, 6+8 = 14 ∈ PAIR

Définition récursive de l'ensemble POLYNOMIAL

1. Tout nombre appartient à POLYNOMIAL.
2. La variable x appartient à POLYNOMIAL.
3. Si p et q appartiennent à POLYNOMIAL, alors p+q, p – q, (p), et pq appartiennent aussi à POLYNOMIAL.
4. Tous les éléments de POLYNOMIAL sont construits en suivant les règles 1, 2 et 3.

Preuve

Prover le théorème: $5x^2-8x+7 \in$ POLYNOMIAL

- Règle 1: $5 \in$ POLYNOMIAL
- Règle 2: $x \in$ POLYNOMIAL
- Règle 3: $5x \in$ POLYNOMIAL
- Règle 3: $5xx = 5x^2 \in$ POLYNOMIAL
- Règle 1: $8 \in$ POLYNOMIAL
- Règle 3: $8x \in$ POLYNOMIAL
- Règle 3: $5x^2 – 8x \in$ POLYNOMIAL
- Règle 1: $7 \in$ POLYNOMIAL
- Règle 3: $5x^2 – 8x + 7 \in$ POLYNOMIAL

Autres exemples: Donner la définition récursive [1]

- Σ^+, $\Sigma=\{x\}$
  - Règle 1: $x \in \Sigma^+$
  - Règle 2: Si w appartient à Σ^+, alors xw appartient à Σ^+
- Σ^*, $\Sigma=\{x\}$
  - Règle 1: $\Lambda \in \Sigma^*$
  - Règle 2: Si w appartient à Σ^*, alors xw appartient à Σ^*
- S-IMPAIR, $\Sigma=\{x\}$
  - Règle 1: x ∈ S-IMPAIR
  - Règle 2: Si w appartient à S-IMPAIR alors xxw appartient à S-IMPAIR

Autres exemples: Donner la définition récursive [2]

- Fermeture S* d'un langage S
  - **Règle 1:** Λ appartient à S*. Tout mot de S appartient à S*
  - **Règle 2**: Si x et y appartiennent à S*, la concaténation xy appartient à S*
- POSITIVE
  - Règle 1: 1, 2, 3, 4, 5, 6, 7, 8, 9 appartiennent à POSITIVE
  - Règle 2: Si w appartient à POSITIVE, w0, w1, w2, w3, w4, w5, w6, w7, w8, w9 appartiennent à POSITIVE

Autres exemples: EA (Expressions Arithmétiques)

- Avec Σ = {0,1,2,3,4,5,6,7,8,9,+, –,*,/,(,)}
  - Règle 1: Tout nombre appartient à EA
  - Règle 2: Si x appartient à EA, les mots suivants appartiennent à EA:
    1. (x)
    2. -x (sauf si la première lettre de x est -)
  - Règle 3: Si x et y appartiennent à EA, les mots suivants appartiennent à EA:
    1. x + y (sauf si la première lettre de y est -)
    2. x - y (même condition)
    3. x * y
    4. x / y
    5. x ** y

Théorème

- **Théorème**: Il n'y a pas d'élément de EA qui commence ou termine par le symbole /.
- Preuve:
  - Le symbole / ne peut être introduit par la règle 1.
  - Supposons que x ne commence ni termine par /. Alors, (x) et – x ne commence ni termine par /. (Règle 2)
  - Supposons que x et y ne commencent ni terminent par /. En appliquant la règle 3 on obtient une expression qui ne commence ni termine par /.
  - Puisque toute expression de EA est construite en appliquant successivement l'une des règles 1,2,3, alors toute expression dans AE ne peut commencer ni terminer avec le symbole /.

Autres exemples: FORMULES

Ensemble des formules de la logique propositionnelle

- **Règle 1**: Toute lettre de l'alphabet latin appartient à FORMULES
- **Règle 2:** Si p appartient à FORMULES, alors (p) et $\neg$p appartiennent à FORMULES
- **Règle 3:** Si p et q appartiennent à FORMULES, alors p $\rightarrow$ q appartient à FORMULES

Exercices

- Définir les opérations ou fonctions récursivement.
  - Fonction successeur d'un entier.
  - Addition de deux entiers positifs
  - Multiplication de deux entiers positifs.
  - Factorielle

Fonction successeur $\boldsymbol{\sigma} : N \rightarrow N$

- **D** = {0, 1, 2, …, 9}
  - **d** 0 1 2 3 4 5 6 7 8 9
  - **m(d)** 1 2 3 4 5 6 7 8 9 0
- **Règle 1**: Si **v** est dans **D** alors
  - Si **v** ≠ 9 alors **σ (v)** = **m(v)**
  - Si **v** = 9 alors **σ (v)** = 10
- **Règle 2**: Si **v** = **wd** où **d** est dans D alors
  - Si **d** ≠ 9 alors σ (v) = **wm(d)**
  - Si **d** = 9 alors **σ (v)** = **σ (w)**0

- Addition de deux entiers positifs:
  - Règle 1: Pour tout entier positif m,

$$m + 0 = m$$

  - Règle 2: $m + \sigma (n) = \sigma (m+n)$
- Multiplication de deux entiers positifs:
  - Règle 1: Pour tout entier positif m,

$$m * 0 = 0$$

  - Règle 2: $m * \sigma (n) = m*n + m$

- La fonction factorielle
  - Règle 1: $0! = 1$
  - Règle 2: $n! = n(n-1)!$

Question?