Informatique Théorique
TD5
SI3-MAM3

1 Libre ou pas

1. Donnez et prouvez une définition inductive pour l'ensemble des mots binaires palindromes. Est-ce que la définition est libre ?
2. Donner un exemple de schéma inductif, ne comportant qu'une règle et qui cependant n'est pas libre.
3. Donner et prouver une définition inductive pour l'ensemble des mots sur l'alphabet $\Sigma = \{a,\ b\}$ ne comportant pas deux a consécutifs. Votre schéma est-il libre ?

---

1. Une définition possible est
   - Base $\{\epsilon, 0, 1\} \subset E$
   - Règles
     - $R0$: $m \in E \Rightarrow 0m0 \in E$
     - $R1$: $m \in E \Rightarrow 1m1 \in E$

   Soit P l'ensemble des mots binaires palindromes.

   Montrons l'égalité des deux ensembles.

   - $E \subset P$
     La preuve se fait très simplement par induction structurelle
     - Base: ϵ ,0,1 sont bien des palindromes binaires.
     - Propagation: Supposons que m est un palindrome binaire, alors il en est de même de $1m1$ et de $0m0$.
   - $P \subset E$
     La preuve se fait sur la longueur des mots de P
     - Si cette longueur est 0 alors le mot est dans la base de E, donc dans E.
     - Supposons que tout mot de P de longueur $< k$ est aussi dans E.
     - Soit alors m un mot de P de longueur k.
       * Si $|m| = 1$, il est dans la base, on peut donc le supposer de longueur au moins 2
       * Si la dernière lettre de m est un 1 alors la première lettre doit aussi être un 1 et ainsi $m = 1m'1$ et m' est lui aussi un mot de P. Comme $|m'| < k$, m' est dans E (par hypothèse de récurrence) et donc $m = 1m'1$ est aussi dans$E$ (par définition de E).
       * Sinon, la dernière lettre de m est un 0, mais alors sa première lettre est aussi un 0 et on a m=$0m'0$ et m' est lui aussi un mot de P. Comme $|m'| < k$, m' est dans E (par hypothèse de récurrence) et donc $m = 0m'0$ est aussi dans E (par définition de E).

   Le schéma est libre, en effet :

   - aucun mot ne peut être dans la base (donc de longueur de un) et produit par une règle (donc de longueur au moins deux)
   - un mot se terminant (et commençant) soit par un un soit par un zéro, il ne peut être produit que par l'une des deux règles.
   - Un seul antécédent est possible pour une règle donnée, car $1m1$=$1m'1$ entraîne $m = m'$ et de même $0m0 = 0m'0$ entraîne $m = m'$.
2. - Base ϵ, a est inclus dans E

- Règles
  - R1 :Si m et m' sont dans E, alors mm' est dans E

3. Une définition possible est

   - Base ϵ, a est inclus dans E
   - Règles
     - R1 :Si m est dans E, alors mb est dans E
     - R2 :Si m est dans E, alors mba est dans E

   Soit $MBS2AC$ l'ensemble des mots de $\{a,b\}^*$ ne comportant pas deux a consécutifs

   Montrons l'égalité des deux ensembles

   - E est inclus dans $MBS2AC$
     La preuve se fait très simplement par induction structurelle
     - Base: ϵ et a sont bien des mots de $\{a,b\}^*$ et ne comportent pas deux a consécutifs
     - Propagation: Supposons que m ne comporte pas deux a consécutifs, alors il en est de même de mb et de mba.
   - $MBS2AC$ est inclus dans E
     La preuve se fait par récurrence sur la longueur des mots de $MBS2AC$
     - Si cette longueur est 0 alors le mot est dans la base de E, donc dans E
     - Supposons que tout mot de $MBS2AC$ de longueur $< k$ est aussi dans E
     - Soit alors m un mot de $MBS2AC$ de longueur k.
       Si $|m| = 1$, alors soit $m = a$ et m est dans la base, soit $m = b$ et il est produit par la règle $R1$ à partir de ϵ.
       On peut donc supposer k au moins égal à deux.
       * Si la dernière lettre de m est un b, alors $m = m'b$ et m' est lui aussi un mot de $MBS2AC$. Comme $|m'| < k$, m' est dans E (par hypothèse de récurrence) et donc $m = m'b$ est aussi dans E (par définition de E).
       * Sinon, la dernière lettre de m est un a, mais alors son avant dernière lettre est forcément un b et $m = m'ba$, et m' est lui aussi un mot de $MBS2A0C$. Comme $|m'| < k$, m' est dans E (par hypothèse de récurrence) et donc $m = m'ba$ est aussi dans E (par définition de E)

   Le schéma est libre, en effet :

   - aucun mot ne peut être dans la base (donc de longueur un) et produit par une règle (donc de longueur au moins deux)
   - un mot se terminant soit par un b soit par un a, il ne peut être produit que par l'une des deux règles.
   - un seul antécédent est possible pour une règle donnée, car $mb = m'b$ entraîne $m = m'$ et de même $mba = m'ba$ entraine $m = m'$.

   un autre schéma possible

   - Base ϵ , a sont dans E
   - Règle m, m' dans $E \Rightarrow mbm'$ dans E

   Mais ce schéma n'est pas libre

---

2 Pas de jaloux

Soit M le sous ensemble de $\{a,\ b\}^\star$ constitué des mots ayant autant de a que de b.
Soit E l'ensemble défini de manière inductive par

- Base : $B = \{\epsilon\}$
- Règles : $\Omega = \{\omega_g, \omega_d\}$ avec $\omega_g(m) = amb$ et $\omega_d(m) = bma$.

1. Le schéma définissant E est-il libre ?

2. A-t-on $M \subset E$?
3. A-t-on $E \subset M$?
4. Déterminez et prouvez une définition inductive pour M.
5. Donnez une définition non inductive de E.

- Le schéma définissant E est libre, en effet
  - aucun mot ne peut être dans la base (donc de longueur nulle) et produit par une règle (donc de longueur au moins deux)
  - un mot se terminant soit par un a soit par un b, il ne peut être produit que par l'une des deux règles.
  - Un seul antécédent est possible pour une règle donnée, car $amb = am'b$ entraîne $m = m'$ et de même $bma = bm'a$ entraîne $m = m'$.
- M n'est pas inclus dans E. Par exemple, le mot $abba$ appartient à M (il comporte autant de a que de b), mais pas dans E (il n'est pas dans la base et aucune règle ne peut le produire puisque sa première et sa dernière lettre sont identiques).
- E est inclus dans M. La preuve se fait facilement par induction structurelle
  - Base : le mot vide contient autant de a que de b (zéro de chaque)
  - Propagation : Supposons que m contienne autant de a que de b, alors clairement il en est de même de amb et de bma.
- - Soit F l'ensemble défini de manière inductive par
    - Base : $\epsilon \in F$
    - Règles :
      1. $m, m' \in F \Rightarrow mm' \in F$
      2. $m \in F \Rightarrow bma \in F$
      3. $m \in F \Rightarrow amb \in F$
    - Montrons que $M = F$
    - Clairement par induction structurelle on a $F \subset M$
    - Réciproquement, montrons par récurence sur la longueur des mots de M que $M \subset F$
      - Base : Un mot de M de longueur nulle est le mot vide, il est donc dans F
      - Supposons que tout mot de M de longueur au plus n est dans F
      - Soit m un mot de M de longueur $n + 1$. Si m ne commence et ne termine pas par la même lettre, on peut supposer que l'on a $m = am'b$, m' a autant de 'a' que de 'b', il est donc dans M il est de longueur $n - 1$ il est donc dans F. Donc m est dans F. Sinon, c'est que m commence et termine par la même lettre, disons un 'a'. On a $m = am'a$. Soit $prefixe(i)$ le préfixe de m de longueur i. Calculons pour tout i, $f(i)=$ le nombre de 'a' de $prefixe(i)$ moins le nombre de 'b' de $prefixe(i)$.On a $f(1) = 1$ et $f(n) = -1$. Il existe donc un entier k compris entre 1 et n tel que $f(k) = 0$. Posons $m_1 = prefixe(k)$, et soit m_2 tel que $m = m_1m_2$. Les deux mots m_1 et m_2 sont tous les deux dans M, ils sont tous les deux de longueur au plus n, ils sont donc tous les deux dans F, et donc m est dans F.

---

3 Maux de parenthèses

1. Soit LP le langage défini sur l'alphabet $\{(,\ )\}$ par
   - Base : $B = \{\epsilon\}$
   - Règle : $\Omega = \{\omega\}$ avec $\omega(u, v) = (u)v$.
   - Montrer par induction structurelle que les mots de LP ont exactement autant de (que de).

   - Notons $m_k = (^k)^k$. On montre facilement par récurrence sur k que tous les m_k sont dans LP. En effet on a $m_{k+1} = (m_k)\epsilon$

- On montre facilement par récurrence sur k que tous les d^k sont dans LP. En effet on a $d^{k+1} = (\epsilon)d^k$
- $()))() = (m_1)d$ est dans LP
- * tous les mots de la base de LP (c'est à dire ϵ) ont autant parenthèses ouvrantes que de fermantes
  * Si $|u|_(= |u|_) = k$ et $|v|_(= |v|_) = k'$ alors $|(u)v|_(= |(u)v|_) = k + k' + 1$

- Le schéma est il libre ?

2. Considérons LBP l'ensemble des mots m sur l'alphabet $\Sigma = \{(,\ )\}$ tels que $|m|_(= |m|_)$ et dans tout pr éfixe u de m, $|u|_(\geq |u|_)$.
   (a) Montrez que $LBP = LP$
   (b) Montrez que le schéma définissant LP est libre

Notons $P(m)$ la propriété :

$m \in \{(,)\}^*$, $\forall p$, p préfixe de m $|p|_(\geq |p|_)$ et $|m|_(= |m|_)$

Notons que $m \in LBP$ est équivalent à $P(m)$.

- Montrons d'abord par induction structurelle que $LP \subset LBP$
  - Base $P(\epsilon)$ est vrai.
  - Propagation : Montrons que $P(u)$ et $P(v)$ entraînent $P((u)v)$.
    Un préfixe de $(u)v$ est soit
    * ϵ
    * $(p_u$ où p_u est un préfixe de u
    * $(u)p_v$ où p_v est un préfixe de v

    Dans tous les cas, ce préfixe comporte au moins autant d'ouvrantes que de fermantes. De plus $(u)v$ contient exactement autant d'ouvrantes que de fermantes.
    Donc $P((u)v)$ est bien vrai.
- Pour démontrer la réciproque on introduit la fonction suivante:
  Soit $altitude(m, i)$= le nombre de (- le nombre de) dans le préfixe de longueur i de m,
  $P(m) \Leftrightarrow \forall i, 0 \leq i \leq |m|$, $altitude(m, i) \geq 0$ et $altitude(m, |m|) = 0$
  Proposition 1
  Pour tout mot m de $\{(,)\}^*$, $\forall i, 1 \leq i \leq |m| + 1$ $altitude((m, i) = altitude(m, i - 1) + 1$
  Proposition 2
  Pour tous les mots m, m' de $\{(,)\}^*$,
  $\forall i, 0 \leq i \leq |m|$, $altitude(mm', i) = altitude(m, i)$
  $\forall i,\ |m| + 1 \leq i \leq |m| + |m'|$, $altitude(mm', i) = altitude(m, |m|) + altitude(m', i - |m|))$,
  Montrons par récurrence sur n la propriété
  $H(n)$: tout mot m de longueur n vérifiant $P(m)$ est dans LP
  $H(0)$ est vrai car ϵ est le seul mot de longueur 0 et il appartient à LP;
  Supposons $H(p)$ vrai pour tout p<n;
  Soit m un mot vérifiant $P(m)$ et de longueur n. Notons m_i la i-ème lettre de m.
  Soit k le plus petit entier tel que $altitude(m, k) = 0$ (il existe car altitude $(m, |m|)) = 0)$
  On a nécessairement $altitude(m, k - 1) = 1$ donc $m_k =)$
  Posons $u = m_2m_3...m_{k-1}$, et $v = m_{k+1}..m_n$
  On a $m = (u)v$.
  D'après la proposition 2 , on a $altitude((u)v, i) = altitude((u, i)$ $\forall i, 0 \leq i \leq |u| + 1$.
  D'après la proposition 1, on a $altitude((u, i) = altitude(u, i - 1) + 1$ $\forall i, 1 \leq i \leq |u| + 1$
  Donc $\forall i,\ 1 \leq i \leq |u|$ $altitude(u, i) = altitude((u, i+1) - 1 = altitude(m, i+1) - 1 \geq 0$ puisque i <k, et $altitude(u, |u|)) = altitude(m, k - 1) - 1 = 0$
  Le mot u appartient donc à LBP.
  D'après la proposition 2, $altitude(v, i) = altitude(m, i + |(u)|))$ $\forall i, 1 \leq i \leq |v|$. On a donc $altitude(v, i) \geq 0$, $\forall i,\ 1 \leq i \leq |v|$ et $altitude(v, |v|) = 0$ Le mot v appartient donc à LBP.
  Les deux mots u et v sont de longueur strictement inférieure à n, donc par hypothèse ils sont dans LP.
  Par définition de LP, m est donc aussi dans LP.

- Montrez que le schéma définissant LP est libre

  Clairement un mot ne peut pas être à la fois dans la base et construit par les règles.

  Supposons qu'il existe $m = (u)v{=}(u')v'$, avec u, v, u', v' dans LP. Avec la même fonction altitude que pour la question précédente, on a $altitude((u), |u|{+}2) = 0 = altitude((u'), |u'|{+}2)$.

  Supposons que u et u' soient différents, alors ils sont forcement de longueur différente. On peut donc supposer que $|u| < |u'|$. Alors $u)$ est un préfixe de u' puisque $(u)v = (u')v'$ mais $u)$ contient plus de) que de (, en contradiction avec u' dans LP. Donc $u = u'$, et donc $v = v'$.

---

3. $LP2$ est définit inductivement par

   - Base : $B = \{\epsilon\}$
   - Règles : $\Omega = \{\omega_1,\ \omega_2\}$ avec $\omega_1(u) = (u)$ et $\omega_2(u, v) = uv$.

   (a) Montrez que $LP = LP2$.

   (b) Le schéma définissant $LP2$ est-il libre ?

- Montrons tout d'abord que LP est inclus dans $LP2$, par induction structurelle sur LP

  Tout mot de la base de LP est dans $LP2$, car les deux ensembles ont la même base.

  Supposons que les mots u et v de LP soient aussi dans $LP2$, alors par définition de $LP2$, le mot (u) est dans $LP2$ (Règle 2) et puisque (u) et v sont dans $LP2$, $(u)v$ est dams $LP2$ (Règle 1), donc si u et v sont dans $LP2$, $(u)v$ est dans $LP2$.

  On vient de prouver que tous les mots de LP ont la propriété d'appartenir à $LP2$ par induction structurelle sur $LP2$.
- Montrons maintenant que $LP2$ est inclus dans LBP

  A nouveau, faisons une induction structurelle sur $LP = LBP$ cette fois.

  Tout mot de la base de $LP2$ est dans LP

  Si un mot u de $LP2$ est dans LP il en est de même de $u = (u)\epsilon$.

  Soient maintenant u et v mots de $LP2$, supposons qu'ils sont dans LBP, il en est alors de même pour uv est lui aussi dans LBP (d'après la propriété 2, on vérifie que l'altitude n'est jamais négative)
- Ce schéma n'est pas libre : par exemple le mot $()()()$ peut être produit à partir de la première régle avec $u = ()$ et $v = ()()$ ou bien avec $u = ()()$ et $v = ()$

---

4 (mots

Soit A l'alphabet $\{(,\ )\}$ et soit L le sous ensemble de $A^\star$ formé des mots dont tous les préfixes contiennent au moins autant de (que de).

1. Donnez une définition inductive de L et prouvez-la.
2. Montrez que L n'est pas égal à l'ensemble des mots bien parenthésés. Comment peut-on associer à un mot de L, un mot bien parenthésé ?
3. Le schéma que vous avez donné à la première question est-il libre ou ambigu ?

Notons $Q(m)$ la propriété :
$m \in \{(,)\}^*$, $\forall p$, p prefixe de m $|p|_(\geq |p|_)$.
Notons que $m \in L$ est equivalent à $Q(m)$.
En utilisant la fonction *altitude* définie à l'exercice précédent, on a aussi
$Q(m) \Leftrightarrow \forall i, 0 \leq i \leq |m|, altitude(m, i) \geq 0$

1. Une définition possible est

   - Base : ϵ est dans M
   - R égles :Si m et m' sont dans M alors $(m$ et $(m)m'$ sont dans M

   M est inclus dans L par induction structurelle, en effet

- ϵ appartient à L
- Supposons $Q(m)$, on vérifie que l'on a $Q((m)$
- Supposons $Q(m)$ et $Q(m')$, on verifie que l'on a aussi $Q((m)m')$

Réciproquement,montrons par récurrence sur n que tout mot de L de longueur n appartient à M.

Base Si $n = 0$,$m = \epsilon$ et donc m appartient à M

Etape inductive

Supposons que tout mot de L de longueur strictement inférieure ? n appartient à M

Soit u un mot de L, de longueur n

Premier cas : $altitude(u, i) > 0$ pour tout i, $1 \leq i \leq n$

Si u n'est pas le mot vide, alors $u = (u'$, et comme $altitude(u, i) = altitude(u', i - 1) + 1$, $altitude(u', j) \geq 0$, pour tout j $1 \leq j \leq n - 1$, et donc u' appartient à L. Comme u' est de longueur $n - 1$, par hypothèse de récurrence u' appartient à M. Donc u appartient à M par définition de M.

Deuxième cas : il existe k tel que $altitude(u, k) = 0$

Soit j le plus petit entier tel que $altitude(u, j) = 0$. On a donc $altitude(u, i) > 0$, pour tout i $1 \leq i \leq j - 1$.

La jième lettre de u est nécessairement un).

Soit m tel que (m) est le prefixe de longueur j de u.

Comme $altitude(u, k) = 1 + altitude(m, k - 1)$, m appartient à L, et par hypothèse de récurrence à M.

Soit m' le motéventuellement vide tel que $u = (m)m'$, on a $altitude(u, k) = altitude(m', k - j)$, et donc m' aussi appartient à L et par récurrence à M. On a donc $u = (m)m'$, avec m et m' dans M, donc u appartient à M.

2. Une autre définition possible est

   - Base : ϵ et (sont dans M
   - Règles :Si m et m' sont dans M alors (m) et mm' sont dans M

   Les preuves sont similaires

3. Le mot (est dans M mais n'est pas bien parenthésé. En fait pour tout mot de M, il suffit d'ajouter en fin le nombre de) nécessaires pour équilibrer le nombre de) et de (pour rendre ce mot bien parenthésé.

4. Le premier schéma donné n'est pas libre, en effet le mot $(()$ peut être dérivé soit avec la règle $(m$ et $m = ()$, soit avec la régle $(m)m'$ avec $m = ($ et m' vide. Le second schéma n'est pas libre non plus, le mot (((peut être dérivé par la régle mm' avec $m = (\ m' = (($, soit avec $m' = (, = (($.

---

5 Expression polonaise

Soient Var et Op deux alphabets disjoints.
On pose A= $Var \cup Op$.
On appelle langage des expressions polonaises préfixées le langage L sur A défini par le schéma:

- Base $Var \subset L$
- Règle $\omega \in Op, u, v \in L \Rightarrow \omega uv \in L$

1. Montrez qu'un mot w de A^* appartient à L si et seulement si il vérifie les deux conditions suivantes:
   - w contient une variable de plus que d'opérateurs
   - tout préfixe propre p de w contient au moins autant d'opérateurs que de variables
2. Montrez que le schéma définissant L est libre.
3. Écrire une méthode récursive pour calculer la valeur d'une expression polonaise
4. Écrire une méthode récursive pour transformer une expression polonaise en expression complétement parenthésée.

1. Montrons par induction structurelle sur L que tout mot m de L vérifie les 2 propriétés.

   Pour la suite de l'exercice nous introduisons les notations suivantes :

   - Pour un mot m, $var(m)$ désignera le nombre de variables dans m.
   - Pour un mot m, $op(m)$ désignera le nombre d'opérateurs dans m.
   - Pour un mot m, $diff(m) = op(m)$ - $var(m)$.

   Avec cette notation les deux propriétés deviennent:

   i. $diff(w) = -1$

   ii. pour tout préfixe propre p de w $diff(p) \geq 0$.

   Si m est dans la base, alors $var(m) = 1$ et $op(m) = 0$ et m n'a que le mot vide comme préfixe propre, donc les deux propriétés sont vérifiées.

   Supposons que u et v vérifient les deux propriétés, alors ωuv vérifie aussi les deux propriétés. En effet par hypothèse : $var(u) = op(u) + 1$ et $var(v) = op(v) + 1$, donc $var(\omega uv) = var(u) + var(v)$ $= op(u) + 1 + op(v) + 1 = op(\omega uv) + 1$. Et donc uv vérifie la première propriété

   Pour montrer que ωuv vérifie la seconde propriété, examinons p un préfixe propre quelconque de ωuv

   - Cas 1, $p = \omega$, alors p vérifie bien $var(p) = 0 < 1 = op(p)$.
   - Cas 2 $p = \omega p_u$ où p_u est un préfixe propre de u. Par hypothèse d'induction $var(p_u) \leq op(p_u)$, et donc $var(p) \leq op(p)$.
   - Cas 3 $p = \omega u p_v$ où p_v est un préfixe de v ($p_v \neq v$). Par hypothèse d'induction $var(u) = op(u) + 1$ et $var(p_v) \leq op(p_v)$, et donc $var(p) \leq op(p)$.

   Réciproquement, montrons par induction sur la longueur du mot que si un mot vérifie les deux propriétés, alors il appartient à L.

   Soit w un mot vérifiant les deux propriétés, c'est à dire tel que $diff(w) = -1$ et $diff(p) \geq 0$ pour tout préfixe propre p de w.

   Si w est de longueur 1, alors w est une variable (à cause de la propriété 1) et w appartient à L.

   Sinon supposons que tout mot de longueur inférieure où égale à n vérifiant les deux propriétés est dans L.

   Soit w un mot de longueur n+1, vérifiant les deux conditions, alors nécessairement ce mot commence par un opérateur : $w = \omega s$ (remarque : on ne peux pas dire ici que s vérifie les propriétés i et ii parce que ce n'est pas vrai, on va chercher deux mots u et v qui vérifient les propriétés (et qui donc par hypothèse de récurrence seront dans L), tels que s =uv)

   Appelons s_i le préfixe de s comportant i lettres. Puisque ωs_i est un préfixe de w, $diff(\omega s_i) \geq 0$, d'autre part puisque ωs_n =w, on a $diff(\omega s_n) = -1$. Puisque $diff(\omega) = 1$, il existe i, $1 \leq i \leq n-1$ tel que pour tout j<i, $diff(\omega s_j)$>0 et $diff(\omega s_i)$ =0 (i est le plus petit entier tel que $diff(\omega s_i)$ =0). Puisque ω est un opérateur, on a donc pour tout j<i, $diff(s_j) \geq 0$ et $diff(s_i) = -1$.

   Donc s_i vérifie les deux propriétés, comme il est de longueur strictement inférieure à n, il est donc dans L.

   Soit r tel que $w = \omega s_i r$. Puisque $diff(\omega s_i) = 0$, r lui aussi vérifie les deux conditions. Comme r est de longueur strictement inférieure à n, il appartient à L, il en est donc de même pour w.

2. Montrons que le schéma est libre

   Soit w un mot de L, ayant n lettres, et supposons qu'il existe deux opérateurs ω_1 et ω_2 et quatre mots de L u_1, u_2, v_1, v_2 tels que $w = \omega_1 u_1 v_1 = \omega_2 u_2 v_2$.

   Nécessairement $\omega_1 = \omega_2$.

   Si u_1 et u_2 sont de la même longueur, alors ils doivent être égaux. Sinon, supposons $|u_1| < |u_2|$. Alors u_1 étant un mot de L , $diff(u_1)$ = -1, mais par ailleurs u_1 est un préfixe propre de u_2 donc $diff(u_1) \geq 0$, contradiction. Donc nécessairement u_1 et u_2 sont de la même longueur. On a donc $u_1 = u_2$ et $v_1 = v_2$.

3 et 4 run:./PostFixe.java Voir fichier Postfixe.java