

**Exercice 1 - Recherche de grammaires****Solution page 2**

Prouver que les langages suivants sont algébriques en donnant une grammaire qui les engendre.

1. Ensemble des mots sur  $\Sigma = \{a, b\}$  de longueur impaire.
2.  $L = \{a^m b^n ; n \neq m\}$
3.  $L = \{a^n b^m c^{n+m} ; n, m \in \mathbb{N}\}$ .
4.  $L = \{a^m b^n ; m \leq n \leq 2m\}$
5. Le complémentaire de  $\{a^n b^n \mid n \in \mathbb{N}\}$ .
6.  $L = \{a^p b^q c^r ; p \neq q \text{ ou } q \neq r\}$
7.  $L = \{a^p b^q c^r d^s ; p + q = r + s\}$

**Exercice 2 - Langage transposé****Solution page 2**

Prouver que si  $L$  est algébrique alors  $L^T$ , le langage des mots transposés de  $L$  est algébrique.

**Exercice 3 - Tautologies****Solution page 2**

On considère l'alphabet  $\Sigma = \{(\,,\,), \wedge, \vee, \neg, p, \top, \perp\}$  où  $\top$  représente vrai,  $\perp$  représente faux et  $p$  est une variable booléenne. Déterminer une grammaire qui engendre les tautologies.

**Exercice 4****Solution page 2**

On considère la grammaire  $G = (\{a, b\}, \{S\}, \{S \rightarrow SbS|a\}, S)$ .

Déterminer le langage engendré par  $G$ .

Prouver que la grammaire est ambiguë en donnant deux arbres de dérivation pour  $abababa$ .

**Autant de  $a$  que de  $b$** **Exercice 5 -****Solution page 2**

Prouver que le langage  $L_{eq} = \{u \in \{a, b\}^* ; |u|_a = |u|_b\}$  est engendré par la grammaire  $S \rightarrow aSbS|bSaS|\varepsilon$ . Étudier son ambiguïté.

**Exercice 6****Solution page 3**

Prouver que le langage  $L_{eq} = \{u \in \{a, b\}^* ; |u|_a = |u|_b\}$  est engendré par la grammaire  $S \rightarrow aSb|bSa|SS|\varepsilon$ . Étudier son ambiguïté.

**Exercice 7 -****Solution page 4**

Prouver que le langage  $L_{eq} = \{u \in \{a, b\}^* ; |u|_a = |u|_b\}$  est engendré par la grammaire  $S \rightarrow aB|bA|\varepsilon, A \rightarrow aS|bAA, B \rightarrow bS|aBB$ . Étudier son ambiguïté.

**Exercice 8****Solution page 5**

Prouver que le langage  $L_{eq} = \{u \in \{a, b\}^* ; |u|_a = |u|_b\}$  est engendré par la grammaire  $S \rightarrow aDbS|bIaS|\varepsilon, D \rightarrow aDbD|\varepsilon, I \rightarrow bIaI|\varepsilon$ . Étudier son ambiguïté.

**Exercice 9****Solution page 6**

Prouver que le langage  $L_{eq} = \{u \in \{a, b\}^* ; |u|_a = |u|_b\}$  est engendré par la grammaire  $S \rightarrow aAS|bBS|\varepsilon, A \rightarrow aAA|b, B \rightarrow bBB|a$ . Étudier son ambiguïté.

## Solutions

### Exercice 1

1.  $S \rightarrow SSS|a|b$
2.  $S \rightarrow aSb|aA|Bb, A \rightarrow aA|\varepsilon, B \rightarrow Bb|\varepsilon.$
3.  $S \rightarrow aSc|T, T \rightarrow bTc|\varepsilon.$
4.  $S \rightarrow aSb|aSbb|\varepsilon.$
5. Un mot qui n'est pas de la forme  $a^n b^n$  est soit un mot qui commence par  $b$ , soit finit par  $a$ , soit de la forme  $aub$  avec  $u$  qui n'est pas de la forme  $a^n b^n$ . On en déduit la grammaire :  
 $S \rightarrow bX \mid Xa \mid aSb, X \rightarrow aX \mid bX \mid \varepsilon.$
6. C'est une union de quatres produits de langages algébriques :  
 $S \rightarrow S_1|S_2|S_3|S_4, S_1 \rightarrow M_1C, M_1 \rightarrow aM_1b|bB, S_2 \rightarrow P_1C, P_1 \rightarrow aP_1b|aA, S_3 \rightarrow AM_2,$   
 $M_2 \rightarrow bM_2c|cC, S_4 \rightarrow AP_2, P_2 \rightarrow bP_1c|bB, A \rightarrow aA|\varepsilon, B \rightarrow bB|\varepsilon, C \rightarrow cC|\varepsilon.$
7. Après avoir associé les  $a$  et  $d$  il reste un mot de la forme  $a^p b^q c^{p+q}$  ou  $b^{p+q} c^p d^q$ .  
 $S \rightarrow aSd|A|D, A \rightarrow aAc|B, D \rightarrow bDd|B, B \rightarrow bBc|\varepsilon.$

### Exercice 2

On transpose toutes les règles.

### Exercice 3

On utilise 5 variables ;  $T$  génère les proposition vraies,  $F$  les propositions fausses,  $P$  celles qui sont équivalentes à  $p$  et  $N$  à  $\neg p$ ,  $X$  génère toutes les formules.

$T \rightarrow \top|(T \vee X)|(X \vee T)|(P \vee N)|(N \vee P)|(T \wedge T)|\neg F$   
 $F \rightarrow \perp|(F \vee X)|(F \wedge X)|(X \wedge F)|(P \wedge N)|(N \wedge P)|\neg T,$   
 $P \rightarrow p|(P \vee P)|(P \vee F)|(F \vee P)|(P \wedge P)|(P \wedge T)|(T \wedge P)|\neg N,$   
 $N \rightarrow \neg p|(N \vee N)|(N \vee F)|(F \vee N)|(N \wedge N)|(N \wedge T)|(T \wedge N)|\neg P,$   
 $X \rightarrow T|F|P|N.$

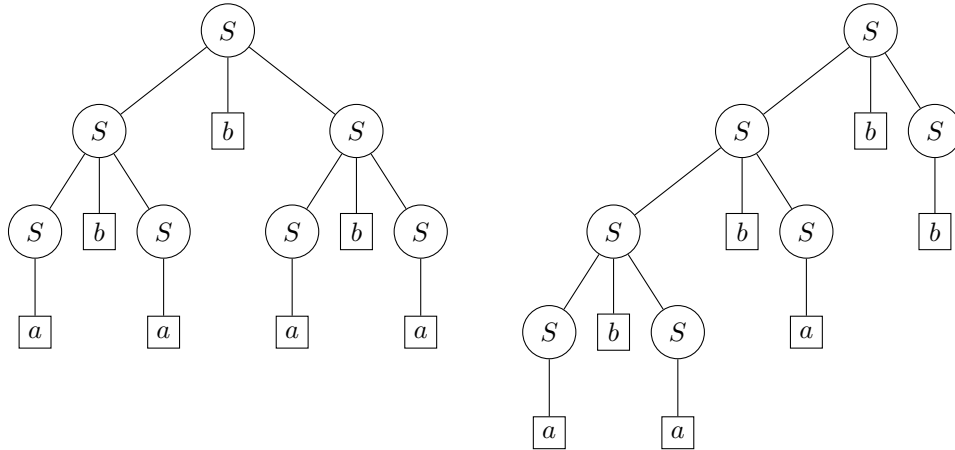
### Exercice 4

On génère  $a$  avec une dérivation  $S \Rightarrow a$  et  $aba$  avec  $S \Rightarrow SbS \Rightarrow abS \Rightarrow ababa$ .

On prouve par récurrence sur la hauteur des arbres de dérivation que tout mot généré par  $G$  est de la forme  $a(ba)^n$ .

C'est vrai pour un arbre de hauteur 1 qui ne peut être que défini par  $S \Rightarrow a$ .

Si la propriété est vraie pour tous les arbres de dérivation de hauteur  $n$  au plus, un arbre de dérivation de hauteur  $n+1 \geq 2$  doit appliquer la règle  $S \rightarrow SbS$  à la racine et les arbres issus des deux nœuds de racine  $S$  sont de hauteur  $n$  au plus donc génèrent des mots  $a(ba)^p$  et  $a(ba)^q$  donc le mot généré est  $a(ba)^p ba(ba)^q = a(ba)^{p+q+1}$ . La récurrence est prouvée.



**Exercice 5**

Chaque règle ajoute autant de  $a$  que de  $b$  donc le langage engendré est inclus dans  $L_{eq}$ .

Soit  $u \in L_{eq}$ . On montre par récurrence généralisée sur  $|u|$  que  $u$  est généré par  $G$ .

Si  $u = \varepsilon$ ,  $u$  est généré par  $S \Rightarrow \varepsilon$ .

On suppose que tous les mots de  $L_{eq}$  de longueur inférieure à  $2n$  sont générés par  $G$ .

Soit  $u$  de longueur  $2n + 2$ .

Si  $u$  commence par  $a$ ,  $u = x_1 x_2 \cdots x_p$  avec  $x_1 = a$ , on note  $v_i = x_1 x_2 \cdots x_i$  le préfixe de  $u$  de longueur  $i$  et  $h_i = |v_i|_a - |v_i|_b$ .

On a  $h_1 = 1$  et  $h_p = 0$  car  $u \in L_{eq}$  et  $h_{i+1} = h_i \pm 1$ .

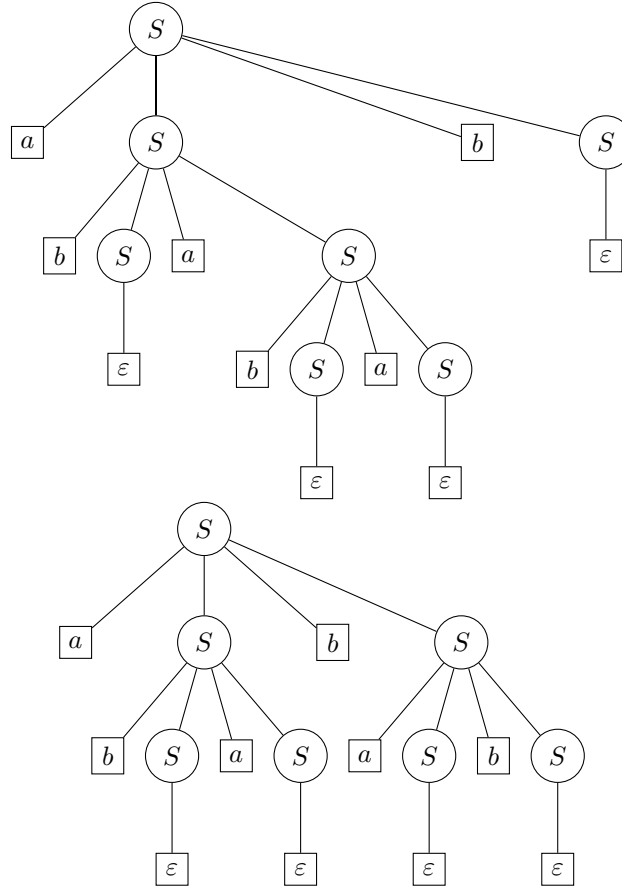
On considère le premier indice  $q$  tel que  $h_q = 0$ , on a alors  $h_i > 0$  pour  $1 \leq i < q$ . On a alors  $h_{q-1} = 1$  donc  $x_q = b$ . On peut donc écrire  $u = avbw$  avec  $v = x_2 \cdots x_{p-1} \in L_{eq}$  et  $w = x_{p+1} \cdots x_{2n+2} \in L_{eq}$ .

Comme  $|v| \leq |u| - 2 = 2n$  et  $|w| \leq 2n$ ,  $v$  et  $w$  sont générés par  $G$  donc  $u$  est généré par  $G$ .

De même si  $u$  commence par  $b$ .

Ainsi tout mot de  $L_{eq}$  est engendré par  $G$ .  $G$  engendre  $L_{eq}$ .

La grammaire est ambiguë, par exemple pour  $ababab$ .

**Exercice 6**

Chaque règle ajoute autant de  $a$  que de  $b$  donc le langage engendré est inclus dans  $L_{eq}$ .

Soit  $u \in L_{eq}$ . On montre par récurrence généralisée sur  $|u|$  que  $u$  est généré par  $G$ .

Si  $u = \varepsilon$ ,  $u$  est généré par  $S \Rightarrow \varepsilon$ .

On suppose que tous les mots de  $L_{eq}$  de longueur inférieure à  $2n$  sont générés par  $G$ .

Soit  $u$  de longueur  $2n + 2$ .

Si  $u$  commence par  $a$  et finit par  $b$  alors  $u = avb$  avec  $v \in L_{eq}$  et  $|v| = 2n$ .  $v$  est généré par  $G$  par hypothèse de récurrence et on peut compléter la dérivation de  $v$ ,  $S \xRightarrow{*} v$ , en introduisant la règle  $S \rightarrow aSb$  au début :  $D \Rightarrow aSb \xRightarrow{*} avb = u$  donc  $u$  est généré par  $G$ .

De même, si  $u$  commence par  $b$  et finit par  $a$  alors  $u$  est généré par  $G$ .

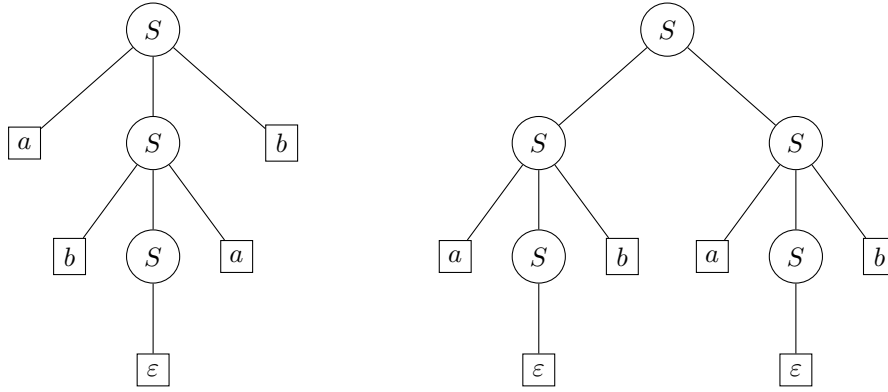
Si  $u$  commence et finit par la même lettre, on note  $v_i = x_1 x_2 \dots x_i$  le préfixe de  $u$  de longueur  $i$  et  $h_i = |v_i|_a - |v_i|_b$ . On a  $h_p = 0$  car  $u \in L_{eq}$  et  $h_{i+1} = h_i \pm 1$ .

On a  $h_1 = 1$  et  $h_{2n+1} = -1$  si  $u$  commence et finit par  $a$  ou  $h_1 = -1$  et  $h_{2n+1} = 1$  si  $u$  commence et finit par  $b$ .

Il existe un indice  $q \in \{2, 3, \dots, 2n\}$  tel que  $h_q = 0$ , on a alors  $u = vw$  avec  $v = x_1 \dots x_p \in L_{eq}$  et  $w = x_{p+1} \dots x_{2n+2} \in L_{eq}$ .  $v$  et  $w$  sont non vides donc  $|v| < |u|$  et  $|w| < |u|$ . Par hypothèse de récurrence,  $v$  et  $w$  sont générés par  $G : S \xRightarrow{*} v$  et  $S \xRightarrow{*} w$  donc  $S \Rightarrow SS \xRightarrow{*} Sw \xRightarrow{*} vw = u$ .

Ainsi tout mot de  $L_{eq}$  est engendré par  $G$ .  $G$  engendre  $L_{eq}$ .

La grammaire est ambiguë, par exemple pour  $abab$ .



### Exercice 7

On note  $G = (\{a, b\}, \{S, A, B\}, P, S)$ ,  $G_a = (\{a, b\}, \{S, A, B\}, P, A)$  et  $G_b = (\{a, b\}, \{S, A, B\}, P, B)$  avec  $P = \{S \rightarrow aB|bA|\varepsilon, A \rightarrow aS|bAA, B \rightarrow bS|aBB\}$ .

On veut montrer que  $G$  engendre  $L_{eq}$ ,  $G_a$  engendre  $L_a = \{u \in \{a, b\}^* ; |u|_a = |u|_b + 1\}$  et  $G_b$  engendre  $L_b = \{u \in \{a, b\}^* ; |u|_b = |u|_a + 1\}$ .

On montre  $L(G) \subset L_{eq}$ ,  $L(G_a) \subset L_a$  et  $L(G_b) = L_b$  par récurrence généralisée sur la longueur des dérivations ; supposera qu'on a des dérivations à droite.

La seule dérivation possible de longueur 1 est  $S \Rightarrow \varepsilon$  qui génère  $\varepsilon \in L_{eq}$ .

On suppose que les dérivations de longueur  $n \geq 1$  au plus engendrent des mots de  $L_{eq}$  si la racine est  $S$ , de  $L_a$  si la racine est  $A$  ou de  $L_b$  si la racine est  $B$ .

Une dérivation,  $X \xRightarrow{n+1} u$  avec  $n+1 \geq 2$ , peut commencer par

- $S \Rightarrow aB$ , dans ce cas elle se poursuit par une dérivation  $aB \xRightarrow{n} av$  donc, par hypothèse de récurrence,  $v \in L_b$  et  $u = av \in L_{eq}$ ,
- $S \Rightarrow bA$ , avec le même raisonnement  $u = av$  avec  $v \in L_b$  donc  $u = av \in L_{eq}$ ,
- $A \Rightarrow aS$ , ici aussi,  $u = av$  avec  $v \in L_{eq}$  donc  $u \in L_a$ ,
- $B \Rightarrow BS$ , ici aussi,  $u = bv$  avec  $v \in L_{eq}$  donc  $u \in L_b$ ,
- $A \Rightarrow bAA$ , elle se poursuit par  $bAA \xRightarrow{p} bvA \xRightarrow{q} bvw$  avec  $p+q=n$ , donc  $p \leq n$  et  $q \leq n$  d'où, d'après l'hypothèse de récurrence,  $v, W \in L_a$  donc  $u = bvw \in L_a$ ,
- $B \Rightarrow aBB$ , elle se poursuit par  $aBB \xRightarrow{p} avB \xRightarrow{q} avw$  avec  $p+q=n$ , donc  $p \leq n$  et  $q \leq n$  d'où, d'après l'hypothèse de récurrence,  $v, W \in L_b$  donc  $u = avw \in L_b$ .

Dans tous les cas,  $u$  appartient bien au langage attendu.

On a donc prouvé la récurrence.

On montre  $L_{eq} \subset L(G)$ ,  $L_a \subset L(G_a)$  et  $L_b \subset L(G_b)$  par récurrence sur la longueur des mots.

Un mot de longueur 0,  $\varepsilon$ , appartient à  $L_{eq}$  et est généré par  $G$ .

On suppose que tous les mots de longueur  $n$  au plus appartenant à  $L_{eq}$  ou  $L_a$  ou  $L_b$  sont générés par  $G$ ,  $G_a$  ou  $G_b$  respectivement.

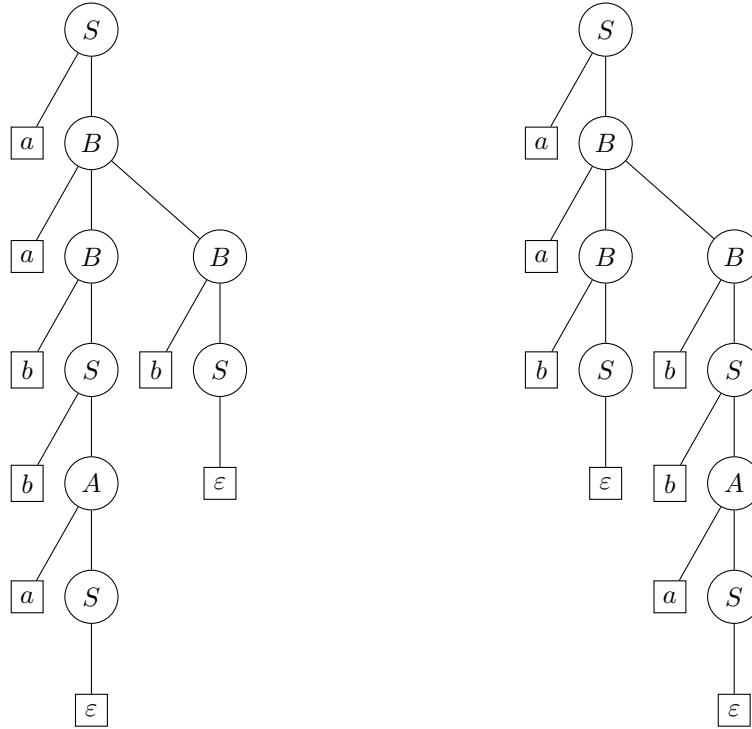
Soient  $u$  de longueur  $n+1$ ,  $u = xv$  avec  $x \in \{a, b\}$ .

- Si  $u \in L_{eq}$  et  $x = a$  alors  $v \in L_b$ . Par hypothèse de récurrence il existe une dérivation  $B \xRightarrow{*} v$  pour  $v$  dans  $G_b$  donc il existe une dérivation  $u \Rightarrow av \xRightarrow{*} av = u$  pour  $u$  dans  $G : u \in L(G)$ .
- Si  $u \in L_{eq}$  et  $x = b$ , le même raisonnement montre que  $u \in L(G)$ .
- Si  $u \in L_a$  et  $x = a$  alors  $v \in L_{eq}$ . Par hypothèse de récurrence il existe une dérivation  $S \xRightarrow{*} v$  pour  $v$  dans  $G$  donc il existe une dérivation  $A \Rightarrow aS \xRightarrow{*} av = u$  pour  $u$  dans  $G : u \in L(G_a)$ .
- Si  $u \in L_b$  et  $x = b$ , le même raisonnement montre que  $u \in L(G_b)$ .
- Si  $u \in L_a$  et  $x = b$  alors  $|v|_a = |v|_b + 2$ . On note  $v = y_1 y_2 \dots y_n$ ,  $v_i = y_1 y_2 \dots y_i$  le préfixe de  $y$  de longueur  $i$  et  $h_i = |v_i|_a - |v_i|_b$ . On a  $h_n = 2$ ,  $h_1 = \pm 1$  et  $h_{i+1} = h_i \pm 1$ .  
Il existe un indice  $p$  tel que  $h_p = 1$ , on a alors  $v' = y_1 \dots y_p \in L_a$  et  $v'' = y_{p+1} \dots y_n \in L_a$ . Par hypothèse de récurrence il existe 2 dérivations  $A \xRightarrow{*} v'$  et  $A \xRightarrow{*} v''$  pour  $v'$  et  $v''$  dans  $G_a$  donc il existe une dérivation  $A \Rightarrow bAA \xRightarrow{*} bv' B \xRightarrow{*} bv' v'' = bv = u$  pour  $u$  dans  $G_a : u \in L(G_a)$ .
- Si  $u \in L_b$  et  $x = a$ , le même raisonnement montre que  $u \in L(G_b)$ .

Dans tous les cas on a  $u \in L(G)$  pour  $u \in L_{eq}$ ,  $u \in L(G_a)$  pour  $u \in L_a$  ou  $u \in L(G_b)$  pour  $u \in L_b$ . On a donc prouvé la récurrence.

Les inclusions donnent les égalités des ensembles

La grammaire est ambiguë, par exemple pour  $aabbab$ .



### Exercice 8

La variable  $D$  suit la grammaire  $D \rightarrow aDbD|\varepsilon$  qui engendre les mots de Dyck : mots possédant autant de  $a$  que de  $b$  et tels que tout préfixe contient plus de  $a$  que de  $b$ .

De même la variable  $I$  permet de générer tous les mots possédant autant de  $a$  que de  $b$  et tels que tout préfixe contient plus de  $a$  que de  $b$ .

Ainsi par récurrence sur la longueur de la dérivation ou sur la longueur du mot généré ou sur la hauteur de l'arbre de dérivation la grammaire ne peut générer que des mots de  $L_{eq}$ .

Inversement, soit  $u \in L_{eq}$ . Si  $u = \varepsilon$ , il est généré par la grammaire.

On suppose  $u \neq \varepsilon$  et, par exemple,  $u = au'$ .

On note  $\delta(v) = |v|_a - |v|_b$  : si  $v_i$  est le préfixe de taille  $i$  de  $u$ , on a  $\delta(v_1) = \delta(a) = 1$  et  $\delta(u) = 0$ . On considère le plus court préfixe de  $u$  tel que  $\delta(v) = 0$  ; on a  $v = av'x$  et  $\delta(av') > 0$  donc  $x = b$  et on peut donc écrire  $v = av'b$  puis  $u = av'bw$ .

On a  $\delta(v') = 0$  et, pour tout préfixe  $v''$  de  $v'$ ,  $\delta(av'') > 0$  donc  $\delta(v'') \geq 0$  :  $v'$  est un mot de Dyck.

De plus  $\delta(w) = \delta(u) - \delta(v) = 0$  donc  $w \in L_{eq}$ .

De même si la première lettre de  $u$  est  $b$ ,  $u = bv'aw$  avec  $v'$  mot de Dyck pour  $b$  et  $a$  et  $w \in L_{eq}$ .

Il existe une dérivation  $D \xrightarrow{*} v'$  ou  $I \xrightarrow{*} v'$

On prouve alors par récurrence sur la taille des mots qu'on peut générer tout mot de  $L_{eq}$  avec  $G$

car une dérivation de  $w$  donne  $S \Rightarrow aDbS \xrightarrow{*} av'bS \xrightarrow{*} av'bw = u$  ou  $S \Rightarrow bIaS \xrightarrow{*} bv'aS \xrightarrow{*} bv'aw = u$ .

On a ainsi  $L_{eq} \subset L(G)$  puis l'égalité.

Comme pour les mots de Dyck la non-ambiguïté provient de l'unicité de la décomposition  $u = av'bw$  avec  $v'$  mot de Dyck pour  $a$  et  $b$  ou  $u = bv'aw$  avec  $v'$  mot de Dyck pour  $b$  et  $a$  car la grammaire utilisée pour les mots de Dyck est non-ambiguë.

### Exercice 9

Cette grammaire est la même que la précédente en remplaçant  $D$  par  $Ab$  et  $I$  par  $Ba$ .