TD 1

Langages rationnels

MPI/MPI*, lycée Faidherbe

I Démonstrations du cours

Exercice 1 - Propriétés du produit

Solution page 6

Prouver les propriétés suivantes.

1. Pour $p \leqslant |u|$ u admet un unique préfixe (resp. suffixe) de longueur p.
2. $(u \cdot v) \cdot w = u \cdot (v \cdot w)$ la loi est associative, on omettra les parenthèses dans les produits.
3. $|u \cdot v| = |u| + |v|$.
4. Si $u \cdot v = \varepsilon$ alors $u = v = \varepsilon$.
5. $u^n \cdot u^m = u^{n+m}$.
6. Si $u \cdot v = u \cdot v'$ alors $v = v'$ (simplification à gauche).
7. Si $u \cdot v = u' \cdot v$ alors $u = u'$ (simplification à droite).

Exercice 2 - Ordre lexicographique

Solution page 6

Prouver que l'ordre lexicographique est une relation d'ordre.

Exercice 3 - Ordre total

Solution page 6

Prouver que l'ordre lexicographique est une relation d'ordre total.

Exercice 4 - Propriétés de l'étoile

Solution page 6

Prouver les propriétés suivantes

1. L^* est l'ensemble des produits de mots de L.
2. Si $L = \{w\}$ alors $L^* = \{w^n \ ; \ n \in \mathbb{N}\}$.
3. $\emptyset^* = \{\varepsilon\}^* = \{\varepsilon\}$.
4. Si L contient un mot non vide alors L^* est infini.

Exercice 5 - Caractérisation

Solution page 7

$\mathrm{Rat}(\Sigma)$ est le plus petit sous ensemble de $\mathcal{P}(\Sigma^*)$ contenant les langages élémentaires pour Σ et stable par produit, union et étoile (propriétés †).

Exercice 6 - Induction structurelle
Solution page 7

Prouver que si une propriété P portant sur des langages est telle que

- $P(\emptyset)$, $P(\{\varepsilon\})$ et $P(\{a\})$ pour $a \in \Sigma$ sont vérifiées
- la vérité de $P(L_1)$ et $P(L_2)$ implique celle de $P(L_1 \cup L_2)$, $P(L_1 \cdot L_2)$ et $P(L_1^*)$

alors $P(L)$ est vraie pour tout langage rationnel.

Exercice 7 - Première équivalence
Solution page 7

Prouver que les langages rationnels sont les langages réguliers.

II Langages

Exercice 8 - Exemples

Prouver les égalités de langages suivantes :

1. $\emptyset \cdot L = L \cdot \emptyset = \emptyset$.
2. $\{\varepsilon\} \cdot L = L \cdot \{\varepsilon\} = L$.
3. $\{u\} \cdot \{v\} = \{u \cdot v\}$.
4. $\Sigma^+ = \Sigma \cdot \Sigma^*$.
5. $\{a\} \cdot \{b, \varepsilon\} \cdot \{a, b\} = \{a^2, ab, aba, ab^2\}$.
6. l'ensemble des mots de Σ qui commencent par a est $\{a\} \cdot \Sigma^*$.
7. $\Sigma^* \cdot \{a\} \cdot \Sigma^*$ est le langage des mots contenant au moins une fois la lettre a.

Exercice 9 - Alphabet à une lettre
Solution page 7

On considère $\Sigma = \{a\}$ et $L_k = \{a^{3n+k}, n \in \mathbb{N}\}$ pour $k \in \mathbb{N}$.
Déterminer $L_k \cap L_j$.
Déterminer $L_k \cdot L_j$.

Exercice 10 - Racine d'un langage
Solution page 7

Pour un langage L, sa racine est $\sqrt{L} = \{u \in \Sigma \ / \ u \cdot u \in L\}$.
Déterminer $\sqrt{L_k}$ où $L_k = \{a^{3n+k}, n \in \mathbb{N}\}$.
Montrer que $\epsilon \in L$ si et seulement si $\epsilon \in \sqrt{L}$.

Exercice 11 - Carré d'un langage et langage des carrés
Solution page 7

On a toujours $\{w \cdot w \ ; \ w \in L\} \subset L \cdot L$.
Pour quels langages a-t-on $L \cdot L = \{w \cdot w \ ; \ w \in L\}$?

Exercice 12 - Lemme d'Arden
Solution page 8

A et B sont deux langages sur un même alphabet. On suppose que L vérifie $L = A \cdot L \cup B$.

1. Montrer que $A^* \cdot B \subset L$.
2. Montrer que si $\epsilon \notin A$ alors $L = A^* \cdot B$.
3. Si $\epsilon \in A$ montrer que les solutions sont les ensembles de la forme $A^* \cdot C$ avec $B \subset C$.

Exercice 13 - Code Solution page 8

Un code sur Σ est un langage L sur Σ tel que l'égalité $u_1 \cdot u_2 \cdots u_p = v_1 \cdot v_2 \cdots v_q$ avec $u_1, \ldots, u_p$, $v_1, \ldots, v_q$ dans L entraîne $p = q$ et $u_i = v_i$ pour tout i.

1. Déterminer les codes parmi les langages finis suivants :
   - $L_1 = \{ab, baa, abba, aabaa\}$,
   - $L_2 = \{b, ab, baa, abaa, aaaa\}$,
   - $L_3 = \{aa, ab, aab, bba\}$,
   - $L_4 = \{a, ba, bba, baab\}$.
2. Soit $u \in \Sigma^*$, montrer que $\{u\}$ est un code si et seulement si $u \neq \varepsilon$.
3. Soit u et v deux mots distincts ; montrer que $\{u, v\}$ est un code si et seulement si $u \cdot v \neq v \cdot u$.
4. Soit L un langage ne contenant pas ϵ et tel qu'aucun mot de L ne soit préfixe d'un autre mot de L. Montrer que L est un code.

III Langages rationnels

Exercice 14 - Exemples de langages rationnels Solution page 9

- Prouver que l'ensemble des mots de longueur n au plus est rationnel.
- Prouver que l'ensemble des mots de longueur n au moins est rationnel.
- Prouver que l'ensemble des mots qui commencent et qui finissent par la même lettre est rationnel.

Exercice 15 - Langage des préfixes Solution page 9

L est un langage rationnel sur un alphabet Σ.
Prouver que le langage des mots préfixes des mots de L est rationnel.

Exercice 16 - Élimination d'une lettre Solution page 9

L est un langage rationnel sur un alphabet Σ et a est une lettre de Σ.
Prouver que chacun des langages suivants est rationnel.

- L'ensemble des mots de L qui ne contiennent pas a.
- L'ensemble des mots de L qui contiennent a.
- L'ensemble des mots obtenus en ôtant le premier a dans les mots de L contenant a.
- L'ensemble des mots obtenus en enlevant un a dans les mots de L contenant a.

IV Expressions régulières

Exercice 17 - 2 lettres Solution page 9

Donner une expression régulière dénotant l'ensemble L des mots sur l'alphabet $\Sigma = \{a, b\}$ tels que deux lettres consécutives soient toujours distinctes.
Même question si on impose de plus que les mots de L sont non vides.

Exercice 18 - 3 lettres

Solution page 9

Donner une expression régulière dénotant l'ensemble L des mots sur l'alphabet $\Sigma = \{a, b, c\}$ tels que deux lettres consécutives soient toujours distinctes.

Exercice 19 - Complémentaire

Solution page 10

Donner une représentation régulière dénotant $\overline{L}$ avec L dénoté par `(a|b)*aba(a|b)*`.

Exercice 20 - Expressions équivalentes

Solution page 10

- Montrer que $(\mathtt{a}|\mathtt{b})*$, `(a*b)*a*`, `a*|a*b(a|b)*` et `b*(aa*bb*)*a*` sont équivalentes.
- Prouver que $(\mathtt{r}_1\mathtt{r}_2)*$ et $\epsilon|\mathtt{r}_1(\mathtt{r}_2\mathtt{r}_1)*\mathtt{r}_2$ sont équivalentes.

Exercice 21 - Détermination d'expressions régulières

Solution page 10

Déterminer une expression régulière dénotant les langages suivants.

- L'ensemble des mots qui contiennent au moins un a,
- L'ensemble des mots qui contiennent au plus un a,
- L'ensemble des mots tels que toute série de a soit de longueur paire,
- L'ensemble des mots dont la longueur n'est pas divisible par 3,
- L'ensemble des mots tels que deux lettres consécutives soient toujours distinctes,
- L'ensemble des mots contenant au moins un a et un b.

Exercice 22 - Application du lemme d'Arden

Solution page 10

Sur $\Sigma = \{a, b\}$, on note L_1 le langage des mots ayant un nombre pair de b et L_2 le langage des mots ayant un nombre impair de b. Écrire deux relations linéaires liant L_1 et L_2 puis utiliser le lemme d'Arden pour donner des expression rationnelles les caractérisant. expression.

Parmi les langages élémentaires nous avons ajouté $\emptyset$ et $\{\varepsilon\}$. Nous allons voir qu'en fait ils servent uniquement à fabriquer l'ensemble vide ou, dans certains cas, à ajouter le mot vide.

Exercice 23 - Élimination de zéro

Solution page 10

Montrer que toute expression régulière est équivalente à $\varnothing$ ou à une expression régulière ne contenant pas $\varnothing$.

Une expression régulière est appelée *réduite* si elle ne contient ni $\varnothing$ ni ϵ.

Exercice 24 - Réduction

Solution page 11

Prouver que tout langage rationnel est dénoté par une expression régulière de la forme $\varnothing$, ϵ, `r` ou $\mathtt{r}|\epsilon$ avec `r` réduite.

V Langages locaux

Exercice 25 - Exemples Solution page 11

Prouver que les langages élémentaires sont locaux.

Exercice 26 - Exemples Solution page 11

Prouver que les langages $\{(ab)^n\ ;\ n \in \mathbb{N}\}$, $\{a^n b\ ;\ n \in \mathbb{N}^*\}$ et $\{a^n b^p\ ;\ n, p \in \mathbb{N}\}$ sont locaux.

Exercice 27 Solution page 11

Prouver que l'étoile d'un langage local sur Σ est un langage local.

Exercice 28 Solution page 11

Prouver que l'intersection de deux langages locaux sur Σ est un langage local.

Exercice 29 Solution page 11

Un langage L sur l'alphabet Σ est local si et seulement si il existe $P \subset \Sigma$, $S \subset \Sigma$ et $N \subset \Sigma^2$ tels que $L \setminus \{\varepsilon\} = \big(P \cdot \Sigma^* \cap \Sigma^* \cdot S\big) \setminus \Sigma^* \cdot N \cdot \Sigma^*$.

Exercice 30 Solution page 12

Donner un exemple de langage rationnel qui n'est pas local.

Exercice 31 Solution page 12

L'alphabet Σ peut s'écrire sous la forme avec $\Sigma = \Sigma_1 \cup \Sigma_2$ avec $\Sigma_1 \cap \Sigma_2 = \emptyset$.
L_1 est un langage local sur Σ_1 et L_2 est un langage local sur Σ_2.

1. Montrer que $L_1 \cup L_2$ est un langage local sur Σ.
2. Montrer que $L_1 \cdot L_2$ est un langage local sur Σ.
3. En déduire que si $\mathtt{r}$ est une expression régulière linéaire alors $L[\mathtt{r}]$ est un langage local.

Solutions

Exercice 1

1. Le préfixe (resp. suffixe) de longueur p de $x_1x_2\cdots x_n$ est $x_1x_2\cdots x_n$ (resp. $x_{n-p+1}x_{n-p+2}\cdots x_n$).
2. On note $u = u_1u_2\cdots u_n$, $v = v_1v_2\cdots v_p$ et $w = w_1w_2\cdots w_q$.
   Les deux produits $(u\cdot v)\cdot w$ et $u\cdot(v\cdot w)$ s'écrivent $t_1t_2\cdots t_{n+p+q}$
   avec $t_i = u_i$ pour $1 \leqslant i \leqslant n$, $t_i = v_{i-n}$ pour $n+1 \leqslant i \leqslant n+p$
   et $t_i = w_{i-n-p}$ pour $n+p+1 \leqslant i \leqslant n+p+q$.
3. $|u\cdot v| = |u| + |v|$ découle de la définition.
4. Si $u\cdot v = \varepsilon$ alors $|u| + |v| = 0$ avec $|u \geqslant 0$ et $|v| \geqslant 0$ donc $|u| = |v| = 0$ d'où $u = v = \varepsilon$.
5. Par récurrence sur m :
   $u^n\cdot u^0 = u^n\cdot\varepsilon = u^n = u^{n+0}$
   $u^n\cdot u^{m+1} = u^n\cdot\left(u^m\cdot u^1\right) = u^n\cdot\left(u^m\cdot u\right) = \underbrace{\left(u^n\cdot u^m\right)\cdot u = u^{n+m}\cdot u}_{\text{hypothèse de récurrence}} = u^{n+m+1}$
6. Si $u\cdot v = u\cdot v' = w$ alors $v = v'$ est l'unique suffixe de w de longueur $|w| - |u|$.
7. Si $u\cdot v = u'\cdot v = w$ alors $v = v'$ est l'unique préfixe de w de longueur $|w| - |u|$.

Exercice 2

Réflexivité u est un préfixe de u donc $u \leqslant u$.

Antisymétrie On suppose qu'on a $u \leqslant v$.

- Si u est un préfixe de v, $v = u\cdot u'$ alors u et v ne peuvent pas différer par une lettre en même position donc $v \leqslant u$ impose que v est un préfixe de u. Dans ce cas $|u| \leqslant |v|$ et $|v| \leqslant |u|$ d'où $|u| = |v|$. u et v sont donc le seul préfixe de v de longueur $|v|$: $u = v$.
- Si $u = w\cdot a\cdot u'$ et $v = w\cdot b\cdot v'$ avec $a \prec v$ dans Σ alors v n'est pas un préfixe de u et on ne peut avoir $v < u$.

Transitivité On suppose qu'on a $u \leqslant v$ et $v \leqslant w$.

- Si u est un préfixe de $v = u\cdot v'$ et si v est un préfixe de $w = v\cdot w'$ alors $w = u\cdot v'\cdot w'$ donc u est un préfixe de w : $u \leqslant w$.
- Si u est un préfixe de $v = u\cdot v'$ et si $v = (x_1x_2\cdots x_ia)\cdot v''$, $w = (x_1x_2\cdots x_ib)\cdot w'$ avec $a \prec b$ dans Σ
  alors soit $|u| \leqslant i$ donc u est un préfixe de w d'où $u \leqslant w$
  soit $i < |u|$ d'où $u = (x_1x_2\cdots x_ia)u'$ et $w = (x_1x_2\cdots x_ib)\cdot w'$ avec $a \prec b$ d'où $u < w$.
- Si $u = (x_1x_2\cdots x_ia)\cdot u'$, $v = (x_1x_2\cdots x_ib)\cdot v''$ avec $a \prec b$ dans Σ et v est un préfixe de w alors $w = (x_1x_2\cdots x_ib)\cdot v''\cdot w'$ avec $a \prec b$ d'où $u < w$.
- Si $u = (x_1x_2\cdots x_ia)\cdot u'$, $v = (x_1x_2\cdots x_ib)\cdot v''$ avec $a \prec b$ dans Σ et
  $v = (y_1y_2\cdots y_jc)\cdot v''$, $w = (y_1y_2\cdots y_jd)\cdot w'$ avec $c \prec d$ dans Σ .
  On note $k = \min\{i, j\}$: u et w commencent par les mêmes k premières lettretypes.
  On note l_u, l_v et l_w les lettres d'indice $k+1$ de u, v et w.
  Si $i = j = k$ alors $l_u = a \prec b = l_v = c \prec d = l_w$,
  si $i = k < j$ alors $l_u = a \prec b = l_v = y_{i+1} = l_w$
  si $i > k = j$ alors $l_u = x_{i+1} = l_v = c \prec d = l_w$.
  Dans les 3 cas on a $l_u \prec l_w$ donc $u \leqslant w$.

Exercice 3

On considère deux mots $u = x_1x_2\cdots x_n$ et $v = y_1y_2\cdots y_m$ sur Σ.
On suppose, par exemple, $n \leqslant m$.

- Si $x_i = y_i$ pour tout $i \in \{1, 2, \ldots, n\}$ alors u est un préfixe de v donc $u \leqslant v$.
- sinon il existe $p \leqslant n$ tel que $x_i = y_i$ pour tout $i \in \{1, 2, \ldots, p-1\}$ et $x_p \neq y_p$. Ainsi, selon qu'on a $x_p \prec y_p$ ou $y_p \prec x_p$, on a $u \leqslant v$ ou $v \leqslant u$.

Les deux mots sont donc comparables.

Exercice 4

1. $u \in L^* \iff \exists n\ ; u \in L_n \iff \exists n \exists u_1, u_2, \ldots, u_n \in L\ ; u = u_1 \cdot u_2 \cdots u_n$.
2. Si $L = \{w\}$ alors $L^n = \{w^n\}$ d'où $L^* = \{w^n\ ;\ n \in \mathbb{N}\}$.
3. $\emptyset^n = \emptyset$ pour $n \geqslant 1$ mais $\emptyset^0 = \{\varepsilon\}$. $\{\varepsilon\}^n = \{\varepsilon\}$.
4. Si L contient un mot non vide, w de longueur p alors L^* contient w^n de longueur np : la taille des mots n'est pas majorée donc l'ensemble est infini.

Exercice 5

$\mathrm{Rat}(\Sigma)$ contient les langages élémentaires (dans $\mathcal{R}_0(\Sigma)$).
Pour L_1 et L_2 appartenant à $\mathcal{R}$, il existe n_1 et n_2 tels que $L_1 \in \mathcal{R}_{n_1}(\Sigma)$ et $L_2 \in \mathcal{R}_{n_2}(\Sigma)$ d'où L_1 et L_2 appartiennent à $\mathcal{R}_n(\Sigma)$ avec $n = \max(n_1, n_2)$.
On a alors $L_1 \cdot L_2$, $L_1 \cup L_2$ et L_1^* qui appartiennent à $\mathcal{R}_{n+1}(\Sigma) \subset \mathrm{Rat}(\Sigma)$.
Ainsi $\mathrm{Rat}(\Sigma)$ vérifie les propriétés †.

Inversement soit $\mathcal{L}$ une partie de $\mathcal{P}(\Sigma^*)$ vérifiant †.
$\mathcal{L}$ contient les langages élémentaires donc $\mathcal{R}_0(\Sigma) \subset \mathcal{L}$.
Si on a $\mathcal{R}_n(\Sigma) \subset \mathcal{L}$ alors, comme $\mathcal{L}$ est stable par union, produit et étoile, $\mathcal{R}_{n+1}(\Sigma) \subset \mathcal{L}$; par récurrence on a ainsi $\mathcal{R}_n(\Sigma) \subset \mathcal{L}$ pour tout n.
On en déduit $\mathrm{Rat}(\Sigma) = \bigcup\limits_{n\in\mathbb{N}} \mathcal{R}_n(\Sigma) \subset \mathcal{L}$.

Exercice 6

On considère l'ensemble $\mathcal{L}$ des langages vérifiant P. $\mathcal{L}$ contient les langages élémentaires et est stable par union, produit et étoile donc contient $\mathrm{Rat}(\Sigma)$.

Exercice 7

Tout langage rationnel est régulier
On peut le faire par induction structurelle en notant $\mathcal{P}$ la propriété $\mathcal{P}(L)$ est régulier.
$\emptyset = L[\boldsymbol{\varnothing}]$, $\{\varepsilon\} = L[\boldsymbol{\epsilon}]$ et $\{x\} = L[x]$ sont réguliers.
Si $L_1 = L[\mathtt{r}_1]$ et $L_2 = L[\mathtt{r}_2]$ sont réguliers alors
$L_1 \cup L_2 = L[\mathtt{r}_1|\mathtt{r}_2]$, $L_1 \cdot L_2 = L[\mathtt{r}_1|\mathtt{r}_2]$ et $L_1^* = L[\mathtt{r}]$ sont réguliers.

Tout langage régulier est rationnel
On peut invoquer une induction structurelle sur la construction des expressions régulières.
On peut aussi effectuer une récurrence (généralisée) sur $|\mathtt{r}|$, le nombres de caractères de $\mathtt{r}$.
Si $|\mathtt{r}| = 1$ alors $\mathtt{r}$ dénote un langage élémentaire donc rationnel ;
Si $\mathtt{r} = \mathtt{r}_1\mathtt{r}_2$ alors $|\mathtt{r}_1| < |\mathtt{r}|$ et $|\mathtt{r}_2| < |\mathtt{r}|$ donc, par hypothèse de récurrence, $L[\mathtt{r}_1]$ et $L[\mathtt{r}_2]$ sont rationnels puis $L[\mathtt{r}] = L[\mathtt{r}_1] \cdot L[\mathtt{r}_2]$ est rationnel.
De même si $\mathtt{r} = \mathtt{r}_1|\mathtt{r}_2$ ou $\mathtt{r} = \mathtt{r}_1*$.

Exercice 9

$L_k \cap L_j = \emptyset$ si $k - j$ n'est pas un multiple de 3, $L_k \cap L_j = L_p$ avec $p = \min(k, j)$ sinon.
$L_k \cdot L_j = L_{k+j}$

Exercice 10

$\sqrt{L_{2p}} = L_p$, $\sqrt{L_{2p+1}} = L_{p+2}$.

Si $\epsilon \in L$ alors $\epsilon \cdot \epsilon = \epsilon \in L$ donc $\epsilon \in \sqrt{L}$.
Si $\epsilon \in \sqrt{L}$ alors $\epsilon = \epsilon \cdot \epsilon \in L$.

Exercice 11
On suppose L non vide.
Soit p la longueur minimale des mots de L vérifiant $L \cdot L \subset \{w \cdot w\ ;\ w \in L\}$.
S'il existait deux mots, u et v, de longueur p alors il devrait exister $w \in L$ tel que $u \cdot v = w \cdot w$.
Dans ce cas $2|w| = |w \cdot w| = |u \cdot v| = 2p$ donc $|w| = p$ et w est le préfixe de longueur p de $u \cdot v$ d'où $w = u$ et, de même, $w = v$ ce qui est impossible.
Soit u_0 l'unique mot de L de longueur minimale p.
Si on avait $L \neq \{u_0\}$, on considère un mot v de $L \setminus \{u_0\}$ de longueur minimale $q > p$. Le produit $u_0 \cdot v \in L \cdot L$ peut s'écrire $u_0 \cdot v = w \cdot w$; on a $2p < p + q = |u_0 \cdot v| = |w \cdot w| = 2|w|$ donc $|w| > p$ d'où $|w| \geqslant q$ puis $2q \leqslant |w \cdot w| = |u_0 \cdot v| = p + q$ ce qui est impossible.
Les langages tels que $L \cdot L = \{w \cdot w\ ;\ w \in L\}$ ne peuvent avoir qu'un élément.
Inversement tout langage singleton vérifie $L \cdot L = \{w \cdot w\ ;\ w \in L\}$, ainsi que le langage vide.

Exercice 12

1. On a $B \subset A \cdot L \cup B = L$. Alors $A \cdot B \subset A \cdot L \subset A \cdot L \cup B = L$.
   Par récurrence, si $A^n \cdot B \subset L$ alors $A^{n+1} \cdot B = A.(A^n \cdot B) \subset A \cdot L \subset A \cdot L \cup B = L$.
   Ainsi $A^n \cdot B \subset L$ pour tout n donc $A^* \cdot B = \bigcup_{n \in \mathbb{N}} A^n \cdot B \subset L$.
2. On suppose que $\epsilon \notin A$ et que $L \neq A^* \cdot B$. On a donc, comme $A^* \cdot B \subset L$, $L \setminus A^* \cdot B \neq \emptyset$.
   Soit u un mot de longueur minimale appartenant à $L \setminus A^* \cdot B$.
   $u \in L = A \cdot L \cup B$ donc
   - soit $u \in B \subset A^* \cdot B$ ce qui est impossible
   - soit $u \in A \cdot L$ donc $u = u_1 \cdot u_2$ avec $u_1 \in A$ et $u_2 \in L$. On a $u_1 \neq \epsilon$ car $\epsilon \notin A$ donc $|u_1| > 0$ puis $|u_2| < |u|$. On a ainsi $u \in L$ de longueur strictement inférieure à celle de $|u|$ donc $u_2 \in A^* \cdot B$ puis $u = u_1 \cdot u_2 \in A.(A^* \cdot B) \subset A^* \cdot B$ ce qui est impossible aussi.

   On doit donc avoir $L \subset A^* \cdot B$ d'où l'égalité.
3. On a $A^0 \cdot L = L \subset L$. Si $A^n \cdot L \subset L$ alors $A^{n+1} \cdot L = A.(A^n \cdot L) \subset AL \subset L$ donc, par récurrence sur n on a $A^n \cdot L \subset L$ pour tout n puis $A^* \cdot L \subset L$.
   L'inclusion inverse est évidente donc $L = A^* \cdot L$ avec $B \subset A \cdot L \cup B = L$.

   On suppose maintenant que $\epsilon \in A$; soit $C \subset \Sigma^*$ tel que $B \subset C$.
   On veut montrer que $L = A^* \cdot C$ vérifie $A \cdot L \cup B = L$.
   On a $A.(A^* \cdot C) = A^+ \cdot C \subset A^* \cdot C$ et, comme $\epsilon \in A$, $A^* \cdot C = \{\epsilon\}.(A^* \cdot C) \subset A.(A^* \cdot C)$ donc $A^* \cdot C = A.(A^* \cdot C)$. De plus $B \subset C \subset A^* \cdot C$ donc
   $A \cdot L \cup B = A.(A^* \cdot C) \cup B = A^* \cdot C \cup B = A^* \cdot C = L$: $A^* \cdot C$ est bien solution.

Exercice 13

1. - $L_1 = \{ab, baa, abba, aabaa\}$: $ab.(baa)^3 = abba.ab.aabaa$, L_1 n'est pas un code
   - $L_2 = \{b, ab, baa, abaa, aaaa\}$:
     on suppose que $u_1 \cdot u_2 \cdots u_p = v_1 \cdot v_2 \cdots v_q$ avec $u_1 \neq v_1$ (sinon on simplifie).
     Si, par exemple, $|u_1| \leqslant |v_1|$ on voit que les seules possibilités sont $u_1 = b$ et $v_1 = baa$ ou $u_1 = ab$ et $v_1 = abaa$ ce qui impose que u_2 commence par aa donc $u_2 = aaaa$.
     v_2 doit alors commencer par aa donc $v_2 = aaaa$ et on voit que tous les termes restant sont $aaaa$ sans jamais pouvoir égaler les longueurs. L_2 est un code
   - $L_3 = \{aa, ab, aab, bba\}$:
     on suppose que $u_1 \cdot u_2 \cdots u_p = v_1 \cdot v_2 \cdots v_q$ avec $u_1 \neq v_1$ (sinon on simplifie).
     Si, par exemple, $|u_1| \leqslant |v_1|$ on voit que $u_1 = aa$ et $v_1 = aab$ donc $u_2 = bba$ et v_2 devrait commencer par ba ce qui est impossible. L_3 est un code
   - $L_4 = \{a, ba, bba, baab\}$: $ba.a.bba = baab.ba$, L_4 n'est pas un code
2. $u^n = u^p$ impose $n|u| = p|u|$ donc $n = p$ si $|u| \neq 0$. Ainsi $\{u\}$ est un code si $u \neq \varepsilon$.
   Par contre $\varepsilon = \varepsilon.\varepsilon$ donc $\{\varepsilon\}$ n'est pas un code.

3. Si $u \cdot v = v \cdot u$ alors on a 2 décompositions d'un même mot : $\{u, v\}$ n'est pas un code.
   On suppose que $\{u, v\}$ n'est pas un code. On peut écrire $u_1 \cdot u_2 . \cdots \cdot u_p = v_1 \cdot v_2 . \cdots \cdot v_q$ avec $u_1 \neq v_1$. Par exemple $u_1 = u$ et $v_1 = v$. Ainsi u et v sont préfixes d'un même mot donc, par exemple, u est préfixe de v : on pose $v = u \cdot v'$ on a $v' \neq \epsilon$ car $u \neq v$.
   Si $v' = u$ alors $v = u^2$ donc $u \cdot v = v \cdot u$.
   Si $u \neq v'$ on est alors ramené, en remplaçant v par $u \cdot v'$ à un mot dans $\{u, v'\}$ qui admet deux écritures distinctes avec u et v' distincts dans $\{u, v'\}^+$.
   On peut alors conclure par récurrence sur $|u| + |v|$.
   En effet On a $|u| + |v'| < |u| + |v|$ donc $u \cdot v' = v' \cdot u$ d'où $u \cdot v = u \cdot u \cdot v' = u \cdot v' \cdot u = v \cdot u$.
4. Si L n'est pas un code alors on peut écrire $u_1 \cdot u_2 . \cdots \cdot u_p = v_1 \cdot v_2 . \cdots \cdot v_q$ avec $u_1 \neq v_1$ (sinon on simplifie). Si, par exemple, $|u_1| \leqslant |v_1|$ on voit que u_1 est un préfixe de v_1. Ainsi, si aucun mot de L n'est préfixe d'un autre mot de L, alors L est un code.

Exercice 14

- L'ensemble des mots de longueur n au plus est fini donc rationnel.
- L'ensemble des mots de longueur n exactement, Σ^n, est fini donc rationnel. L'ensemble des mots de longueur n au moins est le produit des langages rationnels Σ^n et Σ^* donc est rationnel
- L'ensemble des mots qui commencent et qui finissent par la même lettre est l'union des $\{x\} \cdot \Sigma^* \cdot \{x\}$ pour x décrivant Σ. Il est rationnel.

Exercice 15

$f(L)$ est l'ensemble des préfixes des mots de L.
$f(\emptyset) = \emptyset$, $f(\{\epsilon\}) = \{\epsilon\}$, $f(\{x\}) = \{\epsilon, x\}$,
$f(L_1 \cup L_2) = f(L_1) \cup f(L_2)$, $f(L_1 \cdot L_2) = f(L_1) \cup L_1 \cdot f(L_2)$ et $f(L^*) = L^* \cdot f(L)$.
Par induction structurelle $f(L)$ est rationnel pour tout L rationnel.

Exercice 16

- $f_1(L)$ est l'ensemble des mots de L ne contenant pas a. $f_1(\emptyset) = \emptyset$, $f_1(\{\epsilon\}) = \{\epsilon\}$, $f_1(\{x\}) = \{x\}$ si $x \neq a$ et $f_1(\{a\}) = \emptyset$,
  $f_1(L_1 \cup L_2) = f_1(L_1) \cup f_1(L_2)$, $f_1(L_1 \cdot L_2) = f_1(L_1) \cdot f_1(L_2)$, $f_1(L^*) = \big(f_1(L)\big)^*$.
  Par induction structurelle $f_1(L)$ est rationnel pour tout L rationnel.
- $f_2(L)$ est l'ensemble des mots de L ne contenant pas a.
  $f_2(\emptyset) = \emptyset$, $f_2(\{\epsilon\}) = \emptyset$, $f_2(\{x\}) = \emptyset$ si $x \neq a$ et $f_2(\{a\}) = \{a\}$,
  $f_2(L_1 \cup L_2) = f_1(L_1) \cup f_2(L_2)$, $f_2(L_1 \cdot L_2) = f_2(L_1) \cdot L_2 \cup L_1 \cdot f_2(L_2)$,
  $f_2(L^*) = L^* \cdot f_2(L) \cdot L^*$.
  Par induction structurelle $f_2(L)$ est rationnel pour tout L rationnel.
- $f_3(L)$ est l'ensemble des mots obtenus en enlevant le premier a dans les mots de L.
  $f_3(\emptyset) = \emptyset$, $f_3(\{\epsilon\}) = \emptyset$, $f_3(\{x\}) = \emptyset$ si $x \neq a$ et $f_3(\{a\}) = \{\epsilon\}$,
  $f_3(L_1 \cup L_2) = f_3(L_1) \cup f_3(L_2)$, $f_3(L_1 \cdot L_2) = f_3(L_1) \cdot L_2 \cup f_1(L_1) \cdot f_3(L_2)$,
  $f_3(L^*) = \big(f_1(L)\big)^* \cdot f_3(L) \cdot L^*$.
  Par induction structurelle $f_3(L)$ est rationnel pour tout L rationnel.
- $f_4(L)$ est l'ensemble des mots obtenus en enlevant un a dans les mots de L.
  $f_4(\emptyset) = \emptyset$, $f_4(\{\epsilon\}) = \emptyset$, $f_4(\{x\}) = \emptyset$ si $x \neq a$ et $f_4(\{a\}) = \{\epsilon\}$,
  $f_4(L_1 \cup L_2) = f_4(L_1) \cup f_4(L_2)$, $f_4(L_1 \cdot L_2) = f_4(L_1) \cdot L_2 \cup L_1 \cdot f_4(L_2)$,
  $f_4(L^*) = L^* \cdot f_4(L) \cdot L^*$.
  Par induction structurelle $f_4(L)$ est rationnel pour tout L rationnel.

Exercice 17

Les lettres de L doivent alterner. Si on privilégie la période ab, on doit pouvoir ajouter un b au début et un a à la fin d'où `r=(ϵ|b)(ab)*(ϵ|a)`.
Si on impose que les mots sont non vides, on sépare les cas selon la première lettre d'où
`r=(b(ab)*(ϵ|a))|(a(ba)*(ϵ|b))`

Exercice 18
On isole les lettres c : les mots de L sont le la forme
$u_1.c \cdot u_2 \cdot c \cdots u_{n-1} \cdot c \cdot u_n$ où u_i est un mot du langage sur l'alphabet $\{a, b\}$ tels que deux lettres consécutives sont toujours distinctes et u_i non vide pour $2 \leqslant i \leqslant n-1$
D'après l'exercice précédent une expression régulière dénotant le langage est donc $\mathtt{r(cr_1)*cr|r}$ avec `r=(ϵ|b)(ab)*(ϵ|a)` et $\mathtt{r_1} = (\mathtt{b(ab)}*(\epsilon|\mathtt{a}))|(\mathtt{a(ba)}*(\epsilon|\mathtt{b}))$

Exercice 19
Le complémentaire est l'ensemble des mots qui ne contiennent pas aba.
Si on isole les a et les b toute suite non vide de a doit être suivie d'une suite d'au moins deux b donc les facteurs de base sont de la forme $a^n b^p$ avec $n \geqslant 1$ et $p \geqslant 2$. De plus on peut commencer par des b et finir par des a puis des b d'où $\overline{L}$ est dénoté par `b*(aa*bbb*)*a.b*`.

Exercice 20

- Pour la seconde expression on isole le lettres b ; elles sont séparées par des a^k.
  Pour la troisième on a séparé les mots ne contenant pas de b.
  Pour la dernière on isole les lettres.
- Les mots du langage dénoté par $(\mathtt{r1} \cdot \mathtt{r2})*$ sont soit le mot vide soit de la forme
  $u_1 \cdot v_1 \cdot u_2 \cdot v_2. \cdots \cdot u_n \cdot v_n$ avec $u_1 \in L[\mathtt{r_1}]$, $v_n \in L[\mathtt{r_2}]$ et $v_1 \cdot u_2 \cdot v_2. \cdots \cdot u_n \in L[(\mathtt{r_1 r_2})*]$.

Exercice 21

- `(a|b)*a(a|b)*`
- `b*|b*ab*`
- `(aa|b)*`
- `((a|b)(a|b)(a|b))*(a|b|aa|ab|ba|bb)`
- `(ϵ|b)(ab)*(ϵ|a)` ou `(ab)* | a(ab)* | a(ab)*b | (ab)*b`
- S'il y a un a avant un b alors il y a un facteur ab, sinon il y a un facteur ba d'où l'expression régulière `(a|b)*(ab|ba)(a|b)*`
  On peut aussi distinguer la première lettre : `(aa*b|bb*a)(a|b)*`

Exercice 22
Si un mot contient un nombre pair de b alors il peut être ϵ ou est de la forme $a \cdot u$ avec u contenant un nombre pair de b ou est de la forme $b \cdot u$ avec u contenant un nombre impair de b. Ainsi $L_1 = \{\epsilon\} \cup \{a\} \cdot L_1 \cup \{b\} \cdot L_2$.
De même $L_2 = \{a\} \cdot L_2 \cup \{b\} \cdot L_1$.
Comme $\epsilon \notin \{b\}$ la dernière équation impose $L_2 = \{a\}^* \cdot \{b\} \cdot L_1$ d'où
$L_1 = \{\epsilon\} \cup \{a\} \cdot L_1 \cup \{b\} \cdot \{a\}^* \cdot \{b\} \cdot L_1 = (\{a\} \cup \{b\} \cdot \{a\}^* \cdot \{b\}) \cdot L_1 \cup \{\epsilon\}$.
Si on note $A = L[\mathtt{a|ba*b}]$, on a $L_1 = AA \cdot L_1 \cup \{\epsilon\}$ donc, comme $\varepsilon \notin A$,
$L_1 = A^* \cdot \{\epsilon\} = L[(\mathtt{a|ba*b})*]$ et $L_2 = L[\mathtt{a*b(a|ba*b)*}]$.

Exercice 23
$P(L)$ est la propriété :
L est vide ou est dénoté par un expression régulière ne contenant pas $\varnothing$.

- Pour les langages élémentaires la propriété est évidente.
- On suppose que $P(L_1)$ et $P(L_2)$ sont vraies.
  - Si L_1 est vide alors $L_1 \cup L_2 = L_2$ donc $P(L_1 \cup L_2)$ est vraie
  - De même si L_2 est vide
  - Si L_1 et L_2 sont non vides alors ils sont dénotés par des expressions régulières $\mathtt{r_1}$ et $\mathtt{r_2}$ sans $\varnothing$. $L_1 \cup L_2$ est alors dénoté par $(\mathtt{r_1}|\mathtt{r_2})$ qui ne contient pas $\varnothing$.
    Ainsi $P(L_1 \cup L_2)$ est vraie.
- On suppose que $P(L_1)$ et $P(L_2)$ sont vraies.
  - Si L_1 ou L_2 est vide alors $L_1 \cdot L_2$ est vide donc $P(L_1 \cdot L_2)$ est vraie

— Si L_1 et L_2 sont non vides alors ils sont dénotés par des expressions régulières $\mathtt{r}_1$ et $\mathtt{r}_2$ sans $\varnothing$. $L_1 \cdot L_2$ est dénoté par $\mathtt{r}_1\mathtt{r}_2$ qui ne contient pas $\varnothing$.
Ainsi $P(L_1 \cup L_2)$ est vraie.

- On suppose que $P(L)$ est vraie.
  — Si L est vide alors $L^* = \{\varepsilon\}$ est dénoté par ϵ donc $P(L^*)$ est vraie.
  — Si L est non vide alors il est dénoté par une expression régulière $\mathtt{r}$ ne contenant pas $\varnothing$. L^* est alors dénoté par `(r)*` qui ne contient pas $\varnothing$.
  Ainsi $P(L^*)$ est vraie.

Exercice 24
La démonstration se fait par induction structurelle.
La propriété est immédiate pour les langages élémentaires.
Dans le tableau on étudie les cas d'expressions régulières qui dénotent L_1 et L_2 et on donne une expression régulière pour $L_1 \cup L_2$, $L_1 \cdot L_2$ et L_1^*. $\mathtt{r}_1$ et $\mathtt{r}_2$ sont réduites.

L_1	L_2	$L_1 \cup L_2$	$L_1 \cdot L_2$	L_1^*
$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	ϵ
$\varnothing$	ϵ	ϵ	$\varnothing$	ϵ
$\varnothing$	$\mathtt{r}_2$	$\mathtt{r}_2$	$\varnothing$	ϵ
$\varnothing$	$\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_2\vert\epsilon$	$\varnothing$	ϵ
ϵ	$\varnothing$	ϵ	$\varnothing$	ϵ
ϵ	ϵ	ϵ	ϵ	ϵ
ϵ	$\mathtt{r}_2$	$\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_2$	ϵ
ϵ	$\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_2\vert\epsilon$	ϵ
$\mathtt{r}_1$	$\varnothing$	$\mathtt{r}_1$	$\varnothing$	$\mathtt{r}_1*$
$\mathtt{r}_1$	ϵ	$\mathtt{r}_1\vert\epsilon$	$\mathtt{r}_1$	$\mathtt{r}_1*$
$\mathtt{r}_1$	$\mathtt{r}_2$	$\mathtt{r}_1\vert\mathtt{r}_2$	$\mathtt{r}_1\mathtt{r}_2$	$\mathtt{r}_1*$
$\mathtt{r}_1$	$\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_1\vert\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_1\mathtt{r}_2\vert\mathtt{r}_1$	$\mathtt{r}_1*$
$\mathtt{r}_1\vert\epsilon$	$\varnothing$	$\mathtt{r}_1\vert\epsilon$	$\varnothing$	$\mathtt{r}_1*$
$\mathtt{r}_1\vert\epsilon$	ϵ	$\mathtt{r}_1\vert\epsilon$	$\mathtt{r}_1\vert\epsilon$	$\mathtt{r}_1*$
$\mathtt{r}_1\vert\epsilon$	$\mathtt{r}_2$	$\mathtt{r}_1\vert\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_1\mathtt{r}_2\vert\mathtt{r}_2$	$\mathtt{r}_1*$
$\mathtt{r}_1\vert\epsilon$	$\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_1\vert\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_1\mathtt{r}_2\vert\mathtt{r}_1\vert\mathtt{r}_2\vert\epsilon$	$\mathtt{r}_1*$

Exercice 25

- $\emptyset$ est défini par $P = F = S = \emptyset$ et il ne contient pas ε.
- $\{\varepsilon\}$ est défini par $P = F = S = \emptyset$ et il contient ε.
- $L = \{a\}$ est défini par $P = S = \{a\}$, $F = \emptyset$, $\varepsilon \notin L$.

Exercice 26

- $L = \{(ab)^n \ ; \ n \in \mathbb{N}\}$ est défini par $P = \{a\}$, $F = \{ab, ba\}$, $S = \{b\}$ et $\varepsilon \in L$.
  Si on retire ε on obtient $\{(ab)^n \ ; \ n \in \mathbb{N}^*\}$.
- $L = \{a^n b \ ; \ n \in \mathbb{N}^*\}$ est local, défini par $P = \{a\}$, $F = \{ab, aa\}$, $S = \{b\}$ et $\varepsilon \notin L$.
- $L = \{a^n b^p \ ; \ n, p \in \mathbb{N}\}$ est local, défini par
  $P = \{a, b\}$, $F = \{aa, ab, bb\}$, $S = \{a, b\}$ et $\varepsilon \in L$.

Exercice 27
Si $L \setminus \{\varepsilon\}$ est défini par (P, S, F) alors L^* contient ϵ et est défini par $(P, S, F \cup S.F)$.

Exercice 28
Si $L_1 \setminus \{\varepsilon\}$ et $L_2 \setminus \{\varepsilon\}$ sont associés aux triplets (P_1, S_1, F_1) et (P_2, S_2, F_2) alors $(L_1 \cap L_2) \setminus \{\varepsilon\}$ est local avec les paramètres $(P_1 \cap P_2, S_1 \cap S_2, F_1 \cap F_2)$.

Exercice 29
On suppose L local.
On note $N = \Sigma^2 \setminus F$, $L_1 = \big(P \cdot \Sigma^* \cap \Sigma^* \cdot S\big) \setminus \Sigma^* \cdot N \cdot \Sigma^*$

- Si $u = u_1 u_2 \cdots u_n = u_1 \cdot v \in L_1 \subset P \cdot \Sigma^*$ donc $u_1 \in P$.
  De même $u = w \cdot u_n \in L_1 \subset \Sigma^* \cdot S$ donc $u_n \in S$.
  Si on n'avait pas $u_i u_{i+1} \in F$ alors on aurait $u_i u_{i+1} \in N$ donc $u = (u_1 \ldots u_{i-1}) \cdot (u_i u_{i+1}) \cdot (u_{i+2} \ldots u_n) \in \Sigma^* \cdot N \cdot \Sigma^*$ ce qui est impossible. Ainsi $u_i u_{i+1} \in F$ pour tout i d'où $u \in L$. On a prouvé $L_1 \subset L$.
- Si $u = u_1 u_2 \cdots u_n \in L$ alors $u_1 \in P$ donc $u \in P \cdot \Sigma^*$ et $u_n \in S$ donc $u \in \Sigma^* \cdot S$.
  Si on avait $u \in \Sigma^* \cdot N \cdot \Sigma^*$ alors il existerait i tel que $u_i u_{i+1} \in N = \Sigma^2 \setminus F$ ce qui est impossible. Ainsi $u \in \big(P \cdot \Sigma^* \cap \Sigma^* \cdot S\big) \setminus \Sigma^* \cdot N \cdot \Sigma^* = L_2$. On a prouvé $L \subset L_1$.

Tout langage local est donc de la forme $\big(P \cdot \Sigma^* \cap \Sigma^* \cdot S\big) \setminus \Sigma^* \cdot N \cdot \Sigma^*$.
La même démonstration prouve que tout langage de la forme $\big(P \cdot \Sigma^* \cap \Sigma^* \cdot S\big) \setminus \Sigma^* \cdot N \cdot \Sigma^*$ est local en considérant $F = \Sigma^2 \setminus N$.

Exercice 30
Si L est dénoté par `ab(a|b)*`. alors ses préfixes de taille 1 sont $P = \{a\}$, ses suffixes de taille 1 sont $S = \{a, b\}$ et ses facteurs de taille 2 sont $F = \{aa, ab, ba, bb\}$, le langage local associé est dénoté par `a(a|b)*`, distinct de L.

Exercice 31
$L_1 \setminus \{\varepsilon\}$ est défini par (S_1, P_1, F_1) et $L_2 \setminus \{\varepsilon\}$ est défini par (S_2, P_2, F_2).

1. $L_1 \cup L_2 \setminus \{\varepsilon\}$ est inclus dans le langage défini par $(S_1 \cup S_2, P_1 \cup P_2, F_1 \cup F_2)$.
   Soit $u = x_1 x_2 \cdots x_n$ appartenant au langage local défini par $(S_1 \cup S_2, P_1 \cup P_2, F_1 \cup F_2)$.
   Si $x_1 \in \Sigma_1$ alors $x_1 x_2 \in F_1 \cup F_2$ impose $x_1 x_2 \in F_1$ donc $x_2 \in \Sigma_1$ et toutes les autres lettres appartiennent à Σ_1. On peut alors conclure que $u \in L_1$.
   De même si $x_1 \in \Sigma_2$ alors $u \in L_2$. Ainsi $u \in L_1 \cup L_2$ d'où l'égalité demandée.
2. Les facteurs de $L_1 \cdot L_2$ sont $F = F_1 \cup F_2 \cup S_1 \cdot P_2$.
   Les préfixes de longueur 1 de $L_1 \cdot L_2$ sont P_1 si $\epsilon \notin L_1$ et $P_1 \cup P_2$ sinon.
   Les suffixes de longueur 1 de $L_1 \cdot L_2$ sont S_2 si $\epsilon \notin L_2$ et $S_1 \cup S_2$ sinon.
   On vérifie ensuite que ces ensembles définissent dans tous les cas $L_1 \cup L_2 \setminus \{\varepsilon\}$.
3. On effectuer une récurrence (généralisée) sur $|\mathtt{r}|$, le nombres de caractères de `r` .
   Si $|\mathtt{r}| = 1$ alors `r` dénote un langage élémentaire donc local.
   Si $\mathtt{r} = \mathtt{r}_1 r_2$ ou $\mathtt{r} = \mathtt{r}_1 | \mathtt{r}_2$ alors $|\mathtt{r}_1| < |\mathtt{r}|$ et $|\mathtt{r}_2| < |\mathtt{r}|$ avec $\mathtt{r}_1$ et r_2 linéaires donc, d'après l'hypothèse de récurrence, $L[\mathtt{r}_1]$ et $L[\mathtt{r}_2]$ sont locaux.
   De plus $\mathtt{r}_1$ et r_2 peuvent être définis sur des alphabets distincts : les questions précédentes donnent que $L[r]$ est local.
   Si $\mathtt{r} = r_1 *$ alors $|\mathtt{r}_1| < |\mathtt{r}|$ et $\mathtt{r}_1$ est linéaires donc, d'après l'hypothèse de récurrence, $L[\mathtt{r}_1]$ est local. L'exercice 27 permet de conclure que $L[r]$ est local.