Langages Formels – partiel

durée : 2 heures

jeudi 19 mars 2026, 14h

Consignes. Aucun document autorisé. Vous pouvez utiliser les résultats vus en cours ou dans des questions précédentes, si vous les citez de manière claire et non ambiguë. Les résultats vu en TD ou ailleurs qu'en cours, s'ils sont utilisés, doivent êtres redémontrés. Seules les notions et notations du cours et du sujet seront admises dans les réponses. La lisibilité et la concision seront appréciées.

Pour tout ce sujet, on fixe un alphabet fini Σ avec $|\Sigma| \geq 2$.

Définition A. Pour $u, v \in \Sigma^*$, on écrit $u \sqsubseteq v$ si u est un sous-mot de v. C'est-à-dire que si $u = u_1 \dots u_n$, avec $u_1, \dots, u_n \in \Sigma$, alors $u \sqsubseteq v$ si, et seulement si, $v \in \Sigma^* u_1 \Sigma^* \cdots \Sigma^* u_n \Sigma^*$. On écrit $u \sqsubset v$ si $u \sqsubseteq v$ et $u \neq v$. Pour tout $S \subseteq \Sigma^*$, on note $\uparrow S$ la clôture vers le haut de S pour $\sqsubseteq$, c'est-à-dire

$$\uparrow S := \{w \in \Sigma^* \mid \text{ il existe } s \in S \text{ tel que } s \sqsubseteq w\}.$$

Lorsque $S = \{w\}$, on écrit $\uparrow w$ au lieu de $\uparrow \{w\}$.

Question 1. Donner une grammaire linéaire à droite pour le langage $\uparrow abb$, sur l'alphabet $\{a, b\}$.

Définition B. Soit $L \subseteq \Sigma^*$. Le langage L est *positivement testable par morceaux* (ptpm) si L est fermé vers le haut pour $\sqsubseteq$, c'est-à-dire que pour tout $u \in L$ et $v \in \Sigma^*$, si $u \sqsubseteq v$ alors $v \in L$. On note PTPM la classe des langages positivement testable par morceaux (sur l'alphabet fixé Σ).

Définition C. Soit R une relation binaire sur un ensemble S. Une *antichaîne* pour R est un sous-ensemble A de S tel que, pour tous $a, b \in A$, on n'ait pas aRb. On dit que R satisfait la *condition de chaîne descendante* (DCC) s'il n'existe pas de suite $(s_n)_{n \in \mathbb{N}}$ telle que $s_{n+1} R s_n$ pour tout $n \in \mathbb{N}$. On dit que R est un *wqo* (well-quasi-order) si R satisfait DCC et il n'existe pas d'antichaine infinie pour R.

On admet le fait suivant sans preuve.

Lemme D (Higman). *La relation $\sqsubset$ est un wqo.*

Question 2. Prouver que, pour tout langage ptpm L, il existe un ensemble fini F tel que $L = \uparrow F$.

Question 3. Montrer que l'ensemble des langages ptpm est inclus dans l'ensemble des langages sans étoile, et que cette inclusion est stricte.

Définition E. Pour un langage $L \subseteq \Sigma^*$, on définit le *préordre syntaxique* $\preceq_L$ sur Σ^* par, pour $x, y \in \Sigma^*$,

$$x \preceq_L y \;:\Longleftrightarrow\; \text{pour tous } \alpha, \beta \in \Sigma^*, \text{ si } \alpha y \beta \in L \text{ alors } \alpha x \beta \in L.$$

Question 4. Prouver que L est ptpm si, et seulement si, pour tous $u, v \in \Sigma^*$ et $a \in \Sigma$, on a

$$uav \preceq_L uv.$$

Rappelons que la relation de Green $\mathcal{R}$ est définie par $u\mathcal{R}v$ si, et seulement si, il existe $x, y \in M$ tel que $ux = v$ et $vy = u$, la relation $\mathcal{L}$ est définie par $u\mathcal{L}v$ si, et seulement si, il existe $x, y \in M$ tel que $xu = v$ et $yv = u$, et la relation $\mathcal{J}$ est définie par $u\mathcal{J}v$ si, et seulement si, il existe $x, x', y, y' \in M$ tel que $xux' = v$ et $yvy' = u$.

Définition F. Un monoïde fini M est *$\mathcal{R}$-trivial* si, pour tous $u, v \in M$, $u\mathcal{R}v$ implique $u = v$, *$\mathcal{L}$-trivial* si, pour tous $u, v \in M$, $u\mathcal{L}v$ implique $u = v$, et *$\mathcal{J}$-trivial* si, pour tous $u, v \in M$, $u\mathcal{J}v$ implique $u = v$.

Remarquons que la congruence syntaxique $\sim_L$ est $\preceq_L \cap \succeq_L$, donc le monoïde syntaxique M_L peut être décrit comme $\Sigma^*/(\preceq_L \cap \succeq_L)$.

Question 5. Montrer que, pour tout langage ptpm L, le monoïde syntaxique M_L est $\mathcal{R}$-trivial.

Question 6. Soit M un monoïde fini et supposons que $n \in \mathbb{N}_{\geq 1}$ est tel que, pour tout $x \in M$, x^n est idempotent. Montrer que les propriétés suivantes sont équivalentes :

1. M est $\mathcal{R}$-trivial ;
2. pour tous $x, y, u \in M$, si $uxy = u$, alors $ux = u$;
3. pour tous $x, y \in M$, $(xy)^n x = (xy)^n$.

Rappelons que *la logique du premier ordre sur les mots* a un prédicat binaire $<$ et, pour chaque lettre $a \in \Sigma$, un prédicat unaire a.

Définition G. Une *phrase existentielle* est une formule de la logique du premier ordre sur les mots de la forme

$$\exists x_1 \dots \exists x_n \psi$$

où $n \geq 0$ et ψ est une formule ne contenant pas de quantificateurs, et ne contenant que les variables $x_1, \dots, x_n$. Un langage L est *définissable existentiellement* s'il existe une phrase existentielle ϕ telle que

$$L = \{w \in \Sigma^* \mid w \models \phi\}.$$

Remarquons que $\top$ est une phrase existentielle, donc Σ^* est définissable existentiellement.

Question 7. Prouver qu'un langage $L \subseteq \Sigma^*$ est ptpm si, et seulement si, L est définissable existentiellement.

Question 8. Montrer que le complémentaire d'un langage ptpm n'est pas nécessairement ptpm.

Question 9. Calculer l'automate minimal et le monoïde syntaxique du langage $(\uparrow ab)^{\complement}$.

Définition H. Un langage $L \subseteq \Sigma^*$ est *testable par morceaux* (tpm) s'il est une combinaison booléenne finie de langages ptpm. Plus explicitement : pour $\mathcal{K} \subseteq \mathcal{P}(\Sigma^*)$, on note $\mathcal{B}(\mathcal{K})$ le plus petit ensemble de langages fermé par unions finies et complémentation tel que $\mathcal{K} \subseteq \mathcal{B}(\mathcal{K})$. La classe des langages testable par morceaux est alors définie par $\mathsf{TPM} := \mathcal{B}(\mathsf{PTPM})$.

Rappelons qu'un monoïde N *divise* un monoïde M s'il existe un sous-monoïde M' de M et un morphisme surjectif de M' sur N.

Question 10. Soient $\sim_0, \sim_1, \sim_2$ des congruences de monoïde sur Σ^*, et supposons que

$$\sim_1 \cap \sim_2 \subseteq \sim_0 .$$

Montrer que $\Sigma^*/{\sim_0}$ divise $\Sigma^*/{\sim_1} \times \Sigma^*/{\sim_2}$.

Question 11. Montrer que, pour tous langages $L_1, L_2 \subseteq \Sigma^*$, le monoïde syntaxique $M_{L_1 \cup L_2}$ divise $M_{L_1} \times M_{L_2}$.

Question 12. Montrer que, si deux monoïdes finis M_1, M_2 sont $\mathcal{R}$-triviaux, alors leur produit $M_1 \times M_2$ est $\mathcal{R}$-trivial.

Nous admettons sans preuve le lemme suivant.

Lemme I. *Si M est $\mathcal{R}$-trivial et N divise M, alors N est $\mathcal{R}$-trivial.*

Question 13. Montrer que, pour tout langage testable par morceaux L, le monoïde syntaxique M_L est $\mathcal{R}$-trivial.

Nous admettons le fait suivant.

Lemme J. *Un monoïde fini M est $\mathcal{J}$-trivial si, et seulement si, M est à la fois $\mathcal{L}$-trivial et $\mathcal{R}$-trivial.*

Pour le reste de ce sujet, on fixe un monoïde fini $\mathcal{J}$-trivial M et un morphisme $f : \Sigma^* \to M$.

Définition K. Soit $w \in \Sigma^*$ un mot de longueur k. Pour $0 \leq i \leq k$, on note $w^{(i)}$ le préfixe de longueur i de w (en particulier $w^{(0)}$ est le mot vide et $w^{(k)} = w$). On dit qu'un entier $1 \leq i \leq k$ est *stable à droite dans* w si $f(w^{(i-1)}) = f(w^{(i)})$, et qu'il est *instable à droite* sinon. On définit $r(w)$ comme le sous-mot de w constitué uniquement des lettres situées aux positions instables à droite. Formellement, si $w = w_1 \dots w_k$ avec $w_i \in \Sigma$, alors $r(w) := w_{t_1} \dots w_{t_\ell}$, où $1 \leq t_1 < \cdots < t_\ell \leq k$ est la liste ordonnée des positions instables à droite. Le mot w est *réduit à droite* si $r(w) = w$, c'est-à-dire si toute position $1 \leq i \leq k$ est instable à droite. Lorsque $k = 0$, on pose $r(\epsilon) = \epsilon$ et on considère ϵ comme réduit à droite.

On définit de manière analogue, en utilisant les suffixes au lieu des préfixes, les notions de *positions stables à gauche*, *instables à gauche*, de mots *réduits à gauche*, et la fonction de *réduction gauche* ℓ.

Question 14. Soit $k \geq 1$ et $w = w_1 \dots w_k$ avec $w_1, \dots, w_k \in \Sigma$. Montrer que, si $1 \leq i < j \leq k$ est tel que $f(w^{(i)}) = f(w^{(j)})$, alors $i + 1$ est stable à droite dans w.

Question 15. Prouver qu'il existe au plus $|\Sigma|^{|M|}$ mots réduits à droite.

Question 16. Soit $w \in \Sigma^*$ et supposons que $r(w)a \sqsubseteq w$. Montrer que $f(wa) = f(w)$.

Définition L. On définit $\mathcal{F} := \{ra\ell \mid r \in RD,\ a \in \Sigma,\ \ell \in RG\}$, où RD est l'ensemble des mots réduits à droite et RG l'ensemble des mots réduits à gauche. Pour $w \in \Sigma^*$, on définit $F(w) := \{u \in \mathcal{F} \mid u \sqsubseteq w\}$.

Question 17. Soient $u, v \in \Sigma^*$ et $a \in \Sigma$. Montrer que si $f(uv) \neq f(uav)$, alors $F(uv) \subsetneq F(uav)$.

Définition M. Pour $n \geq 1$, une *n-chaîne alternante* pour L est une suite $w_1, \dots, w_{2n} \in \Sigma^*$ telle que pour tout $1 \leq i < 2n$, $w_i \sqsubseteq w_{i+1}$, et pour tout j impair $1 \leq j < 2n$, $w_j \in L$, tandis que pour tout j pair $1 < j \leq 2n$, $w_j \notin L$.

On admet le fait suivant.

Lemme N. *Si un langage L n'est pas testable par morceaux, alors pour tout $n \geq 1$, il existe une n-chaîne alternante pour L.*

Question 18. Montrer que pour tout $m \in M$, l'ensemble $f^{-1}(m)$ est testable par morceaux.

Question 19. En utilisant les résultats précédents, démontrer le théorème de Simon :

Théorème (Simon)**.** *Soit L un langage régulier. Alors L est testable par morceaux si, et seulement si, M_L est $\mathcal{J}$-trivial.*