

Option Informatique en Spé MP et MP*

Devoir à rendre après les vacances de la Toussaint

Complexité des opérations sur les langages rationnels

Résumé

La notion de *complexité* est restée longtemps intuitive : si chacun fait aisément la différence entre un objet «simple» (disons, le pinpon d'une sirène) et un objet «compliqué» (par exemple, un mouvement de quatuor à cordes), ce n'est que récemment que cette différence a reçu une définition rigoureuse (en fait, plusieurs). Ainsi, Gregory CHAITIN a proposé de définir la complexité d'une suite de n bits, comme la taille minimale d'un programme écrivant cette suite.

Dans le cas d'un langage rationnel, plusieurs points de vues sont possibles, puisqu'il existe diverses façons de décrire un tel langage : avec un expression rationnelle, ou avec un automate fini, déterministe ou non.

Dans ce texte, nous définissons la complexité d'un langage rationnel L comme la taille minimale (en nombre d'états) d'un automate fini, déterministe, complet, reconnaissant L .

Nous nous intéresserons à l'effet des opérations usuelles sur cette complexité : par exemple, si l'on connaît les complexités de deux langages rationnels L et M , que peut-on dire de la complexité de $L \cup M$?

Veillez rédiger chaque partie sur une copie séparée.

Table des matières

1	Définitions et notations	2
2	Préliminaires arithmétiques	2
3	Idéaux à droite de Σ^*	2
4	Complexité d'un langage rationnel	3
5	Langages finis sur l'alphabet $\{a\}$	3
6	Langages finis sur l'alphabet $\{a,b\}$	4

1 Définitions et notations

► Dans tout le texte, Σ désigne un alphabet. Les automates finis considérés sont tous déterministes ; lorsqu'ils sont complets, nous noterons $q \cdot x$ au lieu de $\delta(q, x)$ et $q \cdot u$ au lieu de $\delta^*(q, u)$.

► Un langage L sur l'alphabet Σ est *cofini* si son complémentaire $\Sigma^* \setminus L$ est fini.

► Soit \mathcal{A} un automate fini. Un état q de \mathcal{A} est un *puits* si aucune transition ne sort de q ; si \mathcal{A} est complet, nous aurons donc $q \cdot x = q$ pour toute lettre x . Un *état poubelle* de \mathcal{A} est un état puits qui n'est pas final.

► Un automate fini déterministe complet est *réduit* si ses états sont tous accessibles et, à l'exception éventuelle d'un état poubelle, tous coaccessibles. Nous savons que tout langage rationnel L est reconnu par un automate réduit : partant d'un automate quelconque \mathcal{A} reconnaissant L , on détermine \mathcal{A} , on l'émonde et on le complète au besoin.

► Soient u et v deux mots sur un alphabet Σ . Nous dirons que v est *préfixe* de u s'il existe un mot w tel que $u = vw$. Nous dirons que v est *préfixe propre* de u lorsque $w \neq \varepsilon$.

2 Préliminaires arithmétiques

► Dans cette partie, p et q désignent deux éléments de $\mathbb{N} \setminus \{0,1\}$ premiers entre eux.

Question 1 Soit r un naturel vérifiant $0 < r < \min(p, q)$. Montrez qu'il existe des relatifs a et b tels que $ap + bq = r$, $0 < a < q$ et $-p < b < 0$.

Question 2 Notons $A_{p,q} = \{px + qy \mid x \in \mathbb{N} \text{ et } y \in \mathbb{N}\}$. Montrez que $\mathbb{N} \setminus A_{p,q}$ est une partie finie non vide de \mathbb{N} , et que son plus grand élément est $pq - p - q$.

Question 3 Notons $B_{p,q} = \{px + qy \mid x \in \mathbb{N}^* \text{ et } y \in \mathbb{N}\}$. Montrez que $\mathbb{N} \setminus B_{p,q}$ est une partie finie non vide de \mathbb{N} , et que son plus grand élément est $pq - q$.

Question 4 Notons $C_{p,q} = \{px + qy \mid x \in \mathbb{N}^* \text{ et } y \in \mathbb{N}^*\}$. Montrez que $\mathbb{N} \setminus C_{p,q}$ est une partie finie non vide de \mathbb{N} , et que son plus grand élément est pq .

3 Idéaux à droite de Σ^*

► Soit Σ un alphabet. Nous dirons qu'une partie L de Σ^* est un *idéal à droite* lorsqu'elle vérifie la condition suivante : si $u \in L$ et $x \in \Sigma$, alors $ux \in L$.

Question 5 Montrez qu'une partie L de Σ^* est un idéal à droite ssi $L = L \cdot \Sigma^*$.

Question 6 Quels sont les idéaux à droite de Σ^* lorsque $\Sigma = \{a\}$?

► Soient L un idéal à droite et B une partie de L . Nous dirons que B *engendre* L lorsque $L = B \cdot \Sigma^*$. Nous dirons que B est une *base* de L si elle engendre L , et si aucune partie B' strictement contenue dans B n'engendre L .

Question 7 Soient L un idéal à droite et B une partie génératrice de L . Montrez l'équivalence des deux assertions suivantes :

1. B engendre L , et aucun mot de B n'est préfixe propre d'un autre mot de B ;
2. B est une base de L .

Question 8 Soit L un idéal à droite. Notons $\mathcal{B}(L)$ l'ensemble des mots de L dont aucun préfixe propre n'est dans L . Montrez que $\mathcal{B}(L)$ est une base de L , et qu'il n'en existe pas d'autre.

► Nous dirons qu'un idéal L est *de type fini* si sa base est finie.

Question 9 Montrez que tout idéal à droite de type fini est rationnel.

Question 10 Exhibez un idéal à droite non rationnel.

Question 11 Existe-t-il des idéaux à droite rationnels, qui ne sont pas de type fini ?

4 Complexité d'un langage rationnel

► La *complexité* d'un langage rationnel L est le nombre minimal d'états d'un automate réduit reconnaissant L ; nous la noterons $\mathcal{C}(L)$.

► Soient L un langage rationnel et $n \in \mathbb{N}^*$. Nous dirons que L est un n -langage si $\mathcal{C}(L) = n$. Nous noterons $\text{Rat}_n(\Sigma^*)$ l'ensemble des n -langages sur l'alphabet Σ ; ainsi, $\text{Rat}(\Sigma^*) = \bigcup_{n \geq 1} \text{Rat}_n(\Sigma^*)$, cette union étant disjointe.

Question 12 Soit Σ un alphabet quelconque. Quels sont les 1-langages sur Σ ?

Question 13 Quels sont les 2-langages sur l'alphabet $\Sigma = \{a\}$?

Question 14 Soit L un n -langage sur un alphabet Σ . Soit \mathcal{A} un automate réduit à n états reconnaissant L ; montrez que \mathcal{A} possède au plus deux états puits.

Question 15 Soit L un langage rationnel sur l'alphabet Σ . Montrez que L et $\Sigma^* \setminus L$ ont la même complexité.

Question 16 Soit L un langage rationnel cofini autre que Σ^* . Montrez que $\mathcal{C}(L) \geq 2 + \max_{u \notin L} |u|$.

Question 17 Soient L un n -langage rationnel, et \mathcal{A} un automate réduit à n états reconnaissant L . Donnez une condition nécessaire et suffisante simple, portant sur \mathcal{A} , pour que L soit un idéal à droite.

5 Langages finis sur l'alphabet $\{a\}$

► Dans cette partie, Σ désigne l'alphabet à une seule lettre $\{a\}$. Soient L un n -langage fini sur Σ , et $\mathcal{A} = (Q, \delta, i, F)$ un automate fini réduit à n états reconnaissant L . Nous ne restreignons pas la généralité en supposant que l'ensemble Q des états de \mathcal{A} est l'intervalle discret $\llbracket 0, n-1 \rrbracket$, que l'état initial est 0, et que $q \cdot a = q+1$ pour $0 \leq q < n-1$.

Question 18 Que pouvez-vous dire de l'état $n-1$?

Question 19 Si $n \geq 2$, que pouvez-vous dire de l'état $n-2$?

Question 20 Montrez que la complexité d'un langage fini non vide L sur Σ est $\mathcal{C}(L) = 2 + \max_{u \in L} |u|$.

Question 21 Soient L et M deux langages finis sur Σ . Donnez des expressions simples de $\mathcal{C}(L \cup M)$ et $\mathcal{C}(L \cdot M)$, en fonction de $\mathcal{C}(L)$ et $\mathcal{C}(M)$.

Question 22 Soit $n \geq 1$. Montrez que l'ensemble des n -langages finis sur Σ est fini, non vide.

► Le résultat de la question précédente permet de définir la fonction g qui, à $n \geq 1$, associe la complexité maximale de L^* lorsque L décrit l'ensemble $\text{FRat}_n(\sigma^*)$ des n -langages finis :

$$g(n) = \max\{\mathcal{C}(L^*) \mid L \in \text{FRat}_n(\sigma^*)\}$$

Nous nous proposons de déterminer l'expression de $g(n)$ en fonction de n .

Question 23 Calculez $g(n)$ pour $1 \leq n \leq 4$.

► Nous allons établir la formule $g(n) = n^2 - 7n + 13$, pour $n \geq 5$.

Question 24 Vérifiez la validité de cette formule pour $n = 5$.

Question 25 Montrez que, pour $n \geq 5$, le langage $L_n = \{a^{n-3}, a^{n-2}\}$ vérifie $\mathcal{C}(L_n) = n$ et $\mathcal{C}(L_n^*) = n^2 - 7n + 13$.

► Soit L un langage fini, de complexité $n \geq 5$, reconnu par un n -automate \mathcal{A} .

Question 26 Supposons que \mathcal{A} n'a qu'un état final. Montrez que $\mathcal{C}(L^*) = n - 2$

Question 27 Supposons que \mathcal{A} a deux états finals. Montrez que $\mathcal{C}(L^*) \leq n^2 - 7n + 13$

Question 28 Examinez le cas où L est reconnu par un n -automate \mathcal{A} ayant plus de deux états finals.

Question 29 Et maintenant, concluez !

6 Langages finis sur l'alphabet $\{a,b\}$

► Soient $\mathcal{A}_1 = (Q_1, \delta_1, I_1, F_1)$ et $\mathcal{A}_2 = (Q_2, \delta_2, I_2, F_2)$ deux automates finis, déterministes ou non. Chaque partie F de $Q_1 \times Q_2$ permet de définir un *automate produit* $\mathcal{A}_1 \times \mathcal{A}_2$: l'ensemble de ses états est $Q_1 \times Q_2$; l'ensemble de ses états initiaux est $I_1 \times I_2$; l'ensemble de ses états finals est F ; et sa table de transition δ est définie comme suit :

$$((q_1, q_2), x, (q'_1, q'_2)) \in \delta \iff ((q_1, x, q'_1) \in \delta_1 \text{ et } (q_2, x, q'_2) \in \delta_2)$$

► En toute rigueur, la notation $\mathcal{A}_1 \times \mathcal{A}_2$ est ambiguë, puisque l'automate produit dépend de F . Ceci ne pose pas de problème dans la suite.

Question 30 Montrez que, si \mathcal{A}_1 et \mathcal{A}_2 sont déterministes et complets, alors $\mathcal{A}_1 \times \mathcal{A}_2$ l'est aussi.

Question 31 Utilisez la notion d'automate produit pour montrer que $\mathcal{C}(L \cup M) \leq \mathcal{C}(L)\mathcal{C}(M)$.

Question 32 En déduire que $\mathcal{C}(L \cap M) \leq \mathcal{C}(L)\mathcal{C}(M)$.

► Les majorants que nous venons d'établir ne sont pas des bornes supérieures ; toutefois, nous allons montrer qu'elles n'en sont pas très éloignées, en exhibant deux langages L_n et M_n tels que $\mathcal{C}(L_n) = \mathcal{C}(M_n) = 3n + 2$, et $\mathcal{C}(L_n \cap M_n) \geq (n + 1)^2$.

Question 33 Soit L_n l'ensemble des mots u tels que $|u|_a + 2|u|_b = 3n$. Montrez que L_n est rationnel, puis que $\mathcal{C}(L_n) = 3n + 2$.

► Notons de façon analogue M_n l'ensemble des mots u tels que $2|u|_a + |u|_b = 3n$.

Question 34 Caractérisez très simplement les mots du langage $L_n \cap M_n$.

Question 35 Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate réduit reconnaissant $L_n \cap M_n$. Soient (x, y) et (z, t) deux couples de naturels compris entre 0 et n inclus. Supposons $(x, y) \neq (z, t)$. Montrez que $i \cdot a^x b^y \neq i \cdot a^z b^t$ et concluez.

FIN