

<p>info502 : Théorie des langages TD 1 : expressions régulières</p>

Pierre Hyvernât

Laboratoire de mathématiques de l'université Savoie Mont Blanc

bâtiment Chablais, bureau 22, poste : 94 22

email : Pierre.Hyvernât@univ-smb.fr

www : <http://www.lama.univ-smb.fr/~hyvernât/>

Exercice 1 : expressions régulières POSIX

Les expressions régulières POSIX étendues utilisent

- “ x ” pour un caractère donné,
- “.” pour un caractère quelconque (sauf le saut de ligne),
- “[...]” pour “un caractère parmi ...”,
- “[^...]” pour “un caractère qui n’est pas parmi ...”.
- “ reg^* ” pour répéter l’expression reg un nombre arbitraire de fois,
- “ reg^+ ” pour répéter l’expression reg un nombre arbitraire de fois, mais au moins une fois,
- “ $reg?$ ” pour la chaîne vide ou bien l’expression reg ,
- “ $reg|reg$ ” pour la première ou la deuxième expression,
- “ $reg\{m, n\}$ ” pour l’expression reg répétée au moins m fois, et au plus n fois.

Question 1. Écrivez une expression régulière POSIX correspondant aux nombres flottants du langage C.

Question 2. Écrivez une expression régulière POSIX correspondant à une chaîne qui ne contient aucun symbole “” non échappé. Autrement dit, chaque “” doit être précédé d’un “\”.

Exercice 2 : Expressions régulières de Kleene

Rappel : les expressions régulières de Kleene ne peuvent utiliser que les symboles

- “ s ” pour n’importe quel symbole $s \in \Sigma$ de l’alphabet,
- “ \emptyset ” (parfois noté \emptyset) pour l’expression dénotant le langage vide,
- “ 1 ” (parfois noté ε) pour l’expression dénotant le mot vide (noté ε),
- “ $reg + reg$ ” pour une disjonction
- “ $reg \cdot reg$ ” ou une simple juxtaposition pour concaténer des expressions,
- “ reg^* ” pour l’étoile de Kleene.

Question 1. Donnez des expressions régulières sur l’alphabet $\Sigma = \{a\}$ pour les langages suivants :

- les mots de longueur multiple de trois,
- les mots non-vides de longueur multiple de trois,
- les mots de longueur multiple de trois ou de longueur multiple de cinq,
- les mots dont la longueur peut-être obtenue en ajoutant des 3 et des 5.

Question 2. Écrivez une expression régulière sur l’alphabet $\Sigma = \{a, b, c\}$ dont le langage associé est exactement l’ensemble des mots où a est toujours suivi de b et b est toujours suivi de a , sauf éventuellement pour le dernier symbole du mot.

Question 3. Écrivez une expression régulière sur l’alphabet $\Sigma = \{a, b\}$ dont le langage associé est exactement l’ensemble des mots qui ne contiennent jamais deux a consécutifs.

Question 4. Écrivez une expression régulière sur l’alphabet $\Sigma = \{a, b, c\}$ dont le langage associé est exactement l’ensemble des mots qui ne contiennent ni deux a consécutifs ni deux b consécutifs. (Difficile...)

Question 5. Rappelez l'algorithme récursif (vu en cours) utilisé pour tester si une expression régulière contient le mot vide ε .

Est-ce que les expressions régulières suivantes contiennent le mot vide ε ?

- $(a + ba^*)^* + b(a + (b + aba)^*)^*$
- $(1 + b)(aa^* + bb^*a)^*$
- $(1 + a)(1 + b)(1 + c)(1 + d)(e + f)$
- $(a + (b + (c + d)^*)^*)^*$

Question 6. Donnez un algorithme récursif similaire pour tester si le langage d'une expression régulière est vide.

Question 7. Que pensez-vous des langages associés aux expressions régulières suivantes :

- " R^* " et " $(R^*)^*$ ";
- " $(R_1R_2)^*$ " et " $1 + R_1(R_2R_1)^*R_2$ ";
- " $(R_1 + R_2)^*$ " et " $(R_1^*R_2)^*R_1^*$ ";
- " R^* " et " $(1 + R)(RR)^*$ ".