

Compilation : TD n° 2

Devan SOHIER

Exercice 1

Donnez des expressions régulières pour :

1. les entiers en binaire ;
2. les entiers impairs en binaire ;
3. les noms de variables en C ;
4. les entiers divisibles par 5 en base 10 ;
5. les mots sur $\Sigma = \{a, b\}$ contenant deux a consécutifs ;
6. les mots sur $\Sigma = \{a, b\}$ ne contenant pas deux a consécutifs ;
7. les mots sur $\Sigma = \{a, b, c\}$ ne commençant pas par un c ;
8. les entiers divisibles par 4 en base 10 ;
9. les constantes hexadécimales en C ;
10. les constantes C en notation scientifique.

Exercice 2

Montrer que les deux expressions régulières suivantes sont équivalentes : $(a|b)^*$ et $(ab^*)^*|(ba^*)^*$.

Exercice 3

Montrez que tous les langages finis sont réguliers.

Exercice 4 : analyse lexicale

Lecs est un programme créant un analyseur lexical à partir d'expressions régulières associées à des instructions. L'analyseur lexical exécute ces instructions lorsqu'une séquence de caractère est identifiée avec l'expression régulière correspondante.

Lorsque plusieurs expressions régulières peuvent être analysées, c'est celle qui permet d'analyser la plus longue chaîne qui est choisie. Lorsqu'une même chaîne peut être reconnue par plusieurs

expressions régulières, c'est la première dans l'ordre d'écriture qui est choisie. Les chaînes non-reconnues sont reproduites en sortie. Cette chaîne reconnue est stockée dans la variable `texte`.

Ainsi, l'analyseur lexical produit par le programme ci-dessous devra transformer la séquence `for 0a111 fort9a` en `BOUCLE NOMBRE VARIABLE VARIABLE`

```
for {écrire BOUCLE}  
[a-z][a-z0-9]* {écrire VARIABLE}  
[0-9]+ {écrire NOMBRE}
```

- Ecrire un programme `Lecs` supprimant les espaces, tabulations et retour à la ligne d'un texte fourni en entrée.
- Ecrire un programme `Lecs` supprimant les commentaires d'un code C (commentaires *"type C"* et *"type C++"*).
- Ecrire un programme `Lecs` remplaçant dans un code C les noms de variables (ou de fonction) par `VAR`, les constantes numériques et les chaînes de caractère par `CST`, les mots clefs `if`, `for` et `while` par `IF`, `FOR` et `WHILE`, et supprimant les commentaires (attention aux commentaires à l'intérieur d'une chaîne de caractère et réciproquement !)