

TP2 : Réalisation d'un analyseur lexical

Partie1 : Réalisation en langage C

L'objectif de cette partie est de réaliser un analyseur lexical en C du langage décrit par les unités lexicales suivantes:

1. **Les mots clés** : insensibles à la casse : si, sino, is, tq, qt, rpt, jsq, lre, ecr, vrai, Faux
2. **Les entiers naturels** : non signés composés par des suites non vides de chiffres.
3. **Les flottants** : les nombre réels signés en notation scientifique
4. **Les identificateurs** : les suites de lettres ou de chiffres commençant par une lettre
5. **Les opérateurs**: + - * / = < <= > >= == != && || !
6. **Les séparateurs**: ; , () :
7. **Les chaînes de caractères**: elles sont incluses entre double quotes "...". Dans une chaîne une séquence '\c' dénote le caractère c avec les exceptions suivantes :
 \t tab
 \n newline
Une chaîne ne peut contenir le nul (le caractère \0) ni le EOF
8. **Les commentaires** au style du C : sur une seule ligne introduits par "//" ou ceux multi-lignes encadrés par les couples "/*" et "*/". Cette dernière forme peut être imbriquée.

Remarque:

1. **Les blancs** : les suites non vides de : espace, tabulation et fin de ligne sont à ignorés.
2. **ERROR** ; est affiché dans le cas où :
 - Un caractère invalide est rencontré
 - Si une chaîne contient un caractère invalide (le nul par exemple) reporter l'erreur "Chaîne contenant le caractère nul", dans ce cas l'analyse reprendra après la fin de la chaîne
 - Si un commentaire reste ouvert alors qu'un EOF est rencontré reporter l'erreur : "Fin de fichier dans un commentaire". De même, si EOF est rencontré avant la quote de fermeture d'une chaîne reporter "Fin de fichier dans une constante chaîne"

L'analyseur lexical doit recevoir le code source dans un fichier texte et fournir son résultat dans un autre fichier de sortie en respectant le format suivant :

Unité Lexicale	Ligne	lexème
Mot Cle	1	si
Int	2	12
Float	2	1.3
.....		

Partie2

- Ecrire un fichier Lex qui définit la spécification du même langage puis générer l'analyseur lexical correspondant.