

Travaux Dirigés de Compilation n°2

ESIPE - Filière informatique

Analyse lexicale avec antlr

Le but de ce TD est de pratiquer les conventions d'antlr sur les expressions régulières.

► Exercice 1. Caractères spéciaux

1. Faites une grammaire `antlr` dans laquelle les tokens sont les séquences entre apostrophes incluses à l'intérieur d'une ligne du fichier. Votre grammaire doit sauter tout le reste. Faites deux versions :

- (a) une dans laquelle la partie qui reconnaît l'apostrophe est entre apostrophes,
- (b) l'autre dans laquelle la partie qui reconnaît l'apostrophe est entre crochets.

Compilez votre grammaire et testez l'analyseur obtenu sur `weblogic.log`.

2. Mêmes questions, mais les tokens sont les séquences entre crochets incluses à l'intérieur d'une ligne du fichier. Testez les deux analyseurs obtenus sur `error.log`.

► Exercice 2. Séquences d'échappement

Écrivez une grammaire `antlr` dans laquelle les tokens sont les chaînes de caractères entre guillemets qui respectent les normes du langage C. Votre grammaire doit sauter tout le reste. On demande que l'analyseur lexical reconnaisse les chaînes même si elle contiennent les séquences d'échappement suivantes :

- 1. `\n` (pour représenter un caractère fin de ligne),
- 2. `\t` (pour une tabulation),
- 3. `\"` (pour un guillemet),
- 4. `\\` (pour un caractère `"\"`),
- 5. `\` en fin de ligne (pour continuer la chaîne dans la ligne suivante).

Compilez votre grammaire et testez l'analyseur obtenu sur le fichier `c-strings.c`.