

Travaux Dirigés de Compilation n°1

Ingénieurs 2000 - IR2

—2005-2006—

Introduction aux expressions rationnelles

Le but de ce TD est d'écrire des expressions rationnelles caractérisant certains lexèmes, puis d'écrire des programmes Java permettant de les trouver.

► Exercice 1. *La commande grep*

Écrire une commande *shell* utilisant `grep` qui recherche dans la sous-arborescence de `/etc` :

- les lignes contenant la chaîne de caractères `IPV6` sans tenir compte de la casse ;
- les lignes qui commencent par une variable en majuscule suivie de `=` ;
- les lignes qui ne contiennent qu'une chaîne de caractères, éventuellement entourée d'espaces ;

► Exercice 2. *Expressions rationnelles*

Écrire des expressions rationnelles pour reconnaître :

1. les identificateurs du C (ils commencent par une lettre ou un `_`, puis une suite de chiffres, lettres ou `_`) ;
2. les chaînes de caractères du Pascal (`' ... '` avec `''` pour faire une apostrophe) ;
3. les chaînes de caractères du C ;
4. les entiers (décimal, hexadécimal `0x...` et octal `0...`) ;
5. les commentaires du C (`/* ... */` et `// ...`) ;

Pour tester vos expressions rationnelles, vous pouvez utiliser le programme `visual_regex`.

► Exercice 3. *Le paquetage java.util.regex*

Écrire le programme qui reconnaît dans un fichier passé en argument les URL HTTP (par exemple `http://user:passwd@machine:port/path`) et qui affiche, s'ils existent, l'utilisateur, le mot de passe, la machine, le port et le chemin.