

TD numéro 1 :

Analyse lexicale et syntaxique

Assembleur – Compilation, ENSIIE

Semestre 3, 2021–22

Exercice 1 : Analyse lexicale

Une constante entière est donnée par un signe – éventuel suivi soit d’une séquence non vide de chiffres compris entre 0 et 9, soit par un préfixe `0x` ou `0X` suivi d’une séquence non vide de chiffres entre 0 et 9 ou entre A et F (en majuscule ou minuscule), soit par un préfixe `0o` ou `0O` suivi d’une séquence non vide de chiffres entre 0 et 7, soit enfin par un préfixe `0b` ou `0B` suivi d’une séquence non vide de chiffres 0 ou 1.

Une constante réelle est donnée soit en forme décimale : un signe – éventuel suivi d’une séquence non vide de chiffres compris entre 0 et 9, d’un point, d’une séquence éventuellement vide de chiffres compris entre 0 et 9 ; soit en forme avec exposant : une forme décimale suivie d’un exposant constitué d’un préfixe `e` ou `E`, d’un éventuel signe `+` ou `-` et d’une séquence non vide de chiffres compris entre 0 et 9.

1. Donner les expressions régulières permettant de reconnaître ces constantes.
2. Donner la séquence de lexèmes produite par un outil comme `lex` à partir des entrées suivantes, en indiquant les erreurs le cas échéant :
 - a) `45-2.1e-3`
 - b) `0o123456789`
 - c) `0x6626e-37`
 - d) `430xdeadbeef`
 - e) `6.0221417e23.14159`

Exercice 2 : Analyse syntaxique

On considère la grammaire sur $\{a; b; c\}$:

$$N \rightarrow aNb \mid Nb \mid c$$

1. Quel est le langage généré par cette grammaire ?
2. Cette grammaire est-elle ambiguë ? Justifier.
3. Sur un exemple, montrer comment l’analyse ascendante peut mener à un conflit.
4. Construire l’automate LR(0) associé à la grammaire.
5. Y a-t-il un conflit ?
6. Proposer une grammaire non ambiguë reconnaissant le même langage. Est-elle LR(0) ?