

HAI601I : Analyse Syntaxique et interprétation TD1

❖ Analyse lexicale

Exercice 1 (TD Théorie des langages)

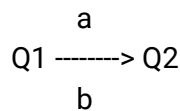
1. Comment peut-on caractériser un langage rationnel (régulier) ?

Un langage rationnel/régulier est la plus petite famille de langages qui contient les langages finis et qui est fermée pour les opérations réunion (U), produit (x) et étoile (*).

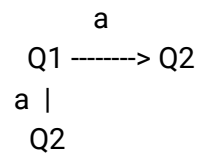
Le langage L est :

- défini par une expression régulière.
- généré par une grammaire régulière
- reconnu par un automate fini déterministe
- reconnu par un automate fini indéterministe.

déterministe :



indéterministe :



2. Les langages de programmation (C, Java, Python) sont-ils réguliers ? Pourquoi ?

Ce ne sont pas des langages réguliers car ils ont des structures emboîtées (ex: ()).

3. Comment peut-on caractériser un langage algébrique ?

4. Soit l'extrait suivant de grammaire EBNF de requête de consultation SQL :

```
SELECT
  [ALL | DISTINCT | DISTINCTROW ]
  select_expr [, select_expr] ...
  [into_option]
  [FROM table_references
    [PARTITION partition_list]]
  [WHERE where_condition]
  [GROUP BY {col_name | expr | position}
    [ASC | DESC], ... [WITH ROLLUP]]
  [HAVING where_condition]
  [ORDER BY {col_name | expr | position}
    [ASC | DESC], ...]
  [LIMIT {[offset,] row_count | row_count OFFSET offset}]
;
```

Expliquer les différentes conventions de notation utilisées

- **MAJUSCULE** : indique une commande, un mot-clé, un nom de paramètre ou un nom de fonction SQL.
- **minuscule** : indique une valeur fournie par l'utilisateur pour un identifiant, un paramètre ou un argument.
- **{ }** : indiquent des groupements d'identifiants, de paramètres ou d'arguments. Utilisées avec des "|" quand il y a plusieurs choix disponibles. Ne doivent pas être saisies.
- **[]** : indiquent les parties facultatives d'une instruction. Ils ne doivent pas être saisis.
- **|** : indique un choix entre deux ou plusieurs éléments ou valeurs, généralement entre [] ou { }.
- **...** : Le point précédent peut être répété un nombre illimité de fois.

5. les requêtes suivantes sont-elles correctes ?

`select 5 ;` **Correct**

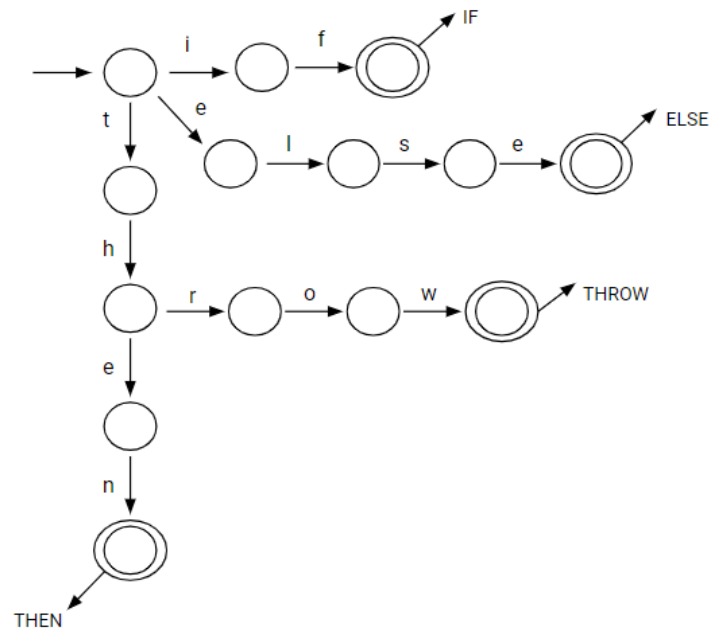
`SELECT distinct nom from etudiant` **Inorrect. Il manque le ";"**

`select ref from articles order by rayon asc, ref asc;` **Correct**

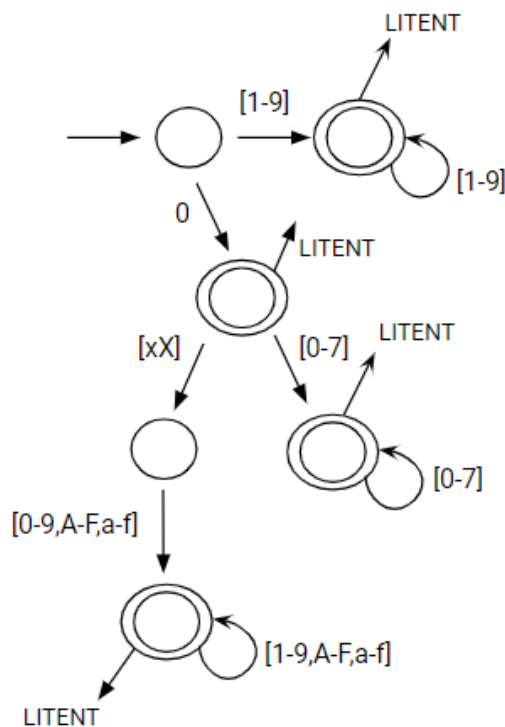
Exercice 2 (TD automate)

Dessiner un Automate à états Finis Déterministe (AFD) distinct pour chacun des langages suivants :

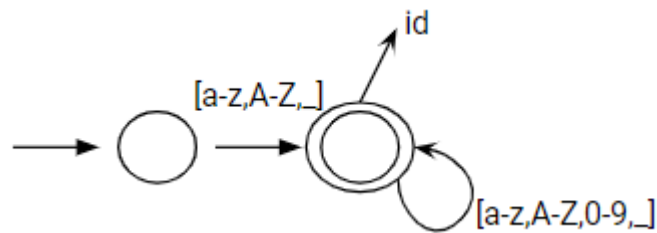
1. Langage de certains mots-clés du C : $L_{key} = \{if, then, else, throw\}$ (sensible à la casse).



2. Langage des littéraux numériques entiers du C (ou C++, ou Java), décimaux Lc10, octaux Lc8, hexadécimaux Lc16.

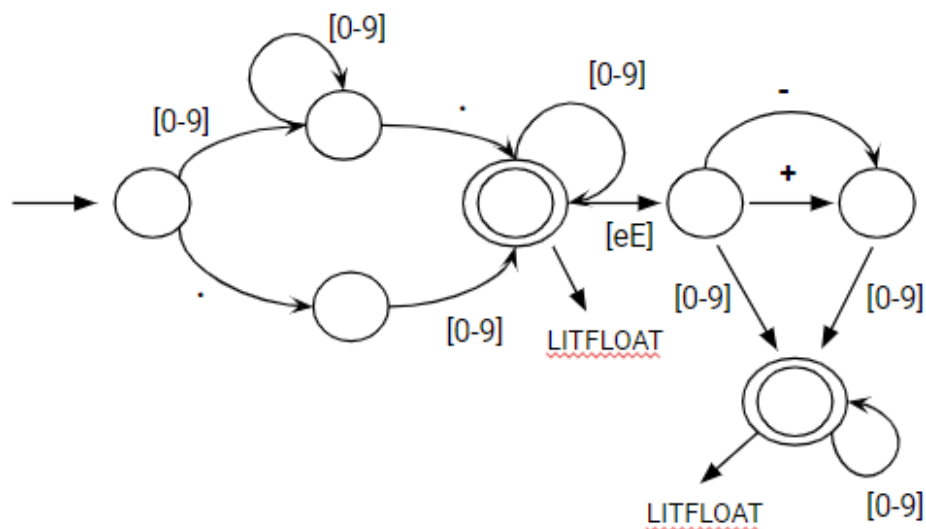


3. Langage Lid des identificateurs composés d'une lettre au moins, éventuellement suivie de chiffres, de lettres et de "_".

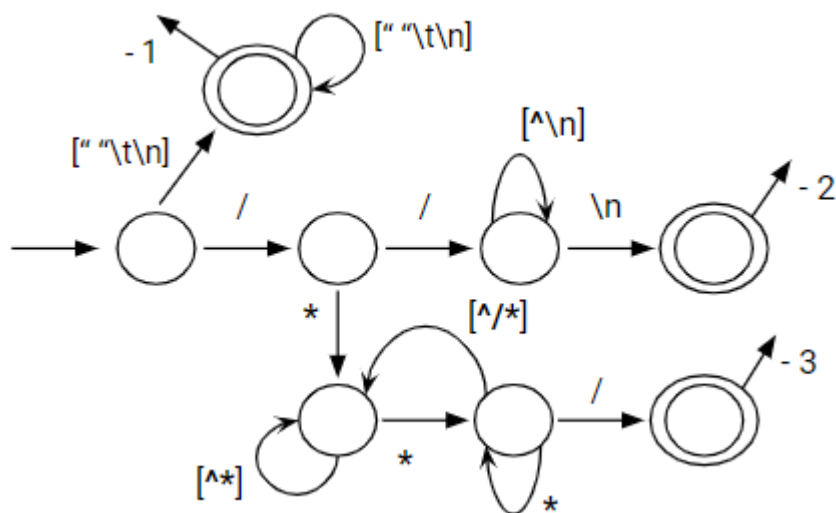


4. Langage des littéraux numériques flottants décimaux Lf. La suite de chiffres à gauche ou bien à droite du point décimal pouvant être vide. L'exposant entier n'est pas obligatoire.

Exemples : 13., 1.7e23, .89E-34



5. Langage Lsep des séparateurs composés de blancs (Espace, \t, \n), des commentaires à la C et à la C++.



Dessiner un unique AFD à jeton reconnaissant une partie de ces langages. Vous reconnaîtrez notamment : le mot-clé if, les identificateurs, les entiers décimaux, les flottants sans exposant, les séparateurs.

Utiliser des jetons négatifs pour les lexèmes à filtrer (séparateurs). flottants sans exposant, les séparateurs. Utiliser des jetons négatifs pour les lexèmes à filtrer (séparateurs).

