

Université de Rouen, UFR Sciences et
Techniques
Master 1 SSI – Compilation
M2L, traducteur L^AT_EX en MathML

Boubekri, Abdelmalek `abdelmalek.boubekri@etu.univ-rouen.fr`
Mezheri, Bilal `bilal.mezheri@etu.univ-rouen.fr`

2 janvier 2017

1 Introduction

Le projet consiste à réaliser un traducteur \LaTeX en MathML, le programme devra prendre les formules entourées du symbole \$ et produire le code MathML correspondant; et doit prendre en compte:

- les opérateurs;
- les identifiants;
- les constantes numériques;
- les fractions;
- les exposants et indices;
- les racines (sqrt et root);
- le groupement avec accolades;
- la somme et le produit itérés;

et tout ça sera intégré à un service web, idéalement la lecture et la sortie se feront sur l'entrée et la sortie standard.

2 Grammaire

AXIOME \rightarrow \$ FORMULE \$

FORMULE \rightarrow \sqrt RACINE FORMULE
| \frac FRACTION FORMULE
| UNDEROVER FORMULE
| CARACTERE FORMULE
| OPERATION FORMULE
| ϵ

RACINE \rightarrow [FORMULE]{FORMULE}
| {FORMULE}

FRACTION \rightarrow {FORMULE}{FORMULE}

OPERATION \rightarrow \wedge FORMULE
| _ FORMULE

UNDEROVER \rightarrow \sum_{FORMULE} {FORMULE} \wedge {FORMULE}
| \prod_{FORMULE} {FORMULE} \wedge {FORMULE}

CARACTERE \rightarrow IDENTIFIER | OPERATOR | NUMBER
| \pm | \alpha | \beta | \infty | \mp
| \cap | \cup | \subset | \supset | \subseteq | \supseteq

3 Fonctionnement

3.1 Analyse lexicale avec Flex

Le rôle de **Flex** est de découper la chaîne entrée en lexèmes (tokenizer) à reconnus grâce aux motifs définis, dans notre exemple les nombres, les identifiants ... ; puis les communique à **Yacc**

3.2 Analyse syntaxique avec Yacc

Yacc quant à lui, permet de vérifier la syntaxe de la chaîne de caractères à travers la grammaire décrite précédemment, puis d'afficher le résultat en MathML.

4 Utilisation

Récupérez l'archive depuis Github: <https://github.com/kabyliano/l2m.git>
Afin de compiler le programme il suffit de lancer le **Makefile**

```
Make
```

Puis vous pourrez lancer le programme **l2m** avec comme argument le flux \LaTeX ; exemple:

```
./l2m "\$\frac{5}{9}\$"
```

5 Problèmes non résolus

- Erreur de segmentation (core dumped) dû à la mauvaise gestion de la mémoire.
- Récursivité mal implémentée.
- Conflit Décalage/Réduction (*Shift/reduce*) dû aux conflits entre $_$ et \wedge (POW / UND) et SUM / PROD.