Antoine BERTRAND
Sébastien TAURAND
SICOM

24/09/2017

PROJET INFORMATIQUE 2ème année

ASSEMBLEUR MIPS 32 BITS

Livrable 1 : Analyse lexicale

Introduction

Ce rapport traite de la première partie du projet informatique en filière SICOM. L'objectif final de ce projet est de concevoir un assembleur en C pour un microprocesseur MIPS.

La première étape est de réaliser l'analyse lexicale du fichier d'entrée. Il faut donc traiter chaque lexème et reconnaitre sa nature.

Organisation du travail

Sur cette phase de début du projet nous avons beaucoup travaillé ensemble sans toutefois arrivé à se répartir de façon efficace les différentes tâches.

Nous avons également eu des difficultés au niveau de la tenue du planning.

Les Outils

Nous avons essayé de documenter le code et les structures de données en utilisant les commentaires respectant la syntaxe de Doxygen. Afin de documenter le diagramme d'état de la machine d'analyse lexicale, nous avons aussi du installer et apprendre à utiliser « dot ». Celui-ci donne un très bon rendu visuel pour l'effort demandé.

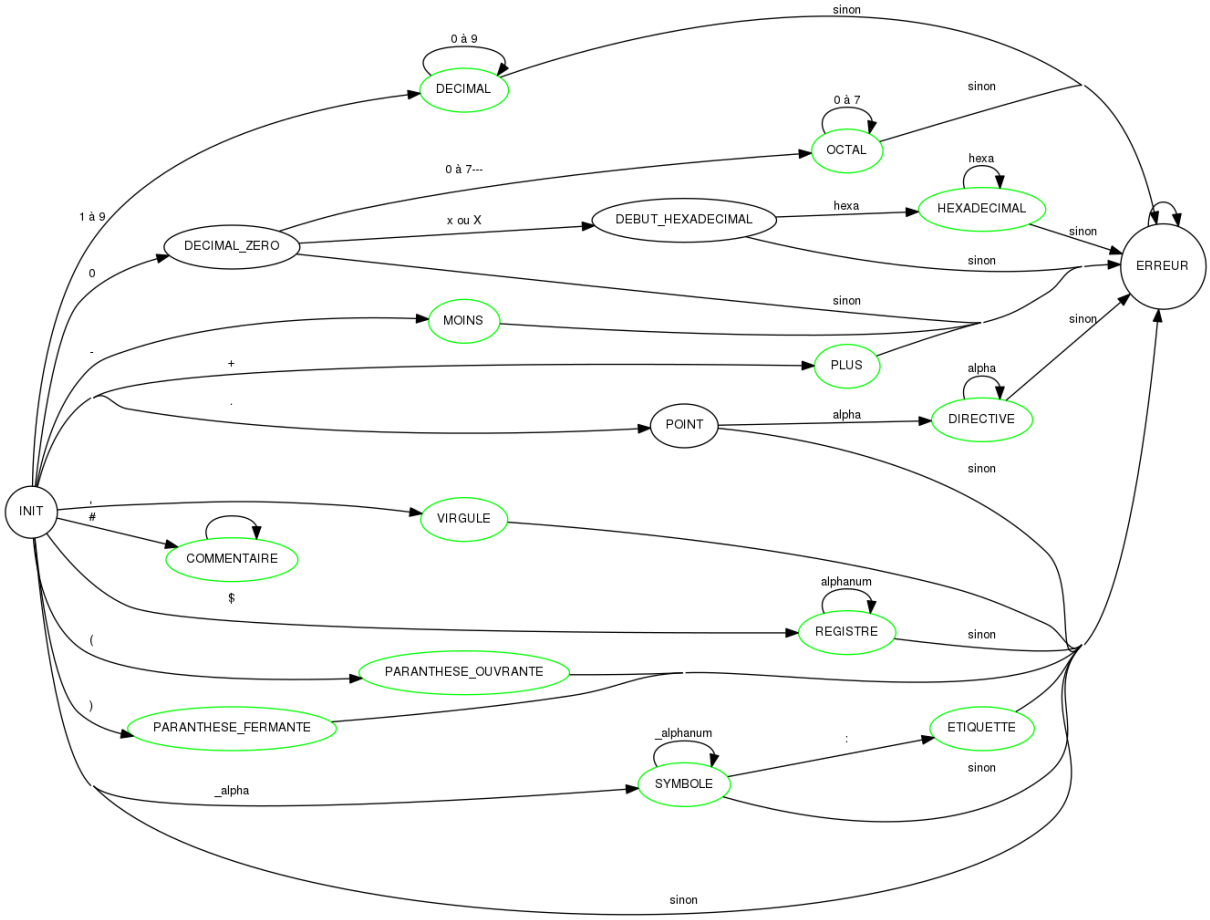
L'analyse lexicale

L'analyse lexicale du fichier assembleur est effectuée ligne par ligne.

Pour chaque ligne, une première phase de canonisation (standardisation) est réalisée :

- Remplacement des espaces par l'espace
- Suppression des espaces avant le ' :'
- Insertion de l'espace avant les caractères '$', ',', '(', ')', '-', '+'
- Insertion de l'espace après le ':' ', ' '-' '(' ')' '+'
- Suppression des espaces après '.'

La deuxième étape est d'effectuer une analyse lexème par lexème, ceux-ci étant séparés par au moins un espace. Les lexèmes sont isolés en utilisant la fonction strtok. On analyse alors la nature du lexème en utilisant la machine à états finis ci-après, chaque transition correspondant à un caractère du lexème analysé.



Machine à états finis d'analyse lexicale

La machine est initialisée pour le premier caractère du lexème. En fin du lexème, on regarde l'état de sortie de la machine :

- Seuls les états représentés en vert correspondent à des lexèmes valides.
- Les états en noir sont remplacés par un état d'erreur.
- Le lexème est alors inséré dans la ligne en cours pour une future analyse syntaxique.

Une fois la ligne terminée, celle-ci sera insérée dans une liste de lignes correspondant à l'intégralité de l'analyse lexicale du fichier source.

Test réalisés

Plusieurs fichiers de tests ont été créés ou modifiés afin de tester la robustesse de l'analyse lexicale. La vérification des résultats de tests est pour l'instant manuelle par analyse de l'opérateur.

Nous avons aussi vérifié à l'aide de valgrind que le programme n'avait pas de fuite mémoire. Pour se faire, nous avons utilisé la commande :

```
valgrind --leak-check=full -v ./as-mips tests/miam.s
```

Problèmes rencontrés

Nous avons commencé à coder avant d'avoir lu l'intégralité des documents recommandés, dont les excellents fichiers sources mis à notre disposition. Nous sommes repartis de ceux-ci par la suite. Cette décision nous aura couté une semaine de planning mais nous a permis de repartir sur des bases saines pour la suite du projet.

La solution au traitement du lexème commentaire ne nous a pas été triviale et nous avons dû discuter de stratégies avec d'autres étudiants.

Besoin d'installer PDFLatex sur la machine virtuelle PHELMA2016. Il a fallu installer Perl ainsi que les sources TextLive et dot.

Résultats

Le programme de test documenté, doté de génération de traces principales est testé avec un échantillon de fichiers tests nous semble pleinement fonctionnel et répondre aux attentes.

Points à améliorer

Créer un programme shell permettant de chainer automatiquement plusieurs fichiers de tests et de vérifier que les résultats sont conformes à ceux attendus (tests de non régression).

Meilleure répartition des tâches dans le binôme avec mise en place d'un suivi de jalons.

Utilisation de git en mode distant afin de faciliter la collaboration et l'efficacité du travail en binôme sur le code et les tests.

Améliorer les informations d'origine des états d'erreurs de l'analyse lexicale.