

Rapport : Projet IPI

GOLDNADEL Nathan

January 9, 2022

1 Introduction

Le projet de programmation impérative de cette année s'intéresse au fonctionnement des automates LR1, qui servent à la reconnaissance d'un langage décrit dans un fichier d'extension *.aut*.

2 Présentation des choix

2.1 Lecture des données

La première difficulté du projet est la lecture des données du fichier. La méthode retenue a été de s'aider de l'emplacement prédéfini et connu à l'avance des données d'un fichier respectant les normes d'un automate LR1. On manipulera à cet effet les indices afin d'accéder aux éléments dans un grand buffer qui correspondra à l'ensemble du fichier dans une chaîne de caractères.

2.2 Le fonctionnement de l'automate

En adéquation avec le principe de l'automate, on a besoin d'une structure de pile, qui a été implémentée à l'aide d'une structure de listes chaînées. On dispose alors des fonctions pour empiler, dépiler et obtenir le haut de la pile. Ainsi, l'état actuel est représenté par le sommet de la pile.

2.3 Les graphes

Pour faire le bonus, on a besoin d'écrire dans un fichier une sortie au format DOT, permettant ainsi de lire un automate sous la forme d'un graphe et comprendre visuellement comment ce dernier marche. On utilise alors essentiellement les informations de l'automate et on agit selon les valeurs des données pour arriver à construire la représentation graphique de notre automate. Ainsi, il fallait implémenter une structure de chaîne de caractères (au lieu d'utiliser celle du module *string*).

2.4 La structure du programme

Voici la structure du programme adopté :

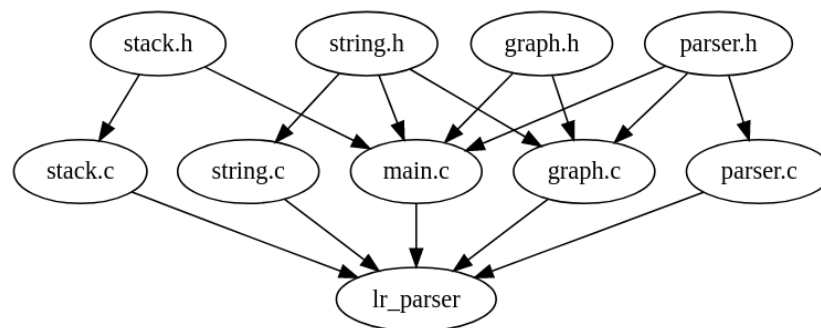


Figure 1: Précision des dépendances

3 Problèmes rencontrés et solutions trouvées

3.1 La lecture du fichier

Il était assez surprenant et déroutant de manipuler des fichiers qui ont des données illisibles à première vue. En effet, mon premier réflexe est de comparer ce que je dois obtenir avec ce que j'obtiens réellement après lecture d'une donnée sur un fichier, mais là il fallait travailler à tâton pour vérifier que la lecture des données était cohérente. Il m'a donc fallu un certain temps et **beaucoup** de tests avant d'être sûr du bon fonctionnement des programmes.

3.2 Les piles

La structure de pile en tant que liste chaînée a été abordée en cours et n'a donc pas créé de problème apparent en elle-même, mais la résolution de certains problèmes tels que la gestion du sommet de la pile en restant cohérent avec le principe de l'automate a demandé un certain temps.

3.3 Les graphes

La création d'une sortie au format DOT a certainement été une des tâches les plus complexes du programme et est celle qui a nécessité le plus de temps. Non seulement il fallait comprendre comment fonctionnait *Graphviz*, mais il fallait adapter le langage utilisé au fonctionnement d'un automate. Plusieurs questions se sont posées : Comment écrire dans la sortie au format DOT ? Comment traduire les données de l'automate pour arriver à concevoir de tels graphiques ? Il fallait être sûr de la compréhension et de la lecture de l'automate avant de se lancer dans la création du graphe. Beaucoup de temps a été passé à la création des noeuds et des arêtes, mais aussi sur les labels. En effet, il faut créer à chaque arête un label qui représente le ou les caractères qui sont nécessaires pour passer par l'arête. L'utilisation du mot clé *strict* a été d'une grande aide dans la simplification des arêtes 'identiques' (joignant deux mêmes noeuds dans le même sens) mais il fallait alors additionner les caractères correspondant aux labels. C'est pourquoi des fonctions sur les chaînes de caractères ont été nécessaires et grandement utiles. Malgré cela, le fichier *arith.aut* n'est pas correctement traduit en graphes.

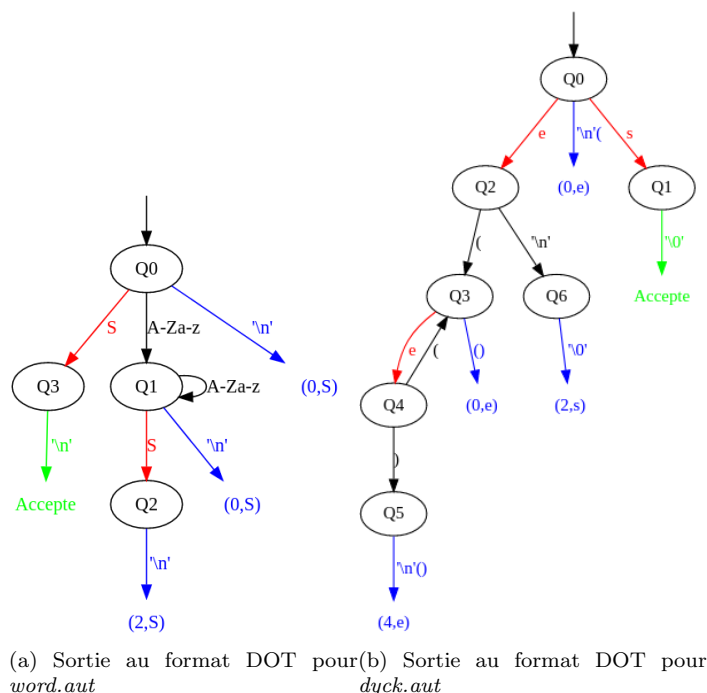


Figure 2: Exemple de sorties au format DOT

4 Limites du programme

A posteriori, mon programme possède une boucle infinie (correspondant à la volonté de pouvoir écrire des entrées sans interruption), il faudrait traiter ce problème afin de pouvoir clairement fermer le fichier avant la fin, sans avoir à appuyer sur CTRL-C.

Enfin, il faut comprendre pourquoi le fichier *arith.aut* pose problème lors de sa réalisation dans une sortie au format DOT.