

Le module d'analyse lexicale

Objectif

Le but de cette séance de TP est de mettre en place le module d'analyse lexicale du compilateur pour le langage Algo. Ce module n'est pas spécifique, il permettra l'analyse lexicale de n'importe quel langage dont les mots sont : les entiers, les réels, les identificateurs (les variables et les mots réservés) et les caractères de ponctuation courants.

La programmation sera guidée (la structure du programme est imposée) et réalisée de façon modulaire (une procédure par fonctionnalité). Chaque procédure (correspondant à une question) devra être testée avant que la suivante ne soit réalisée. L'évaluation se fera donc au fur et à mesure du TP.

ATTENTION : la structure du programme est imposée et une partie du code est fournie ; copier les fichiers (`algo_lexical.c`, `algo_lexical.h` et `test_algo_lexical`) dans votre répertoire de travail.

1. La procédure d'analyse d'un nombre

1.1. Réalisation

Ecrire la procédure `nombre` qui permet de lire un entier ou un réel dans un fichier texte (déjà ouvert : variable `lex_entree`). Cette fonction devra affecter à `lex_val` la valeur du nombre lu et à `lex_sym` son symbole (`s_entier` ou `s_reel`).

1.2. Test

Créer un fichier texte `nombre.txt` avec **un** nombre à l'intérieur (suivi de n'importe quels caractères : espace, lettre, ponctuation, ...) puis exécuter la procédure `exercice1`. Tester différents nombres.

2. Création de la table des mots-clés

2.1. Réalisation

Ecrire la procédure `ajoute_mot_cle(valeur, symbole)` qui permet d'ajouter **un** symbole et une valeur à la table des mots-clés. `nb_mots_cles` est le nombre de mots-clés et `table_mots_cles` est la table contenant tous les mots-clés (voir définition dans le programme fourni).

2.2. Test

Exécuter la procédure `exercice2`.

3. Procédure d'analyse des identificateurs et mots-clés

3.1. Réalisation

Ecrire la procédure `identificateur` qui va permettre de lire un identificateur et de vérifier s'il appartient à la table des mots-clés. Elle devra mettre à jour `lex_val` et `lex_sym`. On supposera comme toujours que le fichier `lex_entree` est ouvert et que le premier caractère est déjà dans `lex_ch`.

3.2. Test

Créer un fichier `identificateur.txt` avec **un** identificateur ou **un** mot-clé à l'intérieur (suivi de n'importe quel caractère : espace ou ponctuation) puis exécuter la procédure `exercice3`. Tester différents identificateurs (mots-clés ou non).

4. Procédure générale d'analyse lexicale

4.1. Réalisation

Ecrire la procédure `lire_symbole_quelconque` qui va reconnaître les caractères de ponctuation (voir constantes dans le programme fourni), les identificateurs (elle les appellera `identificateur`) et les nombres (elle les appellera `nombre`). **ATTENTION :** penser à traiter proprement la fin du fichier (`lex_sym=s_eof`) et les symboles non reconnus (`lex_sym=s_vider`).

4.2. Test

Créer un fichier `lexique.txt` contenant **plusieurs** mots du langage : caractères de ponctuation, nombres, identificateurs et mots-clés. Tester aussi avec des accents dans et au début d'un identificateur. Exécuter la procédure `exercice4` puis regarder ce que fait la procédure `lire_symbole`.