Zsisku Mihai

Grupa 235

07.11.2017

Limbaje formale si tehnici de compilare

Laborator 1

Cerinta:

Realizati specificatia unui mini-limbaj si implementati un analizor lexical pentru acest limbaj.

Minilimbajul trebuie sa contina urmatoarele:

* 2 tipuri de date simple si un tip de data definit de utilizator
* statements
* asignare
* input / output
* instructiune conditionala
* instructiune de ciclare
* restrictii asupra modului de formare a identificatorilor si a constantelor

Analizorul lexical are ca input un fisier text care contine cod sursa scris in minilimbajul definit, si va produce ca output:

* forma interna a programului
* tabela de simboluri

Programul ar trebui sa poata determina erori lexicale, specificand locatia, si preferabil tipul erorii

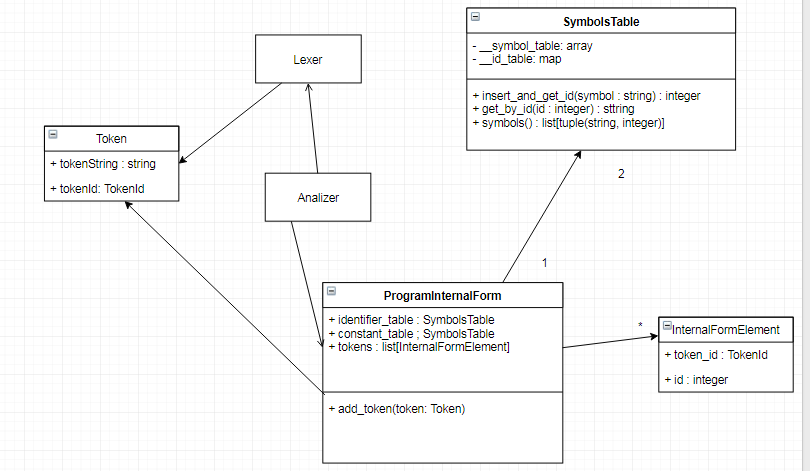
Restrictii:

* identificatorii au lungime maxima de 8 caractere
* tabela de simboluri separata pentru identificatori, respective constant
* tabela de simboluri organizata sub forma unei tabele sortate lexicographic

Implementare:

Specificarea minilimbajului se gaseste in fisierul minilimbajZsisku.txt.

Diagrama de clase:



Functionarea analizorului:

Avem componenta de lexer. Lexerul primeste fisierul de input si produce o lista de tokenuri obtinute prin parsarea fisierului. Lexerul foloseste regexuri pentru identificarea diferitelor tipuri de tokeni lexicali. Lexerul arunca o exceptie de tip LexerException daca intampina o eroare in timpul analizei.

ProgramInternalForm stocheaza tabelele de simboluri si tabela pentru forma interna a programului.

Pentru un fisier de intrare nume\_fisier.extensie analizorul produce 3 fisiere:

1. Tabela de identificatori

Numele fisierului va fi nume\_fisier.extensie.ident

Fisierul va contine un numar aleator de linii, fiecare linie contine un string : identificator, si un id numeric care reprezinta id-ul identificatorului respectiv in forma interna a simbolului

1. Tabela de constante

Numele fisierului va fi nume\_fisier.extensie.const

Analog cu fisierul pentru identificatori

1. Forma interna a programului

Numele fisierului va fi nume\_fisier.extensie.pif

Contine tokenii in ordinea aparitiei lor in fisier. Prima coloana e codificarea tokenului, a doua coloana contine id-ul tokenului, daca acesta e identificator sau constanta si -1 altfel