

Specificarea temei

Sa se specifice un limbaj de programare care permite folosirea unor tipuri de date primitive si compuse, permite declararea de variable, atribuire de valori, instructiuni de ciclare si citiri/scrieri cu intrarea, respectiv iesirea standard. Specificarea se va face in EBNF.

Cerinte functionale

1. se dau un numar de fisiere si se dreste obtinerea a 3 tabele care reprezinta Forma Interna a Programului (FIP), Tabela de Simboluri pentru Identificatori (TSI) si Tabela de Simboluri pentru Constante (TSC).

Detalii de implementare

Specificarea fiecarui atom lexical se face prin automate finite (cate un automat finit per atom lexical). Fiecare tranzitie verifica daca elementul curent (caracterul) este acceptat, dupa fiecare tranzitie pozitia din secventa initiala este incrementat cu 1 (se merge la urmatorul caracter).

Pentru identificarea unui atom lexical se incearca fiecare automat finit. In momentul in care toate automatele finite identifica o secventa de lungime 0 atunci s-a detercat un simbol necunoscut. In caz contrar atomul identificat este secventa cea mai lunga dintre toate secventele identificate de automatele finite. La crearea fiecarui automat finit i se asociaza un cod (codul atomului) prin intermediul unei enumerari.

Structurile de date folosite sunt SortedDictionary (implementat in .NET folosind arbore binar de cautare) si List (implementat in .NET folosind vector dinamic).

Detalii de utilizare

Executabilul **Andrei15193.Kepler** preia ca argumente in linia de comanda cel putin un fisier text. Pentru fiecare fisier text se genereaza 4 fisiere: <numeFisier>.fip.txt unde pe fiecare linie separate prin virgula se afla codul atomului si valoarea in caz de identificator/constanta, <numeFisier>.tsi.txt si <numeFisier>.tsc.txt contin multimea identificatorilor, respectiv multimea constantelor gasite. Fisierul <numeFisier>.BinaryLexicalAnalysis contine forma serializata a FIP, TSI si TSC. Pentru toate fisierele se creaza un fisier errors.txt unde se afla lista erorilor intalnite si numarul acestora.

