Міністерство освіти і науки України

Національний університет “Львівська політехніка”



**Курсова робота**

З дисципліни «Системне програмування»

на тему: "Розробка системних програмних модулів

та компонент систем програмування.

Розробка транслятора з вхідної мови програмування"

Варіант 15

**риВиконав:**

ст. гр. КІ-307

Курило А. О.

**Прийняв:**

Козак Н.Б.

Львів-2024

Анотація

Результатом даної курсової роботи є транслятор, який дозволяє транслювати вхідну мову, яка задається у відповідності з варіантом, у асемблерний код. Транслятор виконує наступні етапи трансляції: лексичний аналіз, синтаксичний аналіз та генерація коду.

Лексичний аналіз: вхідна послідовність символів розбивається на лексеми, які записуються у відповідну таблицю лексем. Кожній лексемі присвоюється число (числа легше порівнювати, ніж символи). Також у таблицю лексем записується додаткова інформація: рядок, у якому розміщується лексема; значення, якщо тип лексеми відповідає числу; та ін..

Синтаксичний аналіз: використовується висхідних метод аналізу без повернення. Призначений призначений для побудови дерева розбору, послідовно рухаючись від листків вгору до кореня дерева розбору.

Генерація коду: відбувається повторне зчитування таблиці лексем та формується відповідний асемблерний код для кожного блоку лексем. Отриманий код записується у результуючий файл та готовий до виконання.

Отриманий у результаті трансляції код, можна відкомпілювати за допомогою відповідних програм (LINK, ML чи ін.).

Зміст

[Анотація 2](#_Toc89096913)

[Завдання до курсової роботи 4](#_Toc89096914)

[Вступ 5](#_Toc89096915)

[1. Огляд методів та способів проектування трансляторів 6](#_Toc89096916)

[2. Формальний опис вхідної мови програмування 8](#_Toc89096917)

[2.1. Деталізований опис вхідної мови в термінах розширеної нотації Бекуса-Наура 8](#_Toc89096918)

[2.2. Опис термінальних символів та ключових слів 10](#_Toc89096919)

[3. Розробка транслятора вхідної мови програмування 12](#_Toc89096920)

[3.1. Вибір технології програмування 12](#_Toc89096921)

[3.2. Проектування таблиць транслятора 13](#_Toc89096922)

[3.3. Розробка лексичного аналізатора 14](#_Toc89096923)

[3.3.1. Розробка блок-схеми алгоритму 15](#_Toc89096924)

[3.3.2. Опис програми реалізації лексичного аналізатора 15](#_Toc89096925)

[3.4. Розробка синтаксичного та семантичного аналізатора 17](#_Toc89096926)

[3.4.1. Розробка дерев граматичного розбору 18](#_Toc89096927)

[3.4.2. Розробка граф-схеми алгоритму 19](#_Toc89096928)

[3.4.3. Опис програми реалізації синтаксичного та семантичного аналізатора 20](#_Toc89096929)

[3.5. Розробка генератора коду 20](#_Toc89096930)

[3.5.1. Розробка граф-схеми алгоритму 21](#_Toc89096931)

[3.5.2. Опис програми реалізації генератора коду 22](#_Toc89096932)

[4. Опис програми 23](#_Toc89096933)

[4.1. Загальна граф-схема програми 27](#_Toc89096934)

[4.2. Опис інтерфейсу та інструкція користувачеві 28](#_Toc89096935)

[5. Відлагодження та тестування програми 29](#_Toc89096936)

[5.1. Виявлення лексичних помилок 29](#_Toc89096937)

[5.2. Виявлення синтаксичних помилок 30](#_Toc89096938)

[5.3. Виявлення семантичних помилок 30](#_Toc89096939)

[5.4. Загальна перевірка коректності роботи транслятора 31](#_Toc89096940)

[6. Висновки 32](#_Toc89096941)

[7. Список використаної літератури 33](#_Toc89096942)

[8. Додатки 34](#_Toc89096943)

Завдання до курсової роботи

Завдання на курсову роботу

1. Цільова мова транслятора – асемблер для 32/64-розрядного процесора (х86).

2. Для отримання виконавчого файлу на виході розробленого транслятора скористатися програмами ml.exe і link.exe (або ml64.exe і link64.exe).

3. Мова розробки транслятора: ANSI C або C++.

4. Реалізувати оболонку або інтерфейс з командного рядка.

5. На вхід розробленого транслятора має подаватися текстовий файл, написаний на заданій мові програмування.

6. На виході розробленого транслятора мають створюватись такі файли:

файл з лексемами;

файл з повідомленнями про помилки (або про їх відсутність);

файл на мові асемблера;

об’єктний файл;

виконавчий файл.

7. Назва вхідної мови програмування утворюється від першої букви у прізвищі студента та останніх двох цифр номера його варіанту. Саме таке розширення повинні мати текстові файли, написані на цій мові програмування.

В моєму випадку це .v04

Опис вхідної мови програмування:

- типи даних: INTEGER32\_t

- оператор вводу: GET ()

- оператор виводу: PUT ()

- оператор: DO – WHILE (Ci)

- регістр ключових слів: Up перший символ Up

- регістр ідентифікаторів: Up8

- операції арифметичні: +, -, Mul, DIV, MOD

- операції порівняння: =, <>, >, <

- операції логічні: NOT, AND, OR

- коментар: @@...

- ідентифікатори змінних, числові константи

- оператор присвоєння: <-

Для отримання виконавчого файлу на виході розробленого транслятора скористатися програмами ml.exe (компілятор мови асемблера) і link.exe (редактор зв’язків).

Вступ

Транслятором називається програма перекладу (трансляції) початкової програми, записаною вхідною мовою, в еквівалентну їй об`єктну програму. Якщо мова високого рівня є вхідною, а мова асемблера чи машинна – вихідною, то такий транслятор називається компілятором.

Транслятори бувають двох типів: компілятори і інтерпретатори. Процес компіляції складається з двох частин: аналізу і синтезу. На етапі аналізу вхідну програму розбивають на окремі елементи (лексеми), перевіряють її на відповідність правилам граматики і створюють проміжне представлення вхідної програми. На етапі синтезу із проміжної форми створюють програму в машинних кодах. Така програма називається об’єктною програмою. В подальшому об’єктну програму можна виконати на комп’ютері без перетрансляції.

На відміну від компіляторів, інтерпретатор не створює нової програми, а просто виконує – інтерпретує кожен оператор вхідної мови програмування. Подібно компілятору, інтерпретатор аналізує вхідну програму, створює проміжне представлення, однак не створює об’єктної програми, а зразу виконує передбачені вхідною програмою команди.

Компілятор перекладає програму з однієї мови на іншу. На вхід компілятора поступає ланцюг символів, який є вхідною програмою на певній мові програмування. На виході компілятора (об’єктна програма) також представляє собою ланцюг символів, який вже належить іншій мові програмування, наприклад машинній мові деякого комп’ютера. При цьому сам компілятор може бути написаний третьою мовою.

1. Огляд методів та способів проектування трансляторів

Транслятор - обслуговуюча програма, що перетворює вихідну програму, надану на вхідній мові програмування, у робочу програму, представлену на об'єктному мовою.

Наведене визначення відноситься до всіх різновидів транслюють програм. Однак у кожної з таких програм можуть бути свої особливості щодо організації процесу трансляції. В даний час транслятори поділяються на три основні групи: асемблери, компілятори та інтерпретатори.

Асемблер - системна обслуговуюча програма, яка перетворить символічні конструкції в команди машинної мови. Специфічною рисою асемблером є те, що вони здійснюють дослівну трансляцію однієї символічної команди в одну машинну.

Компілятор - це обслуговуюча програма, що виконує трансляцію на машинну мову програми, записаної мовою оригіналу програмування. Також як і асемблер, компілятор забезпечує перетворення програми з однієї мови на іншу (найчастіше, в мову конкретного комп'ютера).

Інтерпретатор - програма або пристрій, що здійснює пооператорну трансляцію і виконання вихідної програми. На відміну від компілятора, інтерпретатор не породжує на виході програму на машинній мові. Розпізнавши команду вихідного мови, він тут же виконує її. Як у компіляторах, так і в інтерпретатора використовуються однакові методи аналізу вихідного тексту програми. Але інтерпретатор дозволяє почати обробку даних після написання навіть однієї команди. Це робить процес розробки і налагодження програм більш гнучким.

Процес трансляції складається з наступних фаз: лексичний аналіз, синтаксичний та семантичний аналіз, оптимізація коду, генерація коду.

На фазі лексичного аналізу вхідна програма, що представляє собою потік літер, розбивається на лексеми - слова у відповідності з визначеннями мови. Основними формалізму, які лежать в основі реалізації лексичних аналізаторів, є кінцеві автомати та регулярні вирази. Лексичний аналізатор може працювати в двох основних режимах: або як підпрограма, що викликається синтаксичним аналізатором для отримання чергової лексеми, або як повний прохід, результатом якого є файл лексем.

У процесі виділення лексем лексичний аналізатор може як самостійно будувати таблиці об'єктів (ідентифікаторів, рядків, чисел і т.д.), так і видавати значення для кожної лексеми при черговому до нього зверненні. У цьому випадку таблиці об'єктів будуються в наступних фазах (наприклад, в процесі синтаксичного аналізу).

На етапі лексичного аналізу виявляються деякі (найпростіші) помилки (неприпустимі символи, неправильний запис чисел, ідентифікаторів та ін.)

Основне завдання синтаксичного аналізу - розбір структури програми. Як правило, під структурою розуміється дерево, відповідне розбору в контекстно-вільної граматики мови. В даний час найчастіше використовується або LL (1)-аналіз (і його варіант - рекурсивний спуск), або LR (1) - аналіз та його варіанти (LR (0), SLR (1), LALR (1) та інші) . Рекурсивний спуск частіше використовується при ручному програмуванні синтаксичного аналізатора, LR (1) - при використанні систем автоматичного побудови синтаксичних аналізаторів.

Результатом синтаксичного аналізу є синтаксичне дерево з посиланнями на таблиці об'єктів. У процесі синтаксичного аналізу також виявляються помилки, пов'язані зі структурою програми.

На етапі контекстного аналізу виявляються залежності між частинами програми, які не можуть бути описані контекстно-вільною синтаксисом. Це в основному зв'язку «опис-використання», зокрема, аналіз типів об'єктів, аналіз областей видимості, відповідність параметрів, мітки та інші. У процесі контекстного аналізу таблиці об'єктів поповнюються інформацією про описах (властивостях) об'єктів.

Основним формалізмом, що використовується при контекстному аналізі, є апарат атрибутних граматик. Результатом контекстного аналізу є атрибутовані дерево програми. Інформація про об'єкти може бути як розосереджена в самому дереві, так і зосереджена в окремих таблицях об'єктів. У процесі контекстного аналізу також можуть бути виявлені помилки, пов'язані з неправильним використанням об'єктів.

Потім програма може бути переведена у внутрішнє представлення. Це робиться для цілей оптимізації та / або зручності генерації коду. Ще однією метою перетворення програми у внутрішнє представлення є бажання мати переносимий компілятор. Тоді тільки остання фаза (генерація коду) є машинно-залежною. В якості внутрішнього подання може використовуватися префіксний або постфіксній запис, орієнтований граф, трійки, четвірки та інші.

Фаз оптимізації може бути декілька. Оптимізації зазвичай ділять на машинно-залежні та машинно-незалежні, локальні і глобальні. Частина машинно-залежною оптимізації виконується на фазі генерації коду. Глобальна оптимізація намагається взяти до уваги структуру всієї програми, локальна - тільки невеликих її фрагментів. Глобальна оптимізація грунтується на глобальному потоковий аналіз, який виконується на графі програми і представляє по суті перетворення цього графа. При цьому можуть враховуватися такі властивості програми, як межпроцедурний аналіз, міжмодульних аналіз, аналіз галузей життя змінних і т.д.

Нарешті, генерація коду - остання фаза трансляції. Результатом її є або асемблерний модуль, або об'єктний (або завантажувальний) модуль. У процесі створення коду можуть виконуватися деякі локальні оптимізації, такі як розподіл регістрів, вибір довгих або коротких переходів, облік вартості команд при виборі конкретної послідовності команд. Для генерації коду розроблені різні методи, такі як таблиці рішень, зіставлення зразків, що включає динамічне програмування, різні синтаксичні методи.

Звичайно, ті чи інші фази транслятора можуть або відсутні зовсім, або об'єднуватися. У найпростішому випадку однопрохідного транслятора немає явної фази генерації проміжного представлення та оптимізації, інші фази об'єднані в одну, причому немає і явно побудованого синтаксичного дерева.

1. Формальний опис вхідної мови програмування
   1. Деталізований опис вхідної мови в термінах розширеної нотації Бекуса-Наура

Однією з перших задач, що виникають при побудові компілятора, є визначення вхідної мови програмування. Для цього використовують різні способи формального опису, серед яких я застосував розширену нотацію Бекуса-Наура (Backus/Naur Form - BNF).

<program> -> NAME <ident> BODY DATA<var\_blok>; <code\_blok>END

<var\_blok> -> INTEGER\_4 <ident> [{,<ident>}]

<ident> -> <low\_letter>|<number> [{<low\_letter>|<number>}]

<low\_letter> -> a|b|c|d|e|f|g|h|i|j|k|l|n|m|o|p|q|r|s|t|u|v|w|x|y|z

<up\_letter> -> A|B|C|D|E|F|G|H|I|J|K|L|N|M|O|P|Q|R|S|T|U|V|W|X|Y|Z

<number> -> [-|+]<digit>[{digit}]

<digit> -> 0|1|2|3|4|5|6|7|8|9

<code\_blok> -> <statement> [{<statement>}]

<statement> -> <equation>|<if\_statement>|< loop\_statement >|< read >|< print >

<if\_statement> -> IF (<expression\_l >) BODY <code\_blok> ELSE <code\_blok> END.

<loop\_statement> -> DO <code\_blok> WHILE (<expression\_l >)

<read> -> GET(<ident>); .

<print> -> PUT(<expression\_i>); .

<equation> -> <ident> -> <ident>|<expression\_a>|<expression\_l>|<not>

<expression\_a> -> [(]<term>[)] [<operation\_a> [(]<term>[)]]

<expression\_l> -> [(]<term>[)] [<operation\_l > [(]<term>[)]]

<term> -> <ident>|<number>

<operation\_a> -> <arithmetic>|<mult>

<operation\_l> -> <logic>|<compare>

<arithmetic> -> + | -

<mult> -> \* | DIV | MOD

<logic> -> AND | OR

<compare> -> EQ | NE | < | >

<not> -> NOT [(]{<term>|<expression\_l>}[)]

* 1. Опис термінальних символів та ключових слів

Визначимо окремі термінальні символи та нерозривні набори термінальних символів (ключові слова):

|  |  |
| --- | --- |
| Термінальний символ або ключове слово | Значення |
| NAME | Початок програми |
| BODY | Початок тексту програми |
| DATA | Початок блоку опису змінних |
| END | Кінець розділу операторів |
| GET | Оператор виводу (змінних або рядкових констант) |
| PUT | Оператор вводу змінних |
| -> | Оператор присвоєння |
| IF | Оператор умови |
| ELSE | Оператор умови |
| do | Початок тіла циклу |
| while | Оператор циклу |
| + | Оператор додавання |
| - | Оператор віднімання |
| \* | Оператор множення |
| DIV | Оператор ділення |
| MOD | Оператор знаходження залишку від ділення |
| EQ | Оператор перевірки на рівність |
| NE | Оператор перевірки на нерівність |
| < | Оператор перевірки чи менше |
| > | Оператор перевірки чи більше |
| NOT | Оператор логічного заперечення |
| AND | Оператор кон’юнкції |
| OR | Оператор диз’юнкції |
| INTEGER\_4 | 32‑ох розрядні знакові цілі |
| //…// | Коментар |
| , | Розділювач |
| ; | Ознака кінця оператора |
| ( | Відкриваюча дужка |
| ) | Закриваюча дужка |

До термінальних символів віднесемо також усі цифри (0-9), латинські букви (a-z, A-Z), символи табуляції, символ переходу на нову стрічку, пробілу.

1. Розробка транслятора вхідної мови програмування
   1. Вибір технології програмування

Для ефективної роботи створюваної програми важливу роль відіграє попереднє складення алгоритму роботи програми, алгоритму написання програми і вибір технології програмування.

Тому при складанні транслятора треба брати до уваги швидкість компіляції, якість об’єктної програми. Проект повинен давати можливість просто вносити зміни.

В реалізації мов високого рівня часто використовується специфічний тільки для компіляції засіб “розкрутки”. З кожним транслятором завжди зв`язані три мови програмування: Х – початкова, Y – об`єктна та Z – інструментальна. Транслятор перекладає програми мовою Х в програми, складені мовою Y, при цьому сам транслятор є програмою написаною мовою Z.

При розробці даного курсового проекту був використаний висхідний метод синтаксичного аналізу.

Також був обраний прямий метод лексичного аналізу. Характерною ознакою цього методу є те, що його реалізація відбувається без повернення назад. Його можна сприймати, як один спільний скінченний автомат. Такий автомат на кожному кроці читає один вхідний символ і переходить у наступний стан, що наближає його до розпізнавання поточної лексеми чи формування інформації про помилки. Для лексем, що мають однакові підланцюжки, автомат має спільні фрагменти, що реалізують єдину множину станів. Частини, що відрізняються, реалізуються своїми фрагментами

* 1. Проектування таблиць транслятора

Використання таблиць значно полегшує створення трансляторів, тому у даному випадку використовуються кілька таблиць:

1. Таблиця лексем з елементами, які мають таку структуру:

vector<Lexema> LexTable; // Таблиця лексем

typedef struct Lexem

{

string name; // Ім’я

int value; // Значення

int line; // Рядок

TypeOfLex type; // Тип

}Lexema;

1. Таблиця лексичних класів

vector<LexAndName> LexName =

{

{LAdd,"+"},

{LSub,"-"},

{LMul,"\*" },

{LDiv,"DIV"},

{LMod,"MOD"},

{LEqu,"EQ"},

{LNotEqu,"NE"},

{LGreate,">"},

{LLess,"<"},

{LNot,"NOT"},

{LAnd,"AND"},

{LOr,"OR"},

{LBegProg,""},

{LProgram,"NAME"},

{LProgramName,""},

{LStart,"BODY"},

{LData,"DATA"},

{LVarType,"INTEGER"},

{LIdentifier,""},

{LEnd,"END"},

{LScan,"GET"},

{LPrint,"PUT"},

{LIf,"IF"},

{LElse,"ELSE"},

{LDo,"DO"},

{LWhile,"WHILE"},

{LNewValue,"->"},

{LLBraket,"("},

{LRBraket,")"},

{LDLine,"\_"},

{LNumber,""},

{LSeparator,";"},

{LComma,","},

{LLComentar,"//"},

{LRComentar,"//"},

{LEOF,"EndFile"}

};

1. Таблиця ідентифікаторів з елементами типу Identificator та додатковою змінною цілочисельного типу в якій зберігається кількість визначених змінних. Структура елементів така:

vector<string> IdTable; //Таблиця ідентифікаторів

* 1. Розробка лексичного аналізатора

На фазі лексичного аналізу вхідна програма, що представляє собою потік літер, розбивається на лексеми - слова у відповідності з визначеннями мови. Лексичний аналізатор може працювати в двох основних режимах: або як підпрограма, що викликається синтаксичним аналізатором для отримання чергової лексеми, або як повний прохід, результатом якого є файл лексем.

Для нашої програми виберемо другий варіант. Тобто, спочатку буде виконуватись фаза лексичного аналізу. Результатом цієї фази буде файл з списком лексем. Але лексеми записуються у файл не як послідовність символів. Кожній лексемі присвоюється унікальний номер. Цей номер і записується у файл. Такий підхід дозволяє спростити роботу синтаксичного аналізатора.

Також на етапі лексичного аналізу виявляються деякі (найпростіші) помилки (неприпустимі символи, неправильний запис чисел, ідентифікаторів та ін.)

На вхід лексичного аналізатора надходить текст вихідної програми, а вихідна інформація передається для подальшої обробки компілятором на етапі синтаксичного аналізу.

Існує кілька причин, з яких до складу практично всіх компіляторів включають лексичний аналіз:

* застосування лексичного аналізатора спрощує роботу з текстом вихідної програми на етапі синтаксичного розбору;
* для виділення в тексті та розбору лексем можливо застосовувати просту, ефективну і теоретично добре пророблену техніку аналізу;
  + 1. Розробка блок-схеми алгоритму



Рис. Блок-схема роботи лексичного аналізатора

* + 1. Опис програми реалізації лексичного аналізатора

Основна задача лексичного аналізу – розбити вихідний текст, що складається з послідовності одиночних символів, на послідовність слів, або лексем, тобто виділити ці слова з безперервної послідовності символів. Всі символи вхідної послідовності з цієї точки зору розділяються на символи, що належать яким-небудь лексемам, і символи, що розділяють лексеми. В цьому випадку використовуються звичайні засоби обробки рядків. Вхiдна програма проглядається послідовно з початку до кінця. Базовi елементи, або лексичнi одиницi, роздiляються пробілами, знаками операцiй i спецiальними символами (новий рядок, знак табуляції), i таким чином видiляються та розпізнаються iдентифiкатори, лiтерали i термiнальнi символи (операцiї, ключові слова).

Програма аналізує файл поки не досягне його кінця. В циклі викликається функція GetNextLexem яка повертає наступну лексему.

Функція GetNextLexem зчитує з файлу слово або символ і порівнює з зарезервованими словами якщо є співпадіння то присвоює лексемі відповідний тип, значення якщо це числова константа.

При виділенні лексеми вона розпізнається та записується у таблицю лексем за допомогою відповідного типу лексеми, що є унікальним для кожної лексеми із усього можливого їх набору. Це дає можливість наступним фазам компiляції звертатись до лексеми не як до послідовності символів, а як до унікального типу лексеми, що значно спрощує роботу синтаксичного аналізатора: легко перевіряти належність лексеми до відповідної синтаксичної конструкції та є можливість легкого перегляду програми, як вгору, так і вниз, від текучої позиції аналізу. Також в таблиці лексем ведуться записи, щодо рядка відповідної лексеми – для місця помилки – та додаткова інформація.

При лексичному аналiзі виявляються i вiдзначаються лексичнi помилки (наприклад, недопустимi символи i неправильнi iдентифiкатори). Лексична фаза вiдкидає також i коментарi, оскiльки вони не мають нiякого впливу на виконання програми, отже й на синтаксичний розбір та генерацію коду.

В даній курсовій роботі реалізовано прямий лексичний аналізатор, який виділяє з вхідного тексту програми окремі лексеми і на основі цього формує таблицю.

* 1. Розробка синтаксичного та семантичного аналізатора

Синтаксичний аналізатор - частина компілятора, яка відповідає за виявлення основних синтаксичних конструкцій вхідної мови. У завдання синтаксичного аналізатора входить: знайти і виділити основні синтаксичні конструкції в тексті вхідної програми, встановити тип і перевірити правильність кожної синтаксичної конструкції у вигляді, зручному для подальшої генерації тексту результуючої програми.

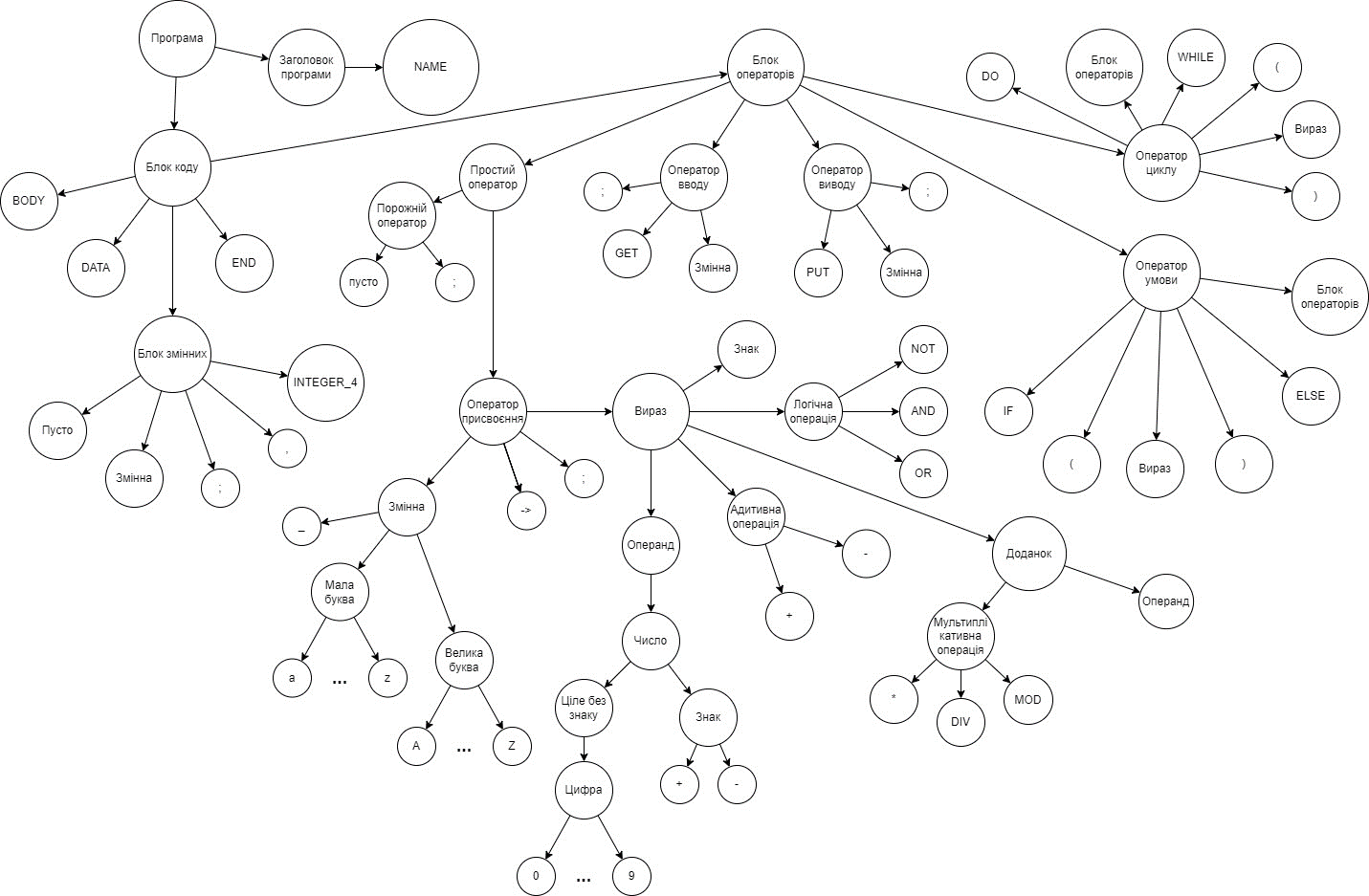
В основі синтаксичного аналізатора лежить Розпізнавач тексту вхідної програми на основі граматики вхідного мови. Як правило, синтаксичні конструкції мов програмування можуть бути описані за допомогою КС-граматик, рідше зустрічаються мови, які можуть бути описані за допомогою регулярних граматик. Найчастіше регулярні граматики застосовні до мов асемблера, а мови високого рівня побудовані на основі КС-мов.

Синтаксичний розбір - це основна частина компіляції на етапі аналізу. Без виконання синтаксичного розбору робота компілятора безглузда, у той час як лексичний аналізатор є зовсім необов'язковим. Усі завдання з перевірки лексики вхідного мови можуть бути вирішені на етапі синтаксичного розбору. Сканер тільки дозволяє позбавити складний за структурою лексичний аналізатор від рішення примітивних завдань з виявлення та запам'ятовування лексем вхідний програми.

В даній курсовій роботі синтаксичний аналіз можна виконувати лише після виконання лексичного аналізу, він являється окремим етапом трансляції.

На вході даного аналізатора є файл лексем, який є результатом виконання лексичного аналізу, на базі цього файлу синтаксичний аналізатор формує таблицю ідентифікаторів та змінних.

* + 1. Розробка дерев граматичного розбору



* + 1. Розробка граф-схеми алгоритму



Рис. Граф-схема роботи синтаксичного аналізатора

* + 1. Опис програми реалізації синтаксичного та семантичного аналізатора

На вхід синтаксичного аналізатора подіється таблиця лексем створена на етапі лексичного аналізу. Аналізатор проходить по ній і перевіряє чи набір лексем відповідає раніше описаним формам нотації Бекуса-Наура. І разі не відповідності у файл з помилками виводиться інформація про помилку і про рядок на якій вона знаходиться.

При знаходженні оператора присвоєння або математичних виразів здійснюється перевірка балансу дужок(кількість відкриваючих дужок має дорівнювати кількості закриваючих). Також здійснюється перевірка чи не йдуть підряд декілька лексем одного типу

Результатом синтаксичного аналізу є синтаксичне дерево з посиланнями на таблиці об'єктів. У процесі синтаксичного аналізу також виявляються помилки, пов'язані зі структурою програми.

В основі синтаксичного аналізатора лежить розпізнавач тексту вхідної програми на основі граматики вхідної мови.

* 1. Розробка генератора коду

Синтаксичне дерево в чистому вигляді несе тільки інформацію про структуру програми. Насправді в процесі генерації коду потрібна також інформація про змінні (наприклад, їх адреси), процедури (також адреси, рівні), мітки і т.д. Для представлення цієї інформації можливі різні рішення. Найбільш поширені два:

* інформація зберігається у таблицях генератора коду;
* інформація зберігається у відповідних вершинах дерева.

Розглянемо, наприклад, структуру таблиць, які можуть бути використані в поєднанні з Лідер-представленням. Оскільки Лідер-представлення не містить інформації про адреси змінних, значить, цю інформацію потрібно формувати в процесі обробки оголошень і зберігати в таблицях. Це стосується і описів масивів, записів і т.д. Крім того, в таблицях також повинна міститися інформація про процедури (адреси, рівні, модулі, в яких процедури описані, і т.д.). При вході в процедуру в таблиці рівнів процедур заводиться новий вхід - вказівник на таблицю описів. При виході вказівник поновлюється на старе значення. Якщо проміжне представлення - дерево, то інформація може зберігатися в вершинах самого дерева.

Генерація коду – це машинно-залежний етап компіляції, під час якого відбувається побудова машинного еквівалента вхідної програми. Зазвичай входом для генератора коду служить проміжна форма представлення програми, а на виході може з’являтися об’єктний код або модуль завантаження.

Генератор асемблерного коду приймає масив лексем без помилок. Якщо на двох попередніх етапах виявлено помилки, то ця фаза не виконується.

В даному курсовому проекті генерація коду реалізується як окремий етап. Можливість його виконання є лише за умови, що попередньо успішно виконався етап синтаксичного аналізу. І використовує результат виконання попереднього аналізу, тобто два файли: перший містить згенерований асемблерний код відповідно операторам які були в програмі, другий файл містить таблицю змінних. Інформація з них зчитується в відповідному порядку, основні константні конструкції записуються в файл asm.

* + 1. Розробка граф-схеми алгоритму



Рис. Блок схема генератора коду

* + 1. Опис програми реалізації генератора коду

У компілятора, реалізованого в даній курсовій роботі, вихідна мова - програма на мові Assembler. Ця програма записується у файл, що має таку ж саму назву, як і файл з вхідним текстом, але розширення “asm”. Генерація коду відбувається одразу ж після синтаксичного аналізу.

В даному трансляторі генератор коду послідовно викликає окремі функції, які записують у вихідний файл частини коду.

Першим кроком генерації коду записується ініціалізація сегменту даних. Далі виконується аналіз коду, та визначаються процедури, зміні, які використовуються.

Проаналізувавши змінні, які є у програмі, генератор формує код даних для асемблерної програми. Для цього з таблиці лексем вибирається ім’я змінної (типи змінних відповідають 4 байтам), та записується 0, в якості початкового значення.

Аналіз наявних процедур необхідний у зв’язку з тим, що процедури введення/виведення, виконання арифметичних та логічних операцій, виконано у вигляді окремих процедур і у випадку їх відсутності немає сенсу записувати у вихідний файл зайву інформацію.

Після цього зчитується лексема з таблиці лексем. Також відбувається перевірка, чи це не остання лексема. Якщо це остання лексема, то функція завершується.

Наступним кроком є аналіз таблиці лексем, та безпосередня генерація коду у відповідності до вхідної програми.

Генератор коду зчитує лексему та генерує відповідний код, який записується у файл. Наприклад, якщо це лексема виведення, то у основну програму записується виклик процедури виведення, попередньо записавши у співпроцесор значення, яке необхідно вивести. Якщо це арифметична операція, так само викликається дана процедура, але як і в попередньому випадку, спочатку у регістри співпроцесора записується інформація, яка вказує над якими значеннями виконувати дії.

Генератор закінчує свою роботу, коли зчитує лексему, що відповідає кінцю файлу.

В кінці своє роботи, генератор формує код завершення ассемблерної програми.

1. Опис програми

Дана програма написана мовою С++ з використанням визначень нових типів, перечислень та класів:

enum TypeOfLex

{

LAdd,

LSub,

LMul,

LDiv,

LMod,

LEqu,

LNotEqu,

LGreate,

LLess,

LNot,

LAnd,

LOr,

LProgram,

LProgramName,

LStart,

LData,

LVarType,

LIdentifier,

LEnd,

LScan,

LPrint,

LIf,

LElse,

LFor,

LStartBlock,

LEndBlock,

LNewValue,

LDLine,

LLBraket,

LRBraket,

LTrLBraket,

LTrRBraket,

LNumber,

LSeparator,

LComma,

LEOF,

LLComentar,

LRComentar,

LUnknown

};

struct LexAndName

{

TypeOfLex type;

std::string name;

};

//DATA

typedef struct Lexem

{

string name;

int value;

int line;

TypeOfLex type;

}Lexema;

typedef struct L

{

vector<Lexema> LexTable; //Таблиця лексем

vector<string> IdTable; //Таблиця ідентифікаторів

string ftype = "v04";

string TName = "\_xxxxxxxx";

vector<LexAndName> LexName =

{

{LAdd,"+"},

{LSub,"-"},

{LMul,"\*" },

{LDiv,"DIV"},

{LMod,"MOD"},

{LEqu,"EQ"},

{LNotEqu,"NE"},

{LGreate,">"},

{LLess,"<"},

{LNot,"NOT"},

{LAnd,"AND"},

{LOr,"OR"},

{LBegProg,""},

{LProgram,"NAME"},

{LProgramName,""},

{LStart,"BODY"},

{LData,"DATA"},

{LVarType,"INTEGER"},

{LIdentifier,""},

{LEnd,"END"},

{LScan,"GET"},

{LPrint,"PUT"},

{LIf,"IF"},

{LElse,"ELSE"},

{LDo,"DO"},

{LWhile,"WHILE"},

{LNewValue,"->"},

{LLBraket,"("},

{LRBraket,")"},

{LDLine,"\_"},

{LNumber,""},

{LSeparator,";"},

{LComma,","},

{LLComentar,"//"},

{LRComentar,"//"},

{LEOF,"EndFile"}

};

bool IsPresentInput, //Ознаки присутності операторів

IsPresentOutput,

IsPresentMod,

IsPresentAnd,

IsPresentOr,

IsPresentNot,

IsPresentEqu,

IsPresentGreate,

IsPresentLess,

IsPresentDiv;

int numberErrors;

std::string InputFileName;

char OutputFileName[50];

}DataType;

bool directory\_exists(const std::string& directory);

В програмі також використовуються багато функцій тому вона є доволі структурованою (граф-схема алгоритму виконання програми розташована в додатку Б).

int main(int argc, char\* argv[]) – Основна підпрограма. З неї починає своє виконання програма. В ній прямо чи опосередковано викликаються наступні функції:

int AnalisisLexems(FILE\* FIn) - Аналізує вхідний файл і розбиває його на лексеми

Lexema\* GetNextLexem(FILE\* f, int ii) - Відділяє наступну лексему

void PrintLexemsInFile() - Друкує лексеми у файл

int ErrorChecking() - функція пошуку помилок у вхідному файлі. Повертає їх кількість

bool IsOperation(TypeOfLex t) – Повертає true якщо лексема є операцією

int IsExpression(int i, FILE\* ef) – Перевірка, чи є дана послідовність лексем виразом

int Balans(int nom, TypeOfLex ends, TypeOfLex ltStart, TypeOfLex ltFinish) – Визначає баланс дужок

void GenerateCode(FILE\* f) - за таблицями ідентифікаторів та лексем генерує вихідний код на мові асемблера

void BeginASMFile(FILE\* fout) - записує до вихідного файлу стандартну „шапку”

void CheckPresent() - Перевіряє наявність операторів

void PrintData(FILE\* f) - Друкує сегмент даних

void BeginCodeSegm(FILE\* fout) - Друкує початок сегменту коду

void PrintCode(FILE\* f) – Видруковує в вихідний файл основний код програми

int ConvertToPostfixForm(int i) - Перетворює вирази в постфіксну форму

bool Prioritet(TypeOfLex t, StackType s) - Перевіряє пріорітет операцій (використовується при перетворенні виразу у посифіксну форму).

void GenASMCode(const char\* str, FILE\* f) - Генерує код для окремих операцій

void PrintEnding(FILE\* f) - Видруковує в вихідний файл стандартне завершення \*.asm-файлу

void PrintAND(FILE\* f) - Видруковує в вихідний файл процедуру And\_

void PrintOR(FILE\* f) - Видруковує в вихідний файл процедуру Or\_

void PrintNOT(FILE\* f) - Видруковує в вихідний файл процедуру Not\_

void PrintEQ(FILE\* f) - Видруковує в вихідний файл процедуру Eq\_

void PrintGE(FILE\* f) - Видруковує в вихідний файл процедуру Greate\_

void PrintLE(FILE\* f) - Видруковує в вихідний файл процедуру Less\_

void PrintInput(FILE\* f) - Видруковує в вихідний файл процедуру вводу

void PrintOutput(FILE\* f) - Видруковує в вихідний файл процедуру виводу

* 1. Загальна граф-схема програми



Рис. Загальна схема програми

* 1. Опис інтерфейсу та інструкція користувачеві

Вхідним файлом для даної програми є звичайний текстовий файл з розширенням v04. У цьому файлі необхідно набрати бажану для трансляції програму та зберегти її. Синтаксис повинен відповідати вхідній мові.

Створений транслятор є консольною програмою, що запускається з командної стрічки з параметром: "CWork\_v04.exe <ім’я програми>.v04"

Якщо обидва файли мають місце на диску та правильно сформовані, програма буде запущена на виконання.

Початковою фазою обробки є лексичний аналіз (розбиття на окремі лексеми). Результатом цього етапу є файл lexems.txt, який містить таблицю лексем. Вміст цього файлу складається з 4 полів – 1 – безпосередньо сама лексема; 2 – тип лексеми; 3 – значення лексеми (необхідне для чисел і ідентифікаторів); 4 – рядок, у якому лексема знаходиться. Наступним етапом є перевірка на правильність написання програми (вхідної). Інформацію про наявність чи відсутність помилок можна переглянути у файлі error.txt. Якщо граматичний розбір виконаний успішно, файл буде містити відповідне повідомлення. Інакше, у файлі будуть зазначені помилки з іх описом та вказанням їх місця у тексті програми.

Останнім етапом є генерація коду. Транслятор переходить до цього етапу, лише у випадку, коли відсутні граматичні помилки у вхідній програмі. Згенерований код записується у файлу <ім’я програми>.asm.

Для отримання виконавчого файлу необхідно скористатись програмою Masm32.exe

1. Відлагодження та тестування програми

Тестування програмного забезпечення є важливим етапом розробки продукту. На цьому етапі знаходяться помилки допущені на попередніх етапах. Цей етап дозволяє покращити певні характеристики продукту, наприклад – інтерфейс. Дає можливість знайти та вподальшому виправити слабкі сторони, якщо вони є.

Відлагодження даної програми здійснюється за допомогою набору кілької програм, які відповідають заданій граматиці. Та перевірці коректності коду, що генерується, коректності знаходження помилок та розбивки на лексеми.

* 1. Виявлення лексичних та синтаксичних помилок

Виявлення лексичних помилок відбувається на стадії лексичного аналізу. Під час розбиття вхідної програми на окремі лексеми відбувається перевірка чи відповідає вхідна лексема граматиці. Якщо ця лексема є в граматиці то вона ідентифікується і в таблиці лексем визначається. У випадку неспівпадіння лексемі присвоюється тип "невпізнаної лексеми". Повідомлення про такі помилки можна побачити лише після виконання процедури перевірки таблиці лексем, яка знаходиться в файлі.

Виявлення синтаксичних помилок відбувається на стадії перевірки програми на коректність окремо від синтаксичного аналізу. При цьому перевіряється окремо кожне твердження яке може бути або виразом, або оператором (циклу, вводу/виводу), або оголошенням, та перевіряється структура програми в цілому.

Приклад виявлення:

***Текст програми з помилками***

//Prog1//

NAME \_prog1;

BODY

DATA INTEGER\_4 \_a,\_b,\_x,\_y;

GET(\_a);

GET(\_b);

PU T(\_a + \_b);

PUT(\_a - \_b);

PUT(\_a \* \_b)

PUT(\_a DIV \_b);

PUT(\_a MOD \_b);

\_x->(\_a \_b) \* 10 + (\_a + \_b) DIV 10;

\_y->\_x + (\_x MOD 10);

PUT(\_x);

PUT(\_y);

END

***Текст файлу з повідомленнями про помилки***

List of bugs in the program

=======================================================================

line 7: Unknown identifier!(PU)

line 7: Unknown identifier!(T)

line 9: After ')' must be ')', operation or ';'!

line 12: After identifier must be ')' or ';' or operation!

* 1. Виявлення семантичних помилок

Суттю виявлення семантичних помилок є перевірка числових констант на відповідність типу INTEGER\_4, тобто знаковому цілому числу з відповідним діапазоном значень і перевірку на коректність використання змінних INTEGER\_4 у цілочисельних і логічних виразах.

* 1. Загальна перевірка коректності роботи транслятора

Для того щоб здійснити перевірку коректності роботи транслятора необхідно завантажити коректну до заданої вхідної мови програму.

***Текст коректної програми***

//Prog1//

NAME \_prog1;

BODY

DATA INTEGER\_4 \_a,\_b,\_x,\_y;

GET(\_a);

GET(\_b);

PUT(\_a + \_b);

PUT(\_a - \_b);

PUT(\_a \* \_b);

PUT(\_a DIV \_b);

PUT(\_a MOD \_b);

\_x->(\_a - \_b) \* 10 + (\_a + \_b) DIV 10;

\_y->\_x + (\_x MOD 10);

PUT(\_x);

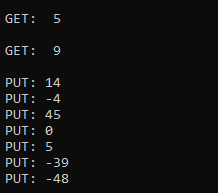
PUT(\_y);

END

Оскільки дана програма відповідає граматиці то результати виконання лексичного, синтаксичного аналізів, а також генератора коду будуть позитивними.

В результаті буде отримано асемблерний файл, який є результатом виконання трансляції з заданої вхідної мови на мову Assembler даної програми (його вміст наведений в Додатку А).

Після виконання компіляції даного файлу на виході отримаєм наступний результат роботи програми:



При перевірці отриманого результату, можна зробити висновок про правильність роботи програми, а отже і про правильність роботи транслятора.

* 1. Тестова програма №1

***Текст програми***

//Prog1//

NAME \_prog1;

BODY

DATA INTEGER\_4 \_a,\_b,\_x,\_y;

GET(\_a);

GET(\_b);

PUT(\_a + \_b);

PUT(\_a - \_b);

PUT(\_a \* \_b);

PUT(\_a DIV \_b);

PUT(\_a MOD \_b);

\_x->(\_a - \_b) \* 10 + (\_a + \_b) DIV 10;

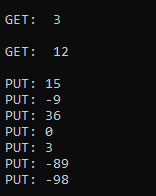
\_y->\_x + (\_x MOD 10);

PUT(\_x);

PUT(\_y);

END

***Результат виконання***



* 1. Тестова програма №2

***Текст програми***

//Prog2//

NAME \_prog2;

BODY

DATA INTEGER\_4 \_a,\_b,\_c;

GET(\_a);

GET(\_b);

GET(\_c);

IF(\_a > \_b)

BODY

IF(\_a < \_c)

BODY

PUT(\_c);

ELSE

PUT(\_a);

END

ELSE

IF(\_b < \_c)

BODY

PUT(\_c);

ELSE

PUT(\_b);

END

END

IF((\_a EQ \_b) AND (\_a EQ \_c) AND (\_b EQ \_c))

BODY

PUT(1);

ELSE

PUT(0);

END

IF((\_a < 0) OR (\_b < 0) OR (\_c < 0))

BODY

PUT(- 1);

ELSE

PUT(0);

END

IF(NOT(\_a < (\_b + \_c)))

BODY

PUT(10);

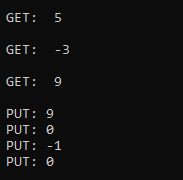
ELSE

PUT(0);

END

END

***Результат виконання***



* 1. Тестова програма №3

***Текст програми***

//Prog3//

NAME \_prog3;

BODY

DATA INTEGER\_4 \_a,\_a2,\_b,\_x,\_c1,\_c2;

GET(\_a);

GET(\_b);

\_a2->\_a;

DO

PUT(\_a2 \* \_a2);

\_a2->\_a2 + 1;

WHILE(NOT(\_a2 > \_b));

\_x->0;

\_c1->1;

DO

\_c2->1;

DO

\_x->\_x + 1;

\_c2->\_c2 + 1;

WHILE(NOT(\_c2 > \_b));

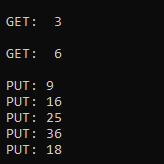
\_c1->\_c1 + 1;

WHILE(NOT(\_c1 > \_a));

PUT(\_x);

END

***Результат виконання***



1. Висновки

В процесі виконання курсової роботи було виконано наступне:

1. Складено формальний опис мови програмування v04, в термінах розширеної нотації Бекуса-Наура, виділено усі термінальні символи та ключові слова.

2. Створено компілятор мови програмування v04, а саме:

2.1. Розроблено прямий лексичний аналізатор, орієнтований на розпізнавання лексем, що є заявлені в формальному описі мови програмування.

2.2. Розроблено синтаксичний аналізатор на основі висхідного методу. Складено деталізований опис вхідної мови в термінах розширеної нотації Бекуса-Наура

2.3. Розроблено генератор коду, відповідні процедури якого викликаються після перевірки синтаксичним аналізатором коректності запису чергового оператора, мови програмування v04. Вихідним кодом генератора є програма на мові Assembler(x86).

3. Проведене тестування компілятора на тестових програмах за наступними пунктами:

3.1. На виявлення лексичних помилок.

3.2. На виявлення синтаксичних помилок.

3.3. Загальна перевірка роботи компілятора.

Тестування не виявило помилок в роботі компілятор, і всі помилки в тестових програмах на мові v04 були успішно виявлені і відповідно оброблені.

В результаті виконання даної курсової роботи було засвоєно методи розробки та реалізації компонент систем програмування.

1. Список використаної літератури
2. Шильдт Г. С++. – Санкт-Петербург: BXV, 2002. – 688 с.
3. Хантер Р. Проектирование и конструирование компиляторов. – М.: Финансы и статистика, 1984. – 232 с.
4. Шамис В. А. С++Builder 4 Техника визуального программирования. – М.: Нолидж, 2000. – 656 с.
5. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции. – М.: Мир, 1978. – т.1, 612 с.
6. Серебряков В.А. Лекции по конструированию компиляторов. – М.: МГУ, 1993.
7. В.С. Проценко, П.Й. Чаленко, А.Б.Ставровський “Техніка програмування мовою Сі”.-Київ “Либідь”, 1993. – 224с.
8. Б.Керниган, Д.Ритчи “Язык программирования Си”. – Москва “Финансы и статистика”, 1992. – 271с.
9. Н.Г.Голубь “Исскуство программирования на асемблере.Лекции и упражнения”. –Киев “DiaSoft”, 2002 – 642с.
10. Л.Бэк “Введение в системное програмирование”. – М.: Мир, 1999
11. Додатки

**Додаток (Код на мові Асемблер)**

**Prog1.asm**

.386

.model flat, stdcall

option casemap :none

include masm32\include\windows.inc

include masm32\include\kernel32.inc

include masm32\include\masm32.inc

include masm32\include\user32.inc

include masm32\include\msvcrt.inc

includelib masm32\lib\kernel32.lib

includelib masm32\lib\masm32.lib

includelib masm32\lib\user32.lib

includelib masm32\lib\msvcrt.lib

.DATA

;===User Data==============

\_a dd 00h

\_b dd 00h

\_x dd 00h

\_y dd 00h

;===Addition Data===========================================

hConsoleInput dd ?

hConsoleOutput dd ?

ErrorMessage db "Error: division by zero", 0

InputBuf db 15 dup (?)

InMessage db "GET: ", 0

OutMessage db "PUT: %d", 0

ResMessage db 20 dup (?)

NumberOfCharsRead dd ?

NumberOfCharsWrite dd ?

msg1310 dw 13,10

buf dd 0

.CODE

start:

invoke AllocConsole

invoke GetStdHandle, STD\_INPUT\_HANDLE

mov hConsoleInput, EAX

invoke GetStdHandle, STD\_OUTPUT\_HANDLE

mov hConsoleOutput, EAX

call Input

mov \_a, eax

call Input

mov \_b, eax

push \_a

push \_b

call Add\_

push eax

call Output

push \_a

push \_b

call Sub\_

push eax

call Output

push \_a

push \_b

call Mul\_

push eax

call Output

push \_a

push \_b

call Div\_

push eax

call Output

push \_a

push \_b

call Mod\_

push eax

call Output

push \_a

push \_b

call Sub\_

push eax

push dword ptr 10

call Mul\_

push eax

push \_a

push \_b

call Add\_

push eax

push dword ptr 10

call Div\_

push eax

call Add\_

push eax

mov \_x, eax

push \_x

push \_x

push dword ptr 10

call Mod\_

push eax

call Add\_

push eax

mov \_y, eax

push \_x

call Output

push \_y

call Output

exit\_label:

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke ReadConsoleA,hConsoleInput,ADDR InputBuf,1,ADDR NumberOfCharsRead,0

invoke ExitProcess, 0

;===Procedure Input==========================================================================

Input PROC

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA,hConsoleOutput,ADDR InMessage,SIZEOF InMessage,ADDR NumberOfCharsWrite,0

invoke ReadConsoleA,hConsoleInput,ADDR InputBuf,11,ADDR NumberOfCharsRead,0

invoke crt\_atoi, addr InputBuf

ret

Input ENDP

;============================================================================================

;===Procedure Output=========================================================================

Output PROC, value:dword

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke wsprintf, addr ResMessage, addr OutMessage, value

invoke WriteConsoleA,hConsoleOutput,ADDR ResMessage,eax,ADDR NumberOfCharsWrite,0

ret

Output ENDP

;============================================================================================

;===Procedure Add=========================

Add\_ PROC

mov eax, [esp+8]

add eax, [esp+4]

ret 8

Add\_ ENDP

;=========================================

;===Procedure Sub=========================

Sub\_ PROC

mov eax, [esp+8]

sub eax, [esp+4]

ret 8

Sub\_ ENDP

;=========================================

;===Procedure Mul=========================

Mul\_ PROC

mov eax, [esp+8]

mov ebx, [esp+4]

imul ebx

ret 8

Mul\_ ENDP

;=========================================

;===Procedure Div=========================

Div\_ PROC

pushf

pop cx

mov eax, [esp+4]

cmp eax, 0

jne end\_check

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA, hConsoleOutput, ADDR ErrorMessage, SIZEOF ErrorMessage, ADDR NumberOfCharsWrite, 0

jmp exit\_label

end\_check:

mov eax, [esp+8]

cmp eax, 0

jge gr

lo:

mov edx, -1

jmp less\_fin

gr:

mov edx, 0

less\_fin:

mov eax, [esp+8]

mov ebx, [esp+4]

idiv ebx

push cx

popf

ret 8

Div\_ ENDP

;=========================================

;===Procedure Mod=========================

Mod\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jne end\_check

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA, hConsoleOutput, ADDR ErrorMessage, SIZEOF ErrorMessage, ADDR NumberOfCharsWrite, 0

jmp exit\_label

end\_check:

mov eax, [esp+8]

cmp eax, 0

jge gr

lo:

mov edx, -1

jmp less\_fin

gr:

mov edx, 0

less\_fin:

mov eax, [esp+8]

mov ebx, [esp+4]

idiv ebx

mov eax, edx

push cx

popf

ret 8

Mod\_ ENDP

;=========================================

;===Procedure And=========================

And\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jnz and\_t1

jz and\_false

and\_t1:

mov eax, [esp+4]

cmp eax, 0

jnz and\_true

and\_false:

mov eax, 0

jmp and\_fin

and\_true:

mov eax, 1

and\_fin:

push cx

popf

ret 8

And\_ ENDP

;=========================================

;===Procedure Or==========================

Or\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jnz or\_true

mov eax, [esp+4]

cmp eax, 0

jnz or\_true

mov eax, 0

jmp or\_fin

or\_true:

mov eax, 1

or\_fin:

push cx

popf

ret 8

Or\_ ENDP

;=========================================

;===Procedure Not=========================

Not\_ PROC

pushf

pop cx

mov eax, [esp+4]

cmp eax, 0

jne not\_false

mov eax, 1

jmp not\_fin

not\_false:

mov eax, 0

not\_fin:

push cx

popf

ret 4

Not\_ ENDP

;=========================================

;===Procedure Eq==========================

Eq\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jne not\_eq

mov eax, 1

jmp eq\_fin

not\_eq:

mov eax, 0

eq\_fin:

push cx

popf

ret 8

Eq\_ ENDP

;=========================================

;===Procedure Greate======================

Greate\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jle lov

mov eax, 1

jmp gr\_fin

lov:

mov eax, 0

gr\_fin:

push cx

popf

ret 8

Greate\_ ENDP

;=========================================

;===Procedure Less========================

Less\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jge gr

lo:

mov eax, 1

jmp less\_fin

gr:

mov eax, 0

less\_fin:

push cx

popf

ret 8

Less\_ ENDP

;=========================================

end start

**Prog2.asm**

.386

.model flat, stdcall

option casemap :none

include masm32\include\windows.inc

include masm32\include\kernel32.inc

include masm32\include\masm32.inc

include masm32\include\user32.inc

include masm32\include\msvcrt.inc

includelib masm32\lib\kernel32.lib

includelib masm32\lib\masm32.lib

includelib masm32\lib\user32.lib

includelib masm32\lib\msvcrt.lib

.DATA

;===User Data==============

\_a dd 00h

\_b dd 00h

\_c dd 00h

;===Addition Data===========================================

hConsoleInput dd ?

hConsoleOutput dd ?

ErrorMessage db "Error: division by zero", 0

InputBuf db 15 dup (?)

InMessage db "GET: ", 0

OutMessage db "PUT: %d", 0

ResMessage db 20 dup (?)

NumberOfCharsRead dd ?

NumberOfCharsWrite dd ?

msg1310 dw 13,10

buf dd 0

.CODE

start:

invoke AllocConsole

invoke GetStdHandle, STD\_INPUT\_HANDLE

mov hConsoleInput, EAX

invoke GetStdHandle, STD\_OUTPUT\_HANDLE

mov hConsoleOutput, EAX

call Input

mov \_a, eax

call Input

mov \_b, eax

call Input

mov \_c, eax

push \_a

push \_b

call Greate\_

push eax

cmp eax, 0

je ifLabel1

push \_a

push \_c

call Less\_

push eax

cmp eax, 0

je ifLabel2

push \_c

call Output

jmp endIf2

ifLabel2:

push \_a

call Output

endIf2:

jmp endIf1

ifLabel1:

push \_b

push \_c

call Less\_

push eax

cmp eax, 0

je ifLabel3

push \_c

call Output

jmp endIf3

ifLabel3:

push \_b

call Output

endIf3:

endIf1:

push \_a

push \_b

call Eq\_

push eax

push \_a

push \_c

call Eq\_

push eax

call And\_

push eax

push \_b

push \_c

call Eq\_

push eax

call And\_

push eax

cmp eax, 0

je ifLabel4

push dword ptr 1

call Output

jmp endIf4

ifLabel4:

push dword ptr 0

call Output

endIf4:

push \_a

push dword ptr 0

call Less\_

push eax

push \_b

push dword ptr 0

call Less\_

push eax

call Or\_

push eax

push \_c

push dword ptr 0

call Less\_

push eax

call Or\_

push eax

cmp eax, 0

je ifLabel5

push dword ptr -1

call Output

jmp endIf5

ifLabel5:

push dword ptr 0

call Output

endIf5:

push \_a

push \_b

push \_c

call Add\_

push eax

call Less\_

push eax

call Not\_

push eax

cmp eax, 0

je ifLabel6

push dword ptr 10

call Output

jmp endIf6

ifLabel6:

push dword ptr 0

call Output

endIf6:

exit\_label:

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke ReadConsoleA,hConsoleInput,ADDR InputBuf,1,ADDR NumberOfCharsRead,0

invoke ExitProcess, 0

;===Procedure Input==========================================================================

Input PROC

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA,hConsoleOutput,ADDR InMessage,SIZEOF InMessage,ADDR NumberOfCharsWrite,0

invoke ReadConsoleA,hConsoleInput,ADDR InputBuf,11,ADDR NumberOfCharsRead,0

invoke crt\_atoi, addr InputBuf

ret

Input ENDP

;============================================================================================

;===Procedure Output=========================================================================

Output PROC, value:dword

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke wsprintf, addr ResMessage, addr OutMessage, value

invoke WriteConsoleA,hConsoleOutput,ADDR ResMessage,eax,ADDR NumberOfCharsWrite,0

ret

Output ENDP

;============================================================================================

;===Procedure Add=========================

Add\_ PROC

mov eax, [esp+8]

add eax, [esp+4]

ret 8

Add\_ ENDP

;=========================================

;===Procedure Sub=========================

Sub\_ PROC

mov eax, [esp+8]

sub eax, [esp+4]

ret 8

Sub\_ ENDP

;=========================================

;===Procedure Mul=========================

Mul\_ PROC

mov eax, [esp+8]

mov ebx, [esp+4]

imul ebx

ret 8

Mul\_ ENDP

;=========================================

;===Procedure Div=========================

Div\_ PROC

pushf

pop cx

mov eax, [esp+4]

cmp eax, 0

jne end\_check

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA, hConsoleOutput, ADDR ErrorMessage, SIZEOF ErrorMessage, ADDR NumberOfCharsWrite, 0

jmp exit\_label

end\_check:

mov eax, [esp+8]

cmp eax, 0

jge gr

lo:

mov edx, -1

jmp less\_fin

gr:

mov edx, 0

less\_fin:

mov eax, [esp+8]

mov ebx, [esp+4]

idiv ebx

push cx

popf

ret 8

Div\_ ENDP

;=========================================

;===Procedure Mod=========================

Mod\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jne end\_check

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA, hConsoleOutput, ADDR ErrorMessage, SIZEOF ErrorMessage, ADDR NumberOfCharsWrite, 0

jmp exit\_label

end\_check:

mov eax, [esp+8]

cmp eax, 0

jge gr

lo:

mov edx, -1

jmp less\_fin

gr:

mov edx, 0

less\_fin:

mov eax, [esp+8]

mov ebx, [esp+4]

idiv ebx

mov eax, edx

push cx

popf

ret 8

Mod\_ ENDP

;=========================================

;===Procedure And=========================

And\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jnz and\_t1

jz and\_false

and\_t1:

mov eax, [esp+4]

cmp eax, 0

jnz and\_true

and\_false:

mov eax, 0

jmp and\_fin

and\_true:

mov eax, 1

and\_fin:

push cx

popf

ret 8

And\_ ENDP

;=========================================

;===Procedure Or==========================

Or\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jnz or\_true

mov eax, [esp+4]

cmp eax, 0

jnz or\_true

mov eax, 0

jmp or\_fin

or\_true:

mov eax, 1

or\_fin:

push cx

popf

ret 8

Or\_ ENDP

;=========================================

;===Procedure Not=========================

Not\_ PROC

pushf

pop cx

mov eax, [esp+4]

cmp eax, 0

jne not\_false

mov eax, 1

jmp not\_fin

not\_false:

mov eax, 0

not\_fin:

push cx

popf

ret 4

Not\_ ENDP

;=========================================

;===Procedure Eq==========================

Eq\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jne not\_eq

mov eax, 1

jmp eq\_fin

not\_eq:

mov eax, 0

eq\_fin:

push cx

popf

ret 8

Eq\_ ENDP

;=========================================

;===Procedure Greate======================

Greate\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jle lov

mov eax, 1

jmp gr\_fin

lov:

mov eax, 0

gr\_fin:

push cx

popf

ret 8

Greate\_ ENDP

;=========================================

;===Procedure Less========================

Less\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jge gr

lo:

mov eax, 1

jmp less\_fin

gr:

mov eax, 0

less\_fin:

push cx

popf

ret 8

Less\_ ENDP

;=========================================

end start

**Prog3.asm**

.386

.model flat, stdcall

option casemap :none

include masm32\include\windows.inc

include masm32\include\kernel32.inc

include masm32\include\masm32.inc

include masm32\include\user32.inc

include masm32\include\msvcrt.inc

includelib masm32\lib\kernel32.lib

includelib masm32\lib\masm32.lib

includelib masm32\lib\user32.lib

includelib masm32\lib\msvcrt.lib

.DATA

;===User Data==============

\_a dd 00h

\_a2 dd 00h

\_b dd 00h

\_x dd 00h

\_c1 dd 00h

\_c2 dd 00h

;===Addition Data===========================================

hConsoleInput dd ?

hConsoleOutput dd ?

ErrorMessage db "Error: division by zero", 0

InputBuf db 15 dup (?)

InMessage db "GET: ", 0

OutMessage db "PUT: %d", 0

ResMessage db 20 dup (?)

NumberOfCharsRead dd ?

NumberOfCharsWrite dd ?

msg1310 dw 13,10

buf dd 0

.CODE

start:

invoke AllocConsole

invoke GetStdHandle, STD\_INPUT\_HANDLE

mov hConsoleInput, EAX

invoke GetStdHandle, STD\_OUTPUT\_HANDLE

mov hConsoleOutput, EAX

call Input

mov \_a, eax

call Input

mov \_b, eax

push \_a

pop eax

mov \_a2, eax

whileStart1:

push \_a2

push \_a2

call Mul\_

push eax

call Output

push \_a2

push dword ptr 1

call Add\_

push eax

mov \_a2, eax

push \_a2

push \_b

call Greate\_

push eax

call Not\_

push eax

cmp eax, 1

jne whileEnd1

jmp whileStart1

whileEnd1:

push dword ptr 0

pop eax

mov \_x, eax

push dword ptr 1

pop eax

mov \_c1, eax

whileStart2:

push dword ptr 1

pop eax

mov \_c2, eax

whileStart3:

push \_x

push dword ptr 1

call Add\_

push eax

mov \_x, eax

push \_c2

push dword ptr 1

call Add\_

push eax

mov \_c2, eax

push \_c2

push \_b

call Greate\_

push eax

call Not\_

push eax

cmp eax, 1

jne whileEnd3

jmp whileStart3

whileEnd3:

push \_c1

push dword ptr 1

call Add\_

push eax

mov \_c1, eax

push \_c1

push \_a

call Greate\_

push eax

call Not\_

push eax

cmp eax, 1

jne whileEnd2

jmp whileStart2

whileEnd2:

push \_x

call Output

exit\_label:

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke ReadConsoleA,hConsoleInput,ADDR InputBuf,1,ADDR NumberOfCharsRead,0

invoke ExitProcess, 0

;===Procedure Input==========================================================================

Input PROC

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA,hConsoleOutput,ADDR InMessage,SIZEOF InMessage,ADDR NumberOfCharsWrite,0

invoke ReadConsoleA,hConsoleInput,ADDR InputBuf,11,ADDR NumberOfCharsRead,0

invoke crt\_atoi, addr InputBuf

ret

Input ENDP

;============================================================================================

;===Procedure Output=========================================================================

Output PROC, value:dword

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke wsprintf, addr ResMessage, addr OutMessage, value

invoke WriteConsoleA,hConsoleOutput,ADDR ResMessage,eax,ADDR NumberOfCharsWrite,0

ret

Output ENDP

;============================================================================================

;===Procedure Add=========================

Add\_ PROC

mov eax, [esp+8]

add eax, [esp+4]

ret 8

Add\_ ENDP

;=========================================

;===Procedure Sub=========================

Sub\_ PROC

mov eax, [esp+8]

sub eax, [esp+4]

ret 8

Sub\_ ENDP

;=========================================

;===Procedure Mul=========================

Mul\_ PROC

mov eax, [esp+8]

mov ebx, [esp+4]

imul ebx

ret 8

Mul\_ ENDP

;=========================================

;===Procedure Div=========================

Div\_ PROC

pushf

pop cx

mov eax, [esp+4]

cmp eax, 0

jne end\_check

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA, hConsoleOutput, ADDR ErrorMessage, SIZEOF ErrorMessage, ADDR NumberOfCharsWrite, 0

jmp exit\_label

end\_check:

mov eax, [esp+8]

cmp eax, 0

jge gr

lo:

mov edx, -1

jmp less\_fin

gr:

mov edx, 0

less\_fin:

mov eax, [esp+8]

mov ebx, [esp+4]

idiv ebx

push cx

popf

ret 8

Div\_ ENDP

;=========================================

;===Procedure Mod=========================

Mod\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jne end\_check

invoke WriteConsoleA,hConsoleOutput,ADDR msg1310,SIZEOF msg1310,ADDR NumberOfCharsWrite,0

invoke WriteConsoleA, hConsoleOutput, ADDR ErrorMessage, SIZEOF ErrorMessage, ADDR NumberOfCharsWrite, 0

jmp exit\_label

end\_check:

mov eax, [esp+8]

cmp eax, 0

jge gr

lo:

mov edx, -1

jmp less\_fin

gr:

mov edx, 0

less\_fin:

mov eax, [esp+8]

mov ebx, [esp+4]

idiv ebx

mov eax, edx

push cx

popf

ret 8

Mod\_ ENDP

;=========================================

;===Procedure And=========================

And\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jnz and\_t1

jz and\_false

and\_t1:

mov eax, [esp+4]

cmp eax, 0

jnz and\_true

and\_false:

mov eax, 0

jmp and\_fin

and\_true:

mov eax, 1

and\_fin:

push cx

popf

ret 8

And\_ ENDP

;=========================================

;===Procedure Or==========================

Or\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, 0

jnz or\_true

mov eax, [esp+4]

cmp eax, 0

jnz or\_true

mov eax, 0

jmp or\_fin

or\_true:

mov eax, 1

or\_fin:

push cx

popf

ret 8

Or\_ ENDP

;=========================================

;===Procedure Not=========================

Not\_ PROC

pushf

pop cx

mov eax, [esp+4]

cmp eax, 0

jne not\_false

mov eax, 1

jmp not\_fin

not\_false:

mov eax, 0

not\_fin:

push cx

popf

ret 4

Not\_ ENDP

;=========================================

;===Procedure Eq==========================

Eq\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jne not\_eq

mov eax, 1

jmp eq\_fin

not\_eq:

mov eax, 0

eq\_fin:

push cx

popf

ret 8

Eq\_ ENDP

;=========================================

;===Procedure Greate======================

Greate\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jle lov

mov eax, 1

jmp gr\_fin

lov:

mov eax, 0

gr\_fin:

push cx

popf

ret 8

Greate\_ ENDP

;=========================================

;===Procedure Less========================

Less\_ PROC

pushf

pop cx

mov eax, [esp+8]

cmp eax, [esp+4]

jge gr

lo:

mov eax, 1

jmp less\_fin

gr:

mov eax, 0

less\_fin:

push cx

popf

ret 8

Less\_ ENDP

;=========================================

end start