Міністерство освіти і науки України

Національний університет "Львівська Політехніка"

Кафедра ЕОМ



**Пояснювальна записка**

до курсового проєкту "Системне програмування"

на тему: “Розробка системних програмних модулів та компонент систем програмування”

Індивідуальне завдання

“Розробка транслятора з вхідної мови програмування”

**Варіант №20**

Виконав:

ст. гр. КІ-307

Маринович Марко

Перевірив:

Козак Н. Б.

Львів-2024

Завдання на курсовий проєкт

1. Цільова мова транслятора – мова програмування С або асемблер для 32/64 розрядного процесора.
2. Для отримання виконуваного файлу на виході розробленого транслятора скористатися середовищем Microsoft Visual Studio або будь-яким іншим.
3. Мова розробки транслятора: C/C++.
4. Реалізувати графічну 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 С або асемблера;*

*об’єктний файл;*

*виконуваний файл.*

1. 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.

**Деталізація завдання на проектування:**

1. В кожному завданні передбачається блок оголошення змінних; змінні зберігають значення цілих чисел і, в залежності від варіанту, можуть бути 16/32 розрядними. За потребою можна реалізувати логічний тип даних.
2. Необхідно реалізувати арифметичні операції – додавання, віднімання, множення, ділення, залишок від ділення; операції порівняння – перевірка на рівність і нерівність, більше і менше; логічні операції – заперечення, “логічне І” і “логічне АБО”.

Пріоритет операцій наступний – круглі дужки (), логічне заперечення, мультиплікативні (множення, ділення, залишок від ділення), адитивні (додавання, віднімання), відношення (більше, менше), перевірка на рівність і нерівність, логічне І, логічне АБО.

1. За допомогою оператора вводу можна зчитати з клавіатури значення змінної; за допомогою оператора виводу можна вивести на екран значення змінної, виразу чи цілої константи.
2. В кожному завданні обов’язковим є оператор присвоєння за допомогою якого можна реалізувати обчислення виразів з використанням заданих операцій і операції круглі дужки (); у якості операндів можуть бути цілі константи, змінні, а також інші вирази.
3. В кожному завданні обов’язковим є оператор типу “блок” (складений оператор), його вигляд має бути таким, як і блок тіла програми.
4. Необхідно реалізувати задані варіантом оператори, синтаксис операторів наведено у таблиці 1.1. Синтаксис вхідної мови має забезпечити реалізацію обчислень лінійних алгоритмів, алгоритмів з розгалуженням і циклічних алгоритмів. Опис формальної мови студент погоджує з викладачем.
5. Оператори можуть бути довільної вкладеності і в будь-якій послідовності.
6. Для перевірки роботи розробленого транслятора, необхідно написати три тестові програми на вхідній мові програмування.

**Деталізований опис власної мови програмування:**

Опис вхідної мови програмування:

* Тип даних: INT\_4
* Блок тіла програми: STARTPROGRAM VARIABLE…; STARTBLOK ENDBLOK
* Оператор вводу: READ ()
* Оператор виводу: WRITE ()
* Оператори: IF ELSE (C)

GOTO (C)

FOR-TO-DO (Паскаль)

FOR-DOWNTO-DO (Паскаль)

WHILE (Бейсік)

REPEAT-UNTIL (Паскаль)

* Регістр ключових слів: Up
* Регістр ідентифікаторів: Low4 перший символ \_
* Операції арифметичні: ADD, SUB, MUL, DIV, MOD
* Операції порівняння: EQ, NE, GT, LT
* Операції логічні: !, &, |
* 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).

Анотація

Цей курсовий проект приводить до розробки транслятора, який здатен конвертувати вхідну мову, визначену відповідно до варіанту, у мову асемблера. Процес трансляції включає в себе лексичний аналіз, синтаксичний аналіз та генерацію коду.

Лексичний аналіз розбиває вхідну послідовність символів на лексеми, які записуються у відповідну таблицю лексем. Кожній лексемі присвоюється числове значення для полегшення порівнянь, а також зберігається додаткова інформація, така як номер рядка, значення (якщо тип лексеми є числом) та інші деталі.

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

Генерація коду включає повторне прочитання таблиці лексем та створення відповідного асемблерного коду для кожного блоку лексем. Отриманий код записується у результуючий файл, готовий для виконання.

Отриманий після трансляції код можна скомпілювати за допомогою відповідних програм (наприклад, LINK, ML і т. д.).

# 

# **ЗМІСТ**

[Завдання на курсовий проєкт 2](#_Toc188006012)

[Анотація 5](#_Toc188006013)

[ЗМІСТ 6](#_Toc188006014)

[Вступ 7](#_Toc188006015)

[1. Огляд методів та способів проєктування трансляторів 8](#_Toc188006016)

[2. Формальний опис вхідної мови програмування 12](#_Toc188006017)

[2.1. Деталізований опис вхідної мови в термінах розширеної нотації Бекуса-Наура. 12](#_Toc188006018)

[3. Розробка транслятора з вхідної мови програмування 16](#_Toc188006019)

[3.1. Вибір технології програмування. 16](#_Toc188006020)

[3.2. Проектування таблиць транслятора та вибір структур даних. 16](#_Toc188006021)

[3.3. Розробка лексичного аналізатора. **Помилка! Закладку не визначено.**](#_Toc188006022)

[3.4. Розробка синтаксичного та семантичного аналізатора. **Помилка! Закладку не визначено.**](#_Toc188006023)

[3.5. Розробка генератора коду. 22](#_Toc188006024)

[4. Налагодження та тестування розробленого транслятора 40](#_Toc188006025)

[4.1. Опис інтерфейсу та інструкції користувачу. 44](#_Toc188006026)

[4.2. Виявлення лексичних і синтаксичних помилок. 45](#_Toc188006027)

[4.3. Перевірка роботи транслятора за допомогою тестових задач. 47](#_Toc188006028)

[Тестова програма №1 49](#_Toc188006029)

[Тестова програма №2 51](#_Toc188006030)

[Тестова програма №3 53](#_Toc188006031)

[Висновки 57](#_Toc188006032)

[Список літературних джерел 58](#_Toc188006033)

[Додатки 59](#_Toc188006034)

Вступ

Термін "транслятор" визначає програму, яка виконує переклад (трансляцію) початкової програми, написаної на вхідній мові, у еквівалентну їй об'єктну програму. У випадку, коли мова високого рівня є вхідною, а мова асемблера або машинна – вихідною, такий транслятор отримує назву компілятора.

Транслятори можуть бути розділені на два основних типи: компілятори та інтерпретатори. Процес компіляції включає дві основні фази: аналіз та синтез. Під час аналізу вхідну програму розбивають на окремі елементи (лексеми), перевіряють її відповідність граматичним правилам і створюють проміжне представлення програми. На етапі синтезу з проміжного представлення формується програма в машинних кодах, яку називають об'єктною програмою. Останню можна виконати на комп'ютері без додаткової трансляції.

У відміну від компіляторів, інтерпретатор не створює нову програму; він лише виконує – інтерпретує – кожну інструкцію вхідної мови програмування. Подібно компілятору, інтерпретатор аналізує вхідну програму, створює проміжне представлення, але не формує об'єктну програму, а негайно виконує команди, передбачені вхідною програмою.

Компілятор виконує переклад програми з однієї мови програмування в іншу. На вхід компілятора надходить ланцюг символів, який представляє вхідну програму на певній мові програмування. На виході компілятора (об'єктна програма) також представляє собою ланцюг символів, що вже відповідає іншій мові програмування, наприклад, машинній мові конкретного комп'ютера. При цьому сам компілятор може бути написаний на третій мові.

1. Огляд методів та способів проєктування трансляторів

Термін "транслятор" визначає обслуговуючу програму, що проводить трансляцію вихідної програми, представленої на вхідній мові програмування, у робочу програму, яка відображена на об'єктній мові. Наведене визначення застосовне до різноманітних транслюють програм. Однак кожна з таких програм може виявляти свої особливості в організації процесу трансляції. В сучасному контексті транслятори поділяються на три основні групи: асемблери, компілятори та інтерпретатори.

Асемблер - це системна обслуговуюча програма, яка перетворює символічні конструкції в команди машинної мови. Типовою особливістю асемблерів є дослівна трансляція однієї символічної команди в одну машинну.

Компілятор - обслуговуюча програма, яка виконує трансляцію програми, написаної мовою оригіналу програмування, в машинну мову. Схоже до асемблера, компілятор виконує перетворення програми з однієї мови в іншу, найчастіше - у мову конкретного комп'ютера.

Інтерпретатор - це програма чи пристрій, що виконує пооператорну трансляцію та виконання вихідної програми. Відмінно від компілятора, інтерпретатор не створює на виході програму на машинній мові. Розпізнавши команду вихідної мови, він негайно її виконує, забезпечуючи більшу гнучкість у процесі розробки та налагодження програм.

Процес трансляції включає фази лексичного аналізу, синтаксичного та семантичного аналізу, оптимізації коду та генерації коду. Лексичний аналіз розбиває вхідну програму на лексеми, що представляють слова відповідно до визначень мови. Синтаксичний аналіз визначає структуру програми, створюючи синтаксичне дерево. Семантичний аналіз виявляє залежності між частинами програми, недосяжні контекстно-вільним синтаксисом. Оптимізація коду та генерація коду спрямовані на оптимізацію та створення машинно-залежного коду відповідно.

Зазначені фази можуть об'єднуватися або відсутні у трансляторах в залежності від їхньої реалізації. Наприклад, у простих однопрохідних трансляторах може відсутні фаза генерації проміжного представлення та оптимізації, а інші фази можуть об'єднуватися.

Під час процесу виділення лексем лексичний аналізатор може виконувати дві основні функції: автоматично побудову таблиць об'єктів (таких як ідентифікатори, рядки, числа і т. д.) і видачу значень для кожної лексеми при кожному новому зверненні до нього. У цьому контексті таблиці об'єктів формуються в подальших етапах, наприклад, під час синтаксичного аналізу.

На етапі лексичного аналізу виявляються деякі прості помилки, такі як неприпустимі символи або невірний формат чисел та ідентифікаторів.

Основним завданням синтаксичного аналізу є розбір структури програми. Зазвичай під структурою розуміється дерево, яке відповідає розбору в контекстно-вільній граматиці мови програмування. У сучасній практиці найчастіше використовуються методи аналізу, такі як LL (1) або LR (1) та їхні варіанти (рекурсивний спуск для LL (1) або LR (1), LR (0), SLR (1), LALR (1) та інші для LR (1)). Рекурсивний спуск застосовується частіше при ручному програмуванні синтаксичного аналізатора, тоді як LR (1) використовується при автоматичній генерації синтаксичних аналізаторів.

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

На етапі контекстного аналізу виявляються взаємозалежності між різними частинами програми, які не можуть бути адекватно описані за допомогою контекстно-вільної граматики. Ці взаємозалежності, зокрема, включають аналіз типів об'єктів, областей видимості, відповідності параметрів, міток та інших аспектів "опис-використання". У ході контекстного аналізу таблиці об'єктів доповнюються інформацією, пов'язаною з описами (властивостями) об'єктів.

В основі контекстного аналізу лежить апарат атрибутних граматик. Результатом цього аналізу є створення атрибутованого дерева програми, де інформація про об'єкти може бути розсіяна в самому дереві чи сконцентрована в окремих таблицях об'єктів. Під час контекстного аналізу також можуть бути виявлені помилки, пов'язані з неправильним використанням об'єктів.

Після завершення контекстного аналізу програма може бути перетворена во внутрішнє представлення. Це здійснюється з метою оптимізації та/або для полегшення генерації коду. Крім того, перетворення програми у внутрішнє представлення може бути використано для створення переносимого компілятора. У цьому випадку, тільки остання фаза (генерація коду) є залежною від конкретної архітектури. В якості внутрішнього представлення може використовуватися префіксний або постфіксний запис, орієнтований граф, трійки, четвірки та інші формати.

Фаза оптимізації транслятора може включати декілька етапів, які спрямовані на покращення якості та ефективності згенерованого коду. Ці оптимізації часто розподіляються за двома головними критеріями: машинно-залежні та машинно-незалежні, а також локальні та глобальні.

Машинно-залежні оптимізації, як правило, проводяться на етапі генерації коду, і вони орієнтовані на конкретну архітектуру машини. Ці оптимізації можуть включати розподіл регістрів, вибір довгих або коротких переходів та оптимізацію вартості команд для конкретних послідовностей команд.

Глобальна оптимізація спрямована на поліпшення ефективності всієї програми і базується на глобальному потоковому аналізі, який виконується на графі програми. Цей аналіз враховує властивості програми, такі як межпроцедурний аналіз, міжмодульний аналіз та аналіз галузей життя змінних.

Фінальна фаза трансляції - генерація коду, результатом якої є або асемблерний модуль, або об'єктний (або завантажувальний) модуль. На цьому етапі можуть застосовуватися деякі локальні оптимізації для полегшення генерації вартісного та ефективного коду.

Важливо відзначити, що фази транслятора можуть бути відсутніми або об'єднаними в залежності від конкретної реалізації. В простіших випадках, таких як у випадку однопроходових трансляторів, може відсутній окремий етап генерації проміжного представлення та оптимізації, а інші фази можуть бути об'єднані в одну, при цьому не створюється 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 = "STARTPROGRAM", variable\_block, ";", code\_block;

variable\_block = "INT\_4", identifier, {comma\_and\_identifier}, ";";

identifier = "\_", low\_letter, {low\_letter | digit}, {3};

comma\_and\_identifier = ",", identifier;

code\_block = "STARTBLOK", {statement}, "ENDBLOK";

statement = write | read | assignment | if\_statement | goto\_statement | label\_rule | for\_to\_or\_downto\_do\_rule | while\_loop | repeat\_until;

write = "WRITE", "(", equation | string\_rule, ")";

read = "READ", "(", identifier, ")";

assignment = identifier, "<==", equation;

if\_statement = "IF", "(", equation, ")", code\_block, [else\_statement];

else\_statement = "ELSE", code\_block;

goto\_statement = "GOTO", identifier;

label\_rule = identifier, ":";

for\_to\_or\_downto\_do\_rule = "FOR", assignment, ("TO" | "DOWNTO"), equation, "DO", code\_block;

while\_loop = "WHILE", "(", equation, ")", code\_block;

repeat\_until = "REPEAT", code\_block, "UNTIL", "(", equation, ")";

equation = signed\_number | identifier | not\_rule, {operation\_and\_identifier\_or\_number | equation};

not\_rule = not\_operation, (signed\_number | identifier | equation);

operation\_and\_identifier\_or\_number = (arithmetic | mult | logic | compare), (signed\_number | identifier | equation);

arithmetic = "ADD" | "SUB";

mult = "MUL" | "DIV" | "MOD";

logic = "&" | "|";

not\_operation = "!";

compare = "EQ" | "NE" | "LT" | "GT";

string\_rule = "\"", string, "\"";

comment = l\_comment, string, r\_comment;

l\_comment = "#\*";

r\_comment = "\*#";

string = {low\_letter | up\_letter | digit};

signed\_number = [sign], digit, {digit};

sign = "+" | "-";

low\_letter = "a" | "b" | "c" | "d" | "e" | "f" | "g" | "h" | "i" | "j" | "k" | "l" | "m" | "n" | "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" | "M" | "N" | "O" | "P" | "Q" | "R" | "S" | "T" | "U" | "V" | "W" | "X" | "Y" | "Z";

digit = "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9";

* 1. **Опис термінальних символів та ключових слів.**

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

|  |  |
| --- | --- |
| Термінальний символ або ключове слово | Значення |
| STARTPROGRAM | Початок програми |
| STARTBLOK | Початок тексту програми |
| VARIABLE | Початок блоку опису змінних |
| ENDBLOK | Кінець розділу операторів |
| READ | Оператор вводу змінних |
| WRITE | Оператор виводу (змінних або рядкових констант) |
| <== | Оператор присвоєння |
| IF | Оператор умови |
| ELSE | Оператор умови |
| GOTO | Оператор переходу |
| LABEL | Мітка переходу |
| FOR | Оператор циклу |
| TO | Інкремент циклу |
| DOWNTO | Декремент циклу |
| DO | Початок тіла циклу |
| WHILE | Оператор циклу |
| REPEAT | Початок тіла циклу |
| UNTIL | Оператор циклу |
| ADD | Оператор додавання |
| SUB | Оператор віднімання |
| MUL | Оператор множення |
| DIV | Оператор ділення |
| MOD | Оператор знаходження залишку від ділення |
| EQ | Оператор перевірки на рівність |
| NE | Оператор перевірки на нерівність |
| LT | Оператор перевірки чи менше |
| GT | Оператор перевірки чи більше |
| ! | Оператор логічного заперечення |
| & | Оператор кон’юнкції |
| | | Оператор диз’юнкції |
| INT\_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 з 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 та вибір структур даних.

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

1. Мульти мапа для лексеми, значення та рядка кожного токена.

std::multimap<int, std::shared\_ptr<IToken>> m\_priorityTokens;

std::string m\_lexeme; //Лексема

std::string m\_value; //Значення

int m\_line = -1; //Рядок

1. Таблиця лексичних класів

Якщо у стовпці «Значення» відсутня інформація про токен, то це означає що його значення визначається користувачем під час написання коду на створеній мові програмування.

Таблиця 2 Опис термінальних символі та ключових слів

|  |  |
| --- | --- |
| **Токен** | **Значення** |
| Program | STARTPROGRAM |
| Start | STARTBLOK |
| Vars | VARIABLE |
| End | ENDBLOK |
| VarType | INT\_4 |
| Read | READ |
| Write | WRITE |
| Assignment | <== |
| If | IF |
| Else | ELSE |
| Goto | GOTO |
| Colon | : |
| Label |  |
| For | FOR |
| To | TO |
| DownTo | DOWNTO |
| Do | DO |
| While | WHILE |
| Repeat | REPEAT |
| Until | UNTIL |
| Addition | ADD |
| Subtraction | SUB |
| Multiplication | MUL |
| Division | DIV |
| Mod | MOD |
| Equal | EQ |
| NotEqual | NE |
| Less | LT |
| Greate | GT |
| Not | ! |
| And | & |
| Or | | |
| Plus | + |
| Minus | - |
| Identifier |  |
| Number |  |
| String |  |
| Undefined |  |
| Unknown |  |
| Comma | , |
| Quotes | “ |
| Semicolon | ; |
| LBraket | ( |
| RBraket | ) |
| LComment | #\* |
| RComment | \*# |
| Comment |  |

* 1. **Розробка лексичного аналізатора**

Основна задача лексичного аналізу — розбити вхідний текст, що складається з послідовності символів, на послідовність лексем (слів), тобто виділити ці слова з безперервної послідовності символів. Всі символи вхідної послідовності можна поділити на дві категорії:

1. Символи, що належать яким-небудь лексемам (ключові слова, ідентифікатори, числові константи, оператори та інші).
2. Символи, що розділяють лексеми (пробіли, знаки операцій, нові рядки тощо).

Лексичний аналізатор працює в два основних режими:

* Як підпрограма, що викликається синтаксичним аналізатором для отримання чергової лексеми.
* Як повний прохід, результатом якого є файл лексем, що містить усі лексеми програми.

Ми обрали другий варіант, де спочатку виконується фаза лексичного аналізу, а результат цієї фази передається для подальшої обробки на етап синтаксичного аналізу.

**3.3.1. Розробка алгоритму роботи лексичного аналізатора**

Лексичний аналізатор працює за принципом скінченного автомату, що містить такі стани:

* Start — початок виділення чергової лексеми.
* Finish — кінець виділення чергової лексеми.
* EndOfFile — кінець файлу, завершення розпізнавання лексем.
* Letter — перший символ є літерою або символом \_, розпізнаються ключові слова та ідентифікатори.
* Digit — перший символ є цифрою, розпізнаються числові константи.
* Separator — обробка роздільників (пробіли, табуляції, нові рядки).
* SComment — початок коментаря.
* Comment — ігнорування коментаря.
* Another — обробка інших символів, зокрема операторів.

Алгоритм лексичного аналізатора передбачає послідовне читання символів з вхідного файлу. Кожен символ порівнюється з набором правил для лексем. Якщо лексема відповідає певному правилу (наприклад, ключове слово або ідентифікатор), вона записується в таблицю лексем разом з її типом та додатковою інформацією.

Лексеми, які не відповідають жодному з правил, позначаються як невизначені (невідомі) лексеми.

Зображення, що містить схема, Креслення, План, текст

Автоматично згенерований опис

*Рис. 3.1. Схема алгоритму роботи лексичного аналізатора.*

**3.3.2. Опис програми реалізації лексичного аналізатора**

У програмі для реалізації лексичного аналізу використовуються такі основні компоненти:

1. Типи лексем (Tokens):
   * Ключові слова: STARTPROGRAM, STARTBLOCK, VARIABLE, ENDBLOCK, INT16, INPUT, OUTPUT, IF, ELSE, FOR, TO, DOWNTO, DO, WHILE, WEND, REPEAT, UNTIL, DIV, MOD.
   * Ідентифікатори: рядок, що починається з символу \_, за яким йдуть літери або цифри, максимум 6 символів.
   * Числові константи: ціле число.
   * Оператори: ==>, +, -, \*, =, !=, >>, <<, !!, &&, ||.
   * Розділювачі: ,, ;.
   * Дужки: (, ).
   * Невідома лексема: символи, що не підпадають під описані правила.
2. Алгоритм роботи лексичного аналізатора:
   * Читання файлу та виділення лексем через функцію tokenize().
   * Визначення типу лексеми за допомогою порівняння з ключовими словами та шаблонами.
   * Формування таблиці лексем m\_tokens, де для кожної лексеми вказується її тип, значення та рядок, на якому вона була знайдена.
   * Виявлення лексичних помилок, таких як недопустимі символи, неправильні ідентифікатори або числові константи.
3. Структура даних для зберігання стану аналізатора:

enum States {

Start, // початковий стан

Finish, // кінцевий стан

Letter, // опрацювання слів (ключові слова та ідентифікатори)

Digit, // опрацювання цифр

Separator, // опрацювання роздільників

Another, // опрацювання інших символів

EndOfFile, // кінець файлу

SComment, // початок коментаря

Comment // ігнорування коментаря

};

1. Функції для лексичного аналізатора:
   * unsigned int getTokens(FILE\* F): основна функція для отримання лексем з файлу.
   * void printTokens(void): друк лексем.
   * void fprintTokens(FILE\* F): запис лексем у файл.
   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. 3азвичай 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 дерева граматичного розбору.**

Схема дерева розбору виглядає наступним чином:

Program

├── "STARTPROGRAM"

├── "STARTBLOCK"

├── {VariableDeclaration ";"}

│ ├── VariableDeclaration

│ │ ├── "INT16"

│ │ └── VariableList

│ │ ├── Identifier

│ │ │ ├── "\_"

│ │ │ └── UpLow {5}

│ │ │ ├── Up

│ │ │ │ └── "A" | "B" | ... | "Z"

│ │ │ ├── Low

│ │ │ │ └── "a" | "b" | ... | "z"

│ │ │ └── Digit

│ │ │ └── "0" | "1" | ... | "9"

│ │ └── {"," Identifier}

├── {Statement}

│ ├── Statement

│ │ ├── InputStatement

│ │ │ ├── "INPUT"

│ │ │ └── Identifier

│ │ ├── OutputStatement

│ │ │ ├── "OUTPUT"

│ │ │ └── ArithmeticExpression

│ │ │ ├── LowPriorityExpression

│ │ │ │ ├── MiddlePriorityExpression

│ │ │ │ │ ├── Identifier

│ │ │ │ │ ├── Number

│ │ │ │ │ │ ├── ["-"]

│ │ │ │ │ │ └── Digit {5}

│ │ │ │ │ └── "(" ArithmeticExpression ")"

│ │ │ │ └── {MiddlePriorityOperator MiddlePriorityExpression}

│ │ │ └── {LowPriorityOperator LowPriorityExpression}

│ │ ├── AssignStatement

│ │ │ ├── ArithmeticExpression

│ │ │ └── "==>" Identifier

│ │ ├── IfElseStatement

│ │ │ ├── "IF"

│ │ │ ├── "(" LogicalExpression ")"

│ │ │ │ ├── AndExpression

│ │ │ │ │ ├── Comparison

│ │ │ │ │ │ ├── ComparisonExpression

│ │ │ │ │ │ │ ├── ArithmeticExpression

│ │ │ │ │ │ │ ├── ComparisonOperator

│ │ │ │ │ │ │ └── ArithmeticExpression

│ │ │ │ │ │ └── [NotOperator] "(" LogicalExpression ")"

│ │ │ │ │ └── {AndOperator AndExpression}

│ │ │ │ └── {OrOperator AndExpression}

│ │ │ ├── Statement

│ │ │ └── ["ELSE" Statement]

│ │ ├── GotoStatement

│ │ │ ├── "GOTO"

│ │ │ └── Identifier

│ │ ├── LabelPoint

│ │ │ ├── Identifier

│ │ │ └── ":"

│ │ ├── ForToStatement

│ │ │ ├── "FOR"

│ │ │ ├── AssignStatement

│ │ │ ├── "TO" | "DOWNTO"

│ │ │ ├── ArithmeticExpression

│ │ │ ├── "DO"

│ │ │ └── Statement

│ │ ├── WhileStatement

│ │ │ ├── "WHILE"

│ │ │ ├── LogicalExpression

│ │ │ ├── {Statement}

│ │ │ └── "END" "WHILE"

│ │ ├── RepeatUntilStatement

│ │ │ ├── "REPEAT"

│ │ │ ├── {Statement}

│ │ │ └── "UNTIL" "(" LogicalExpression ")"

│ │ └── CompoundStatement

│ │ ├── "STARTBLOCK"

│ │ ├── {Statement}

│ │ └── "ENDBLOCK"

└── "ENDBLOCK"

Зображення, що містить малюнок, ескіз, схема, Лінійне мистецтво

Автоматично згенерований опис

*Рис. 3.2. Дерево граматичного розбору.*

* + 1. **Розробка алгоритму роботи синтаксичного і 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.

Процес перевірки EBNF в проекті реалізований через систему правил Backus та складається з наступних компонентів:

Базова структура перевірки:

* Інтерфейс IBackusRule визначає базовий контракт для всіх правил.

Політики перевірки правил:

Enum RuleCountPolicy визначає можливі варіанти входження правил:

* NoPolicy – без політики
* Optional – необов'язкове правило
* OnlyOne – тільки один раз
* Several – декілька разів
* OneOrMore – один або більше разів
* PairStart/PairEnd – парні конструкції

Реєстрація та зберігання правил:

* Клас Controller відповідає за реєстрацію правил.
* BackusRuleStorage зберігає зареєстровані правила.

Визначення правил граматики:

* Правила визначаються через BackusRuleItem з вказанням політики.
* Підтримується ієрархічна структура правил.

Процес перевірки:

* Базовий клас BackusRuleBase реалізує базову перевірку типів.
* Клас BackusRule реалізує складну перевірку правил з урахуванням політик.

Обробка помилок:

* Помилки збираються в multimap з інформацією про тип помилки та контекст.
* Кожне правило може генерувати власні помилки.

Цей механізм дозволяє:

* Перевіряти відповідність коду заданій EBNF граматиці.
* Гнучко налаштовувати правила перевірки.
* Отримувати детальну інформацію про помилки.
* Розширювати граматику новими правилами.

Визначимо назви процедур, що відповідають нетерміналам граматики таким чином:

void program(); // розбір програми

void programBody(); // розбір тіла програми

void variableDeclaration(); // оголошення змінних

void variableList(); // список змінних

void statement(); // оператори

void inputStatement(); // оператор вводу

void outputStatement(); // оператор виводу

void arithmeticExpression(); // арифметичні вирази

void lowPriorityExpression(); // низький пріоритет виразів

void middlePriorityExpression(); // середній пріоритет виразів

void assignStatement(); // оператор присвоєння

void ifStatement(); // оператор if

void logicalExpression(); // логічні вирази

void andExpression(); // вирази з оператором AND

void comparison(); // порівняння

void comparisonExpression(); // порівняння двох арифметичних виразів

void gotoStatement(); // оператор GOTO

void labelPoint(); // мітка

void forStatement(); // оператор FOR

void whileStatement(); // оператор WHILE

void repeatStatement(); // оператор REPEAT

void compoundStatement(); // складний оператор (STARTBLOCK / ENDBLOCK)

Блок-схема алгоритму роботи синтаксичного аналізатора виглядатиме наступним чином:

Зображення, що містить текст, знімок екрана, ряд, схема

Автоматично згенерований опис

*Рис. 3.3. Блок-сема алгоритму роботи синтаксичного аналізатора.*

Верхньорівневий код, який описує блок схема 3.3

auto topRule = controller->addRule("TopRule", {

BackusRuleItem({ Program::Type()}, OnlyOne),

BackusRuleItem({ identRule->type()}, OnlyOne),

BackusRuleItem({ Symbols::Semicolon}, OnlyOne),

BackusRuleItem({ Start::Type()}, OnlyOne),

BackusRuleItem({ Vars::Type()}, OnlyOne),

BackusRuleItem({ varsBlok->type()}, OnlyOne),

BackusRuleItem({ operators->type(), operatorsWithSemicolon->type()}, Optional | OneOrMore),

BackusRuleItem({ End::Type()}, OnlyOne)

});

Код що перевіряє валідність оголошених змінних

std::shared\_ptr<IToken> tryCreateToken(std::string& lexeme) const override

{

if (lexeme.size() > (m\_mask.size() + m\_prefix.size()))

return nullptr;

bool res = true;

if (!lexeme.starts\_with(m\_prefix))

{

return nullptr;

}

std::string\_view ident{ lexeme.begin() + m\_prefix.size(), lexeme.end() };

for (size\_t i = 0; i < ident.size(); i++)

{

if ((isupper(ident[i]) != isupper(m\_mask[i])) && !isdigit(ident[i]))

{

res &= false;

break;

}

}

std::shared\_ptr<IToken> token = nullptr;

if (res)

{

token = clone();

token->setValue(lexeme);

lexeme.clear();

}

return token;

};

І приватні поля що задають формат:

const std::string m\_prefix = "\_";

const std::string m\_mask = "xXXXXXX";

* + 1. **Опис програми реалізації генератора коду.**

У компілятора, реалізованого в даному курсовому проекті, 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. **Розробка алгоритму роботи семантичного аналізатора**

На етапі семантичного аналізу вирішується завдання ідентифікації ідентифікаторів. Алгоритм складається з двох частин:

* Обробка оголошень ідентифікаторів.
* Обробка використання ідентифікаторів.

Коли лексичний аналізатор виявляє чергову лексему, що є ідентифікатором, він формує структуру з атрибутами, такими як ім'я, тип і лексичний клас. Ця інформація передається семантичному аналізатору. Якщо обробляється оголошення ідентифікатора, основним завданням є запис інформації до таблиці ідентифікаторів.

При обробці використання ідентифікатора семантичний аналізатор використовує раніше створену таблицю ідентифікаторів. Для отримання даних про тип ідентифікатора необхідно прочитати відповідне поле цієї таблиці.

* + 1. Опис програмної реалізації семантичного аналізатора

Семантичний аналізатор забезпечує перевірку правильності структури та логіки програми, аналізуючи лексеми та граматику. Реалізація включає кілька ключових функцій.

#### Основні аспекти реалізації:

1. **Лексеми та граматика**  
   Семантичний аналізатор працює з таблицею лексем і граматикою, які є результатом лексичного та синтаксичного аналізу. Типи лексем визначаються полем type, а функція GetTypeName використовується для отримання назви типу.
2. **Перевірка конфліктів**  
   Виявляються помилки в ідентифікаторах, щоб уникнути неоднозначностей і забезпечити коректність виконання програми.
3. **Обробка помилок**  
   Усі знайдені помилки виводяться до консолі за допомогою механізму semErr.
4. **Рекурсивна перевірка правил**  
   Аналізатор підтримує рекурсивну перевірку граматичних правил, обробку необов’язкових конструкцій, парних елементів і використання політик для визначення кількості правил (RuleCountPolicy).
5. **Виконання**  
   Функція CheckSemantic відповідає за запуск семантичного аналізу, використовуючи об'єкт Context для зберігання поточного стану.

bool CheckSemantic(std::ostream& out, std::list<std::shared\_ptr<T>>& tokens)

{

auto endOfFileType = tokens.back()->type();

std::list<std::shared\_ptr<IBackusRule>> rules;

for (auto token : tokens)

{

if (auto rule = std::dynamic\_pointer\_cast<IBackusRule>(token))

rules.push\_back(rule);

}

auto it = rules.begin();

auto end = rules.end();

std::multimap<int, std::pair<std::string, std::vector<std::string>>> errors;

auto res = Controller::Instance()->topRule()->check(errors, it, end);

rules.erase(++std::find\_if(it, rules.end(), [&endOfFileType](const auto& rule) { return rule->type() == endOfFileType; }), rules.end());

end = --rules.end();

std::multimap<int, std::string> errorsMsg;

int lexErr = 0;

int synErr = 0;

int semErr = 0;

tokens.clear();

for (auto rule : rules)

{

tokens.push\_back(std::dynamic\_pointer\_cast<T>(rule));

if (rule->type() == Undefined::Type())

{

res = false;

std::string err;

if (auto erMsg = rule->customData("error"); !erMsg.empty())

{

semErr++;

err = "Semantic error: " + erMsg;

}

else

{

semErr++;

err = std::format("Semantic error: Undefined token: {}", rule->value());

}

errorsMsg.emplace(rule->line(), err);

}

else if (rule->type() == token::Unknown::Type())

{

lexErr++;

res = false;

errorsMsg.emplace(rule->line(), std::format("Lexical error: Unknown token: {}", rule->value()));

}

}

Код який опрацьовує оголошення та використання ідентифікаторів, додає інформацію про ідентифікатор у таблицю ідентифікаторів

identRule->setPostHandler([context](BackusRuleList::iterator&,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

static bool isFirstIdentChecked = !context->IsFirstProgName();

auto isVarBlockChecked = context->IsVarBlockChecked();

auto& identTable = context->IdentTable();

auto identIt = std::prev(it, 1);

if (isVarBlockChecked)

{

if (!identTable.contains((\*identIt)->value()))

{

auto undef = std::make\_shared<Undefined>();

undef->setValue((\*identIt)->value());

undef->setLine((\*identIt)->line());

undef->setCustomData((\*identIt)->customData());

\*identIt = undef;

}

}

else

{

if (isFirstIdentChecked)

{

identTable.insert((\*identIt)->value());

}

else

{

auto progName = std::make\_shared<ProgramName>();

progName->setValue((\*identIt)->value());

progName->setLine((\*identIt)->line());

progName->setCustomData((\*identIt)->customData());

\*identIt = progName;

isFirstIdentChecked = true;

}}

(\*identIt)->setCustomData((\*identIt)->value() + "\_");

});

return identRule;}

* 1. Розробка генератора коду

Синтаксичне дерево само по собі відображає лише структуру програми, але для генерації коду необхідна додаткова інформація про змінні (наприклад, їхні адреси), процедури (адреси, рівні) та мітки. Для цього існують два основні підходи:

1. Зберігання інформації у таблицях генератора коду.
2. Зберігання інформації у відповідних вершинах синтаксичного дерева.

Розглянемо приклад використання таблиць у поєднанні з Лідер-представленням. Лідер-представлення не містить даних про адреси змінних, тому цю інформацію потрібно формувати під час обробки оголошень і записувати в таблиці. Це ж стосується описів масивів, записів тощо.

Таблиці також повинні включати інформацію про процедури: їхні адреси, рівні, модулі, в яких вони описані тощо. При вході у процедуру в таблиці рівнів процедур створюється новий запис — вказівник на таблицю описів. При виході цей вказівник повертається до попереднього значення. У разі використання дерева як проміжного представлення додаткова інформація може зберігатися безпосередньо у вершинах цього дерева.

* + 1. **Опис програми реалізації генератора коду**

### Основні особливості реалізації:

1. **Архітектура:**
   * Використовується патерн Singleton через клас Generator.
   * Базується на шаблоні Visitor: кожен токен або правило має метод genCode().
   * Використовує GeneratorDetails для зберігання налаштувань і допоміжних даних.
2. **Етапи генерації:**
   * Генерація сегмента даних.
   * Генерація сегмента коду.
   * Створення процедур.
   * Завершення програми.
3. **Технічні особливості:**
   * Обчислення виконуються за стековою архітектурою.
   * Підтримується постфіксна форма виразів.
   * Для управління потоком виконання використовується система міток:
     + Унікальні мітки для циклів та умов.
     + Іменовані мітки для операторів GOTO.
4. **Оптимізації:**
   * Мінімізується використання регістрів через стекову модель.
   * Процедури перевикористовуються завдяки механізму реєстрації.
   * Генерація коду оптимізується для простих конструкцій.
5. **Обробка даних:**
   * Підтримуються числові й рядкові типи.
   * Для введення/виведення використовується Windows API.
   * Передбачена система форматування для різних типів даних.
6. **Розширюваність:**
   * Легке додавання нових операторів через систему токенів.
   * Реєстрація користувацьких процедур.
   * Гнучкі налаштування через GeneratorDetails.

Програма має вигляд:

Program

/ \

var statement

Оголошення змінних:

var

/ \

Id var

/ \

Id null

Тіло програми:

Statement

/ \

Statement Оператор

/ \

statement Оператор

Оператор вводу:

Input

/ \

Id null

Оператор виводу:

Output

/ \

Id null

Також оператор виводу може мати за лівого нащадка різні арифметичні вирази, наприклад:

Output

/ \

Add null

/ \

Id num

Умовний оператор (IF() оператор;):

If

/ \

Умова оператор

Умовний оператор (IF() оператор1; else оператор2;):

If

/ \

Умова else

/ \

Оператор1 оператор2

Оператор безумовного переходу:

Goto

/ \

Id null

Оператор циклу for:

For

/ \

(to | downto) оператор

/ \

Оп. Прис. ариф. вир.

Оператор циклу while:

While

/ \

Умова statement

/ \

Statement оператор

/ \

Оператор оператор

Оператор циклу repeat:

Repeat

/ \

Statement умова

/ \

Оператор оператор

Оператор присвоєння:

<-

/ \

Id арифметичний вираз

Арифметичний вираз:

(ADD або MUL)

/ \

Id id

Доданок:

(MUL, DIV або MOD)  
/ \  
множник множник

Множник:

фактор  
/ \

id або number або (арифм. вираз) null

Складений оператор:

compount

/ \

statement null

Дана програма написана мовою С++ з при розробці якої було створено структури BackusRule та BackusRuleItem за допомогою яких можна чітко описати нотатки Бекуса-Наура, які використовуються для семантично-лексичного аналізу написаної програми для заданої мови програмування

auto assingmentRule = BackusRule::MakeRule("AssignmentRule", {

BackusRuleItem({ identRule->type()}, OnlyOne),

BackusRuleItem({Assignment::Type()}, OnlyOne),

BackusRuleItem({ equation->type()}, OnlyOne)

});

auto read = BackusRule::MakeRule("ReadRule", {

BackusRuleItem({ Read::Type()}, OnlyOne),

BackusRuleItem({ LBraket::Type()}, OnlyOne),

BackusRuleItem({ identRule->type()}, OnlyOne),

BackusRuleItem({ RBraket::Type()}, OnlyOne)

});

auto write = BackusRule::MakeRule("WriteRule", {

BackusRuleItem({ Write::Type()}, OnlyOne),

BackusRuleItem({ LBraket::Type()}, OnlyOne | PairStart),

BackusRuleItem({ stringRule->type(), equation->type() }, OnlyOne),

BackusRuleItem({ RBraket::Type()}, OnlyOne | PairEnd)

});

auto codeBlok = BackusRule::MakeRule("CodeBlok", {

BackusRuleItem({ Start::Type()}, OnlyOne),

BackusRuleItem({ operators->type(), operatorsWithSemicolon->type()}, Optional | OneOrMore),

BackusRuleItem({ End::Type()}, OnlyOne)

});

auto topRule = BackusRule::MakeRule("TopRule", {

BackusRuleItem({ Program::Type()}, OnlyOne),

BackusRuleItem({ identRule->type()}, OnlyOne),

BackusRuleItem({ Semicolon::Type()}, OnlyOne),

BackusRuleItem({ Vars::Type()}, OnlyOne),

BackusRuleItem({ varsBlok->type()}, OnlyOne),

BackusRuleItem({ codeBlok->type()}, OnlyOne)

});

Вище наведено приклад опису нотаток Бекуса-Наура за допомогою цих структур. Наприклад topRule це правило, що відповідає за правильну структуру написаної програми, тобто якими лексемами вона повинна починатись та які операції можуть бути використанні всередині виконавчого блоку програми.

Всередині структури BackusRule описаний порядок токенів для певного правила. А в структурі BackusRuleItem описані токени, які при перевірці трактуються програмою як «АБО», тобто повинен бути лише один з описаних токенів. Наприклад для write послідовно необхідний токен Write після якого йде ліва дужка, далі може бути або певний вираз або рядок тексту який необхідно вивести. І закінчується правило токеном правої дужки.

Основна частина програми складається з 3 компонентів: парсера лексем, правил Бекуса-Наура та генератора асемблерного коду. Кожен з цих компонентів працює зі власним інтерфейсом на певному етапі виконання програми.

Кожен токен це окремий клас що наслідує 3 інтерфейси:

* IToken
* IBackusRule
* IGeneratorItem

Наявність наслідування цих інтерфейсів кожним токеном дозволяє без проблем звертатись до кожного віддільного токена на усіх етапах виконання програми

Для процесу парсингу програми використовується інтерфейс IToken. Що дозволяє простіше з точки зору реалізації звертатись до токенів при аналізі вхідної програми.

Правила Бекуса-Наура для своєї роботи використовують інтерфейс IBackusRule. Це дозволяє викликати функцію перевірки check до кожного прописаного у коді правила запису як програми в цілому так і кожного віддільної операції, що спрощує подальший пошук ймовірних помилок у коді програми, яка буде транслюватись у асемблерний код.

Інтерфейс IGeneratorItem використовується генератором асемблерного коду при трансляції вхідної програми. Оскільки кожен токен є віддільним класом, то у ньому була реалізована функція genCode яка використовується генератором, що дозволяє записати необхідний асемблерний код який буде згенерований певним токеном. Наприклад:

Для класу та токену Greate що визначає при порівнянні який елемент більший, функція генерації відповідного коду виглядає наступним чином:

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Greate\_\n";

};

За допомогою функції RegPROC токен за потреби реєструє процедуру у генераторі.

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Greate\_", PrintGreate);

SetRegistered();

}

}

static void PrintGreate(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Greate=========================================================================\n";

out << "Greate\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tjle greate\_false\n";

out << "\tmov " << args.regPrefix << "ax, 1\n";

out << "\tjmp greate\_fin\n";

out << "greate\_false:\n";

out << "\tmov " << args.regPrefix << "ax, 0\n";

out << "greate\_fin:\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Greate\_ ENDP\n";

out << ";============================================================================================\n";

}

Така структура програми дозволяє без проблем аналізувати великі програми, написані на вхідній мові програмування. Також використання правил Бекуса-Наура дозволяє ефективно анадізувати програми великого обсягу.

Генератор у свою чергу буде більш оптимізовано генерувати асемблерний код, створюючи код лише тих операцій, що буди використані у вхідній програмі.

1. Налагодження та тестування розробленого транслятора

Дана програма написана мовою С++ з при розробці якої було створено структури BackusRule та BackusRuleItem за допомогою яких можна чітко описати нотатки Бекуса-Наура, які використовуються для семантично-лексичного аналізу написаної програми для заданої мови програмування

auto assingmentRule = BackusRule::MakeRule("AssignmentRule", {

BackusRuleItem({ identRule->type()}, OnlyOne),

BackusRuleItem({Assignment::Type()}, OnlyOne),

BackusRuleItem({ equation->type()}, OnlyOne)

});

auto read = BackusRule::MakeRule("ReadRule", {

BackusRuleItem({ Read::Type()}, OnlyOne),

BackusRuleItem({ LBraket::Type()}, OnlyOne),

BackusRuleItem({ identRule->type()}, OnlyOne),

BackusRuleItem({ RBraket::Type()}, OnlyOne)

});

auto write = BackusRule::MakeRule("WriteRule", {

BackusRuleItem({ Write::Type()}, OnlyOne),

BackusRuleItem({ LBraket::Type()}, OnlyOne | PairStart),

BackusRuleItem({ stringRule->type(), equation->type() }, OnlyOne),

BackusRuleItem({ RBraket::Type()}, OnlyOne | PairEnd)

});

auto codeBlok = BackusRule::MakeRule("CodeBlok", {

BackusRuleItem({ Start::Type()}, OnlyOne),

BackusRuleItem({ operators->type(), operatorsWithSemicolon->type()}, Optional | OneOrMore),

BackusRuleItem({ End::Type()}, OnlyOne)

});

auto topRule = BackusRule::MakeRule("TopRule", {

BackusRuleItem({ Program::Type()}, OnlyOne),

BackusRuleItem({ identRule->type()}, OnlyOne),

BackusRuleItem({ Semicolon::Type()}, OnlyOne),

BackusRuleItem({ Vars::Type()}, OnlyOne),

BackusRuleItem({ varsBlok->type()}, OnlyOne),

BackusRuleItem({ codeBlok->type()}, OnlyOne)

});

Вище наведено приклад опису нотаток Бекуса-Наура за допомогою цих структур. Наприклад topRule це правило, що відповідає за правильну структуру написаної програми, тобто якими лексемами вона повинна починатись та які операції можуть бути використанні всередині виконавчого блоку програми.

Всередині структури BackusRule описаний порядок токенів для певного правила. А в структурі BackusRuleItem описані токени, які при перевірці трактуються програмою як «АБО», тобто повинен бути лише один з описаних токенів. Наприклад для write послідовно необхідний токен Write після якого йде ліва дужка, далі може бути або певний вираз або рядок тексту який необхідно вивести. І закінчується правило токеном правої дужки.

Основна частина програми складається з 3 компонентів: парсера лексем, правил Бекуса-Наура та генератора асемблерного коду. Кожен з цих компонентів працює зі власним інтерфейсом на певному етапі виконання програми.

Кожен токен це окремий клас що наслідує 3 інтерфейси:

* IToken
* IBackusRule
* IGeneratorItem

Наявність наслідування цих інтерфейсів кожним токеном дозволяє без проблем звертатись до кожного віддільного токена на усіх етапах виконання програми

Для процесу парсингу програми використовується інтерфейс IToken. Що дозволяє простіше з точки зору реалізації звертатись до токенів при аналізі вхідної програми.

Правила Бекуса-Наура для своєї роботи використовують інтерфейс IBackusRule. Це дозволяє викликати функцію перевірки check до кожного прописаного у коді правила запису як програми в цілому так і кожного віддільної операції, що спрощує подальший пошук ймовірних помилок у коді програми, яка буде транслюватись у асемблерний код.

Інтерфейс IGeneratorItem використовується генератором асемблерного коду при трансляції вхідної програми. Оскільки кожен токен є віддільним класом, то у ньому була реалізована функція genCode яка використовується генератором, що дозволяє записати необхідний асемблерний код який буде згенерований певним токеном. Наприклад:

Для класу та токену Greate що визначає при порівнянні який елемент більший, функція генерації відповідного коду виглядає наступним чином:

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Greate\_\n";

};

За допомогою функції RegPROC токен за потреби реєструє процедуру у генераторі.

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Greate\_", PrintGreate);

SetRegistered();

}

}

static void PrintGreate(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Greate=========================================================================\n";

out << "Greate\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tjle greate\_false\n";

out << "\tmov " << args.regPrefix << "ax, 1\n";

out << "\tjmp greate\_fin\n";

out << "greate\_false:\n";

out << "\tmov " << args.regPrefix << "ax, 0\n";

out << "greate\_fin:\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Greate\_ ENDP\n";

out << ";============================================================================================\n";

}

Така структура програми дозволяє без проблем аналізувати великі програми, написані на вхідній мові програмування. Також використання правил Бекуса-Наура дозволяє ефективно анадізувати програми великого обсягу.

Генератор у свою чергу буде більш оптимізовано генерувати асемблерний код, створюючи код лише тих операцій, що буди використані у вхідній програмі.

* 1. Опис інтерфейсу та інструкції користувачу.

Вхідним файлом для даної програми є звичайний текстовий файл з розширенням m20. 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\_m20.exe <ім’я програми>.m20"

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. Виявлення лексичних і 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\*#

STARTPROGRAM

VARI ABLE INT\_4 \_a aaa,\_bbbb,\_xxxx,\_yyyy;

STARTBLOK

WRITdE("Input A: ");

READ(\_aaaa);

WRITE("Input B: ");

READ(\_bbbb);

WRITE("A + B: ");

WRITE(\_aaaa ADD \_bbbb);

WRITE("\nA - B: ");

WRITE(\_aaaa SUB \_bbbb);

WRITE("\nA \* B: ");

WRITE(\_aaaa MUL \_bbbb);

WRITE("\nA / B: ");

WRITE(\_aaaa DIV \_bbbb);

WRITE("\nA % B: ");

WRITE(\_aaaa MOD \_bbbb);

\_xxxx<-(\_aaaa SUB \_bbbb) MUL 10 ADD (\_aaaa ADD \_bbbb) DIV 10;

\_yyyy<-\_xxxx ADD (\_xxxx MOD 10);

WRITE("\nX = (A - B) \* 10 + (A + B) / 10\n");

WRITE(\_xxxx);

WRITE("\nY = X + (X MOD 10)\n");

WRITE(\_yyyy);

ENDBLOK

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

List of errors

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

There are 5 lexical errors.

There are 1 syntax errors.

There are 0 semantic errors.

Line 3: Lexical error: Unknown token: VARI

Line 3: Lexical error: Unknown token: ABLE

Line 3: Lexical error: Unknown token: \_a

Line 3: Lexical error: Unknown token: aaa

Line 3: Syntax error: Expected: Vars before VARI

Line 5: Lexical error: Unknown token: WRITdE

* 1. Перевірка роботи транслятора за допомогою тестових задач.

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

***Текст коректної програми***

#\*Prog1\*#

STARTPROGRAM

VARIABLE INT\_4 \_aaaa,\_bbbb,\_xxxx,\_yyyy;

STARTBLOK

WRITE("Input A: ");

READ(\_aaaa);

WRITE("Input B: ");

READ(\_bbbb);

WRITE("A + B: ");

WRITE(\_aaaa ADD \_bbbb);

WRITE("\nA - B: ");

WRITE(\_aaaa SUB \_bbbb);

WRITE("\nA \* B: ");

WRITE(\_aaaa MUL \_bbbb);

WRITE("\nA / B: ");

WRITE(\_aaaa DIV \_bbbb);

WRITE("\nA % B: ");

WRITE(\_aaaa MOD \_bbbb);

\_xxxx<-(\_aaaa SUB \_bbbb) MUL 10 ADD (\_aaaa ADD \_bbbb) DIV 10;

\_yyyy<-\_xxxx ADD (\_xxxx MOD 10);

WRITE("\nX = (A - B) \* 10 + (A + B) / 10\n");

WRITE(\_xxxx);

WRITE("\nY = X + (X MOD 10)\n");

WRITE(\_yyyy);

ENDBLOK

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:

Зображення, що містить текст, знімок екрана, Шрифт, дизайн

Автоматично згенерований опис

Рис. 4.1 Результат виконання коректної програми

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

Тестова програма №1

***Текст програми***

#\*Prog1\*#

STARTPROGRAM

VARIABLE INT\_4 \_aaaa,\_bbbb,\_xxxx,\_yyyy;

STARTBLOK

WRITE("Input A: ");

READ(\_aaaa);

WRITE("Input B: ");

READ(\_bbbb);

WRITE("A + B: ");

WRITE(\_aaaa ADD \_bbbb);

WRITE("\nA - B: ");

WRITE(\_aaaa SUB \_bbbb);

WRITE("\nA \* B: ");

WRITE(\_aaaa MUL \_bbbb);

WRITE("\nA / B: ");

WRITE(\_aaaa DIV \_bbbb);

WRITE("\nA % B: ");

WRITE(\_aaaa MOD \_bbbb);

\_xxxx<-(\_aaaa SUB \_bbbb) MUL 10 ADD (\_aaaa ADD \_bbbb) DIV 10;

\_yyyy<-\_xxxx ADD (\_xxxx MOD 10);

WRITE("\nX = (A - B) \* 10 + (A + B) / 10\n");

WRITE(\_xxxx);

WRITE("\nY = X + (X MOD 10)\n");

WRITE(\_yyyy);

ENDBLOK

***Результат виконання***

Зображення, що містить текст, знімок екрана, Шрифт, дизайн

Автоматично згенерований опис

Рис. 4.2 Результат виконання тестової програми №1

Тестова програма №2

***Текст програми***

#\*Prog2\*#

STARTPROGRAM

VARIABLE INT\_4 \_aaaa,\_bbbb,\_cccc;

STARTBLOK

WRITE("Input A: ");

READ(\_aaaa);

WRITE("Input B: ");

READ(\_bbbb);

WRITE("Input C: ");

READ(\_cccc);

IF(\_aaaa GT \_bbbb)

STARTBLOK

IF(\_aaaa GT \_cccc)

STARTBLOK

GOTO \_temp;

ENDBLOK

ELSE

STARTBLOK

WRITE(\_cccc);

GOTO \_outi;

\_temp:

WRITE(\_aaaa);

GOTO \_outi;

ENDBLOK

ENDBLOK

IF(\_bbbb LT \_cccc)

STARTBLOK

WRITE(\_cccc);

ENDBLOK

ELSE

STARTBLOK

WRITE(\_bbbb);

ENDBLOK

\_outi:

WRITE("\n");

IF((\_aaaa EQ \_bbbb) & (\_aaaa EQ \_cccc) & (\_bbbb EQ \_cccc))

STARTBLOK

WRITE(1);

ENDBLOK

ELSE

STARTBLOK

WRITE(0);

ENDBLOK

WRITE("\n");

IF((\_aaaa LT 0) | (\_bbbb LT 0) | (\_cccc LT 0))

STARTBLOK

WRITE(- 1);

ENDBLOK

ELSE

STARTBLOK

WRITE(0);

ENDBLOK

WRITE("\n");

IF(!(\_aaaa LT (\_bbbb ADD \_cccc)))

STARTBLOK

WRITE(10);

ENDBLOK

ELSE

STARTBLOK

WRITE(0);

ENDBLOK

ENDBLOK

***Результат виконання***

Зображення, що містить знімок екрана, текст, Шрифт, дизайн

Автоматично згенерований опис

Рис. 4.3 Результат виконання тестової програми №2

Тестова програма №3

***Текст програми***

#\*Prog3\*#

STARTPROGRAM

VARIABLE INT\_4 \_aaaa,\_aaa2,\_bbbb,\_xxxx,\_ccc1,\_ccc2;

STARTBLOK

WRITE("Input A: ");

READ(\_aaaa);

WRITE("Input B: ");

READ(\_bbbb);

WRITE("FOR TO DO");

FOR \_aaa2<-\_aaaa TO \_bbbb DO

STARTBLOK

WRITE("\n");

WRITE(\_aaa2 MUL \_aaa2);

ENDBLOK

WRITE("\nFOR DOWNTO DO");

FOR \_aaa2<-\_bbbb DOWNTO \_aaaa DO

STARTBLOK

WRITE("\n");

WRITE(\_aaa2 MUL \_aaa2);

ENDBLOK

WRITE("\nWHILE A MUL B: ");

\_xxxx<-0;

\_ccc1<-0;

WHILE(\_ccc1 LT \_aaaa)

STARTBLOK

\_ccc2<-0;

WHILE (\_ccc2 LT \_bbbb)

STARTBLOK

\_xxxx<-\_xxxx ADD 1;

\_ccc2<-\_ccc2 ADD 1;

ENDBLOK

\_ccc1<-\_ccc1 ADD 1;

ENDBLOK

WRITE(\_xxxx);

WRITE("\nREPEAT UNTIL A MUL B: ");

\_xxxx<-0;

\_ccc1<-1;

REPEAT

\_ccc2<-1;

REPEAT

\_xxxx<-\_xxxx ADD 1;

\_ccc2<-\_ccc2 ADD 1;

UNTIL(!(\_ccc2 GT \_bbbb))

\_ccc1<-\_ccc1 ADD 1;

UNTIL(!(\_ccc1 GT \_aaaa))

WRITE(\_xxxx);

ENDBLOK

***Результат виконання***

Зображення, що містить текст, знімок екрана, Шрифт

Автоматично згенерований опис

Рис. 4.4 Результат виконання тестової програми №3

Висновки

В процесі виконання курсового проекту було виконано наступне:

1. Складено формальний опис мови програмування m20, в термінах розширеної нотації Бекуса-Наура, виділено усі термінальні символи та ключові слова.

2. Створено компілятор мови програмування m20, 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 низхідного 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 m20. 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 m20 були успішно виявлені і відповідно оброблені.

В результаті виконання даної курсового проекту було засвоєно методи розробки та реалізації компонент систем програмування.

Список літературних джерел

1. Основи проектування трансляторів: Конспект лекцій : [Електронний ресурс] : навч. посіб. для студ. спеціальності 123 - «Комп’ютерна інженерія» / О. І. Марченко ; КПІ ім. Ігоря Сікорського. - Київ: КПІ ім. Ігоря Сікорського, 2021. - 108 с.
2. Формальні мови, граматики та автомати: Навчальний посібник / Гавриленко С.Ю. - Харків: НТУ «ХПІ», 2021. - 133 с.
3. Сопронюк Т.М. Системне програмування. Частина І. Елементи теорії формальних мов: Навчальний посібник у двох частинах. - Чернівці: ЧНУ, 2008. - 84 c.
4. Сопронюк Т.М. Системне програмування. Частина ІІ. Елементи теорії компіляції: Навчальний посібник у двох частинах. - Чернівці: ЧНУ, 2008. - 84 c.
5. Alfred V. Aho, Monica S. Lam, Ravi Seth, Jeffrey D. Ullma. Compilers, principles, techniques, and tools, Second Edition, New York, 2007. - 1038 c.
6. Системне програмування (курсовий проект) [Електронний ресурс] - Режим доступу до ресурсу: <https://vns.lpnu.ua/course/view.php?id=11685>.

MIT OpenCourseWare. Computer Language Engineering [Електронний ресурс] - Режим доступу до ресурсу: [https://ocw.mit.edu/courses/6-035-computer- language-engineering-spring-2010](https://ocw.mit.edu/courses/6-035-computer-language-engineering-spring-2010).

Додатки

**Додаток А (Таблиці лексем)**

Програма 1

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

| # | SYMBOL | TYPE | VALUE | LINE |

|=====|==============|================|=====================================|======|

| 1 | #\* | LComment | #\* | 1 |

| 2 | | Comment | Prog1 | 1 |

| 3 | \*# | RComment | \*# | 1 |

| 4 | STARTPROGRAM | Program | STARTPROGRAM | 2 |

| 5 | VARIABLE | Vars | VARIABLE | 3 |

| 6 | INT\_4 | VarType | INT\_4 | 3 |

| 7 | | Identifier | \_aaaa | 3 |

| 8 | , | Comma | , | 3 |

| 9 | | Identifier | \_bbbb | 3 |

| 10 | , | Comma | , | 3 |

| 11 | | Identifier | \_xxxx | 3 |

| 12 | , | Comma | , | 3 |

| 13 | | Identifier | \_yyyy | 3 |

| 14 | ; | Semicolon | ; | 3 |

| 15 | STARTBLOK | Start | STARTBLOK | 4 |

| 16 | WRITE | Write | WRITE | 5 |

| 17 | ( | LBraket | ( | 5 |

| 18 | " | Quotes | " | 5 |

| 19 | | String | Input A: | 5 |

| 20 | " | Quotes | " | 5 |

| 21 | ) | RBraket | ) | 5 |

| 22 | ; | Semicolon | ; | 5 |

| 23 | READ | Read | READ | 6 |

| 24 | ( | LBraket | ( | 6 |

| 25 | | Identifier | \_aaaa | 6 |

| 26 | ) | RBraket | ) | 6 |

| 27 | ; | Semicolon | ; | 6 |

| 28 | WRITE | Write | WRITE | 7 |

| 29 | ( | LBraket | ( | 7 |

| 30 | " | Quotes | " | 7 |

| 31 | | String | Input B: | 7 |

| 32 | " | Quotes | " | 7 |

| 33 | ) | RBraket | ) | 7 |

| 34 | ; | Semicolon | ; | 7 |

| 35 | READ | Read | READ | 8 |

| 36 | ( | LBraket | ( | 8 |

| 37 | | Identifier | \_bbbb | 8 |

| 38 | ) | RBraket | ) | 8 |

| 39 | ; | Semicolon | ; | 8 |

| 40 | WRITE | Write | WRITE | 9 |

| 41 | ( | LBraket | ( | 9 |

| 42 | " | Quotes | " | 9 |

| 43 | | String | A + B: | 9 |

| 44 | " | Quotes | " | 9 |

| 45 | ) | RBraket | ) | 9 |

| 46 | ; | Semicolon | ; | 9 |

| 47 | WRITE | Write | WRITE | 10 |

| 48 | ( | LBraket | ( | 10 |

| 49 | | Identifier | \_aaaa | 10 |

| 50 | ADD | Addition | ADD | 10 |

| 51 | | Identifier | \_bbbb | 10 |

| 52 | ) | RBraket | ) | 10 |

| 53 | ; | Semicolon | ; | 10 |

| 54 | WRITE | Write | WRITE | 11 |

| 55 | ( | LBraket | ( | 11 |

| 56 | " | Quotes | " | 11 |

| 57 | | String | \nA - B: | 11 |

| 58 | " | Quotes | " | 11 |

| 59 | ) | RBraket | ) | 11 |

| 60 | ; | Semicolon | ; | 11 |

| 61 | WRITE | Write | WRITE | 12 |

| 62 | ( | LBraket | ( | 12 |

| 63 | | Identifier | \_aaaa | 12 |

| 64 | SUB | Subtraction | SUB | 12 |

| 65 | | Identifier | \_bbbb | 12 |

| 66 | ) | RBraket | ) | 12 |

| 67 | ; | Semicolon | ; | 12 |

| 68 | WRITE | Write | WRITE | 13 |

| 69 | ( | LBraket | ( | 13 |

| 70 | " | Quotes | " | 13 |

| 71 | | String | \nA \* B: | 13 |

| 72 | " | Quotes | " | 13 |

| 73 | ) | RBraket | ) | 13 |

| 74 | ; | Semicolon | ; | 13 |

| 75 | WRITE | Write | WRITE | 14 |

| 76 | ( | LBraket | ( | 14 |

| 77 | | Identifier | \_aaaa | 14 |

| 78 | MUL | Multiplication | MUL | 14 |

| 79 | | Identifier | \_bbbb | 14 |

| 80 | ) | RBraket | ) | 14 |

| 81 | ; | Semicolon | ; | 14 |

| 82 | WRITE | Write | WRITE | 15 |

| 83 | ( | LBraket | ( | 15 |

| 84 | " | Quotes | " | 15 |

| 85 | | String | \nA / B: | 15 |

| 86 | " | Quotes | " | 15 |

| 87 | ) | RBraket | ) | 15 |

| 88 | ; | Semicolon | ; | 15 |

| 89 | WRITE | Write | WRITE | 16 |

| 90 | ( | LBraket | ( | 16 |

| 91 | | Identifier | \_aaaa | 16 |

| 92 | DIV | Division | DIV | 16 |

| 93 | | Identifier | \_bbbb | 16 |

| 94 | ) | RBraket | ) | 16 |

| 95 | ; | Semicolon | ; | 16 |

| 96 | WRITE | Write | WRITE | 17 |

| 97 | ( | LBraket | ( | 17 |

| 98 | " | Quotes | " | 17 |

| 99 | | String | \nA % B: | 17 |

| 100 | " | Quotes | " | 17 |

| 101 | ) | RBraket | ) | 17 |

| 102 | ; | Semicolon | ; | 17 |

| 103 | WRITE | Write | WRITE | 18 |

| 104 | ( | LBraket | ( | 18 |

| 105 | | Identifier | \_aaaa | 18 |

| 106 | MOD | Mod | MOD | 18 |

| 107 | | Identifier | \_bbbb | 18 |

| 108 | ) | RBraket | ) | 18 |

| 109 | ; | Semicolon | ; | 18 |

| 110 | | Identifier | \_xxxx | 19 |

| 111 | <- | Assignment | <- | 19 |

| 112 | ( | LBraket | ( | 19 |

| 113 | | Identifier | \_aaaa | 19 |

| 114 | SUB | Subtraction | SUB | 19 |

| 115 | | Identifier | \_bbbb | 19 |

| 116 | ) | RBraket | ) | 19 |

| 117 | MUL | Multiplication | MUL | 19 |

| 118 | | Number | 10 | 19 |

| 119 | ADD | Addition | ADD | 19 |

| 120 | ( | LBraket | ( | 19 |

| 121 | | Identifier | \_aaaa | 19 |

| 122 | ADD | Addition | ADD | 19 |

| 123 | | Identifier | \_bbbb | 19 |

| 124 | ) | RBraket | ) | 19 |

| 125 | DIV | Division | DIV | 19 |

| 126 | | Number | 10 | 19 |

| 127 | ; | Semicolon | ; | 19 |

| 128 | | Identifier | \_yyyy | 20 |

| 129 | <- | Assignment | <- | 20 |

| 130 | | Identifier | \_xxxx | 20 |

| 131 | ADD | Addition | ADD | 20 |

| 132 | ( | LBraket | ( | 20 |

| 133 | | Identifier | \_xxxx | 20 |

| 134 | MOD | Mod | MOD | 20 |

| 135 | | Number | 10 | 20 |

| 136 | ) | RBraket | ) | 20 |

| 137 | ; | Semicolon | ; | 20 |

| 138 | WRITE | Write | WRITE | 21 |

| 139 | ( | LBraket | ( | 21 |

| 140 | " | Quotes | " | 21 |

| 141 | | String | \nX = (A - B) \* 10 + (A + B) / 10\n | 21 |

| 142 | " | Quotes | " | 21 |

| 143 | ) | RBraket | ) | 21 |

| 144 | ; | Semicolon | ; | 21 |

| 145 | WRITE | Write | WRITE | 22 |

| 146 | ( | LBraket | ( | 22 |

| 147 | | Identifier | \_xxxx | 22 |

| 148 | ) | RBraket | ) | 22 |

| 149 | ; | Semicolon | ; | 22 |

| 150 | WRITE | Write | WRITE | 23 |

| 151 | ( | LBraket | ( | 23 |

| 152 | " | Quotes | " | 23 |

| 153 | | String | \nY = X + (X MOD 10)\n | 23 |

| 154 | " | Quotes | " | 23 |

| 155 | ) | RBraket | ) | 23 |

| 156 | ; | Semicolon | ; | 23 |

| 157 | WRITE | Write | WRITE | 24 |

| 158 | ( | LBraket | ( | 24 |

| 159 | | Identifier | \_yyyy | 24 |

| 160 | ) | RBraket | ) | 24 |

| 161 | ; | Semicolon | ; | 24 |

| 162 | ENDBLOK | End | ENDBLOK | 25 |

| 163 | | EndOfFile | | -1 |

Програма 2

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

| # | SYMBOL | TYPE | VALUE | LINE |

|=====|==============|============|===============|======|

| 1 | #\* | LComment | #\* | 1 |

| 2 | | Comment | Prog2 | 1 |

| 3 | \*# | RComment | \*# | 1 |

| 4 | STARTPROGRAM | Program | STARTPROGRAM | 2 |

| 5 | VARIABLE | Vars | VARIABLE | 3 |

| 6 | INT\_4 | VarType | INT\_4 | 3 |

| 7 | | Identifier | \_aaaa | 3 |

| 8 | , | Comma | , | 3 |

| 9 | | Identifier | \_bbbb | 3 |

| 10 | , | Comma | , | 3 |

| 11 | | Identifier | \_cccc | 3 |

| 12 | ; | Semicolon | ; | 3 |

| 13 | STARTBLOK | Start | STARTBLOK | 4 |

| 14 | WRITE | Write | WRITE | 5 |

| 15 | ( | LBraket | ( | 5 |

| 16 | " | Quotes | " | 5 |

| 17 | | String | Input A: | 5 |

| 18 | " | Quotes | " | 5 |

| 19 | ) | RBraket | ) | 5 |

| 20 | ; | Semicolon | ; | 5 |

| 21 | READ | Read | READ | 6 |

| 22 | ( | LBraket | ( | 6 |

| 23 | | Identifier | \_aaaa | 6 |

| 24 | ) | RBraket | ) | 6 |

| 25 | ; | Semicolon | ; | 6 |

| 26 | WRITE | Write | WRITE | 7 |

| 27 | ( | LBraket | ( | 7 |

| 28 | " | Quotes | " | 7 |

| 29 | | String | Input B: | 7 |

| 30 | " | Quotes | " | 7 |

| 31 | ) | RBraket | ) | 7 |

| 32 | ; | Semicolon | ; | 7 |

| 33 | READ | Read | READ | 8 |

| 34 | ( | LBraket | ( | 8 |

| 35 | | Identifier | \_bbbb | 8 |

| 36 | ) | RBraket | ) | 8 |

| 37 | ; | Semicolon | ; | 8 |

| 38 | WRITE | Write | WRITE | 9 |

| 39 | ( | LBraket | ( | 9 |

| 40 | " | Quotes | " | 9 |

| 41 | | String | Input C: | 9 |

| 42 | " | Quotes | " | 9 |

| 43 | ) | RBraket | ) | 9 |

| 44 | ; | Semicolon | ; | 9 |

| 45 | READ | Read | READ | 10 |

| 46 | ( | LBraket | ( | 10 |

| 47 | | Identifier | \_cccc | 10 |

| 48 | ) | RBraket | ) | 10 |

| 49 | ; | Semicolon | ; | 10 |

| 50 | IF | If | IF | 11 |

| 51 | ( | LBraket | ( | 11 |

| 52 | | Identifier | \_aaaa | 11 |

| 53 | GT | Greate | GT | 11 |

| 54 | | Identifier | \_bbbb | 11 |

| 55 | ) | RBraket | ) | 11 |

| 56 | STARTBLOK | Start | STARTBLOK | 12 |

| 57 | IF | If | IF | 13 |

| 58 | ( | LBraket | ( | 13 |

| 59 | | Identifier | \_aaaa | 13 |

| 60 | GT | Greate | GT | 13 |

| 61 | | Identifier | \_cccc | 13 |

| 62 | ) | RBraket | ) | 13 |

| 63 | STARTBLOK | Start | STARTBLOK | 14 |

| 64 | GOTO | Goto | GOTO | 15 |

| 65 | | Identifier | \_temp | 15 |

| 66 | ; | Semicolon | ; | 15 |

| 67 | ENDBLOK | End | ENDBLOK | 16 |

| 68 | ELSE | Else | ELSE | 17 |

| 69 | STARTBLOK | Start | STARTBLOK | 18 |

| 70 | WRITE | Write | WRITE | 19 |

| 71 | ( | LBraket | ( | 19 |

| 72 | | Identifier | \_cccc | 19 |

| 73 | ) | RBraket | ) | 19 |

| 74 | ; | Semicolon | ; | 19 |

| 75 | GOTO | Goto | GOTO | 20 |

| 76 | | Identifier | \_outi | 20 |

| 77 | ; | Semicolon | ; | 20 |

| 78 | | Identifier | \_temp | 21 |

| 79 | : | Colon | : | 21 |

| 80 | WRITE | Write | WRITE | 22 |

| 81 | ( | LBraket | ( | 22 |

| 82 | | Identifier | \_aaaa | 22 |

| 83 | ) | RBraket | ) | 22 |

| 84 | ; | Semicolon | ; | 22 |

| 85 | GOTO | Goto | GOTO | 23 |

| 86 | | Identifier | \_outi | 23 |

| 87 | ; | Semicolon | ; | 23 |

| 88 | ENDBLOK | End | ENDBLOK | 24 |

| 89 | ENDBLOK | End | ENDBLOK | 25 |

| 90 | IF | If | IF | 26 |

| 91 | ( | LBraket | ( | 26 |

| 92 | | Identifier | \_bbbb | 26 |

| 93 | LT | Less | LT | 26 |

| 94 | | Identifier | \_cccc | 26 |

| 95 | ) | RBraket | ) | 26 |

| 96 | STARTBLOK | Start | STARTBLOK | 27 |

| 97 | WRITE | Write | WRITE | 28 |

| 98 | ( | LBraket | ( | 28 |

| 99 | | Identifier | \_cccc | 28 |

| 100 | ) | RBraket | ) | 28 |

| 101 | ; | Semicolon | ; | 28 |

| 102 | ENDBLOK | End | ENDBLOK | 29 |

| 103 | ELSE | Else | ELSE | 30 |

| 104 | STARTBLOK | Start | STARTBLOK | 31 |

| 105 | WRITE | Write | WRITE | 32 |

| 106 | ( | LBraket | ( | 32 |

| 107 | | Identifier | \_bbbb | 32 |

| 108 | ) | RBraket | ) | 32 |

| 109 | ; | Semicolon | ; | 32 |

| 110 | ENDBLOK | End | ENDBLOK | 33 |

| 111 | | Identifier | \_outi | 34 |

| 112 | : | Colon | : | 34 |

| 113 | WRITE | Write | WRITE | 35 |

| 114 | ( | LBraket | ( | 35 |

| 115 | " | Quotes | " | 35 |

| 116 | | String | \n | 35 |

| 117 | " | Quotes | " | 35 |

| 118 | ) | RBraket | ) | 35 |

| 119 | ; | Semicolon | ; | 35 |

| 120 | IF | If | IF | 36 |

| 121 | ( | LBraket | ( | 36 |

| 122 | ( | LBraket | ( | 36 |

| 123 | | Identifier | \_aaaa | 36 |

| 124 | EQ | Equal | EQ | 36 |

| 125 | | Identifier | \_bbbb | 36 |

| 126 | ) | RBraket | ) | 36 |

| 127 | & | And | & | 36 |

| 128 | ( | LBraket | ( | 36 |

| 129 | | Identifier | \_aaaa | 36 |

| 130 | EQ | Equal | EQ | 36 |

| 131 | | Identifier | \_cccc | 36 |

| 132 | ) | RBraket | ) | 36 |

| 133 | & | And | & | 36 |

| 134 | ( | LBraket | ( | 36 |

| 135 | | Identifier | \_bbbb | 36 |

| 136 | EQ | Equal | EQ | 36 |

| 137 | | Identifier | \_cccc | 36 |

| 138 | ) | RBraket | ) | 36 |

| 139 | ) | RBraket | ) | 36 |

| 140 | STARTBLOK | Start | STARTBLOK | 37 |

| 141 | WRITE | Write | WRITE | 38 |

| 142 | ( | LBraket | ( | 38 |

| 143 | | Number | 1 | 38 |

| 144 | ) | RBraket | ) | 38 |

| 145 | ; | Semicolon | ; | 38 |

| 146 | ENDBLOK | End | ENDBLOK | 39 |

| 147 | ELSE | Else | ELSE | 40 |

| 148 | STARTBLOK | Start | STARTBLOK | 41 |

| 149 | WRITE | Write | WRITE | 42 |

| 150 | ( | LBraket | ( | 42 |

| 151 | | Number | 0 | 42 |

| 152 | ) | RBraket | ) | 42 |

| 153 | ; | Semicolon | ; | 42 |

| 154 | ENDBLOK | End | ENDBLOK | 43 |

| 155 | WRITE | Write | WRITE | 44 |

| 156 | ( | LBraket | ( | 44 |

| 157 | " | Quotes | " | 44 |

| 158 | | String | \n | 44 |

| 159 | " | Quotes | " | 44 |

| 160 | ) | RBraket | ) | 44 |

| 161 | ; | Semicolon | ; | 44 |

| 162 | IF | If | IF | 45 |

| 163 | ( | LBraket | ( | 45 |

| 164 | ( | LBraket | ( | 45 |

| 165 | | Identifier | \_aaaa | 45 |

| 166 | LT | Less | LT | 45 |

| 167 | | Number | 0 | 45 |

| 168 | ) | RBraket | ) | 45 |

| 169 | | | Or | | | 45 |

| 170 | ( | LBraket | ( | 45 |

| 171 | | Identifier | \_bbbb | 45 |

| 172 | LT | Less | LT | 45 |

| 173 | | Number | 0 | 45 |

| 174 | ) | RBraket | ) | 45 |

| 175 | | | Or | | | 45 |

| 176 | ( | LBraket | ( | 45 |

| 177 | | Identifier | \_cccc | 45 |

| 178 | LT | Less | LT | 45 |

| 179 | | Number | 0 | 45 |

| 180 | ) | RBraket | ) | 45 |

| 181 | ) | RBraket | ) | 45 |

| 182 | STARTBLOK | Start | STARTBLOK | 46 |

| 183 | WRITE | Write | WRITE | 47 |

| 184 | ( | LBraket | ( | 47 |

| 185 | - | Minus | - | 47 |

| 186 | | Number | 1 | 47 |

| 187 | ) | RBraket | ) | 47 |

| 188 | ; | Semicolon | ; | 47 |

| 189 | ENDBLOK | End | ENDBLOK | 48 |

| 190 | ELSE | Else | ELSE | 49 |

| 191 | STARTBLOK | Start | STARTBLOK | 50 |

| 192 | WRITE | Write | WRITE | 51 |

| 193 | ( | LBraket | ( | 51 |

| 194 | | Number | 0 | 51 |

| 195 | ) | RBraket | ) | 51 |

| 196 | ; | Semicolon | ; | 51 |

| 197 | ENDBLOK | End | ENDBLOK | 52 |

| 198 | WRITE | Write | WRITE | 53 |

| 199 | ( | LBraket | ( | 53 |

| 200 | " | Quotes | " | 53 |

| 201 | | String | \n | 53 |

| 202 | " | Quotes | " | 53 |

| 203 | ) | RBraket | ) | 53 |

| 204 | ; | Semicolon | ; | 53 |

| 205 | IF | If | IF | 54 |

| 206 | ( | LBraket | ( | 54 |

| 207 | ! | Not | ! | 54 |

| 208 | ( | LBraket | ( | 54 |

| 209 | | Identifier | \_aaaa | 54 |

| 210 | LT | Less | LT | 54 |

| 211 | ( | LBraket | ( | 54 |

| 212 | | Identifier | \_bbbb | 54 |

| 213 | ADD | Addition | ADD | 54 |

| 214 | | Identifier | \_cccc | 54 |

| 215 | ) | RBraket | ) | 54 |

| 216 | ) | RBraket | ) | 54 |

| 217 | ) | RBraket | ) | 54 |

| 218 | STARTBLOK | Start | STARTBLOK | 55 |

| 219 | WRITE | Write | WRITE | 56 |

| 220 | ( | LBraket | ( | 56 |

| 221 | | Number | 10 | 56 |

| 222 | ) | RBraket | ) | 56 |

| 223 | ; | Semicolon | ; | 56 |

| 224 | ENDBLOK | End | ENDBLOK | 57 |

| 225 | ELSE | Else | ELSE | 58 |

| 226 | STARTBLOK | Start | STARTBLOK | 59 |

| 227 | WRITE | Write | WRITE | 60 |

| 228 | ( | LBraket | ( | 60 |

| 229 | | Number | 0 | 60 |

| 230 | ) | RBraket | ) | 60 |

| 231 | ; | Semicolon | ; | 60 |

| 232 | ENDBLOK | End | ENDBLOK | 61 |

| 233 | ENDBLOK | End | ENDBLOK | 62 |

| 234 | | EndOfFile | | -1 |

**Додаток Б (Лістинги основного програмного коду)**

Main.cpp

#include "stdafx.h"

#include "Controller.h"

#include "Core/Parser/TokenRegister.h"

#include "Core/Parser/TokenParser.h"

#include "Core/Generator/Generator.h"

int main(int argc, std::string\* argv)

{

try

{

std::filesystem::path file;

const std::string extention = ".m20";

const std::string longLine = "~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~";

std::cout << longLine << std::endl;

std::cout << "TRANSLATOR (" << extention << "->ASSEMBLER)" << std::endl;

std::cout << longLine << std::endl;

if (argc != 2)

{

printf("Input file name\n");

std::cin >> file;

}

else

{

file = argv->c\_str();

}

Init();

if (file.extension() != extention)

{

std::cout << longLine << std::endl;

std::cout << "Wrong file extension" << std::endl;

system("pause");

return 0;

}

std::string fileName = file.replace\_extension("").string();

std::string errorFileName = fileName + "\_errors.txt";

std::string lexemsFileName = fileName + "\_lexems.txt";

std::string tokensFileName = fileName + "\_tokens.txt";

std::string asmFileName = fileName + ".asm";

std::cout << longLine << std::endl;

std::cout << "Breaking into lexems are starting..." << std::endl;

std::fstream inputFile{ fileName + extention, std::ios::in };

auto tokens = TokenParser::Instance()->tokenize(inputFile);

inputFile.close();

std::cout << "Breaking into lexems completed. There are " << tokens.size() << " lexems" << std::endl;

std::fstream lexemsFile(lexemsFileName, std::ios::out);

TokenParser::PrintTokens(lexemsFile, tokens);

lexemsFile.close();

std::cout << "Report file: " << lexemsFileName << std::endl;

std::cout << longLine << std::endl;

std::cout << "Error checking are starting... " << std::endl;

std::fstream errorFile(errorFileName, std::ios::out);

auto semanticCheckRes = CheckSemantic(errorFile, tokens);

errorFile.close();

if (semanticCheckRes)

{

std::cout << "There are no errors in the file" << std::endl;

std::cout << longLine << std::endl;

}

else

{

std::cout << "There are errors in the file. Check " << errorFileName << " for more information" << std::endl;

std::cout << longLine << std::endl;

}

std::fstream tokensFile(tokensFileName, std::ios::out);

TokenParser::PrintTokens(tokensFile, tokens);

tokensFile.close();

std::cout << "There are " << tokens.size() << " tokens." << std::endl;

std::cout << "Report file: " << tokensFileName << std::endl;

if (semanticCheckRes)

{

std::cout << longLine << std::endl;

std::cout << "Code generation is starting..." << std::endl;

std::fstream asmFile(asmFileName, std::ios::out);

Generator::Instance()->generateCode(asmFile, tokens);

asmFile.close();

if (std::filesystem::is\_directory("masm32"))

{

std::cout << "Code generation is completed" << std::endl;

std::cout << longLine << std::endl;

system(std::string("masm32\\bin\\ml /c /coff " + fileName + ".asm").c\_str());

system(std::string("masm32\\bin\\Link /SUBSYSTEM:WINDOWS " + fileName + ".obj").c\_str());

}

else

{

std::cout << "WARNING!" << std::endl;

std::cout << "Can't compile asm file, because masm32 doesn't exist" << std::endl;

}

}

}

catch (const std::exception& ex)

{

std::cout << "Error: " << ex.what() << std::endl;

}

catch (...)

{

std::cout << "Unknown internal error. Better call Saul" << std::endl;

}

system("pause");

return 0;

}

BackusRule.h

#pragma once

#include "stdafx.h"

#include "BackusRule.h"

std::shared\_ptr<IBackusRule> BackusRule::MakeRule(std::string name, std::list<BackusRuleItem> items)

{

struct EnableMakeShared : public BackusRule { EnableMakeShared(const std::string& name, const std::list<BackusRuleItem>& items) : BackusRule(name, items) {} };

return std::make\_shared<EnableMakeShared>(name, items);

}

bool BackusRule::check(std::multimap<int, std::pair<std::string, std::vector<std::string>>>& errorsInfo,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end)

{

bool res = true;

bool pairItem = false;

auto ruleBegin = it;

for (auto item = m\_backusItem.begin(); item != m\_backusItem.end(); ++item)

{

if (it == end || !pairItem && HasFlag(item->policy(), RuleCountPolicy::PairEnd))

{

if (!HasFlag(item->policy(), RuleCountPolicy::Optional) || item != m\_backusItem.end())

{

std::vector<std::string> types;

for (const auto& rule : item->rules())

types.push\_back(rule->type());

errorsInfo.emplace((\*it)->line(), std::make\_pair((\*it)->value(), types));

res = false;

}

break;

}

if (pairItem && HasFlag(item->policy(), RuleCountPolicy::PairEnd) || !HasFlag(item->policy(), RuleCountPolicy::PairEnd))

{

bool resItem = true;

auto startIt = it;

if (HasFlag(item->policy(), RuleCountPolicy::Several))

resItem = oneOrMoreCheck(errorsInfo, it, end, \*item);

else

resItem = checkItem(errorsInfo, it, end, \*item);

if (!resItem && (!HasFlag(item->policy(), RuleCountPolicy::Optional) || startIt != it))

{

res &= resItem;

break;

}

if (resItem && HasFlag(item->policy(), RuleCountPolicy::PairStart))

{

pairItem = true;

}

if (resItem && pairItem && HasFlag(item->policy(), RuleCountPolicy::PairEnd))

{

pairItem = false;

}

}

}

if (res && m\_handler)

m\_handler(ruleBegin, it, end);

return res;

}

bool BackusRule::oneOrMoreCheck(std::multimap<int, std::pair<std::string, std::vector<std::string>>>& errorsInfo,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end,

const BackusRuleItem& item) const

{

bool res = true;

bool resItem = true;

while (resItem && it != end && HasFlag(item.policy(), RuleCountPolicy::Several))

{

auto startIt = it;

res &= resItem;

resItem = checkItem(errorsInfo, it, end, item);

if (!resItem && startIt != it)

res = false;

}

return res;

}

bool BackusRule::checkItem(std::multimap<int, std::pair<std::string, std::vector<std::string>>>& errorsInfo,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end,

const BackusRuleItem& item) const

{

bool res = false;

std::vector<std::string> types;

auto startIt = it;

auto maxIt = it;

if (it != end)

{

std::multimap<int, std::pair<std::string, std::vector<std::string>>> errors;

for (auto rule : item.rules())

{

types.push\_back(rule->type());

if (!res && startIt == it)

{

res = rule->check(errors, it, end);

}

if (res)

{

break;

}

else if (!res && startIt != it)

{

if(std::distance(maxIt, end) > std::distance(it, end))

maxIt = it;

it = startIt;

errorsInfo.insert(errors.begin(), errors.end());

}

}

}

if (std::distance(maxIt, end) < std::distance(it, end))

it = maxIt;

if (!res)

errorsInfo.emplace((\*startIt)->line(), std::make\_pair((\*it)->value(), types));

else

errorsInfo.clear();

return res;

}

bool BackusRule::HasFlag(RuleCountPolicy policy, RuleCountPolicy flag)

{

return (policy & flag) == flag;

}

BackusRuleBase.h

#pragma once

#include "stdafx.h"

#include "Core/Backus/IBackusRule.h"

template <class T>

class BackusRuleBase : public IBackusRule

{

public:

bool check(std::multimap<int, std::pair<std::string, std::vector<std::string>>>& errorsInfo,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end) final

{

auto res = type() == (\*it)->type();

if (res)

it++;

return res;

}

void setPostHandler(const std::function<void(std::list<std::shared\_ptr<IBackusRule>>::iterator& ruleBegin,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end)>& handler) final { };

};

BackusRuleItem.h

#pragma once

#include "stdafx.h"

#include "Core/Backus/IBackusRule.h"

#include "BackusRuleStorage.h"

#include "Symbols.h"

#include "Utils/magic\_enum.hpp"

class BackusRuleItem

{

public:

BackusRuleItem(const std::vector<std::variant<std::string, Symbols>>& rules, RuleCountPolicy policy) : m\_policy(policy)

{

for (auto rule : rules)

{

if (std::holds\_alternative<std::string>(rule))

m\_ruleNames.push\_back(std::get<std::string>(rule));

else

m\_ruleNames.emplace\_back(magic\_enum::enum\_name(std::get<Symbols>(rule)));

}

}

std::vector<std::shared\_ptr<IBackusRule>> rules() const

{

if (m\_rules.empty())

m\_rules = BackusRuleStorage::Instance()->getRules(m\_ruleNames);

return m\_rules;

};

RuleCountPolicy policy() const { return m\_policy; };

private:

std::vector<std::string> m\_ruleNames;

mutable std::vector<std::shared\_ptr<IBackusRule>> m\_rules;

RuleCountPolicy m\_policy = NoPolicy;

};

BackusRuleStorage.h

#pragma once

#include "stdafx.h"

#include "Utils/singleton.hpp"

#include "Core/Backus/IBackusRule.h"

class BackusRuleStorage : public singleton<BackusRuleStorage>

{

public:

void regRule(std::shared\_ptr<IBackusRule> rule)

{

auto [it, inserted] = m\_rules.try\_emplace(rule->type(), rule);

if (!inserted)

{

throw std::runtime\_error("BackusRuleStorage::regRule: A rule with the type " + rule->type() + " already exists.");

}

}

std::vector<std::shared\_ptr<IBackusRule>> getRules(const std::vector<std::string>& ruleTypes) const

{

std::vector<std::shared\_ptr<IBackusRule>> rules;

for (const auto& ruleType : ruleTypes)

{

auto it = m\_rules.find(ruleType);

if (it == m\_rules.end())

throw std::runtime\_error("BackusRuleStorage::regRule: A rule with the type " + ruleType + " not found.");

rules.push\_back(it->second);

}

return rules;

};

private:

std::map<std::string, std::shared\_ptr<IBackusRule>> m\_rules;

};

IBackusRule.h

#pragma once

#include "stdafx.h"

#include "Core/IItem.h"

enum RuleCountPolicy : std::uint16\_t

{

NoPolicy = 0,

Optional = 1 << 0,

OnlyOne = 1 << 1,

Several = 1 << 2,

OneOrMore = OnlyOne | Several,

PairStart = 1 << 3,

PairEnd = 1 << 4,

};

DEFINE\_ENUM\_FLAG\_OPERATORS(RuleCountPolicy)

\_\_interface IBackusRule : public IItem

{

virtual bool check(std::multimap<int, std::pair<std::string, std::vector<std::string>>>& errorsInfo,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end) = 0;

virtual void setPostHandler(const std::function<void(std::list<std::shared\_ptr<IBackusRule>>::iterator& ruleBegin,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end)>& handler) = 0;

};

BackusRule.cpp

#pragma once

#include "stdafx.h"

#include "BackusRule.h"

std::shared\_ptr<IBackusRule> BackusRule::MakeRule(std::string name, std::list<BackusRuleItem> items)

{

struct EnableMakeShared : public BackusRule { EnableMakeShared(const std::string& name, const std::list<BackusRuleItem>& items) : BackusRule(name, items) {} };

return std::make\_shared<EnableMakeShared>(name, items);

}

bool BackusRule::check(std::multimap<int, std::pair<std::string, std::vector<std::string>>>& errorsInfo,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end)

{

bool res = true;

bool pairItem = false;

auto ruleBegin = it;

for (auto item = m\_backusItem.begin(); item != m\_backusItem.end(); ++item)

{

if (it == end || !pairItem && HasFlag(item->policy(), RuleCountPolicy::PairEnd))

{

if (!HasFlag(item->policy(), RuleCountPolicy::Optional) || item != m\_backusItem.end())

{

std::vector<std::string> types;

for (const auto& rule : item->rules())

types.push\_back(rule->type());

errorsInfo.emplace((\*it)->line(), std::make\_pair((\*it)->value(), types));

res = false;

}

break;

}

if (pairItem && HasFlag(item->policy(), RuleCountPolicy::PairEnd) || !HasFlag(item->policy(), RuleCountPolicy::PairEnd))

{

bool resItem = true;

auto startIt = it;

if (HasFlag(item->policy(), RuleCountPolicy::Several))

resItem = oneOrMoreCheck(errorsInfo, it, end, \*item);

else

resItem = checkItem(errorsInfo, it, end, \*item);

if (!resItem && (!HasFlag(item->policy(), RuleCountPolicy::Optional) || startIt != it))

{

res &= resItem;

break;

}

if (resItem && HasFlag(item->policy(), RuleCountPolicy::PairStart))

{

pairItem = true;

}

if (resItem && pairItem && HasFlag(item->policy(), RuleCountPolicy::PairEnd))

{

pairItem = false;

}

}

}

if (res && m\_handler)

m\_handler(ruleBegin, it, end);

return res;

}

bool BackusRule::oneOrMoreCheck(std::multimap<int, std::pair<std::string, std::vector<std::string>>>& errorsInfo,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end,

const BackusRuleItem& item) const

{

bool res = true;

bool resItem = true;

while (resItem && it != end && HasFlag(item.policy(), RuleCountPolicy::Several))

{

auto startIt = it;

res &= resItem;

resItem = checkItem(errorsInfo, it, end, item);

if (!resItem && startIt != it)

res = false;

}

return res;

}

bool BackusRule::checkItem(std::multimap<int, std::pair<std::string, std::vector<std::string>>>& errorsInfo,

std::list<std::shared\_ptr<IBackusRule>>::iterator& it,

std::list<std::shared\_ptr<IBackusRule>>::iterator& end,

const BackusRuleItem& item) const

{

bool res = false;

std::vector<std::string> types;

auto startIt = it;

auto maxIt = it;

if (it != end)

{

std::multimap<int, std::pair<std::string, std::vector<std::string>>> errors;

for (auto rule : item.rules())

{

types.push\_back(rule->type());

if (!res && startIt == it)

{

res = rule->check(errors, it, end);

}

if (res)

{

break;

}

else if (!res && startIt != it)

{

if(std::distance(maxIt, end) > std::distance(it, end))

maxIt = it;

it = startIt;

errorsInfo.insert(errors.begin(), errors.end());

}

}

}

if (std::distance(maxIt, end) < std::distance(it, end))

it = maxIt;

if (!res)

errorsInfo.emplace((\*startIt)->line(), std::make\_pair((\*it)->value(), types));

else

errorsInfo.clear();

return res;

}

bool BackusRule::HasFlag(RuleCountPolicy policy, RuleCountPolicy flag)

{

return (policy & flag) == flag;

}

Generator.h

#pragma once

#include "stdafx.h"

#include "Utils/singleton.hpp"

#include "Core/Generator/GeneratorItemBase.h"

class Generator : public singleton<Generator>

{

public:

template<class T>

void generateCode(std::ostream& out, std::list<std::shared\_ptr<T>>& items) const

{

if (!m\_details)

throw std::runtime\_error("Generator details is not set");

std::list<std::shared\_ptr<IGeneratorItem>> generatorItems;

for (auto item : items)

{

generatorItems.push\_back(std::dynamic\_pointer\_cast<IGeneratorItem>(item));

}

auto it = generatorItems.begin();

auto end = generatorItems.end();

std::stringstream code;

genCode(code, \*m\_details, it, end);

PrintBegin(out, \*m\_details);

PrintData(out, \*m\_details);

PrintBeginCodeSegment(out, \*m\_details);

out << code.str();

PrintEnding(out, \*m\_details);

}

void setDetails(const GeneratorDetails& details) { m\_details = std::make\_shared<GeneratorDetails>(details); }

protected:

Generator() = default;

private:

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const;

private:

static void PrintBegin(std::ostream& out, GeneratorDetails& details);

static void PrintData(std::ostream& out, GeneratorDetails& details);

static void PrintBeginCodeSegment(std::ostream& out, GeneratorDetails& details);

static void PrintEnding(std::ostream& out, GeneratorDetails& details);

private:

std::shared\_ptr<GeneratorDetails> m\_details;

};

GeneratorDetails.h

#pragma once

#include "stdafx.h"

class GeneratorDetails

{

friend class Generator;

public:

struct GeneratorArgs

{

std::string regPrefix;

std::string numberType;

std::string numberTypeExtended;

size\_t argSize;

size\_t posArg0;

size\_t posArg1;

std::string numberStrType;

};

public:

explicit GeneratorDetails(const GeneratorArgs& args) : m\_args(args)

{

m\_args.posArg0 = m\_kRetAddrSize + m\_args.argSize;

m\_args.posArg1 = m\_kRetAddrSize;

}

const GeneratorArgs& args() const { return m\_args; }

void registerNumberData(const std::string& name)

{

throwIfDataExists(name);

m\_userNumberData[name] = '\t' + name + '\t' + m\_args.numberType + '\t' + "0";

}

void registerStringData(const std::string& name, const std::string& data)

{

throwIfDataExists(name);

std::string item;

size\_t start = 0;

size\_t end;

std::string delimiter = "\\n";

m\_userStringData[name] = '\t' + name + "\tdb\t";

while ((end = data.find(delimiter, start)) != std::string::npos)

{

item = data.substr(start, end - start);

if (!item.empty())

m\_userStringData[name] += "\"" + item + "\", ";

m\_userStringData[name] += "13, 10, ";

start = end + delimiter.length();

}

item = data.substr(start);

if (!item.empty())

m\_userStringData[name] += "\"" + item + "\", ";

m\_userStringData[name] += "0";

}

void registerRawData(const std::string& name, const std::string& rawData)

{

throwIfDataExists(name);

m\_userRawData[name] = '\t' + name + '\t' + rawData;

}

void registerProc(const std::string& type, const std::function<void(std::ostream& out, const GeneratorArgs&)>& generator)

{

if (!m\_procGenerators.contains(type))

m\_procGenerators[type] = generator;

else

throw std::runtime\_error("Proc for type " + type + " already exists");

}

private:

void throwIfDataExists(const std::string& name) const

{

if (m\_userNumberData.contains(name) || m\_userStringData.contains(name) || m\_userRawData.contains(name))

throw std::runtime\_error("Data with name " + name + " already exists");

}

private:

GeneratorArgs m\_args;

std::map<std::string, std::string> m\_userNumberData;

std::map<std::string, std::string> m\_userStringData;

std::map<std::string, std::string> m\_userRawData;

std::map<std::string, std::function<void(std::ostream& out, const GeneratorArgs&)>> m\_procGenerators;

static constexpr size\_t m\_kRetAddrSize = 4;

};

GeneratorItemBase.h

#pragma once

#include "stdafx.h"

#include "Core/Generator/GeneratorUtils.h"

template <class T>

class GeneratorItemBase : public IGeneratorItem

{

public:

virtual ~GeneratorItemBase() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const override {};

protected:

std::string customData\_imp(const std::string& id) const { return m\_customData[id]; }

void setCustomData\_imp(const std::string& data, const std::string& id) { m\_customData[id] = data; }

static bool IsRegistered() { return registered; }

static void SetRegistered() { registered = true; }

static bool registered;

private:

mutable std::map<std::string, std::string> m\_customData{ {"default",""} };

};

template<class T>

bool GeneratorItemBase<T>::registered = false;

GeneratorUtils.h

#pragma once

#include "stdafx.h"

#include "Utils/singleton.hpp"

#include "Core/Generator/IGeneratorItem.h"

class GeneratorUtils : public singleton<GeneratorUtils>

{

public:

void RegisterOperation(const std::string& type, size\_t priority)

{

m\_operations[type] = priority;

}

void RegisterOperand(const std::string& type)

{

m\_operands.insert(type);

}

void RegisterEquationEnd(const std::string& type)

{

m\_equationEnd.insert(type);

}

void RegisterLBraket(const std::string& type)

{

m\_lBraketType = type;

}

void RegisterRBraket(const std::string& type)

{

m\_rBraketType = type;

}

std::list<std::shared\_ptr<IGeneratorItem>> ConvertToPostfixForm(

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const

{

std::list<std::shared\_ptr<IGeneratorItem>> postfixForm;

std::list<std::shared\_ptr<IGeneratorItem>> stack;

while (it != end)

{

auto item = \*it;

auto itemType = item->type();

if (IsOperand(item))

{

postfixForm.push\_back(item);

}

else if (IsOperation(item))

{

while (!stack.empty() && !Prioritet(item, stack.back()) && stack.back()->type() != m\_lBraketType)

{

postfixForm.push\_back(stack.back());

stack.pop\_back();

}

stack.push\_back(item);

}

else if (itemType == m\_lBraketType)

{

stack.push\_back(item);

postfixForm.push\_back(item);

}

else if (itemType == m\_rBraketType)

{

while (stack.back()->type() != m\_lBraketType)

{

postfixForm.push\_back(stack.back());

stack.pop\_back();

}

stack.pop\_back();

postfixForm.push\_back(item);

}

if (IsNextEndOfEquation(it, end))

{

break;

}

++it;

}

while (!stack.empty())

{

postfixForm.push\_back(stack.back());

stack.pop\_back();

}

return postfixForm;

}

static void PrintResultToStack(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << "\tmov [esp + " << args.posArg0 << "], " << args.regPrefix << "ax\n";

out << "\tpop ecx\n";

out << "\tpop " << args.regPrefix << "ax\n";

out << "\tpush ecx\n";

}

static bool IsNextTokenIs(const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end,

const std::string& type)

{

auto res = false;

if (it != end && std::next(it) != end && (\*std::next(it))->type() == type)

res = true;

return res;

}

private:

inline bool IsOperand(const std::shared\_ptr<IGeneratorItem>& item) const

{

return m\_operands.contains(item->type());

}

inline bool IsOperation(const std::shared\_ptr<IGeneratorItem>& item) const

{

return m\_operations.contains(item->type());

}

bool Prioritet(const std::shared\_ptr<IGeneratorItem>& left, const std::shared\_ptr<IGeneratorItem>& right) const

{

size\_t leftPriority = 0;

size\_t rightPriority = 0;

if (IsOperation(left))

leftPriority = m\_operations.at(left->type());

if (IsOperation(right))

rightPriority = m\_operations.at(right->type());

return leftPriority > rightPriority;

}

bool IsNextEndOfEquation(const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const

{

auto res = true;

if (it != end && std::next(it) != end)

{

auto next = \*std::next(it);

res = m\_equationEnd.contains(next->type()) || IsNextTokenOnNextLine(it, end);

}

return res;

}

static bool IsNextTokenOnNextLine(const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end)

{

auto res = false;

if (it != end && std::next(it) != end && ((\*it)->line() + 1) == (\*std::next(it))->line())

res = true;

return res;

}

private:

std::map<std::string, size\_t> m\_operations;

std::set<std::string> m\_operands;

std::set<std::string> m\_equationEnd;

std::string m\_lBraketType;

std::string m\_rBraketType;

};

IGeneratorItem.h

#pragma once

#include "stdafx.h"

#include "Utils/singleton.hpp"

#include "Core/Generator/IGeneratorItem.h"

class GeneratorUtils : public singleton<GeneratorUtils>

{

public:

void RegisterOperation(const std::string& type, size\_t priority)

{

m\_operations[type] = priority;

}

void RegisterOperand(const std::string& type)

{

m\_operands.insert(type);

}

void RegisterEquationEnd(const std::string& type)

{

m\_equationEnd.insert(type);

}

void RegisterLBraket(const std::string& type)

{

m\_lBraketType = type;

}

void RegisterRBraket(const std::string& type)

{

m\_rBraketType = type;

}

std::list<std::shared\_ptr<IGeneratorItem>> ConvertToPostfixForm(

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const

{

std::list<std::shared\_ptr<IGeneratorItem>> postfixForm;

std::list<std::shared\_ptr<IGeneratorItem>> stack;

while (it != end)

{

auto item = \*it;

auto itemType = item->type();

if (IsOperand(item))

{

postfixForm.push\_back(item);

}

else if (IsOperation(item))

{

while (!stack.empty() && !Prioritet(item, stack.back()) && stack.back()->type() != m\_lBraketType)

{

postfixForm.push\_back(stack.back());

stack.pop\_back();

}

stack.push\_back(item);

}

else if (itemType == m\_lBraketType)

{

stack.push\_back(item);

postfixForm.push\_back(item);

}

else if (itemType == m\_rBraketType)

{

while (stack.back()->type() != m\_lBraketType)

{

postfixForm.push\_back(stack.back());

stack.pop\_back();

}

stack.pop\_back();

postfixForm.push\_back(item);

}

if (IsNextEndOfEquation(it, end))

{

break;

}

++it;

}

while (!stack.empty())

{

postfixForm.push\_back(stack.back());

stack.pop\_back();

}

return postfixForm;

}

static void PrintResultToStack(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << "\tmov [esp + " << args.posArg0 << "], " << args.regPrefix << "ax\n";

out << "\tpop ecx\n";

out << "\tpop " << args.regPrefix << "ax\n";

out << "\tpush ecx\n";

}

static bool IsNextTokenIs(const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end,

const std::string& type)

{

auto res = false;

if (it != end && std::next(it) != end && (\*std::next(it))->type() == type)

res = true;

return res;

}

private:

inline bool IsOperand(const std::shared\_ptr<IGeneratorItem>& item) const

{

return m\_operands.contains(item->type());

}

inline bool IsOperation(const std::shared\_ptr<IGeneratorItem>& item) const

{

return m\_operations.contains(item->type());

}

bool Prioritet(const std::shared\_ptr<IGeneratorItem>& left, const std::shared\_ptr<IGeneratorItem>& right) const

{

size\_t leftPriority = 0;

size\_t rightPriority = 0;

if (IsOperation(left))

leftPriority = m\_operations.at(left->type());

if (IsOperation(right))

rightPriority = m\_operations.at(right->type());

return leftPriority > rightPriority;

}

bool IsNextEndOfEquation(const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const

{

auto res = true;

if (it != end && std::next(it) != end)

{

auto next = \*std::next(it);

res = m\_equationEnd.contains(next->type()) || IsNextTokenOnNextLine(it, end);

}

return res;

}

static bool IsNextTokenOnNextLine(const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end)

{

auto res = false;

if (it != end && std::next(it) != end && ((\*it)->line() + 1) == (\*std::next(it))->line())

res = true;

return res;

}

private:

std::map<std::string, size\_t> m\_operations;

std::set<std::string> m\_operands;

std::set<std::string> m\_equationEnd;

std::string m\_lBraketType;

std::string m\_rBraketType;

};

Generator.cpp

#include "stdafx.h"

#include "Generator.h"

void Generator::PrintBegin(std::ostream& out, GeneratorDetails& details)

{

out << ".386\n";

out << ".model flat, stdcall\n";

out << "option casemap :none\n";

out << std::endl;

out << "include masm32\\include\\windows.inc\n";

out << "include masm32\\include\\kernel32.inc\n";

out << "include masm32\\include\\masm32.inc\n";

out << "include masm32\\include\\user32.inc\n";

out << "include masm32\\include\\msvcrt.inc\n";

out << "includelib masm32\\lib\\kernel32.lib\n";

out << "includelib masm32\\lib\\masm32.lib\n";

out << "includelib masm32\\lib\\user32.lib\n";

out << "includelib masm32\\lib\\msvcrt.lib\n";

}

void Generator::PrintData(std::ostream& out, GeneratorDetails& details)

{

out << std::endl;

out << ".DATA\n";

out << ";===User Data================================================================================\n";

for (const auto& [\_, data] : details.m\_userNumberData)

{

out << data << std::endl;

}

if (!details.m\_userNumberData.empty())

out << std::endl;

for (const auto& [\_, data] : details.m\_userStringData)

{

out << data << std::endl;

}

if (!details.m\_userStringData.empty())

out << std::endl;

out << ";===Addition Data============================================================================\n";

out << "\thConsoleInput\tdd\t?\n";

out << "\thConsoleOutput\tdd\t?\n";

out << "\tendBuff\t\t\tdb\t5 dup (?)\n";

out << "\tmsg1310\t\t\tdb\t13, 10, 0\n";

if (!details.m\_userRawData.empty())

out << std::endl;

for (const auto& [\_, data] : details.m\_userRawData)

{

out << data << std::endl;

}

}

void Generator::PrintBeginCodeSegment(std::ostream& out, GeneratorDetails& details)

{

out << std::endl;

out << ".CODE\n";

out << "start:\n";

out << "invoke AllocConsole\n";

out << "invoke GetStdHandle, STD\_INPUT\_HANDLE\n";

out << "mov hConsoleInput, eax\n";

out << "invoke GetStdHandle, STD\_OUTPUT\_HANDLE\n";

out << "mov hConsoleOutput, eax\n";

}

void Generator::PrintEnding(std::ostream& out, GeneratorDetails& details)

{

out << "exit\_label:\n";

out << "invoke WriteConsoleA, hConsoleOutput, ADDR msg1310, SIZEOF msg1310 - 1, 0, 0\n";

out << "invoke ReadConsoleA, hConsoleInput, ADDR endBuff, 5, 0, 0\n";

out << "invoke ExitProcess, 0\n";

for (const auto& [\_, proc] : details.m\_procGenerators)

{

out << std::endl << std::endl;

proc(out, details.args());

}

out << "end start\n";

}

void Generator::genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const

{

for (; it != end; ++it)

{

(\*it)->genCode(out, details, it, end);

}

}

TokenParser.h

#pragma once

#include "stdafx.h"

#include "Utils/singleton.hpp"

#include "Core/Tokens/IToken.hpp"

#include "Utils/TablePrinter.h"

class TokenParser : public singleton<TokenParser>

{

public:

static constexpr int NoPriority = std::numeric\_limits<int>::min();

public:

std::list<std::shared\_ptr<IToken>> tokenize(std::istream& input);

void regToken(std::shared\_ptr<IToken> token, int priority = NoPriority);

void regUnchangedTextToken(std::shared\_ptr<IToken> target, std::shared\_ptr<IToken> lBorder, std::shared\_ptr<IToken> rBorder);

template<class T>

static void PrintTokens(std::ostream& out, const std::list<std::shared\_ptr<T>>& tokens)

{

auto getNumCount = [](int k) { return std::to\_string(k).size(); };

size\_t maxLemexeLen = 0;

size\_t maxTypeLen = 0;

size\_t maxValueLen = 0;

for (auto token : tokens)

{

maxLemexeLen = std::max(maxLemexeLen, token->lexeme().size());

maxTypeLen = std::max(maxTypeLen, token->type().size());

maxValueLen = std::max(maxValueLen, token->value().size());

}

const std::string kHeaderColumn0 = "#";

const std::string kHeaderColumn1 = "SYMBOL";

const std::string kHeaderColumn2 = "TYPE";

const std::string kHeaderColumn3 = "VALUE";

const std::string kHeaderColumn4 = "LINE";

size\_t colPadding = 1;

auto widthColumn0 = std::max(kHeaderColumn0.size(), getNumCount(tokens.size())) + 2 \* colPadding;

auto widthColumn1 = std::max(kHeaderColumn1.size(), maxLemexeLen) + 2 \* colPadding;

auto widthColumn2 = std::max(kHeaderColumn2.size(), maxTypeLen) + 2 \* colPadding;

auto widthColumn3 = std::max(kHeaderColumn3.size(), maxValueLen) + 2 \* colPadding;

auto widthColumn4 = std::max(kHeaderColumn4.size(), getNumCount(tokens.back()->line())) + 2 \* colPadding;

if ((kHeaderColumn0.size() % 2) != (widthColumn0 % 2)) widthColumn0++;

if ((kHeaderColumn1.size() % 2) != (widthColumn1 % 2)) widthColumn1++;

if ((kHeaderColumn2.size() % 2) != (widthColumn2 % 2)) widthColumn2++;

if ((kHeaderColumn3.size() % 2) != (widthColumn3 % 2)) widthColumn3++;

if ((kHeaderColumn4.size() % 2) != (widthColumn4 % 2)) widthColumn4++;

size\_t index = 1;

auto getIndex = [&index](const std::shared\_ptr<T>&) { return std::to\_string(index++); };

auto getLemexe = [](const std::shared\_ptr<T>& token) { return token->lexeme(); };

auto getType = [](const std::shared\_ptr<T>& token) { return token->type(); };

auto getValue = [](const std::shared\_ptr<T>& token) { return token->value(); };

auto getLine = [](const std::shared\_ptr<T>& token) { return std::to\_string(token->line()); };

TablePrinter::PrintTable(out,

{ kHeaderColumn0, kHeaderColumn1, kHeaderColumn2, kHeaderColumn3, kHeaderColumn4 },

{ widthColumn0, widthColumn1, widthColumn2, widthColumn3, widthColumn4 },

{ TablePrinter::CENTRE, TablePrinter::RIGHT, TablePrinter::RIGHT , TablePrinter::RIGHT , TablePrinter::RIGHT },

tokens,

{ getIndex, getLemexe, getType, getValue, getLine },

colPadding);

}

private:

void throwIfTokenRegistered(std::shared\_ptr<IToken> token);

void recognizeToken(std::string& token, int curLine);

bool isUnchangedTextTokenLast();

private:

static bool IsNewLine(const char& ch);

static bool IsTabulation(const char& ch);

static bool IsAllowedSymbol(const char& ch);

static bool IsAllowedSpecialSymbol(const char& ch);

private:

struct PriorityCompare

{

bool operator()(const int& a, const int& b) const

{

return a > b;

}

};

private:

std::multimap<int, std::shared\_ptr<IToken>, PriorityCompare> m\_priorityTokens;

std::map<std::string, std::tuple<std::shared\_ptr<IToken>, std::shared\_ptr<IToken>, std::shared\_ptr<IToken>>> m\_unchangedTextTokens;

std::list<std::shared\_ptr<IToken>> m\_tokens;

std::function<std::shared\_ptr<IToken>(std::string)> m\_getTokenByType = [this](const std::string& type) {

auto start = m\_priorityTokens.lower\_bound(static\_cast<int>(type.size()));

auto mapItem = std::find\_if(start, m\_priorityTokens.end(), [&type](const auto& pair) { return pair.second->type() == type; });

if (mapItem == m\_priorityTokens.end())

throw std::runtime\_error("TokenParser::getTokenByType: Token with type " + type + " not found");

return mapItem->second;

};

};

TokenParser.cpp

#include "stdafx.h"

#include "Core/Parser/TokenParser.h"

#include "Utils/StringUtils.h"

#include "Tokens/Common/EndOfFile.h"

std::list<std::shared\_ptr<IToken>> TokenParser::tokenize(std::istream& input)

{

m\_tokens.clear();

int curLine = 1;

std::string token;

for (char ch; input.get(ch);)

{

if (!token.empty() && ((IsAllowedSymbol(token.front()) != IsAllowedSymbol(ch)) || IsTabulation(ch)))

recognizeToken(token, curLine);

if (IsNewLine(ch))

++curLine;

if (isUnchangedTextTokenLast())

{

std::string unchangedTextTokenValue{ token };

token.clear();

int unchangedTextTokenLine{ curLine };

const auto& [target, left, right] = m\_unchangedTextTokens[m\_tokens.back()->lexeme()];

auto rBorderLex = right ? right->lexeme() : "\n";

do

{

if (IsNewLine(ch))

++curLine;

unchangedTextTokenValue += ch;

}

while (!StringUtils::Compare(unchangedTextTokenValue, rBorderLex, StringUtils::EndWith) && input.get(ch));

unchangedTextTokenValue = unchangedTextTokenValue.substr(0, unchangedTextTokenValue.size() - rBorderLex.size());

m\_tokens.push\_back(target->tryCreateToken(unchangedTextTokenValue));

m\_tokens.back()->setLine(unchangedTextTokenLine);

if (right)

{

m\_tokens.push\_back(right->tryCreateToken(rBorderLex));

m\_tokens.back()->setLine(curLine);

}

continue;

}

if (!IsTabulation(ch))

token += ch;

}

if (!token.empty())

recognizeToken(token, curLine);

m\_tokens.push\_back(std::make\_shared<EndOfFile>());

return m\_tokens;

}

void TokenParser::regToken(std::shared\_ptr<IToken> token, int priority)

{

throwIfTokenRegistered(token);

if (priority == NoPriority)

priority = static\_cast<int>(token->lexeme().size());

m\_priorityTokens.insert(std::make\_pair(priority, token));

}

void TokenParser::regUnchangedTextToken(std::shared\_ptr<IToken> target, std::shared\_ptr<IToken> lBorder, std::shared\_ptr<IToken> rBorder)

{

if(rBorder)

throwIfTokenRegistered(rBorder);

regToken(lBorder);

throwIfTokenRegistered(target);

m\_unchangedTextTokens.try\_emplace(lBorder->lexeme(), target, lBorder, rBorder);

}

void TokenParser::throwIfTokenRegistered(std::shared\_ptr<IToken> token)

{

auto start = m\_priorityTokens.lower\_bound(static\_cast<int>(token->lexeme().size()));

auto priorToken = std::find\_if(start, m\_priorityTokens.end(),

[&token](const auto& pair) {

return token->type() == pair.second->type();

});

auto unchTextToken = std::ranges::find\_if(m\_unchangedTextTokens,

[&token](const auto& pair) {

auto type = token->type();

const auto& [main, left, right] = pair.second;

return type == main->type() ||

type == left->type() ||

right && type == right->type();

});

if(priorToken != m\_priorityTokens.end() || unchTextToken != m\_unchangedTextTokens.end())

throw std::runtime\_error("TokenParser: Token with type " + token->type() + " already registered");

}

void TokenParser::recognizeToken(std::string& token, int curLine)

{

if(m\_priorityTokens.empty())

throw std::runtime\_error("TokenParser: No tokens registered");

auto start = m\_priorityTokens.lower\_bound(static\_cast<int>(token.size()));

for (auto it = start; it != m\_priorityTokens.end(); ++it)

{

auto curRegToken = it->second;

if (auto newToken = curRegToken->tryCreateToken(token); newToken)

{

m\_tokens.push\_back(newToken);

m\_tokens.back()->setLine(curLine);

break;

}

}

if (!token.empty() && !isUnchangedTextTokenLast())

recognizeToken(token, curLine);

}

bool TokenParser::isUnchangedTextTokenLast()

{

if (!m\_tokens.empty() && m\_unchangedTextTokens.contains(m\_tokens.back()->lexeme()))

{

auto const& [target, left, right] = m\_unchangedTextTokens[m\_tokens.back()->lexeme()];

if (m\_tokens.size() >= 2)

{

if (target->type() != (\*(++m\_tokens.rbegin()))->type())

return true;

}

else

return true;

}

return false;

}

bool TokenParser::IsNewLine(const char& ch)

{

return ch == '\n';

}

bool TokenParser::IsTabulation(const char& ch)

{

return ch == ' ' || ch == '\t' || IsNewLine(ch);

}

bool TokenParser::IsAllowedSymbol(const char& ch)

{

return !!isalpha(ch) || !!isdigit(ch) || IsAllowedSpecialSymbol(ch);

}

bool TokenParser::IsAllowedSpecialSymbol(const char& ch)

{

std::set<char> allowedSymblos{ '\_' };

return allowedSymblos.contains(ch);

}

TokenRegister.h

#pragma once

#include "stdafx.h"

#include "Controller.h"

#include "Rules/IdentRule/Undefined.h"

#include "Tokens/Common/Unknown.h"

void Init();

template <typename T>

bool CheckSemantic(std::ostream& out, std::list<std::shared\_ptr<T>>& tokens)

{

auto endOfFileType = tokens.back()->type();

std::list<std::shared\_ptr<IBackusRule>> rules;

for (auto token : tokens)

{

if (auto rule = std::dynamic\_pointer\_cast<IBackusRule>(token))

rules.push\_back(rule);

}

auto it = rules.begin();

auto end = rules.end();

std::multimap<int, std::pair<std::string, std::vector<std::string>>> errors;

auto res = Controller::Instance()->topRule()->check(errors, it, end);

rules.erase(++std::find\_if(it, rules.end(), [&endOfFileType](const auto& rule) { return rule->type() == endOfFileType; }), rules.end());

end = --rules.end();

std::multimap<int, std::string> errorsMsg;

int lexErr = 0;

int synErr = 0;

int semErr = 0;

tokens.clear();

for (auto rule : rules)

{

tokens.push\_back(std::dynamic\_pointer\_cast<T>(rule));

if (rule->type() == Undefined::Type())

{

res = false;

std::string err;

if (auto erMsg = rule->customData("error"); !erMsg.empty())

{

semErr++;

err = "Semantic error: " + erMsg;

}

else

{

semErr++;

err = std::format("Semantic error: Undefined token: {}", rule->value());

}

errorsMsg.emplace(rule->line(), err);

}

else if (rule->type() == token::Unknown::Type())

{

lexErr++;

res = false;

errorsMsg.emplace(rule->line(), std::format("Lexical error: Unknown token: {}", rule->value()));

}

}

for (auto it = errors.rbegin(); it != errors.rend(); ++it)

{

auto types = it->second.second;

std::stringstream ss;

for (size\_t i = 0; i < types.size(); ++i)

{

if (!types[i].empty())

{

ss << types[i];

if (i != types.size() - 1)

ss << " or ";

}

}

auto ssStr = ss.str();

if (!ssStr.empty())

{

synErr++;

std::string msg = "Syntax error: Expected: " + ssStr;

if (!it->second.first.empty())

msg += " before " + it->second.first;

errorsMsg.emplace(it->first, msg);

}

}

out << "List of errors" << std::endl;

out << "======================================================================" << std::endl;

out << "There are " << lexErr << " lexical errors." << std::endl;

out << "There are " << synErr << " syntax errors." << std::endl;

out << "There are " << semErr << " semantic errors." << std::endl << std::endl;

for (auto const& [line, msg] : errorsMsg)

{

out << "Line " << line << ": " << msg << std::endl;

}

return res;

}

TokenRegister.cpp

#include "stdafx.h"

#include "Core/Parser/TokenRegister.h"

#include "Controller.h"

#include "Tokens/Common.h"

#include "Rules/Operators/If/IfRule.h"

#include "Rules/Operators/Goto/GotoRule.h"

#include "Rules/Operators/For/ForRule.h"

#include "Rules/Operators/WhileC/WhileRule.h"

#include "Rules/Operators/RepeatUntil/RepeatUntilRule.h"

void Init()

{

Controller::Instance()->regOperatorRule(MakeIf);

Controller::Instance()->regOperatorRule(MakeGoto, true);

Controller::Instance()->regOperatorRule(MakeLabel);

Controller::Instance()->regOperatorRule(MakeFor);

Controller::Instance()->regOperatorRule(MakeWhile);

Controller::Instance()->regOperatorRule(MakeRepeatUntil);

Controller::Instance()->regItem<token::Unknown>(ItemType::TokenAndRule, -2);

Controller::Instance()->regUnchangedTextToken(std::make\_shared<Comment>(), std::make\_shared<LComment>(), nullptr);

Controller::Instance()->init();

}

TokenOperators:

Loops:

Do.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Do : public TokenBase<Do>, public BackusRuleBase<Do>, public GeneratorItemBase<Do>

{

BASE\_ITEM

public:

Do() { setLexeme("DO"); };

virtual ~Do() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

out << "\tpop " << details.args().regPrefix << "ax" << std::endl;

out << "\tcmp " << details.args().regPrefix << "ax, 0" << std::endl;

out << "\tje " << customData("endLabel") << std::endl;

};

};

DownTo.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

#include "Rules/EquationRule/Greate.h"

#include "Rules/EquationRule/Not.h"

#include "Rules/EquationRule/Subtraction.h"

class DownTo : public TokenBase<DownTo>, public BackusRuleBase<DownTo>, public GeneratorItemBase<DownTo>

{

BASE\_ITEM

public:

DownTo() { setLexeme("DOWNTO"); };

virtual ~DownTo() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

Greate::RegPROC(details);

Not::RegPROC(details);

Subtraction::RegPROC(details);

out << customData("startLabel") << ":" << std::endl;

it++;

auto postForm = GeneratorUtils::Instance()->ConvertToPostfixForm(it, end);

auto postIt = postForm.begin();

auto postEnd = postForm.end();

for (const auto& item : postForm)

item->genCode(out, details, postIt, postEnd);

out << "\tpush " << customData("ident") << std::endl;

out << "\tcall Greate\_" << std::endl;

out << "\tcall Not\_" << std::endl;

};

};

For.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class For : public TokenBase<For>, public BackusRuleBase<For>, public GeneratorItemBase<For>

{

BASE\_ITEM

public:

For() { setLexeme("FOR"); };

virtual ~For() = default;

};

ForRule.h

#pragma once

#include "stdafx.h"

#include "Controller.h"

BackusRulePtr MakeFor(std::shared\_ptr<Controller> controller);

ForRule.cpp

#include "stdafx.h"

#include "ForRule.h"

#include "Rules/Operators/For/For.h"

#include "Rules/Operators/For/To.h"

#include "Rules/Operators/For/DownTo.h"

#include "Rules/Operators/For/Do.h"

BackusRulePtr MakeFor(std::shared\_ptr<Controller> controller)

{

using enum ItemType;

controller->regItem<For>();

controller->regItem<To>(TokenAndRule | EquationEnd);

controller->regItem<DownTo>(TokenAndRule | EquationEnd);

controller->regItem<Do>(TokenAndRule | EquationEnd);

auto context = controller->context();

static const auto [lStart, lCodeBlok, lEnd] = context->CodeBlockTypes();

auto forToOrDownToDoRule = controller->addRule("ForToOrDownToDoRule", {

BackusRuleItem({ For::Type()}, OnlyOne),

BackusRuleItem({ "AssignmentRule"}, OnlyOne),

BackusRuleItem({ To::Type(), DownTo::Type()}, OnlyOne),

BackusRuleItem({ context->EquationRuleName()}, OnlyOne),

BackusRuleItem({ Do::Type()}, OnlyOne),

BackusRuleItem({ lCodeBlok}, OnlyOne)

});

forToOrDownToDoRule->setPostHandler([context](BackusRuleList::iterator& ruleBegin,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

static size\_t index = 0;

index++;

std::string startLabel = std::format("forPasStart{}", index);

std::string endLabel = std::format("forPasEnd{}", index);

auto ident = \*std::next(ruleBegin, 1);

bool increment = false;

for (auto itr = ruleBegin; itr != it; ++itr)

{

auto type = (\*itr)->type();

if ((type == To::Type() || type == DownTo::Type()))

{

if (type == To::Type())

increment = true;

(\*itr)->setCustomData(startLabel, "startLabel");

(\*itr)->setCustomData(ident->customData(), "ident");

}

else if (type == Do::Type())

{

(\*itr)->setCustomData(endLabel, "endLabel");

break;

}

}

std::string code;

code += std::format("\tpush {}\n", ident->customData());

code += std::format("\tpush {} ptr 1\n", context->Details().args().numberTypeExtended);

code += std::format("\tcall {}\n", increment ? "Add\_" : "Sub\_");

code += std::format("\tpop {}\n", ident->customData());

code += std::format("\tjmp {}\n", startLabel);

code += std::format("{}:", endLabel);

(\*std::prev(it, 1))->setCustomData(code);

});

return forToOrDownToDoRule;

}

To.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

#include "Rules/EquationRule/Less.h"

#include "Rules/EquationRule/Not.h"

#include "Rules/EquationRule/Addition.h"

class To : public TokenBase<To>, public BackusRuleBase<To>, public GeneratorItemBase<To>

{

BASE\_ITEM

public:

To() { setLexeme("TO"); };

virtual ~To() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

Less::RegPROC(details);

Not::RegPROC(details);

Addition::RegPROC(details);

out << customData("startLabel") << ":" << std::endl;

it++;

auto postForm = GeneratorUtils::Instance()->ConvertToPostfixForm(it, end);

auto postIt = postForm.begin();

auto postEnd = postForm.end();

for (const auto& item : postForm)

item->genCode(out, details, postIt, postEnd);

out << "\tpush " << customData("ident") << std::endl;

out << "\tcall Less\_" << std::endl;

out << "\tcall Not\_" << std::endl;

};

};

Goto:

Goto.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Goto : public TokenBase<Goto>, public BackusRuleBase<Goto>, public GeneratorItemBase<Goto>

{

BASE\_ITEM

public:

Goto() { setLexeme("GOTO"); };

virtual ~Goto() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

it++;

out << "\tjmp " << (\*it)->customData() << std::endl;

};

};

GotoRule.h

#pragma once

#include "stdafx.h"

#include "Controller.h"

BackusRulePtr MakeGoto(std::shared\_ptr<Controller> controller);

BackusRulePtr MakeLabel(std::shared\_ptr<Controller> controller);

GotoRule.cpp

#include "stdafx.h"

#include "GotoRule.h"

#include "Rules/Operators/Goto/Goto.h"

#include "Rules/Operators/Goto/Label.h"

#include "Rules/IdentRule/Identifier.h"

#include "Rules/IdentRule/Undefined.h"

static std::map<std::string, std::optional<BackusRuleList::iterator>> labelTable;

BackusRulePtr MakeLabel(std::shared\_ptr<Controller> controller)

{

using enum ItemType;

controller->regItem<Label>(Rule);

auto context = controller->context();

static const auto [lStart, lCodeBlok, lEnd] = context->CodeBlockTypes();

auto labelRule = controller->addRule("LabelRule", {

BackusRuleItem({ context->IdentRuleName()}, OnlyOne),

BackusRuleItem({ Symbols::Colon}, OnlyOne)

});

labelRule->setPostHandler([context](BackusRuleList::iterator&,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

it = std::prev(it, 2);

auto identIt = it;

auto identVal = (\*identIt)->value();

std::shared\_ptr<IToken> label;

if (context->IdentTable().contains((\*identIt)->value()))

{

label = std::make\_shared<Undefined>();

label->setCustomData("Redefinition", "error");

}

else

label = std::make\_shared<Label>();

label->setValue((\*identIt)->value() + (\*(++it))->value());

end = std::remove(it, end, \*it);

label->setLine((\*identIt)->line());

label->setCustomData((\*identIt)->customData());

\*identIt = std::dynamic\_pointer\_cast<IBackusRule>(label);

if (!labelTable.contains(identVal))

{

labelTable.try\_emplace(identVal, std::optional<BackusRuleList::iterator>());

}

else

{

if (auto optIt = labelTable[identVal]; optIt.has\_value())

{

auto gotoIdentIt = optIt.value();

if ((\*gotoIdentIt)->type() == Undefined::Type())

{

auto labelName = std::make\_shared<Identifier>();

labelName->setValue((\*gotoIdentIt)->value());

labelName->setLine((\*gotoIdentIt)->line());

labelName->setCustomData((\*gotoIdentIt)->customData());

\*gotoIdentIt = labelName;

}

}

}

});

return labelRule;

}

BackusRulePtr MakeGoto(std::shared\_ptr<Controller> controller)

{

controller->regItem<Goto>();

auto context = controller->context();

static const auto [lStart, lCodeBlok, lEnd] = context->CodeBlockTypes();

auto gotoStatement = controller->addRule("GotoStatement", {

BackusRuleItem({ Goto::Type()}, OnlyOne),

BackusRuleItem({context->IdentRuleName()}, OnlyOne)

});

gotoStatement->setPostHandler([](BackusRuleList::iterator&,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

it = std::prev(it, 1);

auto identIt = it;

if (!labelTable.contains((\*identIt)->value()))

{

if ((\*identIt)->type() != Undefined::Type())

{

auto undef = std::make\_shared<Undefined>();

undef->setValue((\*identIt)->value());

undef->setLine((\*identIt)->line());

undef->setCustomData((\*identIt)->customData());

\*identIt = undef;

}

labelTable.try\_emplace((\*identIt)->value(), identIt);

}

else

{

auto ident = std::make\_shared<Identifier>();

ident->setValue((\*identIt)->value());

ident->setLine((\*identIt)->line());

ident->setCustomData((\*identIt)->customData());

\*identIt = ident;

}

it = std::next(it);

});

return gotoStatement;

}

Label.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Label : public TokenBase<Label>, public BackusRuleBase<Label>, public GeneratorItemBase<Label>

{

BASE\_ITEM

public:

Label() { setLexeme(""); };

virtual ~Label() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

out << customData() << ":" << std::endl;

};

};

IfElse:

Else.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Else : public TokenBase<Else>, public BackusRuleBase<Else>, public GeneratorItemBase<Else>

{

BASE\_ITEM

public:

Else() { setLexeme("ELSE"); };

virtual ~Else() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

out << "\tjmp " << customData("endLabel") << std::endl;

out << customData("elseLabel") << ":\n";

};

};

If.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class If : public TokenBase<If>, public BackusRuleBase<If>, public GeneratorItemBase<If>

{

BASE\_ITEM

public:

If() { setLexeme("IF"); };

virtual ~If() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

it++;

auto postForm = GeneratorUtils::Instance()->ConvertToPostfixForm(it, end);

auto postIt = postForm.begin();

auto postEnd = postForm.end();

for (const auto& item : postForm)

item->genCode(out, details, postIt, postEnd);

out << "\tpop " << details.args().regPrefix << "ax" << std::endl;

out << "\tcmp " << details.args().regPrefix << "ax, 0" << std::endl;

out << "\tje " << customData("label") << std::endl;

};

};

IfRule.h

#pragma once

#include "stdafx.h"

#include "Controller.h"

BackusRulePtr MakeIf(std::shared\_ptr<Controller> controller);

IfRule.cpp

#include "stdafx.h"

#include "IfRule.h"

#include "Rules/Operators/If/If.h"

#include "Rules/Operators/If/Else.h"

BackusRulePtr MakeIf(std::shared\_ptr<Controller> controller)

{

controller->regItem<If>();

controller->regItem<Else>();

auto context = controller->context();

static const auto [lStart, lCodeBlok, lEnd] = context->CodeBlockTypes();

auto elseStatement = controller->addRule("ElseStatement", {

BackusRuleItem({ Else::Type()}, OnlyOne),

BackusRuleItem({ lCodeBlok}, OnlyOne),

});

auto ifStatement = controller->addRule("IfStatement", {

BackusRuleItem({ If::Type()}, OnlyOne),

BackusRuleItem({ Symbols::LBraket}, OnlyOne),

BackusRuleItem({ context->EquationRuleName()}, OnlyOne),

BackusRuleItem({ Symbols::RBraket}, OnlyOne),

BackusRuleItem({ lCodeBlok}, OnlyOne),

BackusRuleItem({ elseStatement->type()}, Optional)

});

ifStatement->setPostHandler([](BackusRuleList::iterator& ruleBegin,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

static size\_t index = 0;

index++;

std::string elseLabel = std::format("elseLabel{}", index);

std::string endLabel = std::format("endIf{}", index);

bool hasElse = false;

size\_t count = 0;

for (auto itr = ruleBegin; itr != it; ++itr)

{

auto type = (\*itr)->type();

if (type == lStart)

{

count++;

}

else if (type == Else::Type() && count == 0)

{

(\*itr)->setCustomData(elseLabel, "elseLabel");

(\*itr)->setCustomData(endLabel, "endLabel");

hasElse = true;

}

else if (type == lEnd && count == 1 && (\*std::next(itr))->type() != Else::Type())

{

(\*itr)->setCustomData(endLabel + ':');

break;

}

else if (type == lEnd && count > 0)

{

count--;

}

}

(\*ruleBegin)->setCustomData(hasElse ? elseLabel : endLabel, "label");

});

return ifStatement;

}

RepeatUntil:

Repeat.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Repeat : public TokenBase<Repeat>, public BackusRuleBase<Repeat>, public GeneratorItemBase<Repeat>

{

BASE\_ITEM

public:

Repeat() { setLexeme("REPEAT"); };

virtual ~Repeat() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

out << customData("startLabel") << ":" << std::endl;

};

};

RepeatUntilRule.h

#pragma once

#include "stdafx.h"

#include "Controller.h"

BackusRulePtr MakeRepeatUntil(std::shared\_ptr<Controller> controller);

RepeatUntilRule.cpp

#include "stdafx.h"

#include "RepeatUntilRule.h"

#include "Rules/Operators/RepeatUntil/Repeat.h"

#include "Rules/Operators/RepeatUntil/Until.h"

BackusRulePtr MakeRepeatUntil(std::shared\_ptr<Controller> controller)

{

controller->regItem<Repeat>();

controller->regItem<Until>();

auto context = controller->context();

static const auto [lStart, lCodeBlok, lEnd] = context->CodeBlockTypes();

auto operatorsRuleName = context->OperatorsRuleName();

auto repeatUntilRule = controller->addRule("RepeatUntilRule", {

BackusRuleItem({ Repeat::Type()}, OnlyOne),

BackusRuleItem({operatorsRuleName}, OnlyOne),

BackusRuleItem({ Until::Type()}, OnlyOne),

BackusRuleItem({ Symbols::LBraket}, OnlyOne),

BackusRuleItem({ context->EquationRuleName()}, OnlyOne),

BackusRuleItem({ Symbols::RBraket}, OnlyOne)

});

repeatUntilRule->setPostHandler([](BackusRuleList::iterator& ruleBegin,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

static size\_t index = 0;

index++;

std::string startLabel = std::format("repeatStart{}", index);

std::string endLabel = std::format("repeatEnd{}", index);

(\*ruleBegin)->setCustomData(startLabel, "startLabel");

size\_t count = 0;

for (auto itr = ruleBegin; itr != it; ++itr)

{

auto type = (\*itr)->type();

if (type == Repeat::Type())

{

count++;

}

else if (type == Until::Type() && count == 1)

{

count--;

(\*itr)->setCustomData(startLabel, "startLabel");

(\*itr)->setCustomData(endLabel, "endLabel");

break;

}

else if (type == Until::Type() && count > 0)

{

count--;

}

}

});

return repeatUntilRule;

}

Until.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Until : public TokenBase<Until>, public BackusRuleBase<Until>, public GeneratorItemBase<Until>

{

BASE\_ITEM

public:

Until() { setLexeme("UNTIL"); };

virtual ~Until() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

it++;

auto postForm = GeneratorUtils::Instance()->ConvertToPostfixForm(it, end);

auto postIt = postForm.begin();

auto postEnd = postForm.end();

for (const auto& item : postForm)

item->genCode(out, details, postIt, postEnd);

out << "\tpop " << details.args().regPrefix << "ax" << std::endl;

out << "\tcmp " << details.args().regPrefix << "ax, 0" << std::endl;

out << "\tje " << customData("endLabel") << std::endl;

out << "\tjmp " << customData("startLabel") << std::endl;

out << customData("endLabel") << ":" << std::endl;

};

};

While:

While.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class While : public TokenBase<While>, public BackusRuleBase<While>, public GeneratorItemBase<While>

{

BASE\_ITEM

public:

While() { setLexeme("WHILE"); };

virtual ~While() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

if (customData("noGenerateCode") == "true")

{

return;

}

if (customData("ContinueWhile") == "true")

{

out << "\tjmp " << customData("startLabel") << std::endl;

return;

}

if (customData("ExitWhile") == "true")

{

out << "\tjmp " << customData("endLabel") << std::endl;

return;

}

it++;

auto postForm = GeneratorUtils::Instance()->ConvertToPostfixForm(it, end);

out << customData("startLabel") << ":" << std::endl;

auto postIt = postForm.begin();

auto postEnd = postForm.end();

for (const auto& item : postForm)

item->genCode(out, details, postIt, postEnd);

out << "\tpop " << details.args().regPrefix << "ax" << std::endl;

out << "\tcmp " << details.args().regPrefix << "ax, 0" << std::endl;

out << "\tje " << customData("endLabel") << std::endl;

};

};

WhileRule.h

#pragma once

#include "stdafx.h"

#include "Controller.h"

BackusRulePtr MakeWhile(std::shared\_ptr<Controller> controller);

WhileRule.cpp

#include "stdafx.h"

#include "WhileRule.h"

#include "Rules/Operators/WhileC/While.h"

#include "SimpleTokens.h"

SimpleToken(ExitWhile, "EXIT")

SimpleToken(ContinueWhile, "CONTINUE")

BackusRulePtr MakeWhile(std::shared\_ptr<Controller> controller)

{

controller->regItem<While>();

controller->regItem<ExitWhile>();

controller->regItem<ContinueWhile>();

auto context = controller->context();

static const auto [lStart, lCodeBlok, lEnd] = context->CodeBlockTypes();

auto operatorsRuleName = context->OperatorsRuleName();

auto operatorsName = context->OperatorsName();

auto operatorsWithSemicolonsName = context->OperatorsWithSemicolonsName();

auto whileExitStatement = controller->addRule("WhileExitStatement", {

BackusRuleItem({ ExitWhile::Type()}, OnlyOne),

BackusRuleItem({ While::Type()}, OnlyOne)

});

auto whileContinueStatement = controller->addRule("WhileContinueStatement", {

BackusRuleItem({ ContinueWhile::Type()}, OnlyOne),

BackusRuleItem({ While::Type()}, OnlyOne)

});

auto whileCStatement = controller->addRule("WhileStatement", {

BackusRuleItem({ While::Type()}, OnlyOne),

BackusRuleItem({ Symbols::LBraket}, OnlyOne),

BackusRuleItem({ context->EquationRuleName()}, OnlyOne),

BackusRuleItem({ Symbols::RBraket}, OnlyOne),

BackusRuleItem({ Start::Type()}, OnlyOne),

BackusRuleItem({ operatorsName, operatorsWithSemicolonsName,whileContinueStatement->type(), whileExitStatement->type()}, Optional | OneOrMore),

BackusRuleItem({ End::Type()}, OnlyOne),

BackusRuleItem({ While::Type()}, OnlyOne),

});

whileCStatement->setPostHandler([](BackusRuleList::iterator& ruleBegin,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

static size\_t index = 0;

index++;

std::string startLabel = std::format("whileStart{}", index);

std::string endLabel = std::format("whileEnd{}", index);

(\*ruleBegin)->setCustomData(startLabel, "startLabel");

(\*ruleBegin)->setCustomData(endLabel, "endLabel");

size\_t count = 0;

for (auto itr = ruleBegin; itr != it; ++itr)

{

auto type = (\*itr)->type();

if (type == lEnd && itr != it && (\*std::next(itr, 1))->type() == While::Type())

{

(\*std::next(itr, 1))->setCustomData("true", "noGenerateCode");

}

if (type == lStart)

{

count++;

}

else if (type == lEnd && count == 1)

{

(\*itr)->setCustomData(std::format("\tjmp {}\n{}:", startLabel, endLabel));

break;

}

else if (type == ExitWhile::Type() && count == 1 && itr != it && (\*std::next(itr, 1))->type() == While::Type())

{

itr++;

(\*itr)->setCustomData("true", "ExitWhile");

(\*itr)->setCustomData(endLabel, "endLabel");

}

else if(type == ContinueWhile::Type() && count == 1 && itr != it && (\*std::next(itr, 1))->type() == While::Type())

{

itr++;

(\*itr)->setCustomData("true", "ContinueWhile");

(\*itr)->setCustomData(startLabel, "startLabel");

}

else if (type == lEnd && count > 0)

{

count--;

}

}

});

return whileCStatement;

}

Tokens:

Comment.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Generator/GeneratorItemBase.h"

class Comment : public TokenBase<Comment>, public GeneratorItemBase<Comment>

{

BASE\_ITEM

public:

Comment() { setLexeme(""); };

virtual ~Comment() = default;

std::shared\_ptr<IToken> tryCreateToken(std::string& lexeme) const override

{

auto token = clone();

token->setValue(lexeme);

lexeme.clear();

return token;

};

};

KWords:

Program.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Program : public TokenBase<Program>, public BackusRuleBase<Program>, public GeneratorItemBase<Program>

{

BASE\_ITEM

public:

Program() { setLexeme("NAME"); };

virtual ~Program() = default;

};

Vars.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Vars : public TokenBase<Vars>, public BackusRuleBase<Vars>, public GeneratorItemBase<Vars>

{

BASE\_ITEM

public:

Vars() { setLexeme("DATA"); };

virtual ~Vars() = default;

};

General:

EndOfFile.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class EndOfFile : public TokenBase<EndOfFile>, public BackusRuleBase<EndOfFile>, public GeneratorItemBase<EndOfFile>

{

BASE\_ITEM

public:

EndOfFile() { setLexeme(""); };

virtual ~EndOfFile() = default;

};

**Rules:**

AssigmentRules:

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Assignment : public TokenBase<Assignment>, public BackusRuleBase<Assignment>, public GeneratorItemBase<Assignment>

{

BASE\_ITEM

public:

Assignment() { setLexeme("<-"); };

virtual ~Assignment() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

auto ident = \*std::prev(it);

it++;

auto postForm = GeneratorUtils::Instance()->ConvertToPostfixForm(it, end);

auto postIt = postForm.begin();

auto postEnd = postForm.end();

for (const auto& item : postForm)

item->genCode(out, details, postIt, postEnd);

out << "\tpop " << ident->customData() << std::endl;

};

};

#include "stdafx.h"

#include "AssignmentRule.h"

#include "Rules/AssignmentRule/Assignment.h"

BackusRulePtr MakeAssignmentRule(std::shared\_ptr<Controller> controller)

{

controller->regItem<Assignment>();

auto context = controller->context();

auto assingmentRule = controller->addRule(context->AssignmentRuleName(), {

BackusRuleItem({ context->IdentRuleName()}, OnlyOne),

BackusRuleItem({ Assignment::Type()}, OnlyOne),

BackusRuleItem({ context->EquationRuleName()}, OnlyOne)

});

return assingmentRule;

}

EquationRules:

Arithmetic:

Adittion.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Addition : public TokenBase<Addition>, public BackusRuleBase<Addition>, public GeneratorItemBase<Addition>

{

BASE\_ITEM

public:

Addition() { setLexeme("ADD"); };

virtual ~Addition() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Add\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Add\_", PrintAdd);

SetRegistered();

}

}

private:

static void PrintAdd(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Add============================================================================\n";

out << "Add\_ PROC\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tadd " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Add\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Subtraction.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Subtraction : public TokenBase<Subtraction>, public BackusRuleBase<Subtraction>, public GeneratorItemBase<Subtraction>

{

BASE\_ITEM

public:

Subtraction() { setLexeme("SUB"); };

virtual ~Subtraction() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Sub\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Sub\_", PrintSub);

SetRegistered();

}

}

private:

static void PrintSub(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Sub============================================================================\n";

out << "Sub\_ PROC\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tsub " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Sub\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Compare:

Equal.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Equal : public TokenBase<Equal>, public BackusRuleBase<Equal>, public GeneratorItemBase<Equal>

{

BASE\_ITEM

public:

Equal() { setLexeme("EQ"); };

virtual ~Equal() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Equal\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Equal\_", PrintEqual);

SetRegistered();

}

}

private:

static void PrintEqual(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Equal==========================================================================\n";

out << "Equal\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tjne equal\_false\n";

out << "\tmov " << args.regPrefix << "ax, 1\n";

out << "\tjmp equal\_fin\n";

out << "equal\_false:\n";

out << "\tmov " << args.regPrefix << "ax, 0\n";

out << "equal\_fin:\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Equal\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Greate.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Greate : public TokenBase<Greate>, public BackusRuleBase<Greate>, public GeneratorItemBase<Greate>

{

BASE\_ITEM

public:

Greate() { setLexeme(">="); };

virtual ~Greate() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Greate\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Greate\_", PrintGreate);

SetRegistered();

}

}

private:

static void PrintGreate(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Greate=========================================================================\n";

out << "Greate\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tjle greate\_false\n";

out << "\tmov " << args.regPrefix << "ax, 1\n";

out << "\tjmp greate\_fin\n";

out << "greate\_false:\n";

out << "\tmov " << args.regPrefix << "ax, 0\n";

out << "greate\_fin:\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Greate\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Less.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Less : public TokenBase<Less>, public BackusRuleBase<Less>, public GeneratorItemBase<Less>

{

BASE\_ITEM

public:

Less() { setLexeme("<="); };

virtual ~Less() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Less\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Less\_", PrintLess);

SetRegistered();

}

}

private:

static void PrintLess(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Less===========================================================================\n";

out << "Less\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tjge less\_false\n";

out << "\tmov " << args.regPrefix << "ax, 1\n";

out << "\tjmp less\_fin\n";

out << "less\_false:\n";

out << "\tmov " << args.regPrefix << "ax, 0\n";

out << "less\_fin:\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Less\_ ENDP\n";

out << ";============================================================================================\n";

}

};

NotEqual.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

#include "Rules/EquationRule/Equal.h"

#include "Rules/EquationRule/Not.h"

class NotEqual : public TokenBase<NotEqual>, public BackusRuleBase<NotEqual>, public GeneratorItemBase<NotEqual>

{

BASE\_ITEM

public:

NotEqual() { setLexeme("NE"); };

virtual ~NotEqual() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

Equal::RegPROC(details);

Not::RegPROC(details);

out << "\tcall Equal\_\n";

out << "\tcall Not\_\n";

};

};

Logic:

And.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class And : public TokenBase<And>, public BackusRuleBase<And>, public GeneratorItemBase<And>

{

BASE\_ITEM

public:

And() { setLexeme("AND"); };

virtual ~And() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall And\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("And\_", PrintAnd);

SetRegistered();

}

}

private:

static void PrintAnd(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure And============================================================================\n";

out << "And\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjnz and\_t1\n";

out << "\tjz and\_false\n";

out << "and\_t1:\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjnz and\_true\n";

out << "and\_false:\n";

out << "\tmov " << args.regPrefix << "ax, 0\n";

out << "\tjmp and\_fin\n";

out << "and\_true:\n";

out << "\tmov " << args.regPrefix << "ax, 1\n";

out << "and\_fin:\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "And\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Not.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Not : public TokenBase<Not>, public BackusRuleBase<Not>, public GeneratorItemBase<Not>

{

BASE\_ITEM

public:

Not() { setLexeme("NOT"); };

virtual ~Not() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Not\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Not\_", PrintNot);

SetRegistered();

}

}

private:

static void PrintNot(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Not============================================================================\n";

out << "Not\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjnz not\_false\n";

out << "not\_t1:\n";

out << "\tmov " << args.regPrefix << "ax, 1\n";

out << "\tjmp not\_fin\n";

out << "not\_false:\n";

out << "\tmov " << args.regPrefix << "ax, 0\n";

out << "not\_fin:\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

out << "\tmov [esp + " << args.posArg1 << "], " << args.regPrefix << "ax\n";

out << "\tret\n";

out << "Not\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Or.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Or : public TokenBase<Or>, public BackusRuleBase<Or>, public GeneratorItemBase<Or>

{

BASE\_ITEM

public:

Or() { setLexeme("OR"); };

virtual ~Or() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Or\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Or\_", PrintOr);

SetRegistered();

}

}

private:

static void PrintOr(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Or=============================================================================\n";

out << "Or\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjnz or\_true\n";

out << "\tjz or\_t1\n";

out << "or\_t1:\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjnz or\_true\n";

out << "or\_false:\n";

out << "\tmov " << args.regPrefix << "ax, 0\n";

out << "\tjmp or\_fin\n";

out << "or\_true:\n";

out << "\tmov " << args.regPrefix << "ax, 1\n";

out << "or\_fin:\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Or\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Mult:

Division.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Division : public TokenBase<Division>, public BackusRuleBase<Division>, public GeneratorItemBase<Division>

{

BASE\_ITEM

public:

Division() { setLexeme("DIV"); };

virtual ~Division() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Div\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerStringData("DivErrMsg", "\\n" + Type() + ": Error: division by zero");

details.registerProc("Div\_", PrintDiv);

SetRegistered();

}

}

private:

static void PrintDiv(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Div============================================================================\n";

out << "Div\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjne end\_check\n";

out << "\tinvoke WriteConsoleA, hConsoleOutput, ADDR DivErrMsg, SIZEOF DivErrMsg - 1, 0, 0\n";

out << "\tjmp exit\_label\n";

out << "end\_check:\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjge gr\n";

out << "lo:\n";

out << "\tmov " << args.regPrefix << "dx, -1\n";

out << "\tjmp less\_fin\n";

out << "gr:\n";

out << "\tmov " << args.regPrefix << "dx, 0\n";

out << "less\_fin:\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tidiv " << args.numberTypeExtended << " ptr [esp + " << args.posArg1 << "]\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Div\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Mod.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Mod : public TokenBase<Mod>, public BackusRuleBase<Mod>, public GeneratorItemBase<Mod>

{

BASE\_ITEM

public:

Mod() { setLexeme("MOD"); };

virtual ~Mod() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Mod\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerStringData("ModErrMsg", "\\n" + Type() + ": Error: division by zero");

details.registerProc("Mod\_", PrintMod);

SetRegistered();

}

}

private:

static void PrintMod(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Mod============================================================================\n";

out << "Mod\_ PROC\n";

out << "\tpushf\n";

out << "\tpop cx\n\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg1 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjne end\_check\n";

out << "\tinvoke WriteConsoleA, hConsoleOutput, ADDR ModErrMsg, SIZEOF ModErrMsg - 1, 0, 0\n";

out << "\tjmp exit\_label\n";

out << "end\_check:\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tcmp " << args.regPrefix << "ax, 0\n";

out << "\tjge gr\n";

out << "lo:\n";

out << "\tmov " << args.regPrefix << "dx, -1\n";

out << "\tjmp less\_fin\n";

out << "gr:\n";

out << "\tmov " << args.regPrefix << "dx, 0\n";

out << "less\_fin:\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\tidiv " << args.numberTypeExtended << " ptr [esp + " << args.posArg1 << "]\n";

out << "\tmov " << args.regPrefix << "ax, " << args.regPrefix << "dx\n";

out << "\tpush cx\n";

out << "\tpopf\n\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Mod\_ ENDP\n";

out << ";============================================================================================\n";

}

};

Multiplication.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Multiplication : public TokenBase<Multiplication>, public BackusRuleBase<Multiplication>, public GeneratorItemBase<Multiplication>

{

BASE\_ITEM

public:

Multiplication() { setLexeme("MUL"); };

virtual ~Multiplication() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

out << "\tcall Mul\_\n";

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerProc("Mul\_", PrintMul);

SetRegistered();

}

}

private:

static void PrintMul(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Mul============================================================================\n";

out << "Mul\_ PROC\n";

out << "\tmov " << args.regPrefix << "ax, [esp + " << args.posArg0 << "]\n";

out << "\timul " << args.numberTypeExtended << " ptr [esp + " << args.posArg1 << "]\n";

GeneratorUtils::PrintResultToStack(out, args);

out << "\tret\n";

out << "Mul\_ ENDP\n";

out << ";============================================================================================\n";

}

};

EquationRule.cpp

#include "stdafx.h"

#include "EquationRule.h"

#include "Rules/EquationRule/Number.h"

#include "Rules/EquationRule/Addition.h"

#include "Rules/EquationRule/Subtraction.h"

#include "Rules/EquationRule/Multiplication.h"

#include "Rules/EquationRule/Division.h"

#include "Rules/EquationRule/Mod.h"

#include "Rules/EquationRule/And.h"

#include "Rules/EquationRule/Or.h"

#include "Rules/EquationRule/Equal.h"

#include "Rules/EquationRule/Greate.h"

#include "Rules/EquationRule/Less.h"

#include "Rules/EquationRule/NotEqual.h"

#include "Rules/EquationRule/Not.h"

BackusRulePtr MakeEquationRule(std::shared\_ptr<Controller> controller)

{

using enum ItemType;

controller->regItem<Number>(TokenAndRule | Operand, 0);

controller->regItem <Addition>(TokenAndRule | Operation, 4);

controller->regItem <Subtraction>(TokenAndRule | Operation, 4);

controller->regItem<Multiplication>(TokenAndRule | Operation, 5);

controller->regItem <Division>(TokenAndRule | Operation, 5);

controller->regItem <Mod>(TokenAndRule | Operation, 5);

controller->regItem <And>(TokenAndRule | Operation, 1);

controller->regItem <Or>(TokenAndRule | Operation, 1);

controller->regItem <Equal>(TokenAndRule | Operation, 2);

controller->regItem <NotEqual>(TokenAndRule | Operation, 2);

controller->regItem <Greate>(TokenAndRule | Operation, 3);

controller->regItem <Less>(TokenAndRule | Operation, 3);

controller->regItem <Not>(TokenAndRule | Operation, 6);

auto context = controller->context();

auto equationRuleName = context->EquationRuleName();

auto sign = controller->addRule("Sign", { BackusRuleItem({ Symbols::Plus, Symbols::Minus }, Optional) });

auto signedNumber = controller->addRule("SignedNumber", {

BackusRuleItem({ sign->type()}, Optional),

BackusRuleItem({Number::Type()}, OnlyOne)

});

signedNumber->setPostHandler([](BackusRuleList::iterator&,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

auto begRuleIt = std::prev(it, 2);

if ((\*begRuleIt)->type() == Symbols::Minus)

{

it = begRuleIt;

end = std::remove(it, end, \*it);

(\*it)->setValue('-' + (\*it)->value());

it++;

}

});

auto arithmetic = controller->addRule("Arithmetic", { BackusRuleItem({ Addition::Type(), Subtraction::Type() }, OnlyOne) });

auto mult = controller->addRule("Mult", { BackusRuleItem({ Multiplication::Type(), Division::Type(), Mod::Type() }, OnlyOne) });

auto logic = controller->addRule("Logic", { BackusRuleItem({ And::Type(), Or::Type() }, OnlyOne) });

auto compare = controller->addRule("Compare", { BackusRuleItem({ Equal::Type(), Greate::Type(), Less::Type(), NotEqual::Type() }, OnlyOne) });

auto operationAndEquation = controller->addRule("OperationAndEquation", {

BackusRuleItem({ mult->type(), arithmetic->type(), logic->type(), compare->type() }, OnlyOne),

BackusRuleItem({ equationRuleName }, OnlyOne)

});

auto notRule = controller->addRule("NotRule", {

BackusRuleItem({ Not::Type()}, OnlyOne),

BackusRuleItem({ equationRuleName}, Optional | OneOrMore)

});

auto equationWithBrakets = controller->addRule("EquationWithBrakets", {

BackusRuleItem({ Symbols::LBraket }, OnlyOne | PairStart),

BackusRuleItem({ equationRuleName }, OnlyOne),

BackusRuleItem({ Symbols::RBraket }, OnlyOne | PairEnd)

});

auto equation = controller->addRule(equationRuleName, {

BackusRuleItem({signedNumber->type(), context->IdentRuleName(), notRule->type(), equationWithBrakets->type()}, OnlyOne),

BackusRuleItem({operationAndEquation->type()}, Optional | OneOrMore)

});

return equation;

}

IdentRule:

Identifier.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

#include "Rules/AssignmentRule/Assignment.h"

#include "Tokens/Common/EndOfFile.h"

class Identifier : public TokenBase<Identifier>, public BackusRuleBase<Identifier>, public GeneratorItemBase<Identifier>

{

BASE\_ITEM

public:

Identifier() { setLexeme(""); };

virtual ~Identifier() = default;

std::shared\_ptr<IToken> tryCreateToken(std::string& lexeme) const override

{

if (lexeme.size() > (m\_mask.size() + m\_prefix.size()))

return nullptr;

bool res = true;

if (!lexeme.starts\_with(m\_prefix))

{

return nullptr;

}

std::string\_view ident{ lexeme.begin() + m\_prefix.size(), lexeme.end() };

for (size\_t i = 0; i < ident.size(); i++)

{

if ((isupper(ident[i]) != isupper(m\_mask[i])) && !isdigit(ident[i]))

{

res &= false;

break;

}

}

std::shared\_ptr<IToken> token = nullptr;

if (res)

{

token = clone();

token->setValue(lexeme);

lexeme.clear();

}

return token;

};

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

if (!GeneratorUtils::IsNextTokenIs(it, end, Assignment::Type()))

{

if ((\*std::prev(end))->type() == EndOfFile::Type())

details.registerNumberData(customData());

else

out << "\tpush " << customData() << std::endl;

}

};

private:

const std::string m\_prefix = "\_";

const std::string m\_mask = "xXXXXXX";

};

IdentRule.cpp

#include "stdafx.h"

#include "IdentRule.h"

#include "Rules/IdentRule/Identifier.h"

#include "Rules/IdentRule/Undefined.h"

SimpleToken(ProgramName, "");

BackusRulePtr MakeIdentRule(std::shared\_ptr<Controller> controller)

{

using enum ItemType;

controller->regItem<Identifier>(TokenAndRule, -1);

GeneratorUtils::Instance()->RegisterOperand(Identifier::Type());

auto context = controller->context();

auto identRule = controller->addRule(context->IdentRuleName(), {

BackusRuleItem({ Identifier::Type()}, OnlyOne)

});

identRule->setPostHandler([context](BackusRuleList::iterator&,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

static bool isFirstIdentChecked = !context->IsFirstProgName();

auto isVarBlockChecked = context->IsVarBlockChecked();

auto& identTable = context->IdentTable();

auto identIt = std::prev(it, 1);

if (isVarBlockChecked)

{

if (!identTable.contains((\*identIt)->value()))

{

auto undef = std::make\_shared<Undefined>();

undef->setValue((\*identIt)->value());

undef->setLine((\*identIt)->line());

undef->setCustomData((\*identIt)->customData());

\*identIt = undef;

}

}

else

{

if (isFirstIdentChecked)

{

identTable.insert((\*identIt)->value());

}

else

{

auto progName = std::make\_shared<ProgramName>();

progName->setValue((\*identIt)->value());

progName->setLine((\*identIt)->line());

progName->setCustomData((\*identIt)->customData());

\*identIt = progName;

isFirstIdentChecked = true;

}

}

(\*identIt)->setCustomData((\*identIt)->value() + "\_");

});

return identRule;

}

Undefined.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Undefined : public TokenBase<Undefined>, public BackusRuleBase<Undefined>, public GeneratorItemBase<Undefined>

{

BASE\_ITEM

public:

Undefined() { setLexeme(""); };

virtual ~Undefined() = default;

std::shared\_ptr<IToken> tryCreateToken(std::string& lexeme) const override

{

auto token = clone();

token->setValue(lexeme);

lexeme.clear();

return token;

};

};

ReadRule:

Read.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class Read :public TokenBase<Read>, public BackusRuleBase<Read>, public GeneratorItemBase<Read>

{

BASE\_ITEM

public:

Read() { setLexeme("SCAN"); };

virtual ~Read() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

RegPROC(details);

it = std::next(it, 2);

out << "\tcall Input\_\n";

out << "\tmov " << (\*it)->customData() << ", " << details.args().regPrefix << "ax\n";

it = std::next(it, 2);

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerRawData("InputBuf", "\tdb\t15 dup (?)");

details.registerRawData("CharsReadNum", "dd\t?");

details.registerProc("Input\_", PrintInput);

SetRegistered();

}

}

private:

static void PrintInput(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Input==========================================================================\n";

out << "Input\_ PROC\n";

out << "\tinvoke ReadConsoleA, hConsoleInput, ADDR InputBuf, 13, ADDR CharsReadNum, 0\n";

out << "\tinvoke crt\_atoi, ADDR InputBuf\n";

out << "\tret\n";

out << "Input\_ ENDP\n";

out << ";============================================================================================\n";

}

};

ReadRule.cpp

#include "stdafx.h"

#include "ReadRule.h"

#include "Rules/ReadRule/Read.h"

BackusRulePtr MakeReadRule(std::shared\_ptr<Controller> controller)

{

controller->regItem<Read>();

auto context = controller->context();

auto read = controller->addRule("ReadRule", {

BackusRuleItem({ Read::Type()}, OnlyOne),

BackusRuleItem({ Symbols::LBraket}, OnlyOne),

BackusRuleItem({ context->IdentRuleName()}, OnlyOne),

BackusRuleItem({ Symbols::RBraket}, OnlyOne)

});

return read;

}

StringRule:

String.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class String : public TokenBase<String>, public BackusRuleBase<String>, public GeneratorItemBase<String>

{

BASE\_ITEM

public:

String() { setLexeme(""); };

virtual ~String() = default;

std::string stringName() const { return m\_stringName; };

std::shared\_ptr<IToken> tryCreateToken(std::string& lexeme) const override

{

auto token = clone();

token->setValue(lexeme);

lexeme.clear();

return token;

};

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

m\_stringName = std::format("String\_{}", index++);

details.registerStringData(m\_stringName, value());

};

private:

mutable std::string m\_stringName;

static size\_t index;

};

StringRule.cpp

#include "stdafx.h"

#include "StringRule.h"

#include "Rules/StringRule/String.h"

SimpleToken(Quotes, "\"");

BackusRulePtr MakeStringRule(std::shared\_ptr<Controller> controller)

{

using enum ItemType;

controller->regUnchangedTextToken(std::make\_shared<String>(), std::make\_shared<Quotes>(), std::make\_shared<Quotes>());

controller->regItem<Quotes>(Rule);

controller->regItem<String>(Rule);

auto stringRule = controller->addRule("StringRule", {

BackusRuleItem({ Quotes::Type()}, OnlyOne),

BackusRuleItem({ String::Type()}, OnlyOne),

BackusRuleItem({ Quotes::Type()}, OnlyOne)

});

stringRule->setPostHandler([](BackusRuleList::iterator&,

BackusRuleList::iterator& it,

BackusRuleList::iterator& end)

{

it = std::prev(it, 3);

end = std::remove(it, end, \*it);

it++;

end = std::remove(it, end, \*it);

});

return stringRule;

}

VarsBlockRule:

VarsBlockRule.cpp

#include "stdafx.h"

#include "VarsBlokRule.h"

#include "Rules/VarsBlokRule/VarType.h"

BackusRulePtr MakeVarsBlokRule(std::shared\_ptr<Controller> controller)

{

controller->regItem<VarType>();

auto context = controller->context();

auto commaAndIdentifier = controller->addRule("CommaAndIdentifier", {

BackusRuleItem({ Symbols::Comma}, OnlyOne),

BackusRuleItem({ context->IdentRuleName()}, OnlyOne)

});

auto varsBlok = controller->addRule("VarsBlok", {

BackusRuleItem({ VarType::Type()}, OnlyOne),

BackusRuleItem({ context->IdentRuleName()}, OnlyOne),

BackusRuleItem({commaAndIdentifier->type()}, Optional | OneOrMore),

BackusRuleItem({ Symbols::Semicolon}, OnlyOne)

});

varsBlok->setPostHandler([context](BackusRuleList::iterator&, BackusRuleList::iterator&, BackusRuleList::iterator&)

{

auto isVarBlockChecked = context->IsVarBlockChecked();

context->SetVarBlockChecked();

});

return varsBlok;

}

VarType.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

class VarType : public TokenBase<VarType>, public BackusRuleBase<VarType>, public GeneratorItemBase<VarType>

{

BASE\_ITEM

public:

VarType() { setLexeme("INTEGER\_2"); };

virtual ~VarType() = default;

};

WriteRule:

Write.h

#pragma once

#include "stdafx.h"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRuleBase.h"

#include "Core/Generator/GeneratorItemBase.h"

#include "Rules/StringRule/String.h"

class Write : public TokenBase<Write>, public BackusRuleBase<Write>, public GeneratorItemBase<Write>

{

BASE\_ITEM

public:

Write() { setLexeme("PRINT"); };

virtual ~Write() = default;

void genCode(std::ostream& out, GeneratorDetails& details,

std::list<std::shared\_ptr<IGeneratorItem>>::iterator& it,

const std::list<std::shared\_ptr<IGeneratorItem>>::iterator& end) const final

{

if (auto string = std::dynamic\_pointer\_cast<String>(\*std::next(it, 2)))

{

it = std::next(it, 2);

string->genCode(out, details, it, end);

it = std::next(it, 2);

out << "\tinvoke WriteConsoleA, hConsoleOutput, ADDR " << string->stringName() << ", SIZEOF " << string->stringName() << " - 1, 0, 0\n";

}

else

{

RegPROC(details);

auto postForm = GeneratorUtils::Instance()->ConvertToPostfixForm(it, end);

auto postIt = postForm.begin();

auto postEnd = postForm.end();

for (const auto& item : postForm)

item->genCode(out, details, postIt, postEnd);

out << "\tcall Output\_\n";

}

};

static void RegPROC(GeneratorDetails& details)

{

if (!IsRegistered())

{

details.registerRawData("OutMessage", "\tdb\t\"" + details.args().numberStrType + "\", 0");

details.registerRawData("ResMessage", "\tdb\t20 dup (?)");

details.registerProc("Output\_", PrintOutput);

SetRegistered();

}

}

private:

static void PrintOutput(std::ostream& out, const GeneratorDetails::GeneratorArgs& args)

{

out << ";===Procedure Output=========================================================================\n";

out << "Output\_ PROC value: " << args.numberTypeExtended.c\_str() << std::endl;

out << "\tinvoke wsprintf, ADDR ResMessage, ADDR OutMessage, value\n";

out << "\tinvoke WriteConsoleA, hConsoleOutput, ADDR ResMessage, eax, 0, 0\n";

out << "\tret " << args.argSize << std::endl;

out << "Output\_ ENDP\n";

out << ";============================================================================================\n";

}

};

WriteRule.cpp

#include "stdafx.h"

#include "WriteRule.h"

#include "Rules/StringRule/StringRule.h"

#include "Rules/WriteRule/Write.h"

BackusRulePtr MakeWriteRule(std::shared\_ptr<Controller> controller)

{

controller->regItem<Write>();

auto context = controller->context();

auto stringRule = MakeStringRule(controller);

auto write = controller->addRule("WriteRule", {

BackusRuleItem({ Write::Type()}, OnlyOne),

BackusRuleItem({ Symbols::LBraket}, OnlyOne | PairStart),

BackusRuleItem({ stringRule->type(), context->EquationRuleName() }, OnlyOne),

BackusRuleItem({ Symbols::RBraket}, OnlyOne | PairEnd)

});

return write;

}

Controller.h

#pragma once

#include "stdafx.h"

#include "Utils/singleton.hpp"

#include "Core/Tokens/TokenBase.hpp"

#include "Core/Backus/BackusRule.h"

#include "Core/Generator/GeneratorItemBase.h"

#include "Core/Generator/GeneratorDetails.h"

#include "Symbols.h"

using BackusRulePtr = std::shared\_ptr<IBackusRule>;

using BackusRuleList = std::list<BackusRulePtr>;

using BackusRuleListIt = BackusRuleList::iterator;

using RuleMaker = std::function<BackusRulePtr(std::shared\_ptr<Controller>)>;

class Context

{

public:

std::string IdentRuleName() const { return "IdentRule"; }

std::string EquationRuleName() const { return "Equation"; }

std::string OperatorsRuleName() const { return "OperatorsRule"; }

std::string OperatorsName() const { return "Operators"; }

std::string OperatorsWithSemicolonsName() const { return "OperatorsWithSemicolon"; }

std::string AssignmentRuleName() const { return "AssignmentRule"; }

std::tuple<std::string, std::string, std::string> CodeBlockTypes() const { return { "Start", "CodeBlok", "End" }; }

bool IsVarBlockChecked() const { return m\_isVarBlockChecked; }

void SetVarBlockChecked() { m\_isVarBlockChecked = true; }

bool IsFirstProgName() const { return true; }

std::set<std::string>& IdentTable() { return m\_identTable; }

const GeneratorDetails& Details() const { return m\_details; }

private:

std::set<std::string> m\_identTable{};

bool m\_isVarBlockChecked = false;

const GeneratorDetails m\_details{ {

.numberType = "dw", .numberTypeExtended = "word",

.argSize = 2,

.numberStrType = "%hd"

} };

};

enum class ItemType : uint32\_t

{

None = 0,

Token = 1 << 0,

Rule = 1 << 1,

TokenAndRule = Token | Rule,

Operand = 1 << 2,

Operation = 1 << 3,

EquationEnd = 1 << 4,

LBracket = 1 << 5,

RBracket = 1 << 6

};

DEFINE\_ENUM\_FLAG\_OPERATORS(ItemType)

class Controller : public singleton<Controller>

{

public:

static constexpr int NoPriority = std::numeric\_limits<int>::min();

public:

void init();

template<typename T>

void regItem(ItemType type = ItemType::TokenAndRule, int priority = NoPriority) const

{

auto item = std::make\_shared<T>();

regItem(item, item, type, priority);

}

void regUnchangedTextToken(std::shared\_ptr<IToken> target, std::shared\_ptr<IToken> lBorder, std::shared\_ptr<IToken> rBorder) const;

void regOperatorRule(const RuleMaker& rule, bool isNeedSemicolon = false);

std::shared\_ptr<IBackusRule> addRule(const std::string& name, const std::list<BackusRuleItem>& items) const;

BackusRulePtr topRule();

std::shared\_ptr<Context> context() { return m\_context; }

protected:

Controller() { m\_context = std::make\_shared<Context>(); }

void regItem(std::shared\_ptr<IToken> token, std::shared\_ptr<IBackusRule> rule, ItemType type, int priority) const;

BackusRulePtr MakeTopRule(std::shared\_ptr<Controller> controller) const;

private:

BackusRulePtr m\_topRule;

std::set<std::string> m\_operatorRuleNames;

std::set<std::string> m\_operatorRuleWithSemicolonNames;

std::shared\_ptr<Context> m\_context;

};

Controller.cpp

#include "stdafx.h"

#include "Controller.h"

#include "Core/Parser/TokenParser.h"

#include "Core/Backus/BackusRuleStorage.h"

#include "Core/Generator/Generator.h"

#include "Core/Generator/GeneratorUtils.h"

#include "SimpleTokens.h"

#include "Tokens/Common/Program.h"

#include "Tokens/Common/Vars.h"

#include "Rules/IdentRule/IdentRule.h"

#include "Rules/VarsBlokRule/VarsBlokRule.h"

#include "Rules/EquationRule/EquationRule.h"

#include "Rules/ReadRule/ReadRule.h"

#include "Rules/WriteRule/WriteRule.h"

#include "Rules/AssignmentRule/AssignmentRule.h"

void Controller::regItem(std::shared\_ptr<IToken> token, std::shared\_ptr<IBackusRule> rule, ItemType type, int priority) const

{

using enum ItemType;

if ((type & Token) == Token)

TokenParser::Instance()->regToken(token, ((type & Operation) == Operation) ? TokenParser::NoPriority : priority);

if ((type & Rule) == Rule)

BackusRuleStorage::Instance()->regRule(rule);

auto tokenType = token->type();

if ((type & Operand) == Operand)

GeneratorUtils::Instance()->RegisterOperand(tokenType);

if ((type & Operation) == Operation)

{

if (priority == TokenParser::NoPriority)

throw std::runtime\_error("Controller::RegItem: Operation " + token->type() + " priority is not set");

GeneratorUtils::Instance()->RegisterOperation(tokenType, priority);

}

if ((type & EquationEnd) == EquationEnd)

GeneratorUtils::Instance()->RegisterEquationEnd(tokenType);

if ((type & LBracket) == LBracket)

GeneratorUtils::Instance()->RegisterLBraket(tokenType);

if ((type & RBracket) == RBracket)

GeneratorUtils::Instance()->RegisterRBraket(tokenType);

}

void Controller::init()

{

m\_topRule = MakeTopRule(Instance());

Generator::Instance()->setDetails(context()->Details());

}

void Controller::regUnchangedTextToken(std::shared\_ptr<IToken> target, std::shared\_ptr<IToken> lBorder, std::shared\_ptr<IToken> rBorder) const

{

TokenParser::Instance()->regUnchangedTextToken(target, lBorder, rBorder);

}

void Controller::regOperatorRule(const RuleMaker& rule, bool isNeedSemicolon)

{

auto ruleName = rule(Instance())->type();

if (ruleName.empty())

throw std::runtime\_error("Controller::RegOperatorRule: Rule name is empty");

if (m\_operatorRuleNames.contains(ruleName) || m\_operatorRuleWithSemicolonNames.contains(ruleName))

throw std::runtime\_error(std::format("Controller::RegOperatorRule: Rule with name {} already registered", ruleName));

if (isNeedSemicolon)

m\_operatorRuleWithSemicolonNames.insert(ruleName);

else

m\_operatorRuleNames.insert(ruleName);

}

std::shared\_ptr<IBackusRule> Controller::addRule(const std::string& name, const std::list<BackusRuleItem>& items) const

{

auto rule = BackusRule::MakeRule(name, items);

BackusRuleStorage::Instance()->regRule(rule);

return rule;

}

BackusRulePtr Controller::topRule()

{

if (!m\_topRule)

throw(std::runtime\_error("Controller is not inited"));

return m\_topRule;

}

BackusRulePtr Controller::MakeTopRule(std::shared\_ptr<Controller> controller) const

{

using enum ItemType;

controller->regItem<Program>();

controller->regItem<Vars>();

controller->regItem<Start>(TokenAndRule | EquationEnd);

controller->regItem<End>();

controller->regItem<Comma>();

controller->regItem<Colon>();

controller->regItem<Semicolon>(TokenAndRule | EquationEnd);

controller->regItem<LBraket>(TokenAndRule | LBracket);

controller->regItem<RBraket>(TokenAndRule | RBracket);

controller->regItem<Plus>();

controller->regItem<Minus>();

auto identRule = MakeIdentRule(controller);

auto varsBlok = MakeVarsBlokRule(controller);

auto equation = MakeEquationRule(controller);

auto read = MakeReadRule(controller);

auto write = MakeWriteRule(controller);

auto assingmentRule = MakeAssignmentRule(controller);

auto operatorWithSemicolonTypes = std::vector<std::variant<std::string, Symbols>>{ read->type(), write->type(), assingmentRule->type() };

operatorWithSemicolonTypes.insert(operatorWithSemicolonTypes.end(), m\_operatorRuleWithSemicolonNames.begin(), m\_operatorRuleWithSemicolonNames.end());

auto operatorsWithSemicolon = controller->addRule("OperatorsWithSemicolon", {

BackusRuleItem({ operatorWithSemicolonTypes }, OnlyOne),

BackusRuleItem({ Symbols::Semicolon }, OnlyOne)

});

auto operatorTypes = std::vector<std::variant<std::string, Symbols>>{ m\_operatorRuleNames.begin(), m\_operatorRuleNames.end() };

auto operators = controller->addRule("Operators", {

BackusRuleItem({ operatorTypes }, OnlyOne)

});

auto operatorsRule = controller->addRule("OperatorsRule", {

BackusRuleItem({ operators->type(), operatorsWithSemicolon->type()}, Optional | OneOrMore),

});

auto codeBlok = controller->addRule("CodeBlok", {

BackusRuleItem({ Start::Type()}, OnlyOne),

BackusRuleItem({ operators->type(), operatorsWithSemicolon->type()}, Optional | OneOrMore),

BackusRuleItem({ End::Type()}, OnlyOne)

});

auto topRule = controller->addRule("TopRule", {

BackusRuleItem({ Program::Type()}, OnlyOne),

BackusRuleItem({ identRule->type()}, OnlyOne),

BackusRuleItem({ Symbols::Semicolon}, OnlyOne),

BackusRuleItem({ Start::Type()}, OnlyOne),

BackusRuleItem({ Vars::Type()}, OnlyOne),

BackusRuleItem({ varsBlok->type()}, OnlyOne),

BackusRuleItem({ operators->type(), operatorsWithSemicolon->type()}, Optional | OneOrMore),

BackusRuleItem({ End::Type()}, OnlyOne)

});

return topRule;

}

**Додаток В(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================================================================================

\_aaaa\_ dd 0

\_bbbb\_ dd 0

\_xxxx\_ dd 0

\_yyyy\_ dd 0

DivErrMsg db 13, 10, "Division: Error: division by zero", 0

ModErrMsg db 13, 10, "Mod: Error: division by zero", 0

String\_0 db "Input A: ", 0

String\_1 db "Input B: ", 0

String\_2 db "A + B: ", 0

String\_3 db 13, 10, "A - B: ", 0

String\_4 db 13, 10, "A \* B: ", 0

String\_5 db 13, 10, "A / B: ", 0

String\_6 db 13, 10, "A % B: ", 0

String\_7 db 13, 10, "X = (A - B) \* 10 + (A + B) / 10", 13, 10, 0

String\_8 db 13, 10, "Y = X + (X MOD 10)", 13, 10, 0

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

hConsoleInput dd ?

hConsoleOutput dd ?

endBuff db 5 dup (?)

msg1310 db 13, 10, 0

CharsReadNum dd ?

InputBuf db 15 dup (?)

OutMessage db "%d", 0

ResMessage db 20 dup (?)

.CODE

start:

invoke AllocConsole

invoke GetStdHandle, STD\_INPUT\_HANDLE

mov hConsoleInput, eax

invoke GetStdHandle, STD\_OUTPUT\_HANDLE

mov hConsoleOutput, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_0, SIZEOF String\_0 - 1, 0, 0

call Input\_

mov \_aaaaaaaa\_, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_1, SIZEOF String\_1 - 1, 0, 0

call Input\_

mov \_bbbbbbbb\_, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_2, SIZEOF String\_2 - 1, 0, 0

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Add\_

call Output\_

invoke WriteConsoleA, hConsoleOutput, ADDR String\_3, SIZEOF String\_3 - 1, 0, 0

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Sub\_

call Output\_

invoke WriteConsoleA, hConsoleOutput, ADDR String\_4, SIZEOF String\_4 - 1, 0, 0

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Mul\_

call Output\_

invoke WriteConsoleA, hConsoleOutput, ADDR String\_5, SIZEOF String\_5 - 1, 0, 0

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Div\_

call Output\_

invoke WriteConsoleA, hConsoleOutput, ADDR String\_6, SIZEOF String\_6 - 1, 0, 0

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Mod\_

call Output\_

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Sub\_

push dword ptr 10

call Mul\_

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Add\_

push dword ptr 10

call Div\_

call Add\_

pop \_xxxxxxxx\_

push \_xxxxxxxx\_

push \_xxxxxxxx\_

push dword ptr 10

call Mod\_

call Add\_

pop \_yyyyyyyy\_

invoke WriteConsoleA, hConsoleOutput, ADDR String\_7, SIZEOF String\_7 - 1, 0, 0

push \_xxxxxxxx\_

call Output\_

invoke WriteConsoleA, hConsoleOutput, ADDR String\_8, SIZEOF String\_8 - 1, 0, 0

push \_yyyyyyyy\_

call Output\_

exit\_label:

invoke WriteConsoleA, hConsoleOutput, ADDR msg1310, SIZEOF msg1310 - 1, 0, 0

invoke ReadConsoleA, hConsoleInput, ADDR endBuff, 5, 0, 0

invoke ExitProcess, 0

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

Add\_ PROC

mov eax, [esp + 8]

add eax, [esp + 4]

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Add\_ ENDP

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

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

Div\_ PROC

pushf

pop cx

mov eax, [esp + 4]

cmp eax, 0

jne end\_check

invoke WriteConsoleA, hConsoleOutput, ADDR DivErrMsg, SIZEOF DivErrMsg - 1, 0, 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]

idiv dword ptr [esp + 4]

push cx

popf

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Div\_ ENDP

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

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

Input\_ PROC

invoke ReadConsoleA, hConsoleInput, ADDR InputBuf, 13, ADDR CharsReadNum, 0

invoke crt\_atoi, ADDR InputBuf

ret

Input\_ ENDP

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

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

Mod\_ PROC

pushf

pop cx

mov eax, [esp + 4]

cmp eax, 0

jne end\_check

invoke WriteConsoleA, hConsoleOutput, ADDR ModErrMsg, SIZEOF ModErrMsg - 1, 0, 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]

idiv dword ptr [esp + 4]

mov eax, edx

push cx

popf

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Mod\_ ENDP

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

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

Mul\_ PROC

mov eax, [esp + 8]

imul dword ptr [esp + 4]

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Mul\_ ENDP

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

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

Output\_ PROC value: dword

invoke wsprintf, ADDR ResMessage, ADDR OutMessage, value

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

ret 4

Output\_ ENDP

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

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

Sub\_ PROC

mov eax, [esp + 8]

sub eax, [esp + 4]

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Sub\_ 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================================================================================

\_aaaa\_ dd 0

\_bbbb\_ dd 0

\_cccc\_ dd 0

String\_0 db "Input A: ", 0

String\_1 db "Input B: ", 0

String\_2 db "Input C: ", 0

String\_3 db 13, 10, 0

String\_4 db 13, 10, 0

String\_5 db 13, 10, 0

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

hConsoleInput dd ?

hConsoleOutput dd ?

endBuff db 5 dup (?)

msg1310 db 13, 10, 0

CharsReadNum dd ?

InputBuf db 15 dup (?)

OutMessage db "%d", 0

ResMessage db 20 dup (?)

.CODE

start:

invoke AllocConsole

invoke GetStdHandle, STD\_INPUT\_HANDLE

mov hConsoleInput, eax

invoke GetStdHandle, STD\_OUTPUT\_HANDLE

mov hConsoleOutput, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_0, SIZEOF String\_0 - 1, 0, 0

call Input\_

mov \_aaaaaaaa\_, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_1, SIZEOF String\_1 - 1, 0, 0

call Input\_

mov \_bbbbbbbb\_, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_2, SIZEOF String\_2 - 1, 0, 0

call Input\_

mov \_cccccccc\_, eax

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Greate\_

pop eax

cmp eax, 0

je endIf2

push \_aaaaaaaa\_

push \_cccccccc\_

call Greate\_

pop eax

cmp eax, 0

je elseLabel1

jmp \_temporal\_

jmp endIf1

elseLabel1:

push \_cccccccc\_

call Output\_

jmp \_outugoto\_

\_temporal\_:

push \_aaaaaaaa\_

call Output\_

jmp \_outugoto\_

endIf1:

endIf2:

push \_bbbbbbbb\_

push \_cccccccc\_

call Less\_

pop eax

cmp eax, 0

je elseLabel3

push \_cccccccc\_

call Output\_

jmp endIf3

elseLabel3:

push \_bbbbbbbb\_

call Output\_

endIf3:

\_outugoto\_:

invoke WriteConsoleA, hConsoleOutput, ADDR String\_3, SIZEOF String\_3 - 1, 0, 0

push \_aaaaaaaa\_

push \_bbbbbbbb\_

call Equal\_

push \_aaaaaaaa\_

push \_cccccccc\_

call Equal\_

call And\_

push \_bbbbbbbb\_

push \_cccccccc\_

call Equal\_

call And\_

pop eax

cmp eax, 0

je elseLabel4

push dword ptr 1

call Output\_

jmp endIf4

elseLabel4:

push dword ptr 0

call Output\_

endIf4:

invoke WriteConsoleA, hConsoleOutput, ADDR String\_4, SIZEOF String\_4 - 1, 0, 0

push \_aaaaaaaa\_

push dword ptr 0

call Less\_

push \_bbbbbbbb\_

push dword ptr 0

call Less\_

call Or\_

push \_cccccccc\_

push dword ptr 0

call Less\_

call Or\_

pop eax

cmp eax, 0

je elseLabel5

push dword ptr -1

call Output\_

jmp endIf5

elseLabel5:

push dword ptr 0

call Output\_

endIf5:

invoke WriteConsoleA, hConsoleOutput, ADDR String\_5, SIZEOF String\_5 - 1, 0, 0

push \_aaaaaaaa\_

push \_bbbbbbbb\_

push \_cccccccc\_

call Add\_

call Less\_

call Not\_

pop eax

cmp eax, 0

je elseLabel6

push dword ptr 10

call Output\_

jmp endIf6

elseLabel6:

push dword ptr 0

call Output\_

endIf6:

exit\_label:

invoke WriteConsoleA, hConsoleOutput, ADDR msg1310, SIZEOF msg1310 - 1, 0, 0

invoke ReadConsoleA, hConsoleInput, ADDR endBuff, 5, 0, 0

invoke ExitProcess, 0

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

Add\_ PROC

mov eax, [esp + 8]

add eax, [esp + 4]

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Add\_ 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

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

And\_ ENDP

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

;===Procedure Equal==========================================================================

Equal\_ PROC

pushf

pop cx

mov eax, [esp + 8]

cmp eax, [esp + 4]

jne equal\_false

mov eax, 1

jmp equal\_fin

equal\_false:

mov eax, 0

equal\_fin:

push cx

popf

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Equal\_ ENDP

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

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

Greate\_ PROC

pushf

pop cx

mov eax, [esp + 8]

cmp eax, [esp + 4]

jle greate\_false

mov eax, 1

jmp greate\_fin

greate\_false:

mov eax, 0

greate\_fin:

push cx

popf

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Greate\_ ENDP

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

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

Input\_ PROC

invoke ReadConsoleA, hConsoleInput, ADDR InputBuf, 13, ADDR CharsReadNum, 0

invoke crt\_atoi, ADDR InputBuf

ret

Input\_ ENDP

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

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

Less\_ PROC

pushf

pop cx

mov eax, [esp + 8]

cmp eax, [esp + 4]

jge less\_false

mov eax, 1

jmp less\_fin

less\_false:

mov eax, 0

less\_fin:

push cx

popf

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Less\_ ENDP

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

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

Not\_ PROC

pushf

pop cx

mov eax, [esp + 4]

cmp eax, 0

jnz not\_false

not\_t1:

mov eax, 1

jmp not\_fin

not\_false:

mov eax, 0

not\_fin:

push cx

popf

mov [esp + 4], eax

ret

Not\_ ENDP

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

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

Or\_ PROC

pushf

pop cx

mov eax, [esp + 8]

cmp eax, 0

jnz or\_true

jz or\_t1

or\_t1:

mov eax, [esp + 4]

cmp eax, 0

jnz or\_true

or\_false:

mov eax, 0

jmp or\_fin

or\_true:

mov eax, 1

or\_fin:

push cx

popf

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Or\_ ENDP

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

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

Output\_ PROC value: dword

invoke wsprintf, ADDR ResMessage, ADDR OutMessage, value

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

ret 4

Output\_ 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================================================================================

\_aaa2\_ dd 0

\_aaaa\_ dd 0

\_bbbb\_ dd 0

\_ccc1\_ dd 0

\_ccc2\_ dd 0

\_xxxx\_ dd 0

String\_0 db "Input A: ", 0

String\_1 db "Input B: ", 0

String\_2 db "FOR TO DO", 0

String\_3 db 13, 10, 0

String\_4 db 13, 10, "FOR DOWNTO DO", 0

String\_5 db 13, 10, 0

String\_6 db 13, 10, "WHILE A MUL B: ", 0

String\_7 db 13, 10, "REPEAT UNTIL A MUL B: ", 0

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

hConsoleInput dd ?

hConsoleOutput dd ?

endBuff db 5 dup (?)

msg1310 db 13, 10, 0

CharsReadNum dd ?

InputBuf db 15 dup (?)

OutMessage db "%d", 0

ResMessage db 20 dup (?)

.CODE

start:

invoke AllocConsole

invoke GetStdHandle, STD\_INPUT\_HANDLE

mov hConsoleInput, eax

invoke GetStdHandle, STD\_OUTPUT\_HANDLE

mov hConsoleOutput, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_0, SIZEOF String\_0 - 1, 0, 0

call Input\_

mov \_aaaaaaaa\_, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_1, SIZEOF String\_1 - 1, 0, 0

call Input\_

mov \_bbbbbbbb\_, eax

invoke WriteConsoleA, hConsoleOutput, ADDR String\_2, SIZEOF String\_2 - 1, 0, 0

push \_aaaaaaaa\_

pop \_aaaaaaa2\_

forPasStart1:

push \_bbbbbbbb\_

push \_aaaaaaa2\_

call Less\_

call Not\_

pop eax

cmp eax, 0

je forPasEnd1

invoke WriteConsoleA, hConsoleOutput, ADDR String\_3, SIZEOF String\_3 - 1, 0, 0

push \_aaaaaaa2\_

push \_aaaaaaa2\_

call Mul\_

call Output\_

push \_aaaaaaa2\_

push dword ptr 1

call Add\_

pop \_aaaaaaa2\_

jmp forPasStart1

forPasEnd1:

invoke WriteConsoleA, hConsoleOutput, ADDR String\_4, SIZEOF String\_4 - 1, 0, 0

push \_bbbbbbbb\_

pop \_aaaaaaa2\_

forPasStart2:

push \_aaaaaaaa\_

push \_aaaaaaa2\_

call Greate\_

call Not\_

pop eax

cmp eax, 0

je forPasEnd2

invoke WriteConsoleA, hConsoleOutput, ADDR String\_5, SIZEOF String\_5 - 1, 0, 0

push \_aaaaaaa2\_

push \_aaaaaaa2\_

call Mul\_

call Output\_

push \_aaaaaaa2\_

push dword ptr 1

call Sub\_

pop \_aaaaaaa2\_

jmp forPasStart2

forPasEnd2:

invoke WriteConsoleA, hConsoleOutput, ADDR String\_6, SIZEOF String\_6 - 1, 0, 0

push dword ptr 0

pop \_xxxxxxxx\_

push dword ptr 0

pop \_ccccccc1\_

whileStart2:

push \_ccccccc1\_

push \_aaaaaaaa\_

call Less\_

pop eax

cmp eax, 0

je whileEnd2

push dword ptr 0

pop \_ccccccc2\_

whileStart1:

push \_ccccccc2\_

push \_bbbbbbbb\_

call Less\_

pop eax

cmp eax, 0

je whileEnd1

push \_xxxxxxxx\_

push dword ptr 1

call Add\_

pop \_xxxxxxxx\_

push \_ccccccc2\_

push dword ptr 1

call Add\_

pop \_ccccccc2\_

jmp whileStart1

whileEnd1:

push \_ccccccc1\_

push dword ptr 1

call Add\_

pop \_ccccccc1\_

jmp whileStart2

whileEnd2:

push \_xxxxxxxx\_

call Output\_

invoke WriteConsoleA, hConsoleOutput, ADDR String\_7, SIZEOF String\_7 - 1, 0, 0

push dword ptr 0

pop \_xxxxxxxx\_

push dword ptr 1

pop \_ccccccc1\_

repeatStart2:

push dword ptr 1

pop \_ccccccc2\_

repeatStart1:

push \_xxxxxxxx\_

push dword ptr 1

call Add\_

pop \_xxxxxxxx\_

push \_ccccccc2\_

push dword ptr 1

call Add\_

pop \_ccccccc2\_

push \_ccccccc2\_

push \_bbbbbbbb\_

call Greate\_

call Not\_

pop eax

cmp eax, 0

je repeatEnd1

jmp repeatStart1

repeatEnd1:

push \_ccccccc1\_

push dword ptr 1

call Add\_

pop \_ccccccc1\_

push \_ccccccc1\_

push \_aaaaaaaa\_

call Greate\_

call Not\_

pop eax

cmp eax, 0

je repeatEnd2

jmp repeatStart2

repeatEnd2:

push \_xxxxxxxx\_

call Output\_

exit\_label:

invoke WriteConsoleA, hConsoleOutput, ADDR msg1310, SIZEOF msg1310 - 1, 0, 0

invoke ReadConsoleA, hConsoleInput, ADDR endBuff, 5, 0, 0

invoke ExitProcess, 0

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

Add\_ PROC

mov eax, [esp + 8]

add eax, [esp + 4]

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Add\_ ENDP

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

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

Greate\_ PROC

pushf

pop cx

mov eax, [esp + 8]

cmp eax, [esp + 4]

jle greate\_false

mov eax, 1

jmp greate\_fin

greate\_false:

mov eax, 0

greate\_fin:

push cx

popf

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Greate\_ ENDP

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

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

Input\_ PROC

invoke ReadConsoleA, hConsoleInput, ADDR InputBuf, 13, ADDR CharsReadNum, 0

invoke crt\_atoi, ADDR InputBuf

ret

Input\_ ENDP

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

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

Less\_ PROC

pushf

pop cx

mov eax, [esp + 8]

cmp eax, [esp + 4]

jge less\_false

mov eax, 1

jmp less\_fin

less\_false:

mov eax, 0

less\_fin:

push cx

popf

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Less\_ ENDP

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

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

Mul\_ PROC

mov eax, [esp + 8]

imul dword ptr [esp + 4]

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Mul\_ ENDP

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

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

Not\_ PROC

pushf

pop cx

mov eax, [esp + 4]

cmp eax, 0

jnz not\_false

not\_t1:

mov eax, 1

jmp not\_fin

not\_false:

mov eax, 0

not\_fin:

push cx

popf

mov [esp + 4], eax

ret

Not\_ ENDP

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

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

Output\_ PROC value: dword

invoke wsprintf, ADDR ResMessage, ADDR OutMessage, value

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

ret 4

Output\_ ENDP

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

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

Sub\_ PROC

mov eax, [esp + 8]

sub eax, [esp + 4]

mov [esp + 8], eax

pop ecx

pop eax

push ecx

ret

Sub\_ ENDP

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

end start