

Минобрнауки России

Федеральное государственное бюджетное образовательное учреждение   
высшего образования

**«МИРЭА – Российский технологический университет»**

**РТУ МИРЭА**

Институт Искусственного Интеллекта  
Факультет – компьютерной безопасности

**Реферат**

по дисциплине «Ознакомительная Практика»

**Тема реферата:** Изучение основных частей фронтенда большинства компиляторов: лексера, парсера и генератора AST

|  |  |  |
| --- | --- | --- |
| **Студент учебной группы ККСО-01-22** |  | Козлов И.А. |
| **Студент учебной группы ККСО-01-22** |  | Колесников Е.А. |
| **Руководитель** |  | - |

Москва 2023

**Оглавление**

[Введение: 3](#__RefHeading___Toc299_2298221030)

[Теоретическая часть 3](#__RefHeading___Toc251_30324797651)

[1. Лексер 3](#__RefHeading___Toc317_2298221030)

[2. Парсер 3](#__RefHeading___Toc319_2298221030)

[3. Генератор AST 3](#__RefHeading___Toc321_2298221030)

[Практическая часть 4](#__RefHeading___Toc258_3032479765)

[Приложение с исходным кодом 4](#__RefHeading___Toc311_22982210301)

[Заключение: 4](#__RefHeading___Toc323_2298221030)

[Используемая литература 4](#__RefHeading___Toc309_2298221030)

# Введение:

Компиляторы - это программы, которые позволяют переводить исходный код на одном языке программирования в эквивалентный код на другом языке программирования. Фронтенд компилятора отвечает за первую стадию компиляции, которая включает в себя лексический анализ, синтаксический анализ и генерацию абстрактного синтаксического дерева (AST). В данном докладе мы рассмотрим основные части фронтенда большинства компиляторов: лексер, парсер и генератор AST.

# Теоретическая часть

## 1. Лексер

Лексер - это программа, которая разбивает исходный код на маленькие кусочки, называемые токенами. Токены представляют собой минимальные единицы языка программирования, такие как числа, строки, операторы, ключевые слова и т.д.

Лексер производит токенизацию исходного кода, используя регулярные выражения, которые описывают структуру языка программирования. После токенизации лексер передает токены парсеру для дальнейшей обработки.

## 2. Парсер

Парсер - это программа, которая анализирует токены, создавая из них дерево разбора (Parse tree). Дерево разбора представляет собой структуру данных, которая отражает иерархию элементов в исходном коде.

Дерево разбора может быть использовано для проверки корректности исходного кода и генерации кода на другом языке программирования. Однако, дерево разбора может быть очень большим и неудобным для работы, поэтому часто используется абстрактное синтаксическое дерево (AST).

## 3. Генератор AST

Генератор AST - это программа, которая обрабатывает дерево разбора, создавая более удобное и компактное абстрактное синтаксическое дерево (AST). AST является упрощенным представлением дерева разбора, которое содержит только необходимые элементы для дальнейшей работы компилятора.

AST может быть использовано для генерации кода на другом языке программирования, оптимизации кода, анализа кода и т.д.

# Практическая часть

## Приложение с исходным кодом

<https://github.com/AIwanKozlovA/Brainfuck>



# Заключение:

В данном докладе мы рассмотрели основные части фронтенда большинства компиляторов: лексер, парсер и генератор AST. Лексер разбивает исходный код на токены, парсер создает дерево разбора, а генератор AST создает более удобное и компактное абстрактное синтаксическое дерево. Фронтенд компилятора является важной частью компиляции, так как он отвечает за первую стадию перевода исходного кода на одном языке программирования в эквивалентный код на другом языке программирования.

# Используемая литература

«Построение компиляторов» — книга Никлауса Вирта.

Brainfuck — статья на платформе «Википедия».

Hello, Brainfuck! — статья на сайте «Хабр».