

Первый рубежный контроль по курсу «Конструирование компиляторов»

Билет 1

1. Компиляторы и интерпретаторы: основные определения (компилятор, исходный язык, целевой язык, язык реализации, транслятор, интерпретатор).
2. Описать регулярными выражениями все десятичные целые числа, которые делятся на 2, на 4, на 8. Построить диаграммы минимального ДКА. Числа могут начинаться с незначащих нулей, не могут быть пустой строкой.

Билет 2

1. Т-диаграммы.
2. Построить диаграмму минимального конечного автомата, принимающего десятичные числа от 0 до 65535. Число может начинаться с незначащих нулей, не может быть пустой строкой.
Принимаемые числа: 0, 007, 6666, 9999, 0000065535.
Не принимаемые числа: 999999, 65536, 66666.

Билет 3

1. Самоприменимые компиляторы: раскрутка и перенос.
2. Построить диаграмму минимального конечного автомата, принимающего числа, делящиеся на 2, на 3, на 6. Число может начинаться с незначащих нулей, не может быть пустой строкой.

Билет 4

1. Общая схема компиляции.
2. Записать регулярное выражение для строки в двойных кавычках с escape-последовательностями языка C89. Построить для него диаграмму конечного автомата.
Для справки: <https://ru.cppreference.com/w/cpp/language/escape> (C89 не поддерживал \u и \U, их не надо).

Билет 5

1. Фазы анализа.
2. Язык состоит из целых и вещественных десятичных чисел, разделённых пробельными символами. Десятичным разделителем является точка. Следует описать его лексическую структуру и построить лексический распознаватель.

Правильные целые числа: 3, 30, 1234567890987654321, +1, -0.

Правильные вещественные числа: 6.022e23, +6.022e+23, 1.38e-23, 8.31, 3e8, -273.15.

Неправильные вещественные числа: .33e7 (нет цифры перед точкой), 20. (нет цифры после точки), e10 (нет мантииссы), e2e4.

Билет 6

1. Группировка фаз компиляции.
2. Язык состоит из целочисленных констант языка ассемблера (синтаксис MASM), требуется записать его лексическую структуру и построить лексический распознаватель.

Десятичные константы — последовательности десятичных цифр, которые заканчиваются на необязательный знак `d`.

Шестнадцатеричные константы — последовательности шестнадцатеричных цифр, начинающихся на десятичную цифру и заканчивающиеся на `h`.

Восьмеричные константы — последовательности восьмеричных цифр, заканчивающихся на знак `q`.

Двоичные константы — последовательности двоичных цифр, заканчивающихся на `b`.

Для справки: <https://prog-cpp.ru/asm-datatypes/>.

Билет 7

1. Постановка задачи лексического анализа.
2. Описать регулярным выражением все десятичные целые числа, которые делятся на 8. Построить диаграмму ДКА, не обязательно минимальную. Числа могут начинаться с незначащих нулей, не могут быть пустой строкой.

Билет 8

1. Проектирование объектно-ориентированного лексического анализатора.
2. Построить диаграмму минимального конечного автомата, принимающего десятичные числа от 0 до 65535. Число может начинаться с незначащих нулей, не может быть пустой строкой.
Принимаемые числа: 0, 007, 6666, 9999, 0000065535.
Не принимаемые числа: 999999, 65536, 66666.

Билет 9

1. Реализация объектно-ориентированного лексического анализатора вручную.
2. Построить диаграмму минимального конечного автомата, принимающего числа, делящиеся на 2, на 3, на 6. Число может начинаться с незначащих нулей, не может быть пустой строкой.

Билет 10

1. Лексические распознаватели.
2. Записать регулярное выражение для строки в двойных кавычках с escape-последовательностями языка C89. Построить для него диаграмму конечного автомата.
Для справки: <https://ru.cppreference.com/w/cpp/language/escape> (C89 не поддерживал `\u` и `\U`, их не надо).

Билет 11

1. Детерминизация лексического распознавателя.
2. Язык состоит из целых и вещественных десятичных чисел, разделённых пробельными символами. Десятичным разделителем является точка. Следует описать его лексическую структуру и построить лексический распознаватель.

Правильные целые числа: 3, 30, 1234567890987654321, +1, -0.

Правильные вещественные числа: 6.022e23, +6.022e+23, 1.38e-23, 8.31, 3e8, -273.15.

Неправильные вещественные числа: .33e7 (нет цифры перед точкой), 20. (нет цифры после точки), e10 (нет мантиссы), e2e4.

Билет 12

1. Представление лексического распознавателя в программе.
2. Язык состоит из целочисленных констант языка ассемблера (синтаксис MASM), требуется записать его лексическую структуру и построить лексический распознаватель.

Десятичные константы — последовательности десятичных цифр, которые заканчиваются на необязательный знак d.

Шестнадцатеричные константы — последовательности шестнадцатеричных цифр, начинающихся на десятичную цифру и заканчивающиеся на h.

Восьмеричные константы — последовательности восьмеричных цифр, заканчивающихся на знак q.

Двоичные константы — последовательности двоичных цифр, заканчивающихся на b.

Для справки: <https://prog-cpp.ru/asm-datatypes/>.

Билет 13

1. Генератор лексических анализаторов flex.
2. Описать регулярными выражениями все десятичные целые числа, которые делятся на 2, на 4. Построить диаграммы минимального ДКА. Числа могут начинаться с незначащих нулей, не могут быть пустой строкой.