



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н. Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н. Э. Баумана)

ФАКУЛЬТЕТ «Информатика, искусственный интеллект и системы управления»
КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

ОТЧЁТ
по лабораторной работе № 1
по курсу «Конструирование компиляторов»
на тему: «Распознавание цепочек регулярного языка»
Вариант № 3

Студент ИУ7-23М
(Группа)

(Подпись, дата)

Волкова А. А.
(Фамилия И. О.)

Преподаватель

(Подпись, дата)

Ступников А. А.
(Фамилия И. О.)

2022 г.

Описание задания

Цель работы: приобретение практических навыков реализации важнейших элементов лексических анализаторов на примере распознавания цепочек регулярного языка.

Задачи работы:

1. Ознакомиться с основными понятиями и определениями, лежащими в основе построения лексических анализаторов.
2. Прояснить связь между регулярным множеством, регулярным выражением, праволинейным языком, конечно-автоматным языком и недетерминированным конечно-автоматным языком.
3. Разработать, тестировать и отладить программу распознавания цепочек регулярного или праволинейного языка в соответствии с предложенным вариантом грамматики.

Вариант 3

Напишите программу, которая в качестве входа принимает произвольное регулярное выражение, и выполняет следующие преобразования:

1. По регулярному выражению строит НКА.
2. По НКА строит эквивалентный ему ДКА.
3. По ДКА строит эквивалентный ему КА, имеющий наименьшее возможное количество состояний.

Указание. Воспользоваться алгоритмом, приведённым по адресу http://neerc.ifmo.ru/wiki/index.php?title=Минимизация_ДКА,_алгоритм_Хопкрофта_ (сложность $O(n \log n)$)

4. Моделирует минимальный КА для входной цепочки из терминалов исходной грамматики.

Текст программы

С полным текстом программы можно ознакомиться по адресу: <https://github.com/Volkovaan/CD/tree/lab01/lab01/src>.

Тестирование и результаты

Для тестирования использовались примеры 3.21, 3.24 из [3] а также был написан ряд тестов, проверяющих корректное создание НКА по регулярным выражениям.

Регулярные выражения и пары (тест, результат):

- $(a|b)^*abb$ — $(a, 0)$, $(b, 0)$, $(ab, 0)$, $(aab, 0)$, $(abb, 1)$, $(aabb, 1)$
- $a^*|b$ — $(\varepsilon, 1)$, $(a, 1)$, $(aaaa, 1)$, $(b, 1)$, $(bbb, 0)$, $(ab, 0)$
- $((((ab))))$ — $(\varepsilon, 0)$, $(a, 0)$, $(b, 0)$, $(ab, 1)$, $(abb, 0)$
- $((((a))^*)^*)$ — $(\varepsilon, 1)$, $(a, 1)$, $(aaaa, 1)$, $(ab, 0)$, $(b, 0)$
- $((a|bb)^*(a|bb)b^*)|b$ — $(\varepsilon, 0)$, $(a, 1)$, $(b, 1)$, $(aa, 1)$, $(bb, 1)$, $(ab, 1)$, $(ba, 0)$, $(abbabab, 0)$, $(aaa, 1)$, $(abbabbbb, 1)$

Выводы

В данной лабораторной работе были приобретены практические навыки реализации важнейших элементов лексических анализаторов на примере распознавания цепочек регулярного языка. Были проработаны основные понятия и определения, лежащие в основе построения лексических анализаторов и прояснены связи между регулярным множеством, регулярным выражением, праволинейным языком, конечно-автоматным языком и недетерминированным конечно-автоматным языком. Результатом работы является разработанная программа распознавания цепочек регулярного языка в соответствии с предложенным вариантом грамматики.

Список использованной литературы

1. Белоусов А. И., Ткачёв С. Б. Дискретная математика: Учеб. Для вузов / Под ред. В. С. Зарубина, А.П. Крищенко. – М.: Изд-во МГТУ им. Н. Э. Баумана, 2021.
2. Минимизация ДКА, алгоритм Хопкрофта (сложность $O(n \log n)$). Университет ИТМО, URL: [http://neerc.ifmo.ru/wiki/index.php?title=Минимизация_ДКА,_алгоритм_Хопкрофта_\(сложность_O\(n_log_n\)\)](http://neerc.ifmo.ru/wiki/index.php?title=Минимизация_ДКА,_алгоритм_Хопкрофта_(сложность_O(n_log_n)))
3. Ахо А. В, Лам М. С., Сети Р., Ульман Дж. Д. Компиляторы: принципы, технологии и инструменты. – М.: Вильямс, 2008.