

Практикум №2

Контекстно-свободные грамматики

Практикум состоит в программной реализации некоторого алгоритма синтаксического разбора для контекстно-свободных грамматик. В качестве языка программирования взять C++ или Python. Необходимо написать класс, получающий на вход КС-грамматику, осуществляющий её предобработку (приведение к нормальной форме Хомского в случае алгоритма Кока-Янгера-Касами, построение LR-анализатора для LR-грамматики и т.п.) и проверяющий, принадлежит ли слово языку, задаваемому грамматикой. Основная программа должна, помимо этого, прочитать вход и вывести ответ задачи.

Первым параметром, подаваемым на вход, является файл с грамматикой. Данный файл имеет вид

$$\begin{aligned} & n \\ & \langle \text{правило}_1 \rangle \\ & \dots \\ & \langle \text{правило}_n \rangle, \end{aligned}$$

где n —число правил грамматики, а правило $A \rightarrow B_1 \dots B_n$ закодировано в виде $A_B_1_ \dots _B_n$, где $_$ обозначает пробел. Стартовым нетерминалом считается первый символ первого правила. В качестве терминальных символов используются малые буквы латинского алфавита, а в каждый нетерминал имеет вид $A\alpha$, где A — прописная буква латинского алфавита, а α — последовательность цифр (возможно, пустая). То есть допустимыми именами нетерминалов являются, например, $A1, B0, A004, C$. Помимо файла с грамматикой, программа получает на вход натуральное число m и m слов w_1, \dots, w_m над латинским алфавитом. в котором на каждой строчке написано некоторое слово над латинским алфавитом (гарантируется, что других символов не содержится). Разделителем строк служит '\n'. Программа должна вывести на стандартный вывод m строк со словами YES или NO, где YES в i -ой строке означает, что i -ое слово, поданное на вход, принадлежит языку, задаваемому грамматикой. В задаче 6 в качестве первого аргумента перед именем файла с грамматикой подаётся регулярное выражение, задающее регулярный язык.

Номер	Задача
1	Алгоритм Кока-Янгера-Касами
2	Алгоритм Эрли
3	Алгоритм Эрли, МП-автомат
4	SLR-алгоритм (Ахо, Сети, Ульман, стр. 226–234)
5	LR-алгоритм (Ахо, Сети, Ульман, стр. 234–240)
6	пересечение КС-языка с регулярным (алгоритм произвольный)

Пример входных данных

Имя файла	Содержимое файла	Вход	Выход
grammar.in	S_a_S_b_S S	abbab aaababbb abba	NO YES NO