

Задание 12

LR-анализ

Ключевые слова¹: язык, контекстно-свободный язык, магазинный автомат, грамматика, LL(k)-грамматика, LL(1)-анализатор, функции FIRST, FOLLOW, LR(k)-грамматика, LR(1)-анализатор, LR(0)-анализатор.

Это задание, которое не требуется сдавать на проверку!

Для самопроверки используйте онлайн-конструктор анализаторов (см. описание на сайте).

1 LR(0)-анализаторы

Параметр k в LR(k) анализаторах — число предпросматриваемых символов, по которым анализатор принимает решение какое из действий «сдвиг» или «свёртка» совершить на текущем такте. На первый взгляд может показаться, что $k \geq 1$, однако существуют и LR(0)-анализаторы. LR(0)-анализатор принимает решение не зная какой символ находится на входной ленте. Как же такое возможно?

Принимать решение без предпросмотра возможно, поскольку не смотря на независимость решений для произвольного символа, переходы между состояниями LR(0)-автомата Кнута по-прежнему зависят от символа помещённого в стек. В управляющей таблице LR(0)-анализатора каждому состоянию соответствует ровно одно действие (независимо от символа входа). В остальном, LR(0)-анализаторы не отличаются от LR(1)-анализаторов, изученных на семинаре.

LR(0)-анализатор принимает слово, если оказался в конфигурации $(I_0 S I_1, \$, p_n p_{n-1} \dots p_1)$, где p_i — номер правила, а состоянию I_1 соответствует действие **Акцепт**.

Задача 1. Постройте LR(0)-анализатор для грамматики $S \rightarrow aSb \mid b$. Продемонстрируйте работу анализатора на слове $aabbb$ и постройте по результатам работы дерево разбора. Если затрудняетесь, то воспользуйтесь конструктором анализаторов.

¹минимальный необходимый объем понятий и навыков по этому разделу)

Как понять, является ли грамматика LR(0)- или LR(1)-грамматикой? (точнее, какое из чисел 0 или 1 является минимальным k). Есть критерий LR(k)-грамматик, записанный, например, в книге Ахо и Ульмана, но он достаточно сложный и мы не будем приводить его здесь. Мы приведём следующий рецепт.

Постройте LR(1)-анализатор и если в управляющей таблице некоторому состоянию I_m , в зависимости от символа, соответствуют два разных правила свёртки (в том числе и **Accept**) или для по одному символу должна быть операция сдвиг, а по другому — свёртка, то это заведомо не LR(0)-грамматика. Действительно, состояния LR(0)-автомата Кнута отличаются от состояний LR(1)-автомата Кнута только в следующем. Каждому состоянию LR(1)-автомата соответствует состояние LR(0)-автомата; последнее содержит ситуацию² $[A \rightarrow \alpha \cdot X\beta]$ тогда и только тогда, когда в первом есть ситуация вида $[A \rightarrow \alpha \cdot X\beta, u]$. Это свойство легко вытекает из процедуры детерминизации. Таким образом, если в некотором состоянии LR(1)-автомата есть две разных операции (по разным символам), то при построении LR(0)-анализатора, в некотором состоянии (множестве ситуаций) потребуется совершить две разных операции.

2 Задачи

Задача 2. Язык L задан неоднозначной КС-грамматикой

$$G = \{\{S\}, \{(\cdot)\}, \{S \rightarrow (S) \mid SS \mid ()\}, S\}.$$

Написать LL(1)-грамматику для языка L .

Задача 3. Дана грамматика $G = \{\{A, S\}, \{a, b, c\}, \{S \rightarrow Aa \mid b \mid \varepsilon; A \rightarrow Ab \mid c\}, S\}$. Является ли грамматика G LR(k)-грамматикой? При положительном ответе на вопрос найти минимальное k и построить соответствующий анализатор. Построить дерево разбора для цепочки $cbba$.

²Мы опускаем аванпечку в LR(0)-ситуациях, поскольку её длина ноль.

Задача 4. Дана грамматика $G = \{ \{A, S\}, \{a\}, \{ S \rightarrow A; A \rightarrow aAa \mid a \}, S \}$. Является ли грамматика G LR(k)-грамматикой? При положительном ответе на вопрос найти минимальное k и построить соответствующий анализатор. Построить дерево разбора для цепочки $aaaaa$.

Задача 5. Дана грамматика $G = \{ \{A, S\}, \{a, b, c\}, \{S \rightarrow Aa \mid b; A \rightarrow Ab \mid c \}, S \}$. Является ли грамматика G LR(k)-грамматикой? При положительном ответе на вопрос найти минимальное k и построить соответствующий анализатор. Продемонстрировать работу анализатора на цепочке $cbbab$.

2.1 Контрольные вопросы

Задача 6. При построении LR(1)-анализатора для грамматики G в одном множестве оказались ситуации $[A \rightarrow .aA\alpha, b]$ и $[B \rightarrow \beta.a, a]$, где α, β некоторые цепочки из $(N \cup T)^*$. Может ли грамматика G быть LR(0)-грамматикой?