

2024

КЫК и его друзья

15 мая 2024 г.

## Содержание

### I КЯК парсер

1

#### Часть I

## КЯК парсер

Рассмотрим пример грамматики ниже. Возьмём для разбора строку *baaba*.

$$S \rightarrow AB \mid BC$$

$$A \rightarrow BA \mid a$$

$$B \rightarrow CC \mid b$$

$$C \rightarrow AB \mid a$$

Суть КЯК в рассмотрении подстрок от самых маленьких к целой строке. По итогу нам нужно узнать, является ли строка каким-то из нетерминалов (т.е. словом языка, порождённого грамматикой). Если совсем кратко - принадлежит ли строка языку.

Приступим к разбору. Ещё раз напомним, что мы его ведём снизу вверх - от терминалов к нетерминалам и так пока не охватим всю строку.

$$B \rightarrow b$$

$$A \rightarrow a$$

$$C \rightarrow a$$

Идём дальше. Это были подстроки длины 1. *baaba* также имеет подстроки длины 2, которые формируются из подстрок длины 1 ( $\times$  - декартово произведение):

$$ba = b \times a = B \times A = BA = A$$

$$ba = b \times a = B \times C = BC = S$$

Пока у нас есть  $A \rightarrow ba, S \rightarrow ba$ . Которые получаются, как мы выяснили, вот так:

$$A \rightarrow BA \rightarrow bA \rightarrow ba$$

$$S \rightarrow BC \rightarrow bC \rightarrow ba$$

Тупо последовательно заменяем нетерминалы, пока не доберёмся до терминалов. К сожалению, дальше мы упоремся и запутаемся, если будем так продолжать. Это была всего лишь первая 2-подстрока, а я уже устала.

Видимо, кто-то из авторов тоже устал, и придумал лестницу КЯК или как она там. Выглядит она следующим образом (цифры - длины подстрок, НТ - нетерминалы):

5	НТ				
4	НТ	НТ			
3	НТ	НТ	НТ		
2	НТ	НТ	НТ	НТ	
1	НТ	НТ	НТ	НТ	НТ
	b	a	a	b	a

Эта таблица показывает, из каких нетерминалов какую строку мы можем построить. Пока что для нас она выглядит так:

5	НТ				
4	НТ	НТ			
3	НТ	НТ	НТ		
2	НТ	НТ	НТ	НТ	
1	В	А,С	А,С	В	А,С
	b	a	a	b	a

И, учитывая что мы уже успели проанализировать первое  $ba$ :

5	НТ				
4	НТ	НТ			
3	НТ	НТ	НТ		
2	С,А	НТ	НТ	НТ	
1	В	А,С	А,С	В	А,С
	b	a	a	b	a

Теперь давайте заполнять эту таблицу. Для следующих подстрок мы получим:

$$aa = a \times a = A, C \times A, C = AA, CA, AC, CC = B$$

$$ab = a \times b = A, C \times B = AB, CB = S, C$$

$$bb = b \times b = B \times B = \emptyset$$

5	HT				
4	HT	HT			
3	HT	HT	HT		
2	S,A	B	S,C	S,A	
1	B	A,C	A,C	B	A,C
	b	a	a	b	a

Строки длины 3. Мне тоже надоело, но что есть:

$$baa = b \times aa = B \times B = BB = \emptyset$$

$$baa = ba \times a = S, A \times A, C = SA, AA, SC, AC = \emptyset$$

$$aab = a \times ab = A, C \times S, C = AS, CS, AC, CC = B$$

$$aab = aa \times b = B \times B = \emptyset$$

$$aba = a \times ba = A, C \times S, A = AS, CS, AA, CA = \emptyset$$

$$aba = ab \times a = S, C \times A, C = SA, CA, SC, CC = B$$

5	HT				
4	HT	HT			
3	$\emptyset$	B	B		
2	S,A	B	S,C	S,A	
1	B	A,C	A,C	B	A,C
	b	a	a	b	a

Строки длины четыре.

$$baab = b \times aab = B \times B = BB = \emptyset$$

$$baab = ba \times ab = S, A \times S, C = SS, AS, SC, AC = \emptyset$$

$$baab = baa \times b = \emptyset \times B = \emptyset$$

$$aaba = a \times aba = A, C \times B = AB, CB = S, C$$

$$aaba = aa \times ba = B \times S, A = BS, BA = A$$

$$aaba = aab \times a = B \times A, C = BA, BC = A, S, C$$

5	HT				
4	$\emptyset$	S,A,C			
3	$\emptyset$	B	B		
2	S,A	B	S,C	S,A	
1	B	A,C	A,C	B	A,C
	b	a	a	b	a

Строки длины ять. Пять. А вы о чем подумали?

$$baaba = b \times aaba = B \times S, A, C = BS, BA, BC = A, S$$

$$baaba = ba \times aba = S, A \times B = SA, AB = S, C$$

$$baaba = baa \times ba = \emptyset \times S, A = \emptyset$$

$$baaba = baab \times a = \emptyset \times A, C = \emptyset$$

5	S,A,C				
4	$\emptyset$	S,A,C			
3	$\emptyset$	B	B		
2	S,A	B	S,C	S,A	
1	B	A,C	A,C	B	A,C
	b	a	a	b	a

Таким образом, строка может быть получена путем последовательного раскрытия (развертки) нетерминалов S, A, C.

УПРАЖНЕНИЕ. Убедиться, что все указанные нетерминалы преобразуются в иско-  
мую строку.