

Формальные языки, ДЗ 2

Василевский Тимофей

20.09.2021

1 Обязательная часть.

1. Давайте докажем, что любое слово, которое подходит под первое регулярное выражение – также подходит под второе. Заметим, что слово из первого регулярного выражения заканчивается одну из следующих строк:
 - (а) a – единственный случай, когда все слово это просто a , тогда оно подходит и под вторую регулярку (из первой скобки взяли a , а из второй – ничего)
 - (б) aa – все такие подстроки можно последовательно стирать с конца, потому что они будут соответствовать второй скобке второй регулярки, теперь в конце слова стоит либо a – тогда слово подходит (по пункту 1, так как там мы не используем 2 скобку), либо ba , тогда оно подойдет по пункту 3 (смотреть ниже, там мы тоже не используем вторую скобку).
 - (с) ba – любое слово состоящие из a , b и заканчивающееся на ba – подходит, потому что ее можно набрать только из первой скобки (сначала берем всю строку без двух последних символов из $(a|b)^*$, а в конце как раз добавляем ba).

Наоборот:

Все $(aa)^*$, идущие в конце, очевидно, можно покрыть аналогичным количеством $((a|b)b * a)^*$ из первой скобки, значит осталось покрыть только $((a|b) * ba|a)$, теперь заметим, что просто a мы тоже может покрыть первым регулярным выражением, просто не используя ни одного выражения в скобках, то есть осталось покрыть $((a|b) * ba)$, тут опять же заметим, что ba , идущее в конце мы можем покрыть используя одну скобку такого вида: $((a|b)b * a)^*$, осталось получить $(a|b)^*$. Это легко сделать, заметив, что с помощью первой строки мы можем выбрать произвольное начало, потом взять любое количество a , а затем брать подстроки вида: k раз b , затем a (не забудем про окончание ba из второй регулярки, которое убережет нас от того, что строка заканчивается на символ b).

2. Давайте просто заметим, что можно кодировать названия не ключевых переменных в двоичном коде, а все ключевые слова оставлять обычным текстом, то есть алфавит будет выглядеть как *keyWords*, 00000000, 00000001, ..., в случае моего автомата *keyWords* = *Alphabet*, *StartVertex*, *Edge*, *States*, *TerminalVertexes*.
3. Смотреть *lex.py*