

ФЯ, дз на 27.09.2021

Малофеев Михаил

27 сентября 2021 г.

Задача №1

а) Докажем, что язык

$$\{uabv | u \in \{a, b\}^*, v \in \{a, b\}^*, |u| = |v|, u \neq v^R\}$$

не является регулярным. Воспользуемся леммой о накачке. Докажем, что для любого n существует такое слово длины хотя бы n , такое, что для любого разбиения этого слова на части x y z , где длина y ненулевая, а длина xu не превосходит n , существует такое k , что $xy^kz \notin L$. Давайте для каждого n возьмём слово a^nabb^n . Тогда для любого разбиения этого слова на части x , y , z часть y будет состоять из ненулевого количества букв a . Пусть длина y будет равна l . Возьмём $k = 0$. Тогда новое слово будет иметь вид $a^{n-l}abb^n$. Это слово не принадлежит языку L , так как есть противоречие с условием $|u| = |v|$, а здесь $|u| = n - l$, $|v| = n$.

б) Докажем, что язык

$$\{a^k c^m e^n | k \geq 0, n \geq 0, m = n + k + 1\}$$

не является регулярным. Воспользуемся леммой о накачке. Докажем, что для любого n существует такое слово длины хотя бы n , такое, что для любого разбиения этого слова на части x y z , где длина y ненулевая, а длина xu не превосходит n , существует такое k , что $xy^kz \notin L$. Давайте для каждого n возьмём слово $c^{n+1}e^n$. Тогда при любом разбиении на x , y , z , y будет иметь вид c^l . Давайте возьмём $k = 0$. Тогда новое слово будет иметь вид $c^{n+1-l}e^n$. Это слово не принадлежит языку L , так как есть противоречие с тем, что $m = n + k + 1$.

с) Докажем, что язык

$$\{a^n | \exists p \geq n : p \text{ prime and } p + 2 \text{ prime}\}$$

является регулярным. Если гипотеза верна, что простых чисел-близнецов бесконечно много, то в этом языке будут содержаться строки, состоящие из любого количества букв a , так как для любой длины строки n найдутся простые числа-близнецы, большие n . Следовательно, для данного языка можно построить регулярное выражение a^* .

В том случае, если гипотеза о бесконечности простых чисел-близнецов неверна, то всего таких простых чисел-близнецов конечно. Возьмём наибольшую пару из них. Пусть они равны p и $p + 2$. Тогда в нашем языке будут содержаться все слова, составленные из буквы a и имеющие длину не больше p . Тогда их можно перечислить, так как их конечно. Следовательно, такой язык является регулярным.

Задача №2

В `main.cpp` находятся тесты, в `my_test.h` и `my_test.cpp` содержится структура для тестов. В `my_regexr.cpp` и `my_regexr.h` находится реализация класса регулярных выражений `regexr`.

В output.txt находятся результаты тестов и замеры времени до и после оптимизаций. После предложенных оптимизаций проверка на матчинг регуляркой строки стала работать быстрее, особенно это заметно в тесте с большой вложенностью звёздочек. Это происходит оттого, что с увеличением числа звёздочек без оптимизаций итоговая регулярка очень сильно разрастается.