

Дополнительное задание на +3

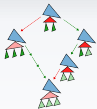
Задание полностью обесценится после 9 октября. 3 балла не входят в стоимость основной л.р. За списанные регулярки в пункте 2 буду нещадно собирать компромат.

1. Выяснить, какой алгоритм парсинга используется в какой-нибудь regex-библиотеке языка из верхней правой четверти графика RedMonk. Описать его на $\frac{1}{2} - \frac{2}{3}$ страницы.
2. Протестировать выбранную библиотеку на следующих регулярных выражениях: $((((a)*a)*) \dots a) * (k \text{ раз итерируется операция приписывания буквы } a \text{ к предыдущему выражению и применение к результату итерации Клини}), (a?) (a?) \dots (a?) aa \dots a (k \text{ раз } a?, \text{ и затем } k \text{ раз просто } a), (.) * b (.) [n] (n \text{ произвольных букв после буквы } b);$ а также двух сериях регулярных выражений на ваш выбор. Построить график зависимости скорости распознавания слова от параметра k (значения параметров лучше брать с шагом, ≥ 3 , или хотя бы ≥ 2).



Коллективная часть (подробности ТА)

- 1 Построение производных и частичных производных, определение длины накачки;
- 2 Построение НКА по классическим алгоритмам (автоматы Глушкова, Томпсона, Антимирова, follow-автомат);
- 3 Базовые алгоритмы над НКА и ДКА (детерминизация, удаление эpsilon-правил, удаление/добавление ловушки, минимизация, обращение, отрицание, пересечение, ...);
- 4 Проверка автоматов на trace equivalence, бисимилярность, эквивалентность их языков (последнее — для контроля). Проверка на синхронизируемость;
- 5 Реализация парсинга слова по НКА (алгоритм Пайка + стандартный алгоритм с откатами);
- 6 Интеграция с лабораторными работами прошлого года (преобразование автомата в regex) и с лабораторными работами этого года (трансформационный моноид, нормализация по списку правил).



Индивидуальные варианты (подробности ТА)

- Вариант $= 0 \bmod 3$: нормализация regex по пользовательским правилам переписывания.
- Вариант $= 1 \bmod 3$: построение регулярных подвыражений для regex.
- Вариант $= 2 \bmod 3$: построение трансформационного моноида.