

Конечные автоматы и алгоритмы поиска образцов в тексте

Во всех задачах, кроме **3** и **4**, языки заданы над алфавитом $\Sigma = \{a, b\}$; в **3** и **4** $\Sigma = \{a, b, c\}$.

Назовём *префикс-функцией* функцию $l : \Sigma^+ \rightarrow \Sigma^*$, которая возвращает самый длинный собственный¹ префикс слова w , являющийся одновременно его суффиксом.

1. Вычислите значение префикс-функции $l(w)$ для w :
а) a^n , $n > 0$; б) $babab$; в) $bababa$; г) bab ; д) baa .
2. Построить КМП-автоматы (ДКА и с суффиксными ссылками), распознающие язык $\Sigma^*aabaaba$.
3. Постройте ДКА для словаря $\{aac, acb, b, ac, c\}$. Добавьте в полученный словарь слово ab и удалите слово ac .
4. Постройте для словаря $S = \{aac, acb, b, ac, c\}$ автомат Ахо–Корасик. Посчитайте с его помощью (или с помощью ДКА Ахо–Корасик) количество различных вхождений слов из словаря S в слово $aacbacb$ в качестве подслов.
5. Докажите, что автоматы Кнута–Морриса–Пратта и Ахо–Корасик имеют минимальное число состояний. То есть, что нет автоматов с меньшим числом состояний, которые распознают соответствующие языки.

¹То есть префикс, не совпадающий со всем словом w .