Отчет по лабораторной работе №1.

Цель: построение различных алгоритмов поиска подстроки в строке, сравнение скорости их работы на различных данных.

В данной лабораторной работе были рассмотрены следующие алгоритмы: Наивны алгоритм, алгоритм Бойера-Мура-Хорспула и алгоритм Кнутта-Мориса-Пратта. Необходимо сравнить их работу и выявить наиболее подходящий алгоритм для различных ситуаций.

Алгоритмы выполняли свою работу на двух типах данных: хорошие (good) и плохие (bad).

1. Проанализируем данные алгоритмы по признаку «Количество операций сравнения».

|  |  |  |  |
| --- | --- | --- | --- |
|  | naive | boyer\_moore\_horspool | knuth\_morris\_pratt |
| good 1 | 633 | 65 | 649 |
| good 2 | 695 | 105 | 779 |
| good 3 | 2074 | 437 | 2446 |
| good 4 | 9614 | 488 | 9705 |
| bad 1 | 18 | 9 | 19 |
| bad 2 | 910 | 99 | 199 |
| bad 3 | 90100 | 999 | 1999 |
| bad 4 | 4001000 | 4999 | 9999 |

На хороших данных алгоритмы Кнутта-Мориса-Пратта и наивный выполняют похожее количество итераций, а алгоритм Бойера-Мура-Хорспула показывает лучшие результаты, выполняя на порядок меньше операций.

На плохих данных ситуация в целом похожая – алгоритм Бойера-Мура-Хорспула занимает лидирующую позицию, однако алгоритм Кнутта-Мориса-Пратта отстает уже не так сильно, чего нельзя сказать про Наивный подход, который в худшей ситуации выполняет в 800 и 400 раз соответственно больше операций, нежели его альтернативы.

1. Проанализируем данные алгоритмы по признаку «Скорость работы».

|  |  |  |  |
| --- | --- | --- | --- |
|  | naive | boyer\_moore\_horspool | knuth\_morris\_pratt |
| good 1 | 0,0002108 | 0,0000254 | 0,0001376 |
| good 2 | 0,0001816 | 0,0000685 | 0,0001979 |
| good 3 | 0,0010713 | 0,0002390 | 0,0004142 |
| good 4 | 0,0031560 | 0,0002674 | 0,0022004 |
| bad 1 | 0,0000066 | 0,0000047 | 0,0000047 |
| bad 2 | 0,0001049 | 0,0000196 | 0,0000244 |
| bad 3 | 0,0115129 | 0,0002064 | 0,0002501 |
| bad 4 | 0,4934772 | 0,0010378 | 0,0016413 |

На хороших данных опять же лучше всего себя показывает Бойера-Мура-Хорспула, работая в несколько раз быстрее других алгоритмов.

На плохих данных ситуация повторяется, однако опять же разница между алгоритмами Бойера-Мура-Хорспула и Кнутта-Мориса-Пратта уже не столь существенна.

**Вывод**. На хороших данных лучшие показатели имеет алгоритм Бойера-Мура-Хорспула, а другие алгоритмы работают хуже. На плохих данных зависимости похожие, однако разница между работами алгоритмов Бойера-Мура-Хорспула и Кнутта-Мориса-Пратта сокращается. Наивный алгоритм является худшим алгоритмам по всем показателям. В целом практические результаты подтверждают теоретические преимущества и недостатки представленных алгоритмов.