

Суффиксные деревья

Занятие 19 в ФМЛ 5 г. Долгопрудного

Задача:

- Даны два текста T1 и T2
- Найти наибольшую общую подстроку
- Как решать?

Задача:

- Даны два текста T_1 и T_2
- Найти наибольшую общую подстроку
- Как решать **эффективно**?

Задача:

- Даны два текста $T1$ и $T2$
- Найти наибольшую общую подстроку
- Как решать:
 - Наивный метод (неэффективный)
 - Алгоритмы расстояния между строками
 - Строковые деревья

Бор(Trie)

- Рассмотрим Бор.

Бор (Trie)

- Зачем?

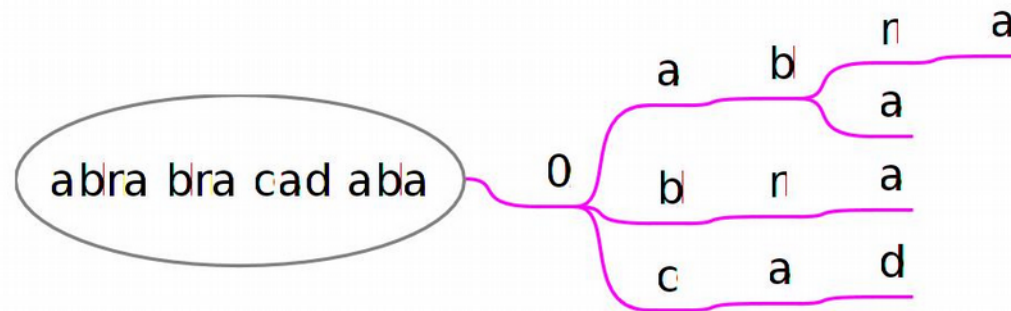
Бор (Trie)

- Зачем?
- Способ хранения похожих строк
- Поиск
- Алгоритмы сжатия (LZ77)

Бор

- Рассмотрим строки
- abra bra cad aba
- Уложим в бор

Строим бор

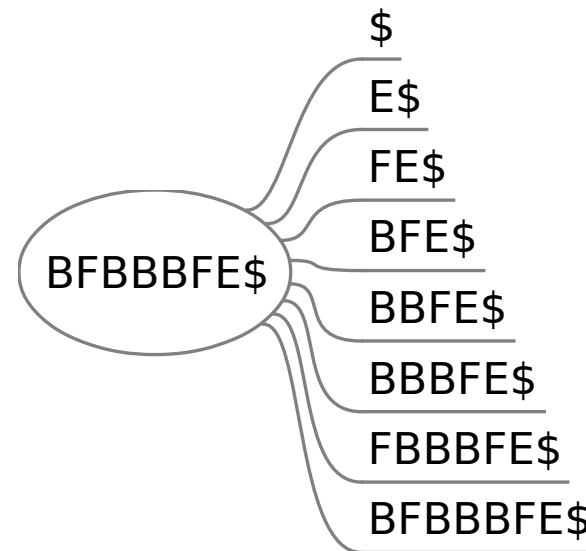


Суффиксное дерево

- Бор суффиксов

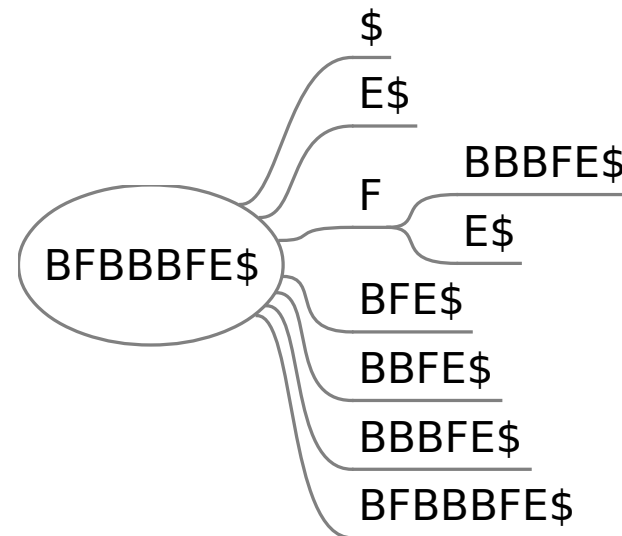
Пример

- BFBVBVFE\$
- 1 строим все



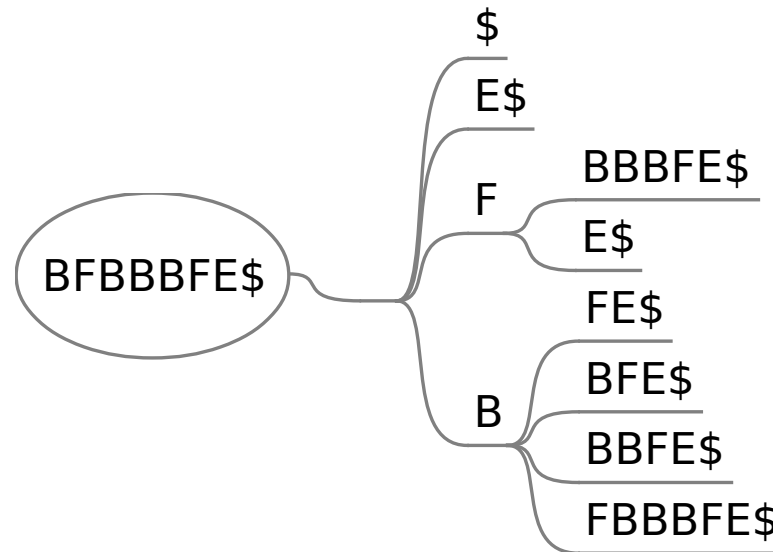
Пример

- BFBVBVFE\$
- 2 оптимизация 1



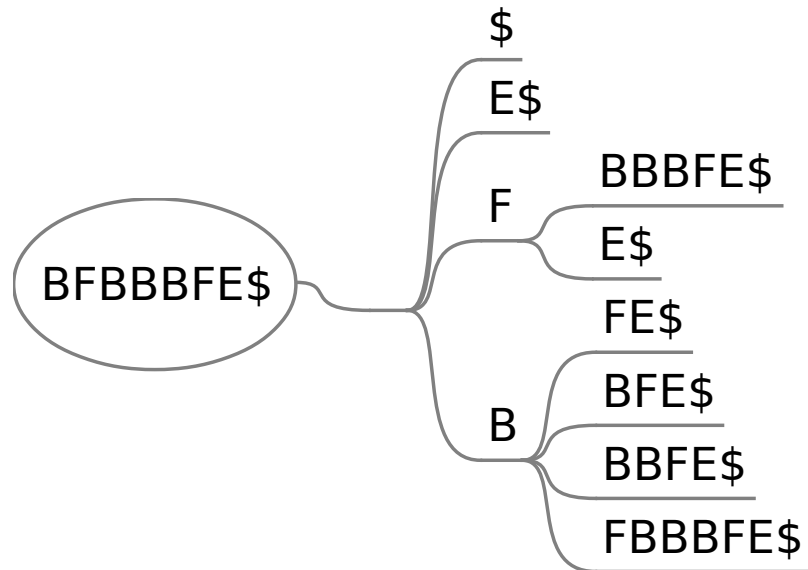
Пример

- BFBVBVFE\$
- 2 оптимизация 1



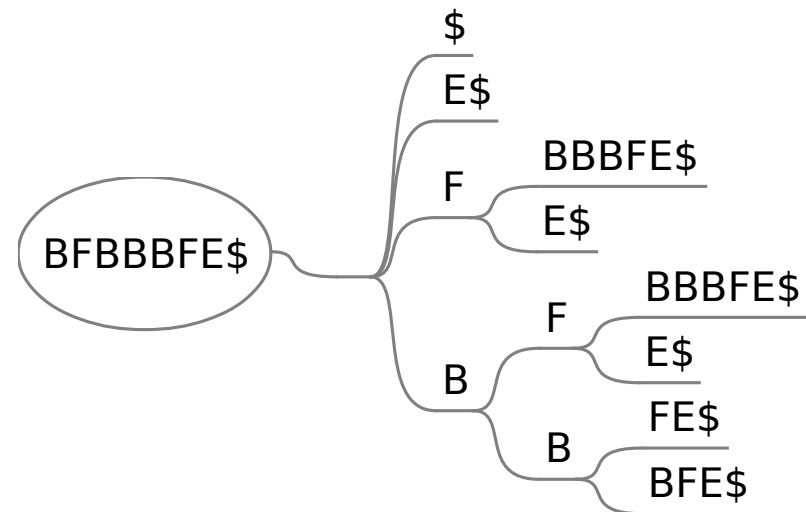
Пример

- VFBVBVFE\$
- 2 оптимизация 2



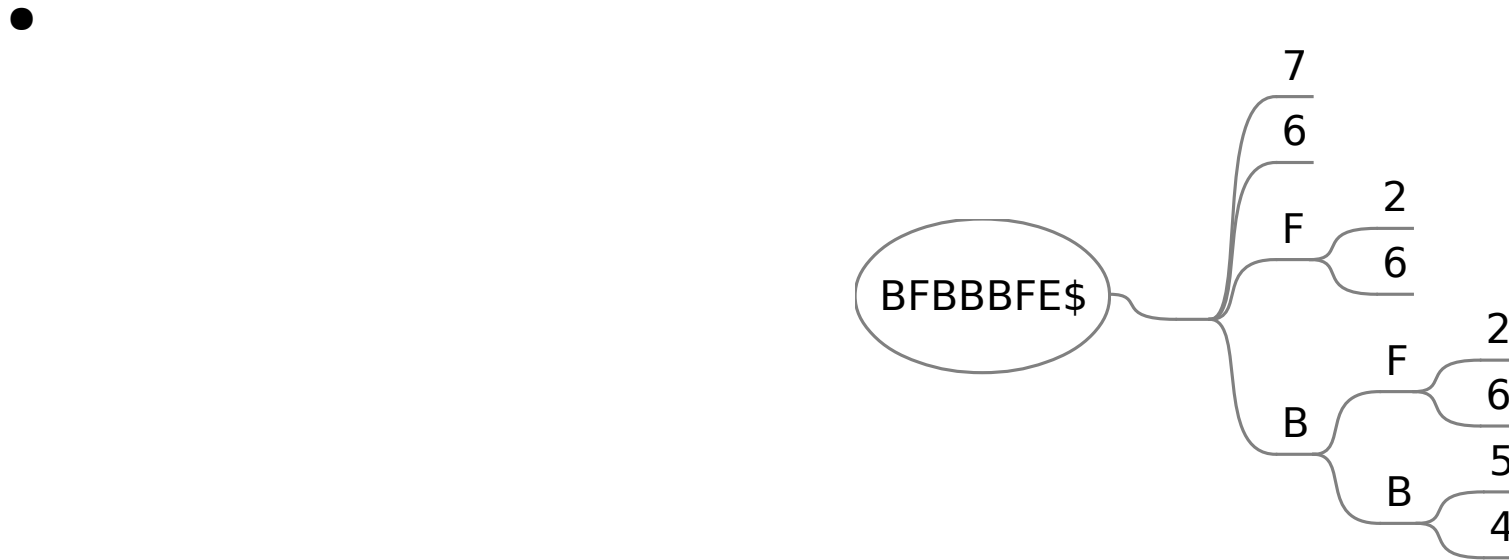
Пример

- BFBVBVFE\$
- 2 оптимизация 3
-



Компрессия

- Заменим листья на индексы с которых они начинаются



Задание

- Построить суффиксное дерево для:
- baracarmacarabaraca

Задание 2

- Реализовать поиск подстроки

Ссылки Дополнительно:

Список задач на бор

<http://codeforces.com/blog/entry/19524>

Алгоритм Укконена на c++

ideone.com/sT8Vd1

Спасибо!

Вернемся к суффиксным
деревьям когда будем говорить
об Ахо и Карасике