

Семинарское задание №1

Булев поиск

Мосиенко Константин Викторович

2017

Необходимо написать две программы:

- Индексатор. Программа должна строить индекс по всем документам *.txt из указанной директории. Директория указывается первым аргументом командной строки. Если директория не указана, необходимо взять значение "./txt/". Формат индекса остаётся на вашем усмотрении. Файлы с индексом должны находиться в текущей директории программы.
- Поисковик. Программа считывает индекс из текущей директории и предлагает пользователю ввести запрос из командной строки. Должны поддерживаться запросы, состоящие из нескольких слов, разделённых пробелом. Такие запросы интерпретируются как «документы, содержащие все слова из запроса». Ответ на запрос программа выводит в виде списка имён файлов, которые релевантны запросу, по одному на строке, но не более 10 имён. Последней строчкой выводится общее количество найденных документов. После чего программа переходит к ожиданию нового запроса.

Назовите ваши программы `make_index` и `search`. Автоматическая проверка не предусмотрена, поэтому не переживайте, из-за того, что вам кажется, что вы как-то не так разбили предложения на слова. Коллекцию документов можно взять со страницы курса: txt.tar.gz.