2016.10.09

Стовпяк Андрей

[Stovpyak@gmail.com](mailto:Stovpyak@gmail.com)

# Описание алгоритма/архитектуры

1. Файлов может быть несколько. Способы задания - какие именно файлы нужно обрабатывать могут быть разные. Поэтому заведена сущность IFileNamesSource.

В текущей версии есть одна реализация FileNameSourceStub.

Дальше можно развивать, добавляя чтение параметров командной строки, openDialog или чтение из конфигурационного файла.

1. Файлов несколько и их обработку можно вести параллельно. Заведена сущность FileAnalizer. Читает один файл построчно и делегирует обработку одной строки сущности LineAnalizer. Результатом анализа одного файла является WordsCount. FileAnalizer`ы работают в отдельных потоках.
2. Для сборки результатов обработки отдельных файлов заведена очередь WordsCountQueue. FileAnalizer после обработки «своего» файла добавляет результат в очередь и завершает свою работу.
3. За очередью «следит» Marger. Он также работает в отдельном потоке. Он ждет появления результата в очереди, извлекает его и добавляет его в «общий» результат. Признаком завершения работы является marge стольких результатов, сколько было передано файлов для обработки.
4. Итоговый результат аккумулируется в IResultSaver.

Создано две реализации:

ResultSaverSpy – собирает результат, никуда его не сохраняет, есть возможность посмотреть результат (используется в тестах).

ResultSaverToFile - собирает результат, сохраняет его в файл.

При дальнейшем развитии можно добавить saver для базы данных.

1. Создание и связывание частей реализовано в сущности Appl

Она парамметризуется IFileNamesSource (в конструкторе) и IResultSaver (в методе Execute).

# Что можно улучшить

При распознавании слов в тексте с переносами отдельные части перенесенных слов считаются отдельными словами.

Проблему можно решить - распознать две различные ситуации: когда слово содержит дефис (чуть-чуть, во-первых) и когда слово перенесено на новую строку можно только подключив словарь. В текущей версии это не сделано.

# Уточняющие вопросы

Получившуюся библиотеку можно использовать в различных приложениях.

## Получение списка файлов

Это может быть, как консольное приложение, получающее параметры в виде параметров командной строки, так и приложение с графическим интерфейсом и выбором файлов в openDialog`е или же сервис, получающий имена файлов каким то другим образом.

## Сохранение результата

Реализовано сохранение результатов в файл.

Сохранение результатов в базу будет заключаться в создании ещё одной реализации IResultSaver.

Необходимо ли дальше развивать приложение или текущего результата достаточно для демонстрации?