Пользовательская документация к проекту:

**Программа для нахождения дублирующих директорий**

**Автор: Устинов Владимир**

**Группа 5103**

**ФНБИК МФТИ**

**2017 год**

Содержание:

1. Цель проекта

2. Концепция работы программы

3. Реализация алгоритма и построение GUI

а) Использованные библиотеки

б) Описание функций

4. Инструкция для пользователя

Цель проекта:

Раньше существенной проблемой являлся малый объем жестких дисков. Объемы современных дисков достигли поразительных объемов, однако зачастую из-за невнимательности пользователей винчестер постепенно "засоряется", в разных папках оказываются одинаковые файлы. На смену вопросу "где хранить данные?" пришёл новый: "как не запутаться в таком количестве данных?".

Рано или поздно, приходит время, когда на жестком диске становится слишком мало места, или же беспорядок в папках начинает мешать пользователю работать.

Дубликаты — полностью идентичные (одинаковые) файлы/директории. Они существенно загрязняют жесткий диск, поэтому от них нужно периодически избавляться, ведь в зависимости от степени "беспорядка" в файловой системе можно освободить до нескольких гигабайтов места.

Удаление дубликатов невозможно произвести без посторонних программ, поскольку разработчики операционных систем не добавили такую функцию, а в ручную заниматься поиском дубликатов просто неудобно, именно поэтому написание подобной программы крайне актуальным и полезным.

Концепция работы программы:

Программа работает в двух режимах: первый проверяет директории на идентичность по количеству одинаковых файлов, второй по суммарному объёму одинаковых файлов.

Должна находить в том числе и частичные дубликаты. Например, если совпадают 95% файлов по количеству или 95% по суммарному объёму (пороговое значение задается пользователем), директории считаются идентичными.

Помимо обычной проверки, программа рекурсивно обрабатывает вложенные в каждую из рабочих директорий поддиректории и выводит группы идентичных директорий.

Реализация алгоритма:

***Используемые библиотеки***

Базовый язык программирования -- C++.

Проект написан в IDE Code::Blocks.

Для реализации поставленной в ТЗ задачи были подключены следующие библиотеки:

* Для реализации GUI использована кросплатформенная библиотека инструментов wxWidgets.
* Для создания форм использовался инструмент wxFormBuilder, инструмент быстрой разработки графической оболочки приложений на основе wxWidgets.

***Описание основных функций:***

*Fill -* функция заполнения ассоциативного массива "директория-файл".

*Binds -* основной обработчик. Получает карту директорий-файлов, перебирает все директории, пустые запоминает отдельно. Среди непустых сравнивает каждую со всеми последующими в выбранном режиме. Если директории эквивалентны, то добавляем пути в набор дубликатов и сохраняем процент соответствия. Отдельно выводит уникальные директории. В этой же функции происходит отображение итоговой таблицы, отображение пустых директорий и вывод результатов в файл.

*CountCompare -* количественное сравнение. Отталкиваясь от порогового значения, определяем максимально допустимое количество различных файлов в идентичных директориях. Если количество файлов в папках отличается на большее - возврат. Далее начинаем сравнивать файлы в директориях. Если размер совпадает, ранее этот файл не был найден и при этом он идентичен проверяемому, добавляем в помеченные и отмечаем соответствие.

*VolumeCompare -* объемное сравнение. Принцип и алгоритм полностью идентичны предыдущей функции.

Инструкция для пользователя:

**Работа с GUI:**

Пользователь выбирает исходную директорию, устанавливает пороговое значение совпадений в процентах (большее или равное 50%), выбирает режим проверки (по объему или по количеству), запускает работу программы.

Если файлы в поддиректориях совпадают на указанное число процентов или более, программа сообщает об этом, если нет -- сообщает, что директории не являются идентичными.

По окончании работы программы в правом нижнем окне появляется список уникальных директорий, в левом – список идентичных. По нажатию на клавишу «Отобразить» в новом окне открывается таблица, показывающая пары идентичных директорий с указанием соответствующего процентного совпадения их содержания.