|  |  |
| --- | --- |
| Автоматическая генерация оглавления HTML-страницы | |
| Внешняя спецификация | |
| Студент | Чернобаев Е. С. |
| Преподаватель |  |
| Сдано |  |

**1. Назначение**

Программа предназначена для осуществления автоматической генерации оглавления HTML-страницы.

**2. Функциональные требования**

**2.1 Действия над объектами**

Программа должна генерировать оглавление HTML-страницы.

Программа должна осуществлять парсинг сайта по ссылке.

Программа должна уметь работать с файлами (входные/выходные данные).

**2.2 Ограничения**

В качестве входных данных программа может принимать HTML код.

В качестве входных данных программа может принимать URL-адрес сайта.

Если в качестве входных данных передается URL-адрес сайта, то он должен быть действующим.

Под “действующим” URL-адресом сайта понимается URL-адрес, который ведет на существующий сайт.

URL-адрес сайта должен обязательно содержать протокол (http:// или https://).

При составлении оглавления программа ориентируется только на теги <h1>, <h2>, <h3> и т.д.

Под генерацией оглавления понимается изменение имеющихся в HTML-коде «Н» заголовков.

Результатом работы программы должен быть HTML-код с оглавлением HTML-страницы, построенное с учетом вложенности «Н» заголовков.

Оглавление HTML-страницы, полученное в результате работы программы, должно визуально отображать вложенность заголовков при помощи отступов.

**3. Входные и выходные данные**

Программа должна получать два параметра командной строки: исполняющий файл и абсолютный путь входного файла с кодом HTML-страницы или действующий URL-адрес сайта.

Пример команды запуска программы (аргументом является абсолютный путь):

app.exe C:\Documents\input.html

Пример команды запуска программы (аргументом является URL-адрес):

app.exe http://google.com

Входные данные представляются в виде текстового файла с расширением .html или URL-адреса, передаваемые в качестве аргумента при запуске программы.

Выходной файл должен быть текстовым с расширением .html. Он должен содержать оглавление HTML-страницы, построенное на основании заголовков «Н» с учетом их вложенности.

Файл с результатом работы программы сохраняется по пути ./Result/ в файл result.html.

**4. Требования к надежности**

В процессе работы программы не должно происходить ее аварийного завершения или зависания.

В случае ошибки во входных данных, пользователь должен получать сообщения, перечисленные в приложении Б, после чего программа должна корректно завершаться.

Сообщения об ошибке выводятся в консоль и выходной файл при этом не создается.

**5. Требования программной совместимости**

Программа будет разработана на языке С++ с использованием библиотеки Qt и стандартных библиотек языка С++.

Используемое ПО:

Qt Creator 9.0.1;

OpenSSL 3.0.8.

Входные файлы могут быть подготовлены в редакторе Блокнот. Выходные файлы могут быть прочитаны с помощью того же ПО.

**Приложение А**

**Примеры входных и выходных данных**

|  |  |
| --- | --- |
| Входные данные (файл/URL-адрес) | Выходные данные (result.html) |
| <http://google.com> | \*код HTML-страницы по заданному URL-адресу с учетом вложенности <Н> заголовков\* |
|  |  |
| <h1>  First Header  <h2>  Second header  <h3>  Hello in SECOND  </h3>  <h3>  World in SECOND  </h3>  </h2>  <h2>  Third header  <h3>  Hello in Third  </h3>  </h2>  </h1> | <h1>  First Header  <h2 style="margin-left: 20px;">  Second header  <h3 style="margin-left: 40px;">  Hello in SECOND  </h3>  <h3 style="margin-left: 40px;">  World in SECOND  </h3>  </h2>  <h2 style="margin-left: 20px;">  Third header  <h3 style="margin-left: 40px;">  Hello in Third  </h3>  </h2>  </h1> |
|  |  |
| <li>Smaller header 1</li>  <li>Smaller header 2</li>  <ul>  <li>Smallest header 1</li>  <li>Smallest header 2</li>  </ul> | Выходной файл не создается и пользователю выводится ошибка:  There are no <h> headers in the code. Perhaps the code is missing. |

**Приложение Б**

Список ошибок с расшифровкой

|  |  |  |
| --- | --- | --- |
| Ситуация | Пример | Сообщение об ошибке |
| Указанный входной файл не существует, нет доступа к указанному файлу. |  | The input data file is specified incorrectly. The file may not exist. |
| Неверный формат входного файла. | input.docx | The format of the input file does not meet the requirements of the software. |
| Файл, указанный в абсолютном пути, не удалось открыть. |  | The file could not be opened at the specified path. |
| Входной файл пустой. |  | Input file is empty. |
| В коде отсутствуют <h> заголовки. Возможно, код отсутствует. | <li>Smaller header 1</li>  <li>Smaller header 2</li>  <ul>  <li>Smallest header 1</li>  <li>Smallest header 2</li>  </ul> | There are no <h> headers in the code. Perhaps the code is missing. |