

Politechnika Śląska
Wydział Informatyki, Elektroniki i Informatyki

Programowanie Komputerów

Mapa Witryny

autor	Michał Pawłowski
prowadzący	mgr inż. Grzegorz Kwiatkowski
rok akademicki	2019/2020
kierunek	informatyka
rodzaj studiów	SSI
semestr	2
termin laboratorium	piątek, 12:00 – 13:30
sekcja	22
termin oddania sprawozdania	2020-07-06

1 Treść zadania

W folderze znajduje się strona internetowa zapisana w kilkunastu plikach o rozszerzeniu html. Strona główna zawiera hiperłącza do podstron. Podstrony zawierają hiperłącza do swoich podstron oraz strony nadrzędnej. Przykładowy kod podstrony `page1a.html`:

```
<html>
<head>
<title> 0 Uczelni </title>
</head>
<body>
<h1>0 Uczelni</h1>
<a href = "page1aa.html"> Informacje </a>
<a href = "page1ab.html"> Historia </a>

<TREŚĆ>

<a href = "page1.html"> Powrót </a>
</body>
</html>
```

W wyniku działania programu zostanie tworzony plik z mapą witryny

```
> Strona Główna
  > Uczelnia
    > 0 Uczelni
      > Informacje
      > Historia
    > Kandydat
...
```

Program uruchamiany jest z linii poleceń z wykorzystaniem następujących przełączników:

- i plik wejściowy ze stroną główną witryny
- o nazwa pliku wyjściowego
- e rozszerzenie pliku wyjściowego

2 Analiza zadania

Zagadnienie przedstawia problem stworzenia mapy witryny w oparciu o drzewko plików witryny dostępnych dla programu.

3 Specyfikacja zewnętrzna

Program jest uruchamiany z linii poleceń. Należy przekazać do programu nazwy plików: wejściowego, wyjściowego oraz rozszerzenie w którym ma być stworzony plik wyjściowy, korzystając z odpowiednich przełączników (odpowiednio: `-i` dla pliku wejściowego, `-o` dla nazwy pliku wyjściowego oraz `-e` dla rozszerzenia pliku wyjściowego), np.

```
-i index.html -o websiteMap -e html  
-e txt -i index.html -o map
```

Plik wejściowy jest plikiem z rozszerzeniem `.html`. Plik wyjściowy może być wygenerowany z rozszerzeniem `.txt` lub `.html`, przy czym dla drugiego rozszerzenia aktywne będą hiperłącza do danych stron witryny. Przełączniki mogą być podane w dowolnej kolejności. Uruchomienie programu bez żadnego parametru lub z niewłaściwą ich liczbą powoduje wyświetlenie krótkiej pomocy oraz komunikatu

```
Program must be started with 3 parameters!
```

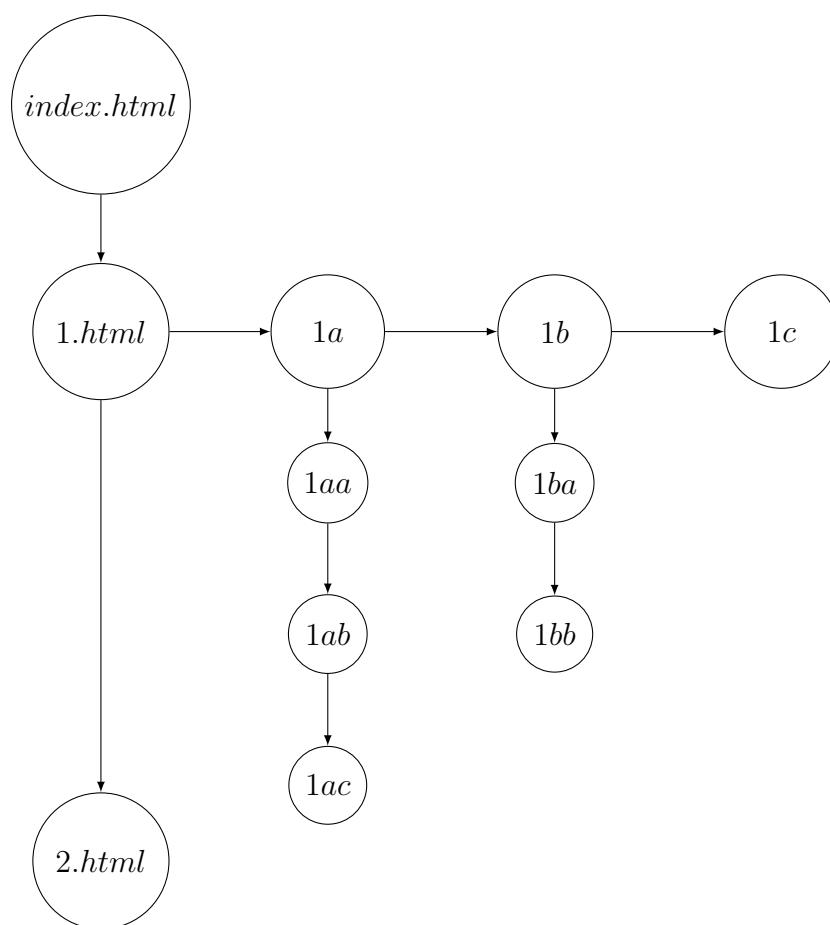
Podanie nieprawidłowej nazwy pliku wejściowego powoduje wyświetlenie odpowiedniego komunikatu:

```
WARNING: File error!
```

4 Specyfikacja wewnętrzna

4.1 Ogólna struktura programu

W funkcji głównej wywoływana jest funkcja `paramsCheck`. Funkcja ta sprawdza, czy program został uruchomiony w prawidłowy sposób. Gdy program nie został wywołany prawidłowo, funkcja `help` wypisuje stosowny komunikat i program się kończy. Następnie wywoływana jest funkcja `createFirst`. Funkcja ta tworzy pierwszy element listy podwieszanej na wcześniej utworzonym wskaźniku. Zaalokowany element posiada jedynie nazwę pliku strony głównej. Nazwa ta została przekazana do programu w jednym z parametrów startowych. Pozostałe wskaźniki pierwszego elementu są wyzerowane.



Rysunek 1: Przykład odwzorowania mapy witryny w pamięci programu poprzez listę wielokrotnie podwieszaną.

Kolejną wywoływaną funkcją jest **generate**. Funkcja otwiera plik strony startowej i kopiuje jego zawartość do bufora. Zamyka plik a następnie za pomocą funkcji **addTitle** wyszukuje tytuł aktualnie rozpatrywanej strony i przypisuje go do odpowiedniego elementu struktury. Ponowne przeszukanie bufora tym razem przy użyciu funkcji **addSubPages** tworzy nowe elementy listy podwieszanej odpowiadające podstronom bieżąco rozpatrywanej strony. Na końcu funkcja zwalnia pamięć bufora oraz rekurencyjnie wywołuje samą siebie dla pierwszego elementu listy odchodzącej oraz kolejnego elementu listy bieżącej. Na koniec, wynik działania programu zostaje wypisany na konsoli przez funkcję **printMap** oraz zapisany do pliku dzięki funkcji **saveMap**. **deleteList** usuwa całą zaalokowaną strukturę.

4.2 Szczegółowy opis typów i funkcji

Szczegółowy opis typów i funkcji zawarty jest w załączniku.

5 Testowanie

Program został przetestowany na różnych plikach wejściowych. Pliki niepoprawne czyli zawierające błędny kod lub też pliki puste są równoważne stronom html nieposiadającym hiperłącz do podstron, zatem wygenerowany plik mapy witryny posiada jedynie stronę wczytaną do programu poprzez parametr startowy. Program został sprawdzony pod kątem wycieków pamięci.

6 Wnioski

Program generujący mapę witryny jest programem dość prostym. Wymagające jednak okazało się alokowanie elementów listy wielokrotnie podwieszanej oraz poprawne ich zwalnianie z pamięci.

Dodatek

Szczegółowy opis typów i funkcji