

計算機網路概論

Lab2

EECS 26' 劉祐廷 111060013

Catalog

A. Self Check.....	P3
B. Functions.....	P3
C. Output Result.....	P7
D. What I Have Learned.....	P9

A. Self Check

- Compiled by gcc ----- done
- Print all hyper links ----- done
- Print the total number of hyperlinks ----- done
- Run on the virtual machine ----- done

B. Functions

1. getResponse()

負責處理輸入的內容，宣告 socket 與 server 端取得連線，發送 request message 並將 response message 寫入 responseFile.txt。

```
47     printf("Please enter the URL:\n");
48     while (1) {
49         c = getchar();
50         if (c == '\n' || c == EOF) break;
51         else if (isHost) {
52             if (c == '/') {
53                 path[0] = '/';
54                 isHost = 0;
55                 idx = 1;
56             }
57             else hostName[idx++] = c;
58         }
59         else path[idx++] = c;
60     }
61
62     fprintf(log, "Host Name: %s\n", hostName);
63     fprintf(log, "Path: %s\n", path);
64
65     struct hostent* entry = gethostbyname(hostName);
66     char* IP_address = inet_ntoa(*(struct in_addr *) entry->h_addr_list[0]);
67
68     fprintf(log, "IP: %s\n", IP_address);
```

圖 1

圖 1 為處理輸入的程式碼，使用 `getchar()` 將輸入一個字一個字放進 `hostName`，若遇到 `'/'` 則開始將讀取到的字元放進 `path`，使用 `fprintf()` 將 `hostName` 與 `path` 輸出在 `log.txt` 以方便 debug。接著使用 `gethostbyname(char*)` 將 `hostName` 轉成 `hostent` 結構的資料，從中提取出 IP address，使用 `inet_ntoa()` 將其轉為字串的形式並存於 `IP_address` 中，接著將 `IP_address` 輸出確認 IP 是否正確。

```

70  int sockfd;
71  struct sockaddr_in server_addr;
72  socklen_t addrlen = sizeof(server_addr);
73  char message[2048] = {0};
74  unsigned char buffer[BUFFER_SIZE] = {'\0'};
75
76  server_addr.sin_family = AF_INET;
77  server_addr.sin_addr.s_addr = inet_addr(IP_address);
78  server_addr.sin_port = htons(PORT);
79
80  // create a socket, return -1 if fail
81  if ((sockfd = socket(AF_INET, SOCK_STREAM, 0)) == -1) {
82      perror("socket()");
83      exit(EXIT_FAILURE);
84  }
85  if (connect(sockfd, (struct sockaddr *)&server_addr, addrlen) == -1) {
86      perror("connect()");
87      fprintf(stderr, "Please start the server first\n");
88      exit(EXIT_FAILURE);
89  }
90  fprintf(log, "Connected to server\n");

```

圖 2

宣告一個結構為 `sockaddr_in` 的變數 `server_addr`，將其設定為 IPV4 (`AF_INET`)，並將 IP address 與 port number 設定好，接著使用 `socket()` 去產生一個 socket，其中 `AF_INET` 代表使用 IPV4，`SOCK_STREAM` 代表 TCP，創建 socket 成功後使用 `connect()` 連上 server，傳入的參數為 (`socket`, `server_addr`, `server_addr` 的 size)，若 `connect` 回傳值不是 -1，則代表連接成功。

```

92  memset(message, 0, sizeof(message));
93  strcpy(message + strlen(message), "GET /");
94  strcpy(message + strlen(message), path);
95  strcpy(message + strlen(message), " HTTP/1.1\r\nHost: ");
96  strcpy(message + strlen(message), hostName);
97  strcpy(message + strlen(message), "\r\nConnection: close\r\n\r\n");
98
99  send(sockfd, message, strlen(message), 0);
100  fprintf(log, "HTTP request sent\n\n");
101  fprintf(socketFile, "Sensing HTTP request\n");

```

圖 3

接著開始製作 request message，利用剛剛已經做好的 `path` 與 `hostName`，使用 `strcpy()` 慢慢合成，並使用 `send()` 將 message 送出去 server。

```
103     while (recv(sockfd, buffer, BUFFER_SIZE, 0) > 0) {
104         fprintf(responseFile, "%s", buffer);
105         memset(buffer, 0, sizeof(buffer));
106     }
107     fprintf(responseFile, "\n");
108     fprintf(log, "recieve html file\n");
109     fprintf(socketFile, "Recieving the response\n");
```

圖 4

最後使用 `recv()` 將 server 傳回來的資料放進 buffer array，並將所收到的 response 存入 `responseFile.txt`。

2. `getHtml()`

```
120     FILE* responseFile = fopen("responseFile.txt", "r");
121     FILE* htmlFile = fopen("htmlFile.txt", "w");
122
123     char c, start = 0;
124     while (~fscanf(responseFile, "%c", &c)) {
125         if (c == '<') start = 1;
126         if (start) fprintf(htmlFile, "%c", c);
127     }
```

圖 5

將 `reponseFile.txt` 裡面的 header 去除，留下 html 的部分，將其存入 `htmlFile.txt`。

3. `getRidOfComment()`

原本是用來去除註解的，但是題目要求註解內的 hyperlinks 也要輸出，因此就沒有用到它了。

4. getHref()

```
185 char* cursor = buffer;
186
187 memset(output, 0, sizeof(output));
188
189 char* aHead, * aTail, * aClose, * hrefHead, * hrefTail;
190
191 int totalHref = 0;
192
193 while ((aHead = strstr(cursor, "<a")) {
194     aTail = strstr(aHead, ">");
195     aClose = strstr(aHead, "</a>");
196     hrefHead = strstr(aHead, "href=");
197     if (hrefHead && hrefHead < aTail) {
198         char quotation[2] = {*(hrefHead + strlen("href=")), 0};
199         hrefTail = strstr(hrefHead + strlen("href=") + 1, quotation);
200         strncpy(output + outputLen, hrefHead + strlen("href=") + 1, hrefTail - (hrefHead + strlen("href=") + 1));
201         outputLen += hrefTail - (hrefHead + strlen("href=") + 1);
202         output[outputLen++] = '\n';
203         totalHref++;
204     }
205     cursor = /*aClose + strlen("</a>")*/ aTail + 1;
206 }
207
208 output[outputLen] = 0;
```

圖 5

這個函式是用來尋找 hyperlinks 的，首先得找出標籤 a 存成指標 aHead，接著找到 '>' 存成 aTail，並尋找 href= 存成 hrefHead，若 hrefHead < aTail 則找到 hyperlinks，接著將引號內的 hyperlink 提取出來複製到 output array，然後將 cursor 移至 aTail 之後，並繼續尋找直到找不到為止。最後將 output array 輸出至 hrefFile.txt。

5. outputToTerminal()

```
223 FILE* socketFile = fopen("socketFile.txt", "r");
224 FILE* hrefFile = fopen("hrefFile.txt", "r");
225
226 char c;
227
228 while (~fscanf(socketFile, "%c", &c)) {
229     printf("%c", c);
230 }
231
232 while (~fscanf(hrefFile, "%c", &c)) {
233     printf("%c", c);
234 }
```

圖 6

將剛剛存好的檔案輸出到終端機。

C. Output Result

```

liuutin9@liuutin9-Modern-14-B11M: ~/桌面/Lab2
liuutin9@liuutin9-Modern-14-B11M:~$ cd 桌面
liuutin9@liuutin9-Modern-14-B11M:~/桌面$ cd Lab2
liuutin9@liuutin9-Modern-14-B11M:~/桌面/Lab2$ ./main.exe
Please enter the URL:
can.cs.nthu.edu.tw
===== Socket =====
Senging HTTP request
Recieving the response
===== Hyperlinks =====
index.php
members.php
http://www.cs.nthu.edu.tw/~jungchuk/home.html
LAB/
gallery.php
contact.php
http://web.cs.nthu.edu.tw/files/14-1015-143485,r109-1.php?Lang=zh-tw
http://www.nthu.edu.tw
http://web.cs.nthu.edu.tw/bin/home.php
http://www.com.nthu.edu.tw/
http://www.highimpact-seo.co.uk/
=====
We have found 11 hyperlinks
liuutin9@liuutin9-Modern-14-B11M:~/桌面/Lab2$

```

```

liuutin9@liuutin9-Modern-14-B11M:~/桌面/Lab2$ ./main.exe
Please enter the URL:
can.cs.nthu.edu.tw/
===== Socket =====
Senging HTTP request
Recieving the response
===== Hyperlinks =====
index.php
members.php
http://www.cs.nthu.edu.tw/~jungchuk/home.html
LAB/
gallery.php
contact.php
http://web.cs.nthu.edu.tw/files/14-1015-143485,r109-1.php?Lang=zh-tw
http://www.nthu.edu.tw
http://web.cs.nthu.edu.tw/bin/home.php
http://www.com.nthu.edu.tw/
http://www.highimpact-seo.co.uk/
=====
We have found 11 hyperlinks
liuutin9@liuutin9-Modern-14-B11M:~/桌面/Lab2$

```

```

liuutin9@liuutin9-Modern-14-B11M:~/桌面/Lab2$ ./main.exe
Please enter the URL:
can.cs.nthu.edu.tw/index.php
===== Socket =====
Senging HTTP request
Recieving the response
===== Hyperlinks =====
index.php
members.php
http://www.cs.nthu.edu.tw/~jungchuk/home.html
LAB/
gallery.php
contact.php
http://web.cs.nthu.edu.tw/files/14-1015-143485,r109-1.php?Lang=zh-tw
http://www.nthu.edu.tw
http://web.cs.nthu.edu.tw/bin/home.php
http://www.com.nthu.edu.tw/
http://www.highimpact-seo.co.uk/
=====
We have found 11 hyperlinks
liuutin9@liuutin9-Modern-14-B11M:~/桌面/Lab2$

```

```

liuutin9@liuutin9-Modern-14-B11M:~/桌面/Lab2$ ./main.exe
Please enter the URL:
can.cs.nthu.edu.tw/contact.php
===== Socket =====
Senging HTTP request
Recieving the response
===== Hyperlinks =====
index.php
members.php
http://www.cs.nthu.edu.tw/~jungchuk/home.html
LAB/
gallery.php
contact.php
http://web.cs.nthu.edu.tw/files/14-1015-143485,r109-1.php?Lang=zh-tw
http://www.nthu.edu.tw
http://web.cs.nthu.edu.tw/bin/home.php
http://www.com.nthu.edu.tw/
http://www.highimpact-seo.co.uk/
=====
We have found 11 hyperlinks
liuutin9@liuutin9-Modern-14-B11M:~/桌面/Lab2$

```

```
canlab@ubuntu:~/Desktop/Lab$ gcc lab.c -o lab.exe
canlab@ubuntu:~/Desktop/Lab$ ./lab.exe
Please enter the URL:
www.cs.nthu.edu.tw/~jungchuk/home.html
===== Socket =====
Senging HTTP request
Recieving the response
===== Hyperlinks =====
home.html
education.html
publication.html
research.html
project.html
professional.html
teaching.html
lab.html
contact.html
http://www.cs.nthu.edu.tw
http://www.nthu.edu.tw
http://www.com.nthu.edu.tw
http://www.nthu.edu.tw
course/Probability_2021Spring/index.html#adding
course/IntroductionToComputerNetworks_2021Fall/index.html#adding
course/IntroductionToComputerNetworks_2021Fall/index.html#distancelearning
course/BroadbandMobileCommunications_2021Fall/index.html#adding
course/BroadbandMobileCommunications_2021Fall/index.html#distancelearning
research.html
publication.html
research.html
contact.html
research.html
contact.html
research.html
research.html
lab.html
contact.html
http://csim.tca.org.tw/award.aspx
research.html
contact.html
=====
We have found 31 hyperlinks
```


D. What I Have Learned

從這次 lab 中，我學到了如何使用 c 去做連線，以前曾經想嘗試自己實作，但無奈當時能力不足，看不懂網路上的教學，剛好這次 lab 有機會讓我接觸這個東西，其實還挺有趣的。從這次 lab 我也可以很清楚地了解到 client 與 server 之間是怎麼運作的，透過有點像一問一答的方式，將 client 需要的內容從 server 一傳進來，這次的 lab 也讓我對於標頭檔的格式以更深的印象，希望這次學到的新技能可以在未來的某一刻用到。