

Kommunikationsnetze – Übung 8

HTTP-Parser

Lernziele:

- Begriff "Parser"
- Zeichenketten durchsuchen
- Die Funktion `strstr()`
- Text aus Zeichenketten extrahieren
- Die Funktion `strncpy()`

Begriff "Parser"

Problem: Web-Server erhält vom Browser ein "Request"-Paket
Web-Server muss diese Anforderung analysieren

1) Welches Kommando hat der Browser gesendet ?

→ Paket nach erlaubten Befehlen durchsuchen
z.B.: nach Befehl GET (siehe 1. Aufgabe)

2) Welche Parameter hat dieses Kommando ?

z.B.: Name der Datei, die zum Browser gesendet werden soll
→ Parameter extrahieren (siehe 2. Aufgabe)

Lösung: ein Parser = ein Programm, das Texte analysiert

Zeichenketten durchsuchen

Problem: Ist das Wort GET im empfangenen Paket enthalten ?

Wenn ja, wo genau befindet sich das Wort (Position?) ?

Lösung: C-Funktion `strstr()` sucht einen Text in einer Zeichenkette

Die Funktion gehört zur String-Bibliothek (`string.h`)

Aufruf: `zeiger = strstr(zeichenkette, gesuchter_text);`

Ergebnis: ein Zeiger auf den Anfang den gesuchten Text

oder: NULL, wenn der Text nicht gefunden wird

Test: `if (zeiger)`

ist wahr, wenn der gesuchte Text gefunden worden ist

Beispiel

Empfangene Zeichenkette (Char-Array empfangen):

G	E	T		/	t	e	s	t	.	h	t	m	l		H	T	T	P	/	1	.	0	
---	---	---	--	---	---	---	---	---	---	---	---	---	---	--	---	---	---	---	---	---	---	---	--

0x2FA567 position

Erstes Zeichen von empfangen steht an der Adresse 0x2FA567

`char *position;` ein Zeiger, für die Adresse des gesuchten Textes

`position = strstr(empfangen, "GET");`

relative Position: `position - empfangen + 1` (ergibt hier 1)

allgemein berechnet man damit, an welcher Stelle der Text steht

Den Dateinamen extrahieren

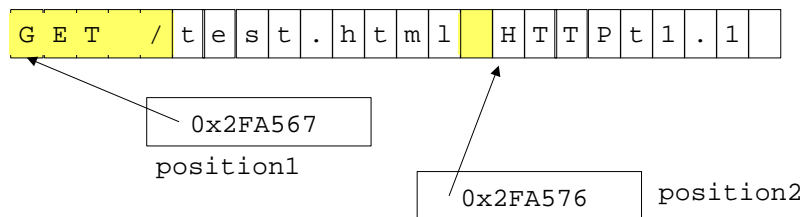
Prinzipielle **Vorgehensweise** bei Aufgabe 2

laut HTTP-Protokoll steht der Dateiname zwischen dem Wort GET und dem Wort HTTP.

- 1) Position des Wortes **GET** suchen (Zeiger position1)
- 2) Position des Wortes **HTTP** suchen (Zeiger position2)
- 3) daraus die **Länge** des Dateinamens berechnen
- 4) Dateiname mit Hilfe der Funktion `strncpy()` aus dem empfangenen Datenpaket herauskopieren

Beispiel

Empfangene Zeichenkette (Char-Array empfangen):



Länge des Dateinamens:

```
laenge = position2 - position1 - 6;
```

Erstes Zeichen des Dateinamens hat die Adresse:

```
position1 + 5
```

Text extrahieren

Problem: Text einer bestimmten Länge aus einer Zeichenkette kopieren

Lösung: C-Funktion `strncpy()`

Die Funktion gehört zur String-Bibliothek (`string.h`)

Aufruf: `strncpy(ziel, quelle, laenge);`

Beispiel:

```
char dateiname[100];  
laenge = position2 - position1 - 6;  
strncpy(dateiname, position1+5, laenge);
```

Musterlösung Aufgabe 2

```
if (position1 = strstr(empfangen, "GET"))  
{  
    printf("\nGET-Befehl beginnt bei Zeichen %d", (position1-empfangen)+1);  
    if (position2 = strstr(empfangen, "HTTP"))  
    {  
        printf("\nHTTP-Befehl beginnt bei Zeichen %d", (position2-empfangen)+1);  
        laenge = position2 - position1 - 6;  
        strncpy(dateiname, position1+5, laenge);  
        dateiname[laenge]=0;  
        printf("\nDateiname = %s", dateiname);  
    }  
    else  
    {  
        printf("\nDer HTTP-Parser hat keinen HTTP-Befehl gefunden");  
    }  
}  
else  
{  
    printf("\nDer HTTP-Parser hat keinen GET-Befehl gefunden");  
}
```