



Systemnahe Software I (WS 2019/2020)

Abgabe bis zum 31. Januar 2020, 14:00 Uhr

Lernziele:

- Umgang mit Dateien, die in den eigenen Adressraum abgebildet worden sind
- Ein- und Ausgabe mit Systemaufrufen

Aufgabe 13: Kommentare extrahieren

Im Rahmen dieser Aufgabe ist ein kleines Werkzeug entwickeln, das nach C-Kommentaren innerhalb einer Datei sucht und diese auf der Standardausgabe ausgibt. Der Dateiname soll als Argument beim Programmstart übergeben werden. Ihre Implementierung darf die *stdio*-Bibliothek nicht verwenden. Stattdessen sind nur Systemaufrufe und die *stralloc*-Bibliothek zulässig.

Hier ist ein beispielhafter Aufruf:

```
theon$ cat testit.c
/* regular comments
   can extend over
   multiple lines */
int i; // this comment is ended by the line terminator
/*
   // within a comment "//" has no special meaning */ int j;
*/ this is still a comment */
**** this a comment with quite a number of asterisks ****/
/* comments cannot be /* nested */ int k;
/* multiple */ int l; /* comments on a */ int m; // line
theon$ extract_comments testit.c
/* regular comments
   can extend over
   multiple lines */
// this comment is ended by the line terminator
```

```
/*  
    // within a comment "//" has no special meaning */  
*/ this is still a comment */  
**** this a comment with quite a number of asterisks ****/  
/* comments cannot be /* nested */  
/* multiple */  
/* comments on a */  
// line  
theon$
```

Die Datei ist mit *open* zum Lesen zu öffnen (*O_RDONLY* aus *fcntl.h*). Danach ist es möglich, diese mit Hilfe des Systemaufrufs *mmap* in den eigenen Adressraum abzubilden. Dies ist vergleichbar zu dem Vorlesungsbeispiel *mcopy.c*. Am Ende ist die Datei mit *munmap* und *close* wieder zu schließen.

Die Programmiersprache C kennt zwei Arten von Kommentaren, die beide zu unterstützen sind:

- „/*“ ... „*/“ (über beliebig viele Zeilen, keine Verschachtelungen)
- „//“ ... \n (bis zum Zeilenende)

Es empfiehlt sich hier, einen endlichen Automaten (ähnlich wie bei der 5. Aufgabe) zu entwickeln und diesen umzusetzen. Sobald Sie einen Kommentar entdecken, ist dieser einschließlich der Begrenzungszeichen auf der Standardausgabe (Dateideskriptor 1) mit einem *write*-Systemaufruf auszugeben. Wenn die erste Form eines Kommentars vorliegt, sollte ein Zeilentrenner hinterher ausgegeben werden – auch wenn dies einen zusätzlichen Systemaufruf kostet. Wenn am Dateiende ein nicht abgeschlossener Kommentar vorliegt, ist dieser ebenfalls auszugeben. Bei fehlschlagenden Systemaufrufen ist das Programm mit einer Fehlermeldung auf der Standardfehlerausgabe (Dateideskriptor 2) und einem Exit-Code von 1 zu beenden.

Reichen Sie bitte Ihre Lösung mit folgendem Kommando ein:

```
theon$ submit ssl 13 extract_comments.c
```

Viel Erfolg!