

Übungen zu Computergrundlagen WS 2018/2019

Übungsblatt 4: Shell-Skripte

16. November 2018

Allgemeine Hinweise

- Abgabetermin für die Lösungen ist **Freitag, 23.11.2017, 11:00 Uhr**
- Schickt die Lösungen bitte per Email an Euren Tutor:
 - Montag 14:00–15:30: Grant Cates (gcates@icp.uni-stuttgart.de)
 - Dienstag 9:45–11:15: Kai Szuttor (kai@icp.uni-stuttgart.de)
 - Dienstag 15:45–17:15: Julian Michalowsky (jmichalowsky@icp.uni-stuttgart.de)
 - Mittwoch 15:45–17:15: Michael Kuron (mkuron@icp.uni-stuttgart.de)
 - Donnerstag 9:45–11:15: Johannes Zeman (zeman@icp.uni-stuttgart.de)
- Die Übungen sollen in Gruppen von jeweils *zwei bis drei* Leuten bearbeitet werden. Abgaben von Einzelpersonen werden nicht akzeptiert. Bitte gebt *nur eine Lösung pro Gruppe* ab und nennt in eurer Abgabe alle Mitglieder eurer Gruppe!

Aufgabe 4.1: Skript (2 Punkte)

Sieh Dir folgendes (sehr) kurzes Bash-Skript an:

```
#!/bin/bash

if [ "$(($RANDOM % 2))" -eq 1 ]; then
    echo "Z"
else
    echo "K"
fi
```

- **4.1.1** (1 Punkt) Welche Bedeutung hat das Prozentzeichen in obigem Skript?
- **4.1.2** (1 Punkt) Welche Aufgabe erfüllt das Skript?

Aufgabe 4.2: Webmaster (5 Punkte)

Ein typischer Webserver erzeugt Logdateien, die jeden Zugriff durch Zeilen der Form

```
127.0.0.1 - - [24/Mar/2018:16:56:12 -100] "GET http://www.google.de HTTP/1.1" 200 2459
IP - - [Timestamp Zone] "Command Webpage Format" Returncode size
```

dokumentieren. Hieraus lässt sich sehr gut analysieren, welche Seite wie oft von wem aufgerufen wurde.

Schreibe ein Skript `analyze_log.sh`, das als einzigen Parameter den Namen einer solchen Logdatei bekommt. Danach soll es zu dieser Logdatei folgendes ausgeben:

- **4.2.1** (1 Punkt) Wie viele *verschiedene* Webseiten angefragt wurden. Die Webseite steht immer in der 7. Spalte.
- **4.2.2** (2 Punkte) Wie oft jede Webseite aufgerufen wurde. Suche dazu zunächst heraus, welche Webseiten es überhaupt gibt. Schreibe anschließend eine Schleife, die für jede Webseite die Anzahl der Aufrufe ausgibt.
- **4.2.3** (2 Punkte) Wie viele Daten insgesamt beim Abrufen aller geloggtten Webseiten übertragen wurden. Die Größe der übertragenen Daten soll in Megabyte ausgegeben werden. In der Logdatei gibt die *letzte* Spalte die Größe der für die jeweilige Webseite übertragenen Daten in Bytes an.

Hinweise:

- Um Dein Skript zu testen, kannst Du die Datei `/group/cgl/2018/04/server.log` verwenden.
- Wenn Ihr in Schritten vorgeht, ist es einfacher.
- Hilfreiche Werkzeuge für diese Aufgabe sind `sort`, `wc`, `grep`, `awk` (und man)
- Mithilfe von `awk` können zum Beispiel einzelne Spalten aus einer Datei gelesen werden. Der folgende Programmausschnitt gibt die 2. Spalte aus der Datei `Textdatei` wieder:

```
awk '{print $2}' Textdatei
```

Als Standard-Trenner für Spalten verwendet `awk` Leerzeichen. Beachte, dass es mit einem geeigneten `awk`-Befehl auch möglich ist die *n*-letzte Spalte auszugeben.

- `awk` ist sehr mächtig und kann beispielsweise auch Berechnungen durchführen.

Aufgabe 4.3: Audiospur aus Online-Video (3 Punkte)

Kopiere Dir die Datei `/group/cgl/2018/04/list.dat`. Diese Datei enthält je Zeile einen erwünschten Dateinamen (ohne Dateiendung), gefolgt von einem Online-Videolink. Schreibe ein Skript, das als einzigen Parameter den Namen einer solchen Datei (Linkliste) bekommt. Das Skript soll die Audiospuren sämtlicher in der Datei aufgeführten Online-Videos herunterladen, in das mp3-Format *konvertieren* und jeweils unter dem erwünschten Dateinamen (mit sinnvoller Dateiendung) abspeichern.

Hinweise:

- Online-Videodateien (bzw. ihre Audiospuren) können mit dem Kommandozeilen-Programm `youtube-dl` heruntergeladen werden. Audiospuren werden in der Regel im m4a-Dateiformat abgespeichert.
- Zur Konvertierung von Audio- und Videodateien kann das Programm `ffmpeg` benutzt werden. Hier kann die Verwendung des Arguments `-nostdin` nötig sein, um zu verhindern, dass `ffmpeg` aus der Standardeingabe liest.
- Wie gewöhnlich helfen `man` und Internetrecherche. Letztere kann beispielsweise helfen herauszufinden, wie man in Bash-Skripten Dateien zeilenweise einlesen kann etc.