

Übung 2: Batch-Verarbeitung

1. Setzen Sie die Variable `FIRSTNAME` auf Ihren Vornamen.
2. Geben Sie den Inhalt der Variablen `FIRSTNAME` aus.
3. Starten Sie nun einen Bash-Subprozess.
4. Geben Sie den Inhalt der Variablen `FIRSTNAME` aus.
5. Beenden Sie den Subprozess.
6. Setzen Sie die Variable `FULLNAME`, indem Sie Ihren Nachnamen an die Variable `FIRSTNAME` anhängen.
7. Schauen Sie sich an, auf was Ihre `$PATH`-Variable gesetzt ist.
8. Ergänzen Sie Ihre `$PATH`-Variable um das Verzeichnis `~/Vorkurs`, das Sie in einer vorherigen Übung angelegt haben. `~/Vorkurs` gehört damit zu Ihrem Suchpfad – ausführbare Programme, die darin liegen, werden dann automatisch gefunden.
9. In dieser Aufgabe sollen Sie nun ein kleines Shell-Skript `overlap.sh` schreiben. Das Skript soll zwei Verzeichnisse als Argumente nehmen. In den beiden Verzeichnissen steht jeweils ein Korpus, der auf verschiedene Dateien verteilt sein kann. Das Skript soll letzten Endes eine Liste der Wörter ausgeben, die in *beiden* Texten vorkommen. Benutzen Sie Ihren Lieblingseditor um das Skript zu schreiben.

Einige Hinweise:

- Der erste Schritt ist, die Dateien aus den beiden Verzeichnissen zusammenzufassen. Das haben wir in vorigen Übungen schon gemacht.
- Danach ist für jeden Text eine Liste der Begriffe zu erstellen, die darin vorkommen. Häufigkeiten sind diesmal nicht von Interesse.
- Benutzen Sie das Programm `comm`, um den tatsächlichen Overlap ermitteln zu lassen. Wie es genau funktioniert, erfahren Sie auf der zugehörigen man-page.
- Zwischendurch lässt es sich nicht vermeiden, Dateien zu erzeugen und zu speichern. Tun Sie das *nicht* in Ihrem Homeverzeichnis, sondern in einem temporär erzeugten. Mit `mktemp` kann man temporäre Dateien und Verzeichnisse erzeugen, ohne sich um Dinge wie Namensgleichheit kümmern zu müssen (auf der man-page von `mktemp` finden sich die Details).
- Um das Skript auszuführen, müssen Sie es einmal ausführbar machen: `$ chmod u+x overlap.sh`. Danach können Sie `overlap.sh` wie jeden anderen Befehl auch benutzen.