

Universität Potsdam  
Institut für Informatik  
Wintersemester 2008–2009  
Rechner- und Netzbetrieb

**Aufgabe zum Selbststudium (Hausaufgabe bis 5. Januar 2009)**

*Wenn Sie diese Aufgabe zum Entschlüsseln eines Textes vollständig bearbeiten, erhalten Sie einen Hinweis zum nächsten Test.*

1. Schreiben Sie ein Shell-Skript **count**, das mit einer ASCII-Datei als Argument auszuführen ist und dann die absoluten Häufigkeiten der Kleinbuchstaben (a bis z ohne Umlaute) ermittelt, die in der als Argument genannten Datei enthalten sind. Diese Häufigkeiten sind in Tabellenform auf **stdout** auszugeben, z.B.:

a	17
b	4
usw.	

2. Schreiben Sie ein weiteres Shell-Skript **relcount**, das wie **count** arbeitet, aber die relativen Häufigkeiten der Buchstaben ermittelt und ausgibt.  
*Hinweis:* Erzeugen Sie Promille-Angaben!

3. Bereiten Sie die Dateien **Schnewittchen.txt** und **Benutzerordnung.txt** auf eine Analyse mit **relcount** vor, indem Sie mit Hilfe von **tr**

- Sonderzeichen und Ziffern entfernen (auch Leer- und Newline-Zeichen)
- Großbuchstaben durch die entsprechenden Kleinbuchstaben ersetzen.

4. Wenden Sie **relcount** auf die beiden so erhaltenen Dateien an und werten Sie die ausgegebenen Häufigkeitsverteilungen aus.

5. Anhand der erhaltenen Häufigkeitsverteilungen von Buchstaben in deutschen Texten versuchen Sie, den Text **hinweise.crypt** zu entschlüsseln.

Es handelt sich um einen Text in deutscher Sprache, bei dem nur Kleinbuchstaben verwendet wurden und Umlaute durch jeweils zwei Buchstaben dargestellt sind (ae, oe, ue). Dieser Text ist monoalphabetisch verschlüsselt (feste Zuordnungstabelle Buchstabe → Buchstabe). Sonderzeichen sind unverschlüsselt geblieben.

*Hinweis:* Beseitigen Sie vor der Häufigkeitsanalyse der Symbole in **hinweise.crypt** alle Sonderzeichen. Sie dienen Ihnen aber beim Entschlüsseln als Orientierung.

Gehen Sie schrittweise (buchstabenweise) vor und setzen Sie, nachdem Sie einige Buchstaben ersetzt haben, auch Ihre Assoziation ein. Experimentieren Sie ein wenig und benutzen Sie **tr**.