

3 Testat Entropie

Übung 10 Fernsehsender

Um die Zuschauerquoten zu ermitteln, soll in Satelliten-Receiver jede Minute der aktuell eingeschaltete Sender gespeichert werden.

1. Wie viele Bits brauchen Sie zur Speicherung der Senderinformation bei 300 verschiedenen Sendern?
2. Nicht alle Sender werden gleich häufig geschaut. Es gelte folgende Verteilung:
 - ARD 8 %
 - ZDF 9 %
 - RTL 7 %
 - SAT 11 %
 - EuroSport 5 %
 - MTV 7 %
 - Arte 3 %
 - die restlichen 50 % verteilen sich gleichmäßig auf alle anderen Sender

Wie viele Bits benötigt man mindestens zur Codierung der Sender bei dieser Häufigkeitsverteilung?

3. Was schlagen Sie vor, um die Gesamtmenge der zu speichernden Daten zu reduzieren?

Übung 11 Entropie-Klasse

Mit der Klasse `Entropie` können Sie den Informationsgehalt von Dateien zu messen. Die Klasse enthält drei Methoden

1. `zaehleDatei(String datei)`: liest alle Bytes in einer Datei einzeln ein und zählt in dem Feld `zaehler` die Häufigkeit mit.
2. `speichereZaehler()`: Schreibt die Zählerstände in eine Datei.
3. `berechneEntropie()`: Hier soll die Entropie berechnet werden.

Übernehmen Sie zunächst die Klasse in ein BlueJ-Projekt. Kopieren Sie außerdem die beiden Dateien `skript.txt` und `skript.zip` von meiner Seite. Sie können jetzt bereits die ersten beiden Methoden ausführen, um die Häufigkeit einzelner Bytes (Zeichen) zu messen. Wichtig: ich musste den vollständigen Dateinamen mit Pfad angeben. Ansonsten wurde die Datei nicht gefunden. Ihre Aufgabe ist es dann:

- Die Methode *berechneEntropie()* zu programmieren
- Die Entropie für die beiden Dateien zu bestimmen.
- Außerdem sollen Sie eine grafische Darstellung der Häufigkeitsverteilungen der Zeichen in den beiden Dateien erstellen (Vorschlag: Excel)
- Wie sind die Ergebnisse zu interpretieren?