**Technische Universität München**
**Fakultät für Informatik**
**Prof. Dr. A. Knoll**

**SS 2006**
**Übungsblatt 1**
**5. Mai 2006**

# Übungen zu Einführung in die Informatik II

## Aufgabe 1 Eigenschaften der Entropie

Gegeben sei eine Shannonsche Nachrichtenquelle S mit einem Alphabet $\mathbf{Z}$, $|\mathbf{Z}| = q$, und einer binären Codierung; beweisen Sie, das für die Entropie $H(S)$ folgendes gilt:

a) $H(S) \geq 0$

b) $H(S) = 0 \Leftrightarrow \exists i \in \{1,..,q\} : p_i = 1 \wedge (\forall j \in \{1,..,q\} : j \neq i \Rightarrow p_j = 0)$

c) $H(S) \leq \mathrm{ld}\ q$

d) $H(S) = \mathrm{ld}\ q \Leftarrow \forall i \in \{1,..,q\} : p_i = \frac{1}{q}$

**Hinweis:** Verwenden Sie für die Teilaufgabe c) folgenden Satz: Sei $\forall i \in \{1,..,q\} : x_i \geq 0 \wedge y_i \geq 0$ und $\sum_i x_i = \sum_i y_i = 1$. Dann gilt:

$$\sum_{i=1}^{q} x_i \mathrm{ld} \frac{1}{x_i} \leq \sum_{i=1}^{q} x_i \mathrm{ld} \frac{1}{y_i}$$

## Aufgabe 2 Korrektur von Übermittlungsfehlern

Wir möchten vierstellige Kontonummern beginnend mit 1000 so vergeben, dass fehlerhafte Nummern erkannt und korrigiert werden können, unter der Annahme, dass genau eine Ziffer falsch ist.

Sei $c(i)$ die i-te Kontonummer. (Es ist also $c : \mathbb{N} \rightarrow \{0,...,9\}^4$.) Wir wählen $c(n)$ für $n > 1$ als die Dezimaldarstellung $\mathrm{dez}(x)$ der kleinsten Zahl $x$ mit der Eigenschaft:

$$\forall i : i < n \Rightarrow (\mathrm{Ham}(c(i), \mathrm{dez}(x)) \geq 3).$$

Hierbei sei $\mathrm{Ham}(x)$ der Hammingabstand, der angibt, an wievielen Stellen sich zwei Wörter aus $\{0,...,9\}^4$ unterscheiden.

Berechnen Sie die ersten 15 Kontonummern, und geben Sie ein Beispiel für eine Fehlerkorrektur an.

## Aufgabe 3 Huffman-Algorithmus

Wir betrachten eine Shannonsche Nachrichtenquelle mit einem aus acht Zeichen bestehenden Alphabet $\mathbf{Z}$; die einzelnen Zeichen treten mit folgenden Wahrscheinlichkeiten auf:

| Zeichen $z$: | $z_1$ | $z_2$ | $z_3$ | $z_4$ | $z_5$ | $z_6$ | $z_7$ | $z_8$ |
|---|---|---|---|---|---|---|---|---|
| Wahrscheinlichkeit $50 \cdot p_z$: | 2 | 3 | 5 | 6 | 6 | 6 | 8 | 14 |

Gesucht ist eine Codierung von **Z** durch Sequenzen über **T** = {O,L}. Die Codierung wird dabei nach folgendem Algorithmus ermittelt:

- Man wähle zwei Zeichen $z'$ und $z''$ mit den Wahrscheinlichkeiten $p_{z'}$ und $p_{z''}$ derart, dass $p_{neu} := p_{z'} + p_{z''}$ minimal wird, und ersetze die beiden Zeichen durch ein neues Zeichen $z_{neu}$ mit der Wahrscheinlichkeit $p_{neu}$.
- Sei $c_{neu}$ eine optimale Huffman-Codierung für $\mathbf{Z}' = \mathbf{Z} \setminus \{z', z''\} \cup \{z_{neu}\}$, so ist

$$c(z) =_{def} \begin{cases} c_{neu}(z) & \text{falls } (z \neq z' \wedge z \neq z'') \\ c_{neu}(z) \circ <\mathrm{O}> & \text{falls } z = z' \\ c_{neu}(z) \circ <\mathrm{L}> & \text{falls } z = z'' \end{cases}$$

  eine optimale Codierung für Z.

a) Ermitteln Sie nach diesem Algorithmus eine Codierung für das Alphabet **Z**.

b) Erfüllt die Codierung die *Fano*-Bedingung ? Begründen Sie Ihre Antwort.

c) Geben Sie die mittlere Länge der Codierung eines Zeichens aus **Z** an.

## Aufgabe 4 Morse-Code

Im Jahr 1884 sendete Samuel Morse eine Nachricht mit einem Fernschreiber von Washington nach Baltimore. Er verwendete dazu einen Code der heute als *Morse-Code* bekannt ist. Dieser Code besteht aus einer Folge von langen und kurzen Tönen. Zwischen zwei Zeichen ist jeweils eine längere Pause.

a) Suchen Sie im Internet nach einer Code-Tabelle für den Morse-Code.

b) Was ist das Alphabet beim Morse-Code?

c) Untersuchen Sie den Morse-Code auf Einhaltung der Fano-Bedingung. Erstellen Sie dazu den kompletten Code-Baum für den Morse-Code.

d) Warum sind die Code-Sequenzen unterschiedlich lang?