Datenkompression (Rote), SS 2003, 2. Übungsblatt
Abgabe: bis Freitag, 2.5.2003, in der Vorlesung

6. Kapazität eines Kanals (5 Punkte)

   Das Morse-Alphabet ist ein Kode über dem Alphabet $\{\cdot, -, \sqcup\}$. Nehmen wir an, dass ein Punkt („$\cdot$") und ein Zwischenraum („$\sqcup$") je 0,2 sec dauert und ein Strich („$-$") 0,4 sec.

   Wie viele verschiedene Nachrichten lassen sich in $n \cdot 0{,}2$ Sekunden übertragen? (Geben Sie eine Rekursionsformel an.) Wie groß ist die Kapazität dieses Kanals (in bit/sec)?

7. (0 Punkte) Finden Sie einen eindeutig entzifferbaren Kode, bei dem es beliebig lange Nachrichten gibt, die man erst dekodieren kann, wenn man das letzte Bit gelesen hat.

8. (0 Punkte) Verallgemeinern Sie den Huffman-Algorithmus, sodass er einen optimalen Kode für ein Zielalphabet mit mehr als zwei Symbolen erzeugt.

9. (0 Punkte)

   (a) Beweisen Sie die folgende Beziehung für die Entropie:

   $$H(p_1\alpha, p_1(1-\alpha), p_2, \ldots, p_n) = H(p_1, p_2, \ldots, p_n) + p_1 \cdot H(\alpha, 1-\alpha),$$

   for $0 \le \alpha \le 1$.

   (b) Beweisen Sie die folgende Ungleichung für die Entropie:

   $$0 \le H(p_1, p_2, \ldots, p_n) \le H(\tfrac{1}{n}, \tfrac{1}{n}, \ldots, \tfrac{1}{n}) = n \log_2 n$$

   Für welche Verteilungen $(p_1, p_2, \ldots, p_n)$ wird in der linken Ungleichung Gleichheit erzielt?

10. Huffman-Nummerierung (15 Punkte)

    Wir nummerieren die Knoten des Kodebaumes, der beim binären Huffman-Algorithmus entsteht, in der Reihenfolge, in der die Knoten gepaart werden. Das heißt, die beiden Blätter mit den kleinsten Gewichten bekommen die Nummern 1 und 2, und die Wurzel hat Nummer $2n-1$. Das *Gewicht* eines Blattes ist seine Wahrscheinlichkeit, und das Gewicht eines inneren Knotens ist die Summe der Gewichte aller Blätter.

    Beweisen Sie, dass diese Nummerierung folgende Eigenschaften hat.

    (a) Knoten mit größerem Gewicht haben größere Nummern.

    (b) Zwei Geschwisterknoten haben aufeinanderfolgende Nummern.

    (c) Alle Knoten auf Tiefe i haben größere Nummern als die Knoten auf Tiefe $i+1$.

    (d) Betrachte zwei Knoten u und v auf Tiefe i. Wenn u eine kleinere Nummer als v hat, dann haben die Kinder von u eine kleinere Nummer als die Kinder von v (sofern diese Kinder vorhanden sind).

    Beweisen Sie, dass ein Baum, dessen Gewichte in den Blättern die gegebenen Wahrscheinlichkeiten p_i sind, und für den eine Nummerierung mit den obigen Eigenschaften existiert, ein Huffman-Baum ist.

11. (10 Punkte) Überlegen Sie, wie man einen optimalen binären Kode möglichst platzsparend abspeichern kann.

    Wieviele Bits benötigen Sie für eine Kode für das Alphabet Σ der 10 Ziffern? Wieviele Bits benötigen Sie, um alle 256 möglichen Bytes zu kodieren?