Prof. Dr. Rolf Drechsler, drechsler@informatik.uni-bremen.de, MZH 3510
Dipl.-Inform. Daniel Große, grosse@informatik.uni-bremen.de, MZH 3490

6. Übungsblatt zur Vorlesung

Technische Informatik 1

Aufgabe 1 (4 Punkte)

Beweise das folgende Lemma:

Lemma 1. *Sei $A = \{a_1, \dots, a_k\}$ ein Alphabet der Größe k und sei $c : A \to \{0, 1, 2, \dots, r-1\}^n$ ein Code der festen Länge n, dann gilt*

$$n \geq \lceil \log_r k \rceil.$$

Aufgabe 2 (8 Punkte)

a) Überführe die Wörter 10010011_2, 11111100_2 und 00001111_2 (dabei steht das Bit mit der höchsten Nummer links) in den in der Vorlesung vorgestellten 1-Bit-fehlerkorrigierenden Hamming-Code.
b) Bei den codierten Wörtern 101101101011_2, 011101000001_2 und 101001010000_2 wurde jeweils ein Bit falsch übertragen. Finde heraus, um welches Bit es sich handelt.
c) Kann es sich bei dem Fehler im ersten Wort aus b) auch um einen Mehrfach-Übertragungsfehler handeln (Begründung und ggf. Beispiel nicht vergessen)?
d) Wann werden Mehrfach-Übertragungsfehler mit dem gegebenen Codierungsschema korrekt erkannt?

Aufgabe 3 (8 Punkte)

Im Folgenden sollen Zeichen effizient codiert werden, um eine möglichst hohe Kompressionsrate zu erzielen. Verwendet wird dazu der Huffman-Code. Eine (repräsentative) Stichprobe der zu codierenden Daten dient der Ableitung des Codes.

a) Als Stichprobe dient der folgende Satz:

ABER HIER, WIE UEBERHAUPT, KOMMT ES ANDERS ALS MAN GLAUBT.

Generiere für diesen Text den Huffman-Code, wobei Leerzeichen, Komma und Punkt ignoriert werden sollen. Damit die Generierung des Codes eindeutig ist, sind noch einige Verabredungen notwendig:
Sortiere die Zeichen von links nach rechts absteigend nach ihren Häufigkeiten (z.B. E,A statt A,E) und bei gleicher Häufigkeit in lexikographischer Reihenfolge (z.B. T,U statt U,T).
Beim Zusammenfassen zweier „Knoten" sollen immer die am weitesten links stehenden Knoten gewählt werden, die für eine Verschmelzung in Frage kommen (z.B. $\{D\} \cup \{G\}$ statt $\{G\} \cup \{K\}$).
Bei der Zuweisung der Codebits bekommt immer der linke Nachfolger eine '1', der rechte eine '0' zugewiesen.
Ein neuer Knoten wird immer oberhalb seines linken Kindes angelegt.
(Zur Kontrolle: ES wird codiert zu 1111010)
Gib in Deiner Lösung auch den Code für jeden einzelnen Buchstaben an. Ferner trenne zur Übersichtlichkeit die einzelnen kodierten Worte durch Leerzeichen.

b) Dekodiere mit Hilfe des in Teilaufgabe a) generierten Codes folgende Nachricht:
0001010001001111011110101011111110001001

c) Gib die Kompressionsrate ((Länge der uncodierten Nachricht)/(Länge der codierten Nachricht)) für die Nachricht aus b) an, wenn ein Zeichen im uncodierten Fall durch 8 Bit repräsentiert wird.

Abgabetermin: zu Beginn der Vorlesung am 06.06.2005