Prof. Dr.-Ing. Dr. h. c. J. Becker

Digitaltechnik

2. Tutorium - Lösung

Institut für Technik der Informationsverarbeitung (ITIV), Karlsruher Institut für Technologie (KIT)

Aufgabe 1: Shannon-FanØ-Codierung

Zur Überwachung der Aktivität von Herz-Kreislauf-Patienten soll durch einen tragbaren Sensor die aktuelle Aktivität eines Patienten via Mobilfunk an eine zentrale Datenbank gesendet werden. Die verschiedenen Patientenaktivitäten treten dabei über den Tag verteilt mit folgenden Wahrscheinlichkeiten auf.

Aktivität	Wahrscheinlichkeit	Codierung
Liegen	41 %	Li: 1/1
Stehen	10 %	St: 0011/001
Sitzen	29 %	Si: 01/01
Gehen	10 %	Ge: 0010/0001
Laufen	6 %	La: 0001/00001
Radfahren	4 %	Ra: 0000/00000

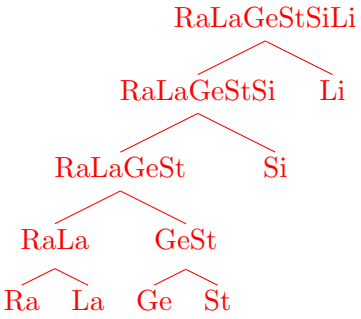
1.1 Damit die Mobilfunkkosten und der Speicherbedarf der Datenbank möglichst gering sind, sollen die verschiedenen Aktivitäten optimal mit einem Shannon-FanØ-Code codiert werden. Bestimmen Sie eine Shannon-FanØ-Codierung.

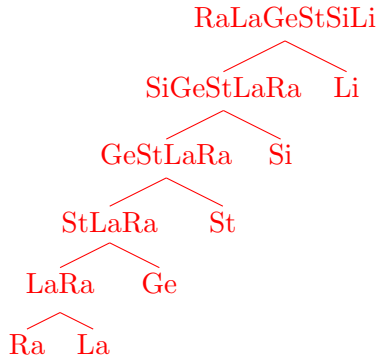
Partitionierungskonvention:

- *Sortieren Sie die Elemente zu Beginn entsprechend den Auftrittshäufigkeiten* ***aufsteigend von links nach rechts****. Falls unterschiedliche Knoten dieselbe Auftrittshäufigkeiten haben, sortieren Sie diese alphabetisch von links nach rechts.*
- *Teilen Sie eine Menge immer so auf, dass die Differenz zwischen den Summen der Auftrittshäufigen der Teilmengen minimiert wird.*
- *Verändern Sie die Reihenfolge der Sortierung/Ordnung während der Anwendung des Verfahrens nicht.*
- *Weisen Sie den linken Ästen des entstehenden Baumes die „0“ zu, den rechten Ästen die „1“.*

© ITIV 2015

Lösung / alternative Lösung





1.2 Wie groß ist die mittlere Codewortlänge?

$mCW = 0,41 \cdot 1 + 0,1 \cdot 4 + 0,29 \cdot 2 + 0,1 \cdot 4 + 0,06 \cdot 4 + 0,04 \cdot 4 = 2,19$

alternativ $mCW = 0,41 \cdot 1 + 0,1 \cdot 3 + 0,29 \cdot 2 + 0,1 \cdot 4 + 0,06 \cdot 5 + 0,04 \cdot 5 = 2,19$

1.3 Anhand welcher informationstechnischer Eigenschaft der Datenquelle können Sie die mittlere Codewortlänge Ihres Codes bewerten? Nennen Sie die Eigenschaft und geben Sie die entsprechende Formel zur Berechnung an.

Entropie der Quelle bzw. durchschnittlicher Informationsgehalt des Zeichenvorrats

$H = \sum(p(x_i) \cdot ld(\frac{1}{p(x_i)})$

(Bei diesem Code $H \approx 2,139$)

1.4 Die Mobilfunkübertragung soll über GPRS erfolgen, wobei der Preis pro über- tragenen 10 KiByte 0,1 € beträgt (1 KiByte $\hat{=}$ 1024 Byte). Der Sensor überträgt jede Sekunde einen Wert für die Aktivität. Geben Sie eine Formel an zur Berechnung der mittleren Übertragungskosten pro Tag und Patient.

$$Kosten = \frac{m \cdot 60 \cdot 60 \cdot 24}{1024 \cdot 8 \cdot 10} \cdot 0,1 \text{ €}$$

(Wären hier 0,23 €pro Tag und Patient.)

Aufgabe 2: Huffman-Codierung

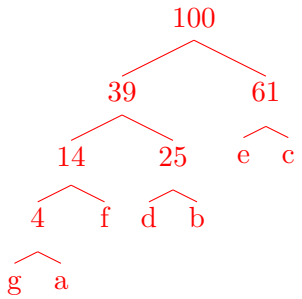
Für die Steuerung eines biologischen Reaktors soll eine Temperaturprotokollierung entworfen werden. Der zu überwachende Prozess durchläuft dabei nacheinander zwei unterschiedliche Temperaturen. Da weniger die genaue Temperatur, als viel mehr die Einhaltung gewisser Grenzen interessiert, wurde eine Einteilung der Reaktortemperatur in folgende Bereiche vorgenommen:

x_i	**Beschreibung**	$p(x_i)$	gefundener Code
a	Reaktor aus	3%	0001
b	zu niedrig	13%	011
c	optimal für Stufe 1	33%	11
d	Übergang	12%	010
e	optimal für Stufe 2	28%	10
f	zu hoch	10%	001
g	GEFAHR	1%	0000

2.1 Um das Messprotokoll möglichst speichereffizient aufzeichnen zu können, soll eine Huffman-Codierung der Reaktortemperatur zum Einsatz kommen. Dazu werden die Auftrittswahrscheinlichkeiten der einzelnen Temperaturbereiche aus dem erwarteten Prozessverlauf ermittelt. Geben Sie eine mögliche Huffman-Codierung der Reaktortemperatur an.

Partitionierungskonvention:

- *Sortieren Sie die Elemente zu Beginn entsprechend den Auftrittshäufigkeiten* ***aufsteigend von links nach rechts****. Falls unterschiedliche Knoten dieselbe Auftrittshäufigkeiten haben, sortieren Sie diese bitte alphabetisch von links nach rechts. Gehen Sie bei jedem nötigen Sortierschritt nach diesem Schema vor.*
- *Weisen Sie den linken Ästen des entstehenden Baumes die „0“ zu, den rechten Ästen die „1“*



2.2 Wie groß ist die mittlere Codewortlänge Ihres Huffman-Codes?
$mCW = 0,04 \cdot 4 + 0,35 \cdot 3 + 0,61 \cdot 2 = 2,43$

2.3 Welche Größe ziehen Sie als Vergleichsmaßstab heran, wenn Sie bewerten möchten, wie effizient Ihr Huffman-Code ist? (Bezeichnung, Begründung und Formel)

Entropie der Quelle bzw. durchschnittlicher Informationsgehalt des Zeichenvorrats;
Die Entropie ist eine geeignete Größe, da sie den Informationsgehalt beschreibt, ein guter Code sollte einen hohen Informationsgehalt haben.
$H = \sum(p(x_i) \cdot ld(\frac{1}{p(x_i)})$
(Bei diesem Code $H \approx 2,34$)

2.4 Welchen Vorteil bietet die Codierung von Information mit Hilfe der Huffman-Codierung oder der Codierung nach Shannon-FanØ gegenüber beispielsweise dem Gray-Code oder einem (n-aus-m)-Code?

Die Zeichen sind nach Auftrittswahrscheinlichkeiten optimiert und verkürzen somit die mittlere Codewortlänge gegenüber anderen Codierungen.

2.5 Beurteilen Sie im Allgemeinen die Huffman-Codierung gegenüber der Shannon-FanØ Codierung im Bezug auf die Qualität der gefundenen Lösung.

Die Huffman- Codierung findet in jedem Fall einen Code mit gleich guter oder besserer mittlerer Codewortlänge als die Codierung nach Shannon-FanØ
Des Weiteren lässt sich Shannon-FanØ nicht so gut formalisieren, da es mehrere richtige Lösungen geben kann (siehe Aufgabe 1). Die Codierung nach Huffman jedoch ist eindeutig.

Aufgabe 3: Blocksicherung und Scrambling

Für die Datenübertragung soll eine Blocksicherung verwendet werden. Damit auch Burstfehler erkennbar sind soll zusätzlich Scrambling eingesetzt werden.

Ein Datenwort hat dabei immer eine Länge von 8 Bit.

3.1 Berechnen Sie wie viele Nutzdatenbits in einem Block übertragen werden müssen, wenn durch das Scrambling Burstfehler der Länge 4 erkennbar sein sollen.

Burstfehler der Länge 4 =>4 Zeilen durch das Scrambling vorgegeben
Blocksicherung bedeutet, dass eine Zeile ein Prüfwort ist.
Daraus ergeben sich 3 Zeilen Nutzdaten je 8 Bit => 24 Bit Nutzdaten

3.2 Berechnen Sie den Overhead der durch die Blocksicherung entsteht. Verwenden Sie die in Aufgabe 3.1 berechnete Blockgröße.

3 Zeilen Nutzdaten =24 Bit
Je ein Paritätsbit pro Spalte => Prüfwort 8 Bit (4.Zeile)
Je ein Paritätsbit pro Zeile => Prüfspalte: 4 Bit
=> 24 Bit Nutzdaten und 12 Bit Parität
Overhead = Bits für Sicherung / Nutzdatenbits = 12/24 = 50%
Sie haben nun folgenden Datenstrom empfangen:

11011000 01011110 01101000 10001110 01001111 00111010 01001000 1011 100

Dieser Datenstrom wurde mit einer Blocksicherung und Scrambling gegen Fehler gesichert.

3.3 Rekonstruieren Sie die gesendeten Datenwörter unter der Annahme, dass 1 Datenwort 8 Bit lang ist.

Länge des Datenstroms: 63 Bit
Länge einer Zeile: 8 Bit Daten + 1 Bit Parität = 9 Bit
Anzahl Zeilen: 63 / 9 = 7 => 6 Bit Daten + 1 Bit Parität pro Spalte
In Spaltenschreibweise ergibt dies: Die Paritätsbits sind grau hinterlegt. Das fehlerhafte Bit ist hellgrau (A3.5)

1	1	0	1	0	1	0	1	1	1
2	1	0	0	0	1	1	1	0	0
3	**0**	**1**	**0**	**0**	**1**	**1**	**1**	**0**	**1**
4	1	0	1	1	0	1	0	1	1
5	1	1	1	0	0	1	1	0	1
6	0	1	0	0	1	0	0	0	0
7	0	1	1	0	0	0	0	0	0

3.4 Welche Länge an Burstfehler lassen sich mit der verwendeten Blockgröße in Verbindung mit dem Scrambling maximal erkennen?

7 Zeilen ergibt eine maximale Länge eines Burstfehlers von 7 Bit.

3.5 In dem übertragenen Datenstrom befindet sich 1 Fehler. Beheben Sie diesen indem Sie die fehlerhafte Zeile und die fehlerhafte Spalte angeben. Markieren Sie das fehlerhafte Datenwort und geben Sie das korrigierte Datenwort an. Für die Blocksicherung wurde eine gerade Parität verwendet.

Anzahl 1er in Zeile 3: 5 $->$ Fehlerhaft
Anzahl 1er in Spalte 4: 1 $->$ Fehlerhaft
Das Fehlerhafte Datenwort (Fett in Aufgabe 3.3) lautet korrigiert: 01011110