

Datenkompression

Sommersemester 2018

Übungsblatt 4

Prof. Ohlebusch

Institut für Theoretische Informatik

Uwe Baier

Ausgegeben am 14.06.2018

Besprechung am 26.06.2018

Aufgabe 4.1

- (a) Bilden Sie die Move-To-Front-Transformierte des Strings $S = \text{yeep\$yaass}$.
- (b) Sei S ein String und S_M die Move-To-Front-Transformierte von S . Seien Σ bzw. Σ_M die Mengen aller in S bzw. S_M vorkommenden Zeichen. Zeigen oder widerlegen Sie folgende Aussagen:
- (1) Es gilt immer $H(S) \leq H(S_M)$. (3) Es gilt immer $|\Sigma| \leq |\Sigma_M|$.
- (2) Es gilt immer $H(S) \geq H(S_M)$. (4) Es gilt immer $|\Sigma| \geq |\Sigma_M|$.
- (c) Sei n die Länge und σ die Alphabetgröße der Eingabe. Ist es möglich, die Move-To-Front-Transformation in besserer Zeitkomplexität als $O(n \cdot \sigma)$ durchzuführen?

Aufgabe 4.2

In dieser Aufgabe soll die Facsimile - Kodierung betrachtet werden. Dazu soll die Kodierung eines Bildes mit Breite 16 dekodiert werden.

- (a) Dekodieren Sie die Bitsequenz 1100111011011011, die der ersten Zeile der Übertragung entspricht. Die Sequenz wurde dabei mit dem "modified Huffman"-Verfahren kodiert, wobei die Laulängen immer als unäre Kodierung der entsprechenden Laulänge kodiert wurden, also $1 \hat{=} 1$, $2 \hat{=} 01$, $3 \hat{=} 001$, $4 \hat{=} 0001$, ...
- Hinweis:* Der erste Lauf ist immer weiß gefärbt.
- (b) Dekodieren Sie die Bitsequenz 101101010001011110101, die der zweiten Zeile der Übertragung entspricht und mit dem "modified READ"-Verfahren kodiert wurde.

Benutzen Sie hierfür die folgende Codetabelle und das Ergebnis aus Aufgabe (a) als Referenzzeile.

Codewort	Abkürzung	Beschreibung	Formel für a_1
0001	P	Zwischenmodus	
1	V(0)	Vertikalmodus	$a_1 = b_1$
011	VR(1)	Vertikalmodus	$a_1 = b_1 + 1$
000011	VR(2)	Vertikalmodus	$a_1 = b_1 + 2$
0000011	VR(3)	Vertikalmodus	$a_1 = b_1 + 3$
010	VL(1)	Vertikalmodus	$a_1 = b_1 - 1$
000010	VL(2)	Vertikalmodus	$a_1 = b_1 - 2$
0000010	VL(3)	Vertikalmodus	$a_1 = b_1 - 3$
001	H	"Horizontalmodus", $ a_1 - b_1 > 3$	

Aufgabe 4.3

(a) Kodieren Sie die nachfolgenden Zahlen mit den angegebenen Verfahren:

- 46 mit Elias-Epsilon-Code
- 121 mit Rice-Code und $p = 16$
- 18 mit Fibonacci-Code

(b) Die folgende Bitsequenz besteht aus den konkatenierten Kodierungen von 5 Zahlen:

00000010110100010011000001101010000000110001011

Dekodieren Sie die 5 Zahlen, die mit den folgenden Verfahren kodiert wurden:

Zahl 1: Elias-Gamma Code

Zahl 4: Comma-Code¹ mit $p = 2$

Zahl 2: Elias-Delta Code

Zahl 5: Fibonacci-Code

Zahl 3: Rice-Code mit $p = 16$

(c) Warum wird für das Kodieren von Runs in einer BWT keiner der oben genannten Codes benutzt? Begründen Sie Ihre Antwort!

¹siehe hierzu die Folien zur Codierung von natürlichen Zahlen, https://www.uni-ulm.de/fileadmin/website_uni_ulm/iui.inst.190/Lehre/SS15/Datenkompression/uninetz/VL_Integer_Compression.pdf