

Datenkompression

Sommersemester 2016

Übungsblatt 5

Prof. Dr. E. Ohlebusch

Institut für Theoretische Informatik

J. Lorenz

Ausgegeben am 14.06.2016

Besprechung am 21.06.2016

Aufgabe 5.1

Implementieren Sie die Move-To-Front-Transformation. Wenden Sie diese auf die Strings $S = \text{abrakadabraabrakadabra\$}$ und $\text{BWT} = \text{arrdda\$kkrrraaaaaaabb}$ an. Vergleichen Sie die mittlere Huffman-Codewortlänge der beiden Move-To-Front-Transformierten.

Aufgabe 5.2

Das Kompressionsschema

$$S \rightarrow \text{BWT} \rightarrow \text{MTF} \rightarrow \text{Huffman} \rightarrow S_c$$

kann man durch eine Lauflängenkodierung (RLE) der 0-Runs im Vektor R (nach der MTF Transformation) erweitern. Wir betrachten die folgenden zwei Varianten:

1. Zuerst wird jeder Wert $R[i]$ um 1 inkrementiert (danach enthält R keine 0-Einträge mehr). Der Code $C(\ell)$ einer Lauflänge ℓ eines 1-Runs ist dann die Binärdarstellung von $\ell + 1$ (man beachte, dass $\ell + 1 \geq 2$ gilt), wobei die führende 1 weggelassen wird.
 2. Zuerst wird jeder Wert $R[i]$ um 2 inkrementiert (danach enthält R weder 0 noch 1 Einträge). Der Code $C(\ell)$ einer Lauflänge $\ell > 1$ eines 2-Runs ist dann die Binärdarstellung von ℓ , wobei die führende 1 weggelassen wird. In dieser Variante werden 2-Runs mit einer Lauflänge $\ell = 1$ nicht codiert (d.h., die 2 bleibt im modifizierten R -Vektor stehen).
- a) Finden Sie zwei Strings s_1, s_2 (nach der MTF Transformation), sodass s_1 durch Variante 1 besser komprimiert wird und s_2 durch Variante 2.
- b) Beschreiben Sie mögliche Vor- und Nachteile der beiden Varianten.

Aufgabe 5.3

Codieren Sie die erste Zeile mit dem MH-Algorithmus und die zweite Zeile mit dem MR-Algorithmus.

		■	■	■				■	■					■	■	■	■	■							■	■	■	■				
	■	■	■	■	■						■	■	■	■						■	■					■	■	■	■	■	■	■

Aufgabe 5.4

Beweisen Sie $H(X, Y) = H(X) + H(Y|X)$ für Zufallsvariablen X und Y . (Die Definition der gemeinsamen Entropie $H(X, Y)$ von X und Y finden Sie im Script.)