ETH Zürich | Prof. Markus Gross
Dept. Informatik | WS 05/06

Informationstheorie

Übung 7

Ausgabe: 19. Dezember 2005
Abgabe: 9. Januar 2005

7.1 Huffman Baum zum Aufwärmen

In der Vorlesung wurde das Tutorial Applet vorgestellt, welches auch auf der Homepage verlinkt ist.

a) Erstellen Sie mit Hilfe des Applets die Huffman Bäume der unterstehenden Verteilungen und berechnen Sie den Erwartungswert der Codierlänge pro Zeichen.

X	P_X	Y	P_Y
a	0.10	a	0.26
b	0.15	b	0.08
c	0.34	c	0.01
d	0.17	d	0.07
e	0.5	e	0.25
f	0.9	f	0.06
g	0.10	g	0.27

b) Wieso erreichen Textkompressionsprogramme wie Zip eine bessere Komprimierung als Huffmancodes über Buchstaben?

7.2 Optimale Codes und Huffman-Algorithmus

a) (Aus dem Vordiplom) Der Bit-String 11100110101000100101111111000 ist das Resultat einer Codierung eines Wortes w über dem Alphabet $\{a, b, c, d, e, f, g, h\}$ mittels eines Huffman Codes. Der Huffman Baum wurde für untenstehende Verteilung konstruiert mit der Konvention, dass beim Zusammenfassen zweier Knoten jeweils dem schwereren die "1" zugeordnet wird. Bestimmen Sie w.

x	a	b	c	d	e	f	g	h
$P_X(x)$	$\frac{7}{128}$	$\frac{8}{128}$	$\frac{9}{128}$	$\frac{10}{128}$	$\frac{16}{128}$	$\frac{18}{128}$	$\frac{28}{128}$	$\frac{32}{128}$

b) Formulieren Sie einen erweiterten Huffman-Algorithmus für D-äre Codes mit $D > 2$. Tip: In jedem Schritt des Algorithmus wird die Anzahl der aktiven Knoten um $D - 1$ reduziert. Am Schluss darf nur noch ein aktiver Knoten übrigbleiben. Führen Sie daher sogenannte Dummy-Knoten mit Wahrscheinlichkeit 0 ein, damit am Schluss ein ausgefüllter Baum entsteht.

c) Finden sie einen optimalen ternären Code für eine WSK-Verteilung mit den sechs Werten 0.3, 0.2, 0.2, 0.1, 0.1 und 0.1.

d) Eine Zufallsvariable Z kann die Werte A, B, C und D annehmen. Betrachten Sie die folgenden beiden binären, präfixfreien Codierungen für Z:

	A	B	C	D
Code 1	00	01	10	11
Code 2	0	10	110	111

Geben Sie eine Verteilung P_Z an, so dass beide Codierungen von Z optimal sind.

7.3 In ein Bit codiert

Wenn die Wahrscheinlichkeit des wahrscheinlichsten Wertes einer Zufallsvariable grösser ist als eine gewisse Schranke, wird dieser Wert in jedem optimalen binären Code mit einem Codewort der Länge eins codiert. Wie gross ist diese Schranke?

7.4 Bedingte Entropie und Fehlerwahrscheinlichkeit

Sei $N \geq 2$ und Y uniform verteilt auf $\{1, \ldots, N\}$. Weiter sei X eine Zufallsvariable auf $\{0, 1\}^N$ mit folgender Verteilung: Bedingt auf $Y = j$ werden an j (fixen) Positionen von X zufällige Bits gewählt und die restlichen $N-j$ Positionen mit Einsen aufgefüllt (für alle $j \in \{1, \ldots, N\}$).

a) Bestimmen Sie $H(X|Y)$.

b) Bestimmen Sie (numerisch) mit Hilfe der Fano-Ungleichung eine untere Schranke an die Fehlerwahrscheinlichkeit P_e bei der Schätzung von X aufgrund von Y im Falle $N = 4$ und $N = 8$. Wie verhält sich P_e im Grenzfall $N \to \infty$?

c) Berechnen Sie nun die Fehlerwahrscheinlichkeit bei der Schätzung von X aufgrund von Y bei Verwendung der bestmöglichen Strategie. Vergleichen Sie das Ergebnis mit den Resultaten aus Aufgabe b).