Dr. Jürg M. Stettbacher
Neugutstrasse 54
CH-8600 Dübendorf

Telefon: +41 43 299 57 23
Email: dsp@stettbacher.ch

Quiz

Quellencodierung: Huffman

Sie sollten in der Lage sein, die folgenden Fragen ohne langes Nachdenken beantworten zu können.

Eine gedächtnislose Quelle erzeugt die Symbole a und b mit den Wahrscheinlichkeiten $P(a) = 0.8$ und $P(b) = 0.2$.

1. Was bedeutet das Prädikat *gedächtnislos* der Quelle?

Die Ergebnisse sind unabhängig von einander.

2. Entwickeln Sie einen Huffman Code, so dass die Redundanz $R < 0.1$ Bit/Symbol ist.

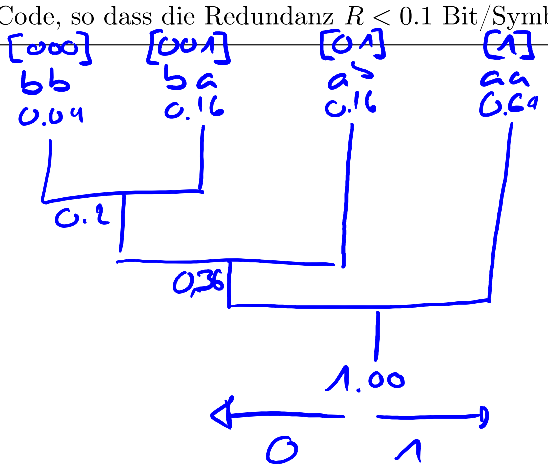
$aa = 0.64$
$ab = 0.16$
$ba = 0.16$
$bb = 0.04$

$L = 1.56$

$H = 1.44$

$R = L - H = 0.12$

$\frac{0.12}{2} = \underline{\underline{0.06}}$



Antworten

1. Das Prädikat *gedächtnislos* heisst, dass die Symbole der Quelle statistisch unabhängig von einander sind.

   Beispiel: Ein Würfel ist ohne Erinnerung. Beim nächsten Wurf weiss er nicht, was das Resultat der vergangenen Würfe war. Daher gibt es keine Abhängigkeit zwischen den Würfen.

2. Der Huffman Code C_1 für zwei Symbole führt zwangsläufig zu den Codeworten (0) und (1). Die mittlere Codewortlänge $L(C_1)$ ist demnach 1 Bit/Symbol.

   Die Entropie H_1 der Quelle ist:

   $$H_1 = P(a) \cdot \log_2 \frac{1}{P(a)} + P(b) \cdot \log_2 \frac{1}{P(b)} = 0.722 \quad \text{Bit/Symbol}$$

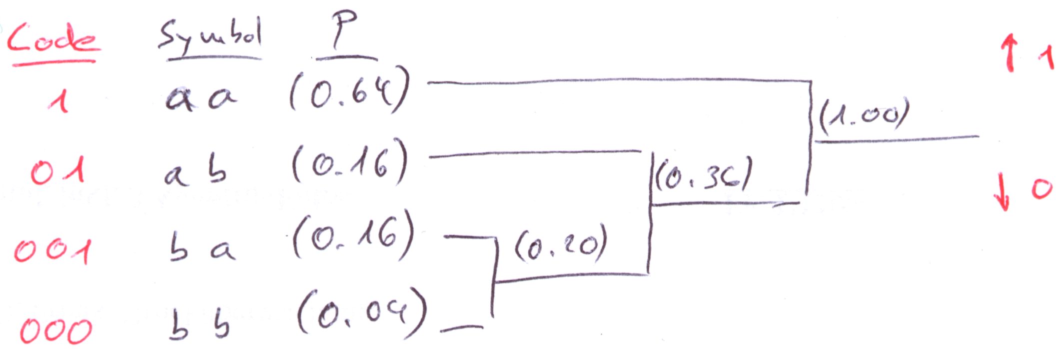
   Damit erreichen wir das Ziel nicht. Wir stellen uns darum die Quelle vor, als würde sie Doppelsymbole liefern. Die Wahrscheinlichkeiten der Doppelsymbole sind, da sie statistisch unabhängig von einander sind:

   $$\begin{aligned} P(aa) &= P(a) \cdot P(a) = 0.64 \\ P(ab) &= P(a) \cdot P(b) = 0.16 \\ P(ba) &= P(b) \cdot P(a) = 0.16 \\ P(bb) &= P(b) \cdot P(b) = 0.04 \end{aligned}$$

   Die Entropie H_2 der Doppelsymbole ist:

   $$H_2 = 2 \cdot H_1 = 1.444 \quad \text{Bit/2 Symbol}$$

   Der Huffman-Baum für die Doppelsymbole sieht so aus:



   Damit haben wir den Code C_2 erhalten. Für die Berechnung der Redundanz benötigen wir die mittlere Codewortlänge aus den Codeworten c_k mit $0 \leq k \leq 3$ und den entsprechenden Codewortlängen λ_k:

   $$L(C_2) = \sum_{k=0}^{3} P(c_k) \cdot \lambda_k = 1.560 \quad \text{Bit/2 Symbol}$$

   Die Redundanz pro Symbol ist folglich:

   $$R = \frac{L(C_2)}{2} - \frac{H_2}{2} = 0.058 \quad \text{Bit/Symbol}$$

   Der Code C_2 erfüllt also die Anforderung.