Prof. Dr. V. Strehl
Informatik 8

18. Dezember 2006

Übungen zur Theoretischen Informatik 3 WS 2006/07

With many profound scientific discoveries it is possible with the aid of hindsight to see that the times were ripe for the breakthrough. Not so with information theory! While of course Shannon was not working in a vacuum in the 1940's, his results were so breathtakingly original that even the communication specialists of the day were at loss to understand their significance.
R. J. McEliece in The Theory of Information and Coding

- **Aufgabe 29: Codewortlängen und Ungleichung von Kraft**

  1. Es sei $(\ell_1, \ell_2, \ldots, \ell_n)$ mit $1 \leq \ell_1 \leq \ell_2 \leq \ldots \leq \ell_n$ eine Folge natürlicher Zahlen, die der Ungleichung von Kraft genügt, d.h. $\sum_{1 \leq i \leq n} 2^{-\ell_i} \leq 1$.
     (a) Zeigen Sie: gilt Gleichheit, d.h. $\sum_{1 \leq i \leq n} 2^{-\ell_i} = 1$, so ist $\ell_{n-1} = \ell_n$.
     (b) Zeigen Sie: gilt strikte Ungleichheit, also $\sum_{1 \leq i \leq n} 2^{-\ell_i} < 1$, so gilt sogar $\sum_{1 \leq i \leq n} 2^{-\ell_i} \leq 1 - 2^{-\ell_n}$.

     Hinweis: benutzen Sie zum Beweis dieses Teils bitte <u>nicht</u> die Aussage, dass die Erfüllung der Ungleichung von Kraft äquivalent zur Existenz eines binären Präfixcodes mit Wortlängen $(\ell_1, \ell_2, \ldots, \ell_n)$ ist.
  2. Überprüfen Sie in den folgenden Fällen, ob die Kraftsche Ungleichung erfüllt ist. wenn ja, konstruieren Sie einen binären Präfixcode mit den entsprechenden Wortlängen:
     (a) $(\ell_1, \ldots, \ell_{10}) = (2, 2, 3, 3, 3, 4, 4, 4, 5, 5)$
     (b) $(\ell_1, \ldots, \ell_{10}) = (2, 2, 3, 3, 4, 4, 4, 4, 5, 5)$
     (c) $(\ell_1, \ldots, \ell_{15}) = (3, 4, 4, 4, 4, 4, 4, 5, 7, 7, 7, 7, 7, 8, 8)$
     (d) $(\ell_1, \ldots, \ell_{15}) = (3, 3, 3, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6)$

- **Aufgabe 30: Schranken für die Summe der Codewortlängen [Klausuraufgabe Frühjahr 2004 und Herbst 2005]**

  1. Zeigen Sie: wird zu einer Quelle mit N Symbolen ein optimaler Präfixcode mit Wortlängen $\ell_1, \ell_2, \ldots, \ell_N$ konstruiert, so gilt stets
     $$N \log N \leq \ell_1 + \ell_2 + \ldots + \ell_N \leq \frac{N^2 + N - 2}{2}$$
     Identifizieren Sie die Fälle, in denen Gleichheit gilt.
  2. Welche Vektoren für die Wortlängen $(\ell_1, \ldots, \ell_5)$ mit $1 \leq \ell_1 \leq \ell_2 \leq \ldots \leq \ell_5$ sind möglich bei *optimaler* binärer Präfixcodierung eines fünfelementigen Quellalphabets? Geben Sie in jedem der Fälle auch eine Wahrscheinlichkeitsverteilung an, für die die optimale Codierung die entsprechenden Wortlängen hat.

- **Aufgabe 31: Optimale Präfixcodes (z.T. Klausuraufgaben)**

1. Bestimmen Sie einen optimalen Präfixcode für die Quelle $\mathcal{Q} = \langle A, \boldsymbol{p} \rangle$ mit

$$A = \{a, b, c, d, e, f\}, \quad p_a = \frac{1}{3}, p_b = \frac{1}{4}, p_c = \frac{1}{6}, p_d = p_e = p_f = \frac{1}{12},$$

   berechnen Sie dessen mittlere Wortlänge und vergleichen sie das Resultat mit den durch das Quellcodierungstheorem gegebenen Schranken.

2. In einem File liegt ein Text $T \in \{a, b, c, d, e, f\}^*$, bei dem die einzelnen Symbole mit den folgenden Häufigkeiten vorkommen

Symbol	a	b	c	d	e	f
Häufigkeit (in %)	45	13	12	16	9	5

   (a) Konstruieren sie einen optimalen Präfixcode Φ für diese Situation.

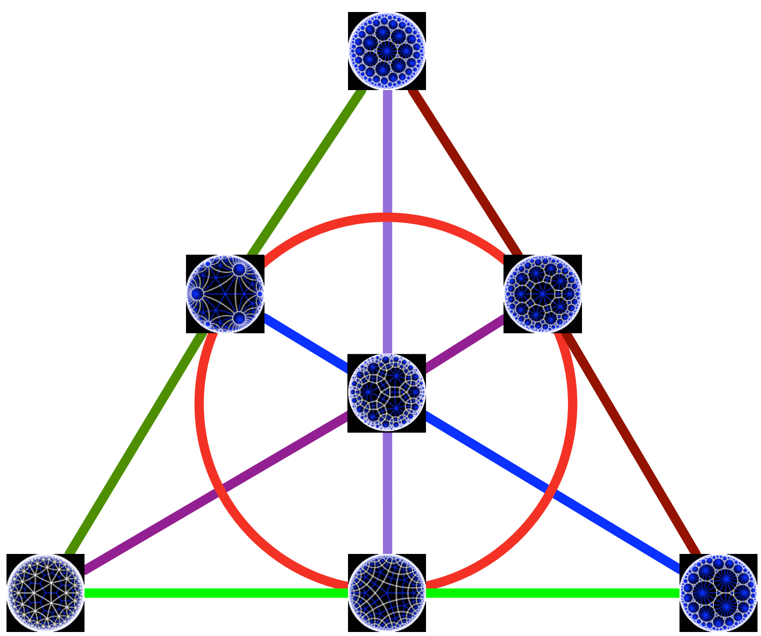
   (b) Wieviel Prozent spart man bei einer Übertragung/Speicherung mittels Φ im Vergleich zu einer Codierung mit fester Wortlänge?

   (c) Geben sie eine Codierung von Φ als bitstring an.

3. Konstruieren sie, ausgehend von den Häufigkeiten der auftretenden Buchstaben, einen optimalen binären Präfix-Code für den Text[1]

   ARIA␣VARIATA␣ALLA␣MANIERA␣ITALIANA

   Dabei soll der Zwischenraum ␣ als eigenes Symbol zählen.
   Es genügt, die Codierung, ohne deren Herleitung, in geeigneter Form (z.B. Liste, Binärbaum) anzugeben.
   Berechnen sie die Länge der entsprechenden optimalen binären Codierung des obigen Textes.



Erholsame Feiertage und einen guten Rutsch ins Neue Jahr!

[1] Titel eines Klavierstücks von J.S. Bach, BWV 989.