Prof. Dr. V. Strehl
Informatik 8

14. Dezember 2005

Übungen zur Theoretischen Informatik 3 WS 2005/06

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 26: Verhalten der Lösungen von divide-and-conquer-Rekursionen**

  1. Bestimmen Sie das Θ-Verhalten der Lösungen der folgenden Rekursionsgleichungen
     (a) $T(n) = 8 \cdot T(n/3) + n^2$
     (b) $T(n) = 9 \cdot T(n/3) + n$
     (c) $T(n) = 9 \cdot T(n/3) + n^2$
     (d) $T(n) = 10 \cdot T(n/3) + n^2$
  2. Zeigen Sie, dass die Lösung der Rekursion $T(n) = 2 \cdot T(\lfloor\sqrt{n}\rfloor) + \log n$ das Verhalten $T(n) \in \Theta(\log n \log\log n)$ hat.
     Hinweis: Das ist leichter, als man beim ersten Hinsehen denkt. Versuchen Sie die Substitution $m = \log n$. Vernachlässigen Sie grosszügig die Rundungseffekte.

- **Aufgabe 27: Schranken für die Summe der Codewortlängen [Klausuraufgabe Frühjahr 2004 und Herbst 2005]**

  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

  $$N \log N \le \ell_1 + \ell_2 + \ldots + \ell_N \le \frac{N^2 + N - 2}{2}$$

  Identifizieren Sie die Fälle, in denen Gleichheit gilt.

- **Aufgabe 28: Optimale Präfixcodes**

  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?

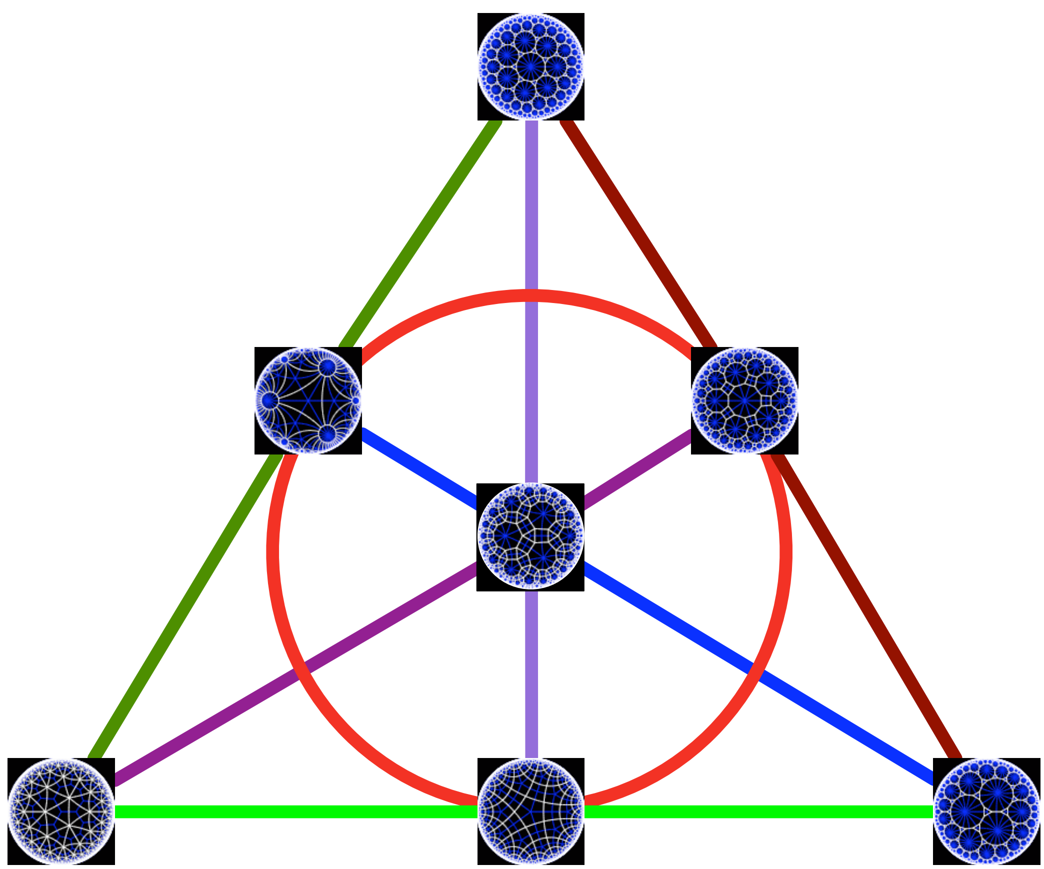
(c) Geben sie eine Codierung von Φ als bitstring an.

- **Aufgabe 29: Probabilistisches zum MaxSat-Problem**

  Es sei $X_n = \{x_1, x_2, \ldots, x_n\}$ eine Menge von booleschen Variablen und $F = C_1 \wedge C_2 \wedge \ldots \wedge C_m$ eine aussagenlogische Formel in Klauselform, d.h. F ist Konjuktion von Klauseln C_i, wobei Klauseln Disjuktionen von Literalen sind. Literale sind Variable oder negierte Variable. eine Variable und ihre Negation sollen nicht in einer Klausel vorkommen. Eine Bewertung ist eine Abbildung $\phi : X_n \mapsto \{\mathtt{t}, \mathtt{f}\}$, die auf die übliche Weise zu einer Bewertung der Literale, Klauseln und Formeln fortgesetzt wird.

  Das Problem MaxSat fragt bei gegebener Formel F danach, wieviele ihrer Klauseln maximal von einer Bewertung ϕ (gleichzeitig) erfüllt werden können. ϕ variiert über die Menge aller Bewertungen, die man mit den 2^n Teilmengen von X_n identifizieren kann. Da die Antwort auf diese Frage auch ein NP-vollständiges Problem löst (ist das Maximium = m?), ist klar, dass dieses ein NP-hartes Problem ist.

  Zeigen Sie: zu einer Formel $F = C_1 \wedge C_2 \wedge \ldots \wedge C_m$ (wie oben) gibt es immer eine Bewertung ϕ, die mindestens $m/2$ der Klauseln wahr macht. Verwenden Sie die “probabilistische Methode”, analog zum MaxCut-Problem.



Erholsame Feiertage und einen guten Rutsch ins Neue Jahr!