Exercise Sheet 1

Due date: 10:30, Apr 30th, to be submitted in Whiteboard.
Late submissions will, like a lazy combinator[1], not count.

You should try to solve all of the exercises below, and clearly mark which two solutions you would like to be graded — each problem is worth 10 points. Starred exercises represent problems that may be a little tougher, should you wish to challenge yourself.

Exercise 1 We say a 4-colouring of the vertices of a k-uniform hypergraph is *rainbow* if every edge has all four colours represented. Prove that all k-uniform hypergraphs H with $e(H) \leq \frac{4^{k-1}}{3^k}$ admit a rainbow 4-colouring.

Exercise 2 A *uniquely decipherable* code is a code $\varphi : A \to \{0,1\}^*$ such that every concatenation of codewords is unique.

(a) Give an example of a uniquely decipherable code that is not prefix-free.

(b) Prove the codeword lengths $\ell_1, \ell_2, \ldots, \ell_n$ of a uniquely decipherable code satisfy

$$\sum_{i=1}^{n} 2^{-\ell_i} \leq 1.$$

[Hint at `http://discretemath.imp.fu-berlin.de/DMIII-2020/hints/S1.html`.]

Exercise 3 A *binary linear code* of *dimension* d and *length* n is a code over the alphabet $A = \mathbb{F}_2^d$ defined through a matrix $M \in \mathbb{F}_2^{n \times d}$. For $\vec{a} \in A$, the corresponding codeword is the vector $M\vec{a}$. The *minimum distance* of the code is the minimum Hamming distance between two distinct codewords; that is, $\min\{\|M\vec{a} - M\vec{a}'\|_1 : \vec{a} \neq \vec{a}' \in \mathbb{F}_2^d\}$.

(a) Prove that if $2^d \mathbb{P}\left(\mathrm{Bin}(n, \frac{1}{2}) < k\right) < 1$, then there is a binary linear code of dimension d, length n and minimum distance at least k.

(b) Deduce that, if $k \geq 3$ is fixed and $d \to \infty$, then there is a binary linear code of dimension d, minimum distance at least k, and length $n = d + (k-1)\log d + O(1)$.

[Hint at `http://discretemath.imp.fu-berlin.de/DMIII-2020/hints/S1.html`.]

[1] One who practices the dark arts of combinatorics is traditionally referred to as a combinatorialist or, occasionally, a combinatorist, but I, ever a fan of brevity, prefer the punchier “combinator.”

Exercise 4 Given constants $a_1, a_2, \ldots, a_k$ and $b_1, b_2, \ldots, b_\ell$ in $\mathbb{N}$, let $\mathcal{L}$ denote the linear system $\sum_{i=1}^{k} a_i x_i = \sum_{j=1}^{\ell} b_j y_j$. We call a set A of natural numbers *$\mathcal{L}$-free* if there is no solution to $\mathcal{L}$ with $x_i, y_j \in A$.

Prove that if $\alpha := \sum_{i=1}^{k} a_i \neq \sum_{j=1}^{\ell} b_j =: \beta$, then there is some constant $c = c(\alpha, \beta) > 0$ such that every set S of n natural numbers contains an $\mathcal{L}$-free subset $A \subseteq S$ with $|A| \geq cn$.

Exercise 5 In lecture we proved the bounds $2^{k+1} - 1 \leq \sigma(k) \leq k^2 2^k(\ln 2 + o(1))$ on the size $\sigma(k)$ of the smallest tournament with the Schütte property S_k. Prove that the lower bound can be improved to $\sigma(k) = \Omega(k 2^k)$.

Exercise 6* Suppose X and Y are independent identically distributed real-valued random variables. Prove that

$$\mathbb{P}(|X - Y| \leq 2) \leq 3\mathbb{P}(|X - Y| \leq 1).$$

[Hint at `http://discretemath.imp.fu-berlin.de/DMIII-2020/hints/S1.html`.]