Homework 1 (corrected) — Scheme review — due Monday 28 January

Palindromes

Given a list of one-letter strings, determine whether it is a palindrome.

Powerset

Represent sets using lists, and write a function that takes a set S and returns its powerset 2^S. Let $f(0) = \emptyset$, $f(k+1) = 2^{f(k)}$. Write a Scheme function for f and compute the sets $f(k)$ for $k = 0, 1, 2, 3, 4$.

Binary trees

For a given n, enumerate all structurally distinct binary trees with a total number of nodes equal to n.

Safe transmission codes and molecules

An alphabet is a set of symbols; for instance $\Sigma_1 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ is the alphabet of decimal digits. A string is a finite sequence of symbols drawn from a specific alphabet; for instance $x_1 = 2718281828$ is a string over the alphabet Σ_1. Given two strings of equal length, the Hamming distance between them is the number of positions in which their symbols differ. For instance, $H(2718281828, 1828182827) = 6$. For strings of length N, the Hamming distance is at least 0 and at most N.

With an alphabet of K symbols, there are K^N strings of length N. We'll say that two strings x and y are distinguishable if $H(x, y) \geq D$, where D is a prescribed minimum Hamming distance. Out of the total K^N strings, how many (at most) can we choose such that each two of them are distinguishable?

Try to work out (mathematicaly) how large the biggest possible subset is, with K, N, and D as parameters. (This is a very difficult problem: no good solutions are known.)

Write a Scheme function that will produce an exemplar of such a subset of distinguishable strings. The function should be called like this: `(maxSafeSubset alphabet wordLength minHammingDistance)`; for instance we might invoke it as: `(maxSafeSubset '(0 1) 3 2)`, and it should then evaluate to `'((0 0 0) (0 1 1) (1 0 1) (1 1 0))`.

Why is this problem really important? If the alphabet is $\{0, 1\}$, we are talking about coding theory used for error detection and correction over telecommunication lines. If the alphabet is $\{A, G, C, T\}$, we are talking about DNA (properly, oligonucleotide) sequences suitable for molecular computation.

How to turn in

Turn in your code by running

~dmykola/handin your-file

on a regular UNM CS machine.

You should use whatever filename is appropriate in place of your-file. You can put multiple files on the command line, or even directories. Directories will have their entire contents handed in, so please be sure to clean out any cruft.