Homework 8 (due November 5, 2019)

(1) Information content of sequences (12 points)

Consider a sequence of N uncorrelated elements, each of which may take on n possible values with constant probabilities, p_i ($i = 1, \ldots, n$). As we scan the sequence, the amount of information that each element contributes, measured in bits, is quantified by Shannon's formula for the information entropy,

$$I/N = -K\sum_{i=1}^{n} p_i \ln(p_i) \text{ with } K = \frac{1}{\ln 2}.$$

a) Show that in a sequence of flips of a balanced coin, every flip adds exactly 1 bit of information.
b) What is the information in bits per base on a fragment of single-stranded DNA, where the relative occurrence of C, T, G, and A is unknown (and so, we assume they all occur with equal frequency)?
c) Chemical analysis establishes the relative occurrence as $p_C = 0.36$, $p_G = 0.24$, $p_T = 0.28$, $p_A = 0.12$. Calculate the information content in a long DNA chain in bits per base.

(2) Thermal equilibration (18 points)

Two thermally insulated tanks filled with an ideal gas contain N_1 and N_2 molecules respectively and are initially at temperatures T_1 and T_2 ($T_1 > T_2$). After they are brought into thermal contact, they reach a new equilibrium state with temperature T_f.

a) Derive T_f as a function of T_1, T_2, N_1, and N_2.
b) For $N_1 = N_2 \equiv N$, calculate the entropy change ΔS upon thermal equilibration.