CS321: Numerical Methods in Comp Mol Bio

Homework 5

Due: Thursday, Sept 29 2005 at the begining of the section

Problem 1

In your 1st project you will be working with the HP model discussed in class.
What is the maximum amount of contacts that a single residue in a chain can make?
What is the lowest possible energy that we can obtain from a sequence. What parameters of the sequence are important for this bound?
Draw a conformation on the lattice with energy ≤ -3 for the sequence:

$$HPPHPHHPHPPH$$

.
Using the same comformation as you did above, show how you can change the energy of the conformation by exchanging 2 of the residues' locations (e.g. exchange locations of residue 1 and 5).
Using the original sequence above, draw a different conformation on the lattice with energy ≤ -3.

Problem 2

Assume that a computer can generate uniformly distributed numbers on the interval $(0, 1)$. Let X_1 be the first number that the computer generates, X_2 be the second such number generated and so on (X_n will of course be the n^{th} such number generated).
What is the probability that the first number generated is more than 0.2 away from the expected value (i.e. $P(|X_1 - 0.5| \geq 0.2)$? What bounds can Chebyshev's inequality provide here?
What is the probability that the avarage of the first two numbers generated is more than 0.2 away from the expected value (i.e. $P(|\frac{X_1+X_2}{2} - 0.5| \geq 0.2)$? What bounds can Chebyshev's inequality provide in this case?
Using Chebyshev's inequality, give a bound to the probability that the avarage of the first ten numbers generated is more than 0.2 away from the expected value (i.e. for $P(|\frac{\sum_{i=1}^{n} X_i}{10} - 0.5| \geq 0.2)$.
How many numbers generated by the computer should you avarage in order to get that the probability of this average being more than 0.2 away from the mean will be less than 0.1% (i.e. $P(|\overline{X}_n - 0.5| \geq 0.2) < 0.1\%$)?