Eidgenössische Technische Hochschule Zürich
Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Institute of Theoretical Computer Science
Lecturer: Prof. Jack Snoeyink
Assistant: Yves Brise

March 28, 2007

Geometric Computations in Molecular Biology SS07
Exercise Set 2

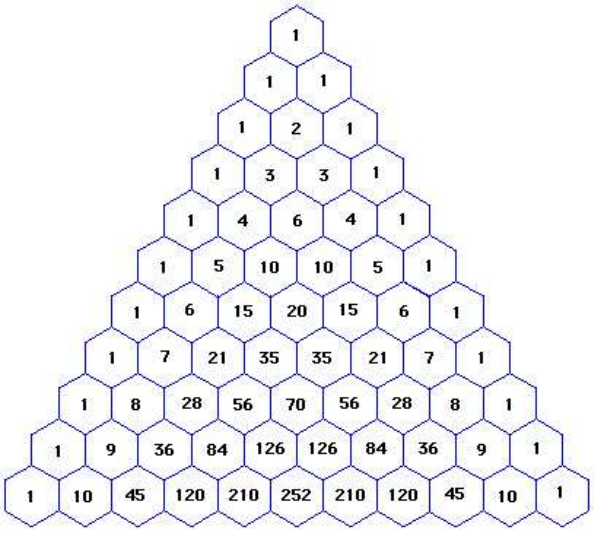
Course Webpage: `http://www.ti.inf.ethz.ch/ew/courses/GCMB07/`

Due date: -

Exercise 1

Suppose that you have an opaque bag containing 10 beads, each colored green or white. How many of each color do you estimate were in the bag if you pull four out and find:
a) (g,g,w,w)
b) (g,g,g,g)
c) (g,g,g,w)

Hint: Here is Pascal's triangle up to $n = 10$. Use it!



Exercise 2

Entropy was defined in today's lecture as $S = k \log(W)$, where k is is Boltzmann's constant and W is the number of states. The common definition in information theory is the entropy of a distribution of symbols that can occur with probability p_i, so that $\sum_i p_i = 1$. It is defined as $H = -\sum_i p_i \log p_i$.

a) Show that the definition of S is a special case of the definition for H, up to the constant k.

b) Assuming the definition for S, derive the definition for H for rational probabilities $p_i = n_i/D$ for integers n_i and D.

Exercise 3

Try folding with the HP model – the goal is to get as many green beads adjacent in the grid as possible: count -1 for each pair of green beads that is adjacent in the grid and not already adjacent in the string. Plain beads don't contribute to the score. What score do you obtain for the following instances?

Instances from Hart and Istrail's HP benchmark set
(http://www.dmi.unict.it/ mpavone/psp.html).

No	Sequence	Length	Score
1	hphpphhphpphphhpphph	20	-9
2	hhpphpphpphpphpphpphh	24	-9
3	pphpphhpppphhpppphhpppphh	25	-8
4	pppphhpphhpppppphhhhhhhpphhpppphhpphpp	36	-14
7	pphhhphhhhhhhhpppphhhhhhhhhhphpppphhhh... hhhhhhhhpppphhhhhhphhphp	60	-36
10	hhhhpppphhhhhhhhhhhhpppppphhhhhhhhh... hhhpppphhhhhhhhhhhpppphhhhhhhhhhhpp... phpphhpphhpphph	85	-53
13	phpphphhhphhphhhhh	18	-9
15	hhpppppphhpppphpphp	18	-4