



Homework 4

cs329e Elements of Biological Data Models
Created on 2/26/07
Version: 1.1

Assigned: 2/19/07

Due: 2/25/07 11:59PM, by email to the T.A

Reading:

- Old: Text 67-89
- New: Test 101-122

Part A:

For part A, same as homework three, (Develop a solution using Rational Rose),

1) Softlist problem, page 91

<http://www.tomjewett.com/dbdesign/dbdesign.php?page=software.php>

2) Pizza Shop problem, page 107

<http://www.tomjewett.com/dbdesign/dbdesign.php?page=pizza.php>

Part B:

First biological data model question. Develop a data model for storing gene sequences and their structural features.

Assume

- 1) the data made available in FASTA format.
 - I.e. be sure your model has a place to store all the information in fasta format
 - The wikipedia entry for FASTA is an acceptable definition.
- 2) use the structural features as they appear in slide 7 of the Genbank lecture, borrowed from:

http://www.cbil.upenn.edu/EpoDB/release/gifs/EpoDB_feats.gif

3) Adopt the protocol that denotes the location of the feature as a starting offset and ending offset in the sequence. E.g. (flank, 1, 97)