EEL851: Major Test (Part A)

May 5, 2014

Note: This part is optional. If you attempt it, and score more than in the Minor Test, then these marks will be counted for your grade instead of those. Otherwise, this will not affect your grade.

Maximum Marks: 20

1. (a) What is DNA phylogeny and why is it important?   **[3]**

   (b) Does the neighbor joining method used in DNA phylogenetic analysis give an unrooted or rooted tree? Why?   **[2]**

2. How is protein-protein interaction data obtained via affinity purification and mass spectrometry (AP/MS) typically converted into network form? What do you think are the advantages and disadvantages of the two approaches? **[4]**

3. Explain concisely the idea of date and party hubs in protein-protein interaction networks. Why do you think this idea had such a large scientific impact? **[5]**

4. Recall the formula for *Newman-Girvan modularity* discussed in class:

$$Q = \frac{1}{2m} \sum_{i,j=1}^{n} (A_{ij} - P_{ij})\delta(g_i, g_j).$$

(a) What does the term P_{ij} represent here? Give two possible choices for defining this term, along with their advantages/disadvantages, according to you. **[4]**

(b) Why is the term P_{ij} needed at all? Would there be a problem if we left it out of the above formulation? **[2]**