Feedback — HW 7 Cool and unusual applications

You submitted this homework on **Sat 17 Nov 2012 12:06 AM PST**. You got a score of **10.00** out of **10.00**.

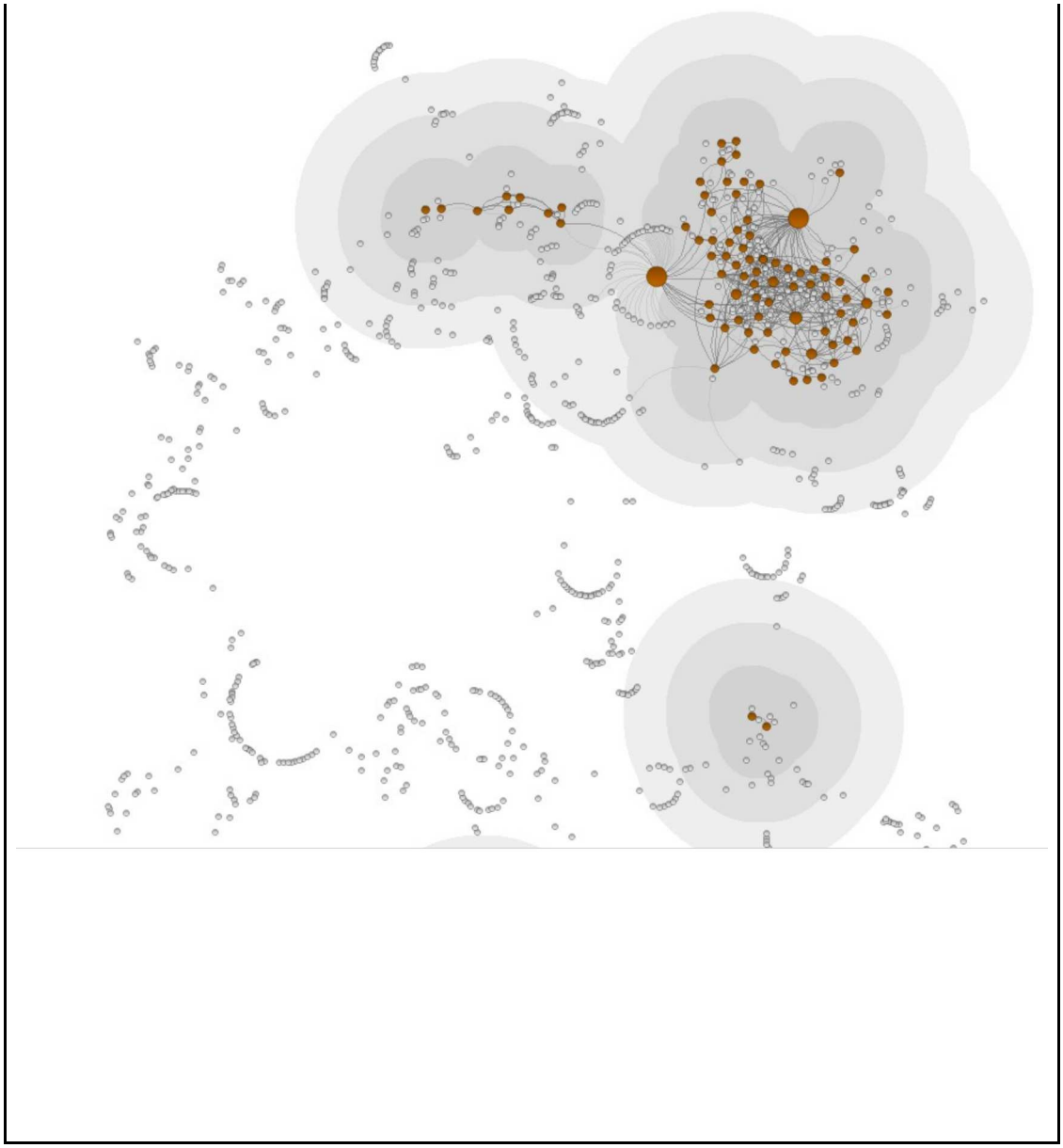
Question 1

Read the paper 'The human disease network' and download the diseasome.gexf file by right clicking on the link (to work with in Gephi). For this question, instead of working in Gephi, you can use the online explorer (Java 6 needed). For those of you working in Gephi, if upon loading the file, you end up with 1 node, right click on it and select ungroup to see all the diseases and genes. Now answer: Diseases of which of the following disease classes are most tightly clustered (proximate to one another) in the gene disease network.

Your Answer		Score	Explanation
◉ Cancer	✓	1.00	The majority of the different cancers are located in the same community, tied by the common genes whose mutations cause them.
Total		1.00 / 1.00	

Question Explanation

Locate the disease nodes associated with each disease class on the network. Here for example are the different cancers highlighted



Question 2

Continuing with the diseasome network, which of the following genes has highest degree centrality?

Your Answer		Score	Explanation
◉ TP53	✓	1.00	TP53 is associated with many different kinds of cancers.
Total		1.00 / 1.00	

Question Explanation

Calculate average degree in Gephi, which will give you the degrees for each node. Then sort by degree in the data laboratory to find the answer.

Question 3

Load the data on social connections between bottlenose dolphins living in Doubtful Sound, New Zealand. What do you observe about the network (check all that apply).

Your Answer		Score	Explanation
☐ the clustering coefficient is no higher than it would be for an equivalent Erdos Renyi random graph.	✓	0.67	to the contrary, clustering is pronounced in the network with many triads present
☐ the average degree is higher than it would be in an ER random graph with the same number of nodes and edges	✓	0.67	since the average degree is 2*E/N, the average degree is the same in the equivalent random graph.
☑ the network is assortative on sex: same sex interactions occur more often than they would if the network were rewired at random	✓	0.67	there are 58 M-M ties, 46 F-F ties, and 44 M-F ties, which contains many more same-sex ties than you would expect
Total		2.00 / 2.00	

Question Explanation

To answer the question, try constructing the equivalent ER random graph in Gephi and also coloring the nodes (in the real graph, not the ER graph) by sex. You can further filter based on intra and inter edges (by sex) using the Filters->Attributes->(Inter Edges or Intra Edges) option.

Question 4

The dolphins were followed for a period of years. At one point, one individual disappeared and the dolphins split up into two separate groups. When the individual reappeared, they again formed one social network. Coincidentally this individual has highest betweenness in the network. Who is it?

Your Answer		Score	Explanation
SN100	✓	1.00	SN100 has highest betweenness and is the individual who disappeared and reappeared in the network.
Total		1.00 / 1.00	

Question Explanation

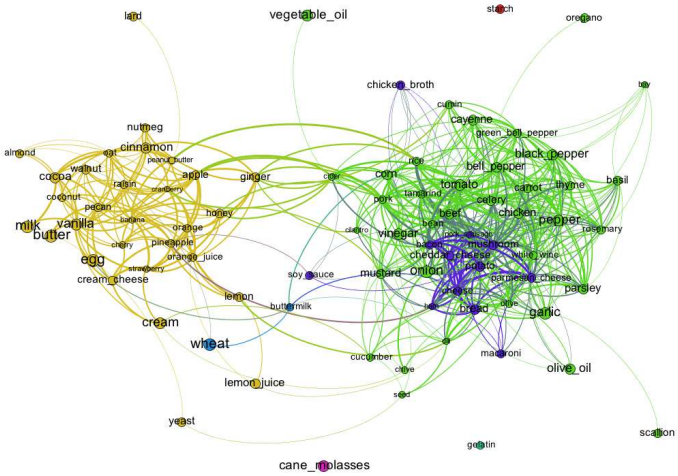
Calculate betweenness using the Avg. Path Length calculation in Gephi, or using other software to answer this question.

Question 5

The images below show a network of ingredients that are paired more often than you would expect at random (using pointwise mutual information), for different cuisines. A layout algorithm with the same parameters was applied to every network. Nodes are sized according to how many recipes the ingredient appears in. They are colored according to assigned community using Gephi's modularity-based community finding algorithm. The thickness of each edge corresponds to the logarithm of the number of shared flavor compounds between the two ingredients. Which of these networks exhibits the most pronounced community structure?

Your Answer		Score	Explanation
North American	✓	1.00	sweet and savory

ingredients are well-separated



network showing co-occurrence of ingredients in north american cuisine.

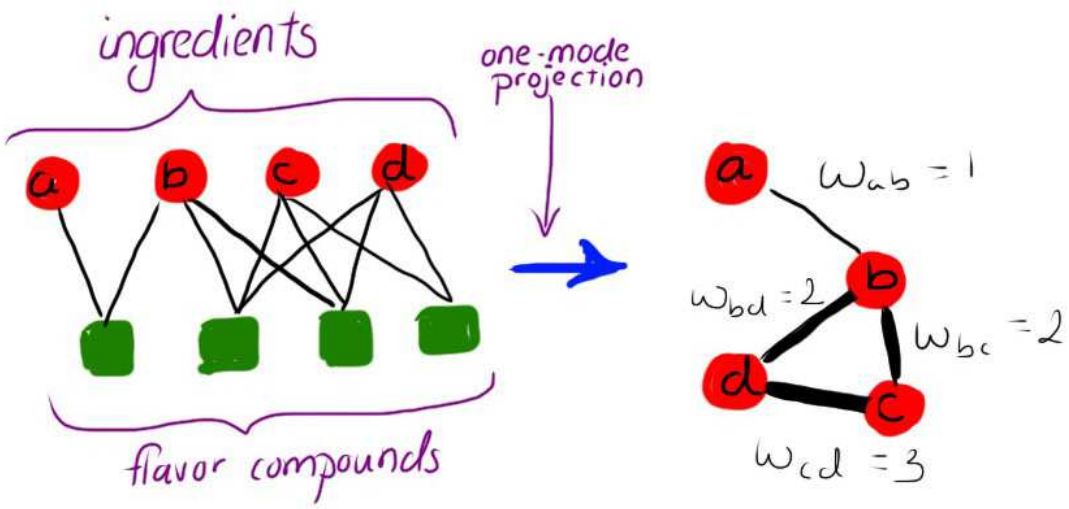
Total	1.00 / 1.00

Question Explanation

Observe whether any sets of nodes are more densely interlinked with each other (meaning that they co-occur often) but less sparsely connected to the rest of the network.

Question 6

The flavor network is a one-mode projection of a bipartite network where ingredients have edges to the flavor compounds they contain. This image shows how the weights are derived.



The networks you were shown in the previous question were flavor networks, with edges filtered out if the ingredients did not co-occur in recipes more often than expected. The average weights for the networks are as follows: Southern European: 18.51, North American: 16.70, East Asian 14.79, Southeast Asian 13.32. Based on this information, which of the following statements is true?

Your Answer		Score	Explanation
◉ Western cuisine tends to pair ingredients that share more flavor compounds.	✓	1.00	Since western cuisine has the highest average edge weight, it means that ingredients placed together in recipes have on average more shared flavor compounds.
Total		1.00 / 1.00	

Question Explanation

Try to determine how the weights are derived.

Question 7

Load the network of Middle Eastern ingredients. Calculate betweenness for all nodes. A high betweenness ingredient in Middle Eastern cuisine is

Your Answer		Score	Explanation
ginger	✓	1.00	
Total		1.00 / 1.00	

Question Explanation

Use the Av. Path Length function under Statistics in Gephi, or use another software of your choice.

Question 8

Based on the Hidalgo et al. paper and Cesar Hidalgo's video presentation, which of the following is true?

Your Answer		Score	Explanation
The product space network can be used to predict which products a country is likely to start exporting next based on its current production capabilities.	✓	2.00	Over time countries tend to start exporting products that are adjacent in the product space.
Total		2.00 / 2.00	

Question Explanation

To explore the work, check out the website devoted to the project.