Cancer Biology 8347 - Cancer Systems Biology - Spring 2023

Solutions to problems on graph theory I

1. Average shortest path length and clustering coefficient are two measures of concentration in a graph. Often graphs that have high clustering coefficient will have low average shortest path length, but that is not always the case. We will look at two graphs on ten vertices with fifteen edges.

Note that you should work out answers as **exact fractions** and also state them as **decimals to 2 decimal places**. You should also **show your working** and **explain your reasoning**. All work should be done by hand (do not use computational tools!); the point is for you to get a feel for things by computing the numbers yourself.

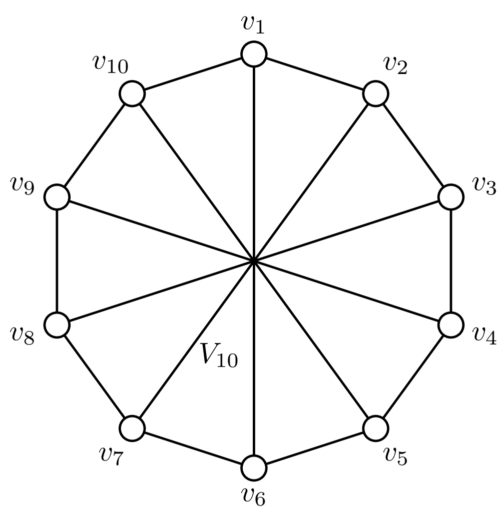
(a) Consider the graph shown at left below (this is known as the 10-*vertex Möbius ladder*, or V_{10}).

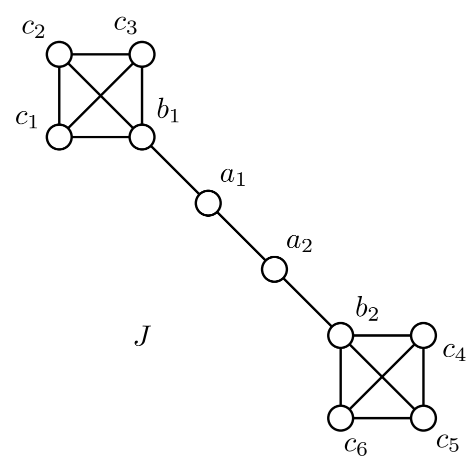
(i) Compute the clustering coefficient of the vertex v_1.

(ii) All vertices of V_{10} are *similar*, i.e., given any two vertices there is a symmetry that moves the first vertex to the second. Given this, what is the clustering coefficient of the whole graph?

(iii) What is the average length (number of edges) of a shortest path from v_1 to one of the other vertices? (Work out the length of a shortest path from v_1 to each of $v_2, v_3, \ldots, v_{10}$, and average these nine numbers.)

(iv) The average shortest path length for the whole graph turns out to be equivalent to the average over all vertices v_i of the average shortest path length from v_i to all other vertices. Given that all vertices of V_{10} are similar, what is its average shortest path length?





(b) Consider the graph shown at right above, which we will call J.

(i) Compute the clustering coefficients for the vertices a_1, b_1 and c_1.

(ii) Given that both vertices a_i are similar to a_1, both vertices b_i are similar to b_1, and all vertices c_i are similar to c_1, what is the clustering coefficient of the whole graph J?

(iii) Work out the average shortest path length from v to all other vertices for $v = a_1$, b_1 and c_1.

(iv) Given that both vertices a_i are similar to a_1, both vertices b_i are similar to b_1, and all vertices c_i are similar to c_1, what is the average shortest path length of the whole graph J?

Solution to 1.(a)(i) Vertex v_1 has three neighbors v_2, v_6, v_{10}, so there are $\binom{3}{2} = 3 \times 2/2 = 3$ potential edges among them. However, no pair of these three vertices is adjacent, so the actual number of edges is 0. Thus, the clustering coefficient is $C_{v_1} = 0/3 = 0$.

(ii) Since all vertices are similar, they all have the same clustering coefficient as v_1, i.e. 0. Therefore, the overall clustering coefficient, which is the mean of the individual vertex clustering coefficients, is also 0.

(iii) 'Shortest path length' is usually just referred to as 'distance' so I will use that term for brevity. Vertices v_2, v_6, v_{10} have distance 1 from v_1. Vertices v_3, v_5, v_7, v_9 have distance 2 from v_1 because there are paths $v_1v_2v_3$, $v_1v_6v_5$, $v_1v_6v_7$, $v_1v_{10}v_9$. Finally, vertices v_4 and v_8 have distance 3 from v_1 because there are paths $v_1v_2v_3v_4$, $v_1v_{10}v_9v_8$. Therefore, the average shortest path length from v_1 is $(1+1+1+2+2+2+2+3+3)/9 = 17/9 \approx 1.89$.

(iv) Since all vertices are similiar, the average shortest path length from any vertex is the same as from v_1, namely 17/9. So the overall average shortest path length, which is the mean of the average shortest path lengths from the individual vertices, is also $17/9 \approx 1.89$.

(b)(i) Vertex a_1 has two neighbors b_1 and a_2, so there is $\binom{2}{2} = 2 \times 1/2 = 1$ potential edge. But b_1 and a_2 are not adjacent, so the actual number of edges is 0. Thus, $C_{a_1} = 0/1 = 0$.

Vertex b_1 has four neighbors, a_1, c_1, c_2 and c_3, so there are $\binom{4}{2} = 4 \times 3/2 = 6$ potential edges. There are three actual edges c_1c_2, c_1c_3 and c_2c_3. Thus, $C_{b_1} = 3/6 = 1/2 = 0.50$.

Vertex c_1 has three neighbors, b_1, c_2 and c_3, so there are $\binom{3}{2} = 3 \times 2/2 = 3$ potential edges. And all three are actual edges b_1c_2, b_1c_3 and c_2c_3. Thus, $C_{c_1} = 3/3 = 1$.

(ii) Since all vertices a_i are similar, both a_1 and a_2 have clustering coefficient 0. Since all vertices b_i are similar, both b_1 and b_2 have clustering coefficient 1/2. Since all vertices c_i are similiar, all of $c_1, c_2, \ldots, c_6$ have clustering coefficient 1. Therefore, taking the mean, the overall clustering coefficient is $(0+0+\frac{1}{2}+\frac{1}{2}+1+1+1+1+1+1)/10 = 7/10 = 0.70$.

(iii) See the table of distances (shortest path lengths) below.

	a_1	a_2	b_1	b_2	c_1	c_2	c_3	c_4	c_5	c_6	av. sh. path length
a_1	–	1	1	2	2	2	2	3	3	3	$\frac{19}{9} \approx 2.11$
b_1	1	2	–	3	1	1	1	4	4	4	$\frac{21}{9} = \frac{7}{3} \approx 2.33$
c_1	2	3	1	4	–	1	1	5	5	5	$\frac{27}{9} = 3$

(iv) Since all vertices a_i are similar, both a_1 and a_2 have average shortest path length 19/9. Since all vertices b_i are similar, both b_1 and b_2 have average shortest path length 7/3. Since all vertices c_i are similiar, all of $c_1, c_2, \ldots, c_6$ have average shortest path length 3. Therefore, taking the mean, the overall average shortest path length is

$$\begin{aligned}
&(\tfrac{19}{9}+\tfrac{19}{9}+\tfrac{7}{3}+\tfrac{7}{3}+3+3+3+3+3+3)/10 \\
&= (\tfrac{38}{9}+\tfrac{14}{3}+18)/10 \\
&= (4\tfrac{2}{9}+4\tfrac{2}{3}+18)/10 = 26\tfrac{8}{9}/10 = \tfrac{242}{9}/10 \\
&= \tfrac{242}{90} = \tfrac{121}{45} = 2\tfrac{31}{45} \approx 2.69.
\end{aligned}$$