COMPSCI 2XB3 L04 Lab Report #7

Mahad Aziz - azizm17 - 400250379 - azizm17@mcmaster.ca

Talha Amjad - amjadt1 - 400203592 - amjadt1@mcmaster.ca

Logan Brown - brownl33 - 400263889 - brownl33@mcmaster.ca

McMaster University

COMPSCI/SFWRENG 2XB3: Binding Theory To Practice

Dr. Vincent Maccio

TA: Amir Afzali

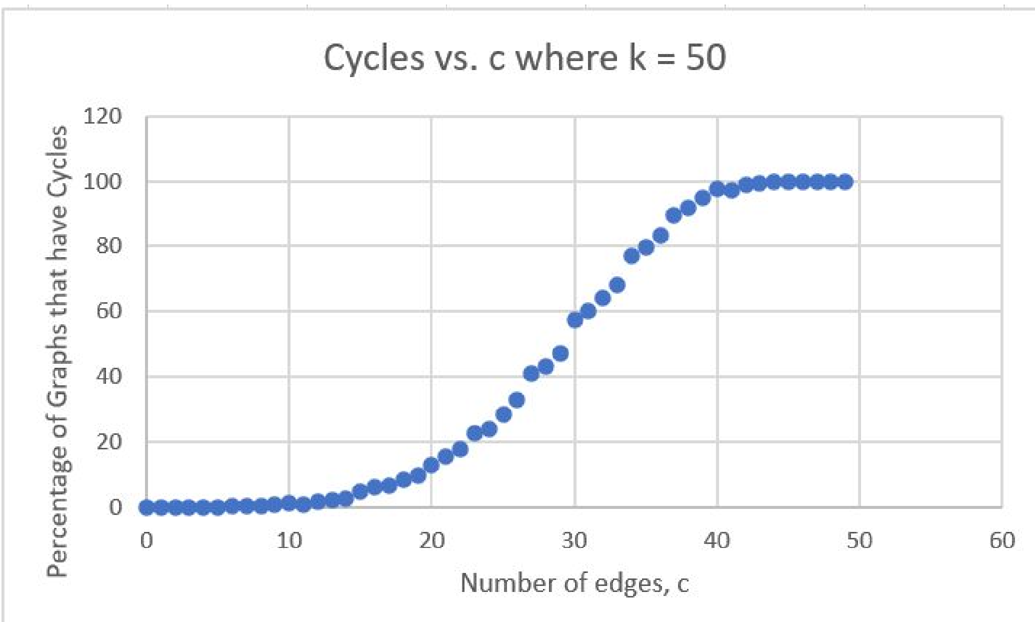
March 12, 2021

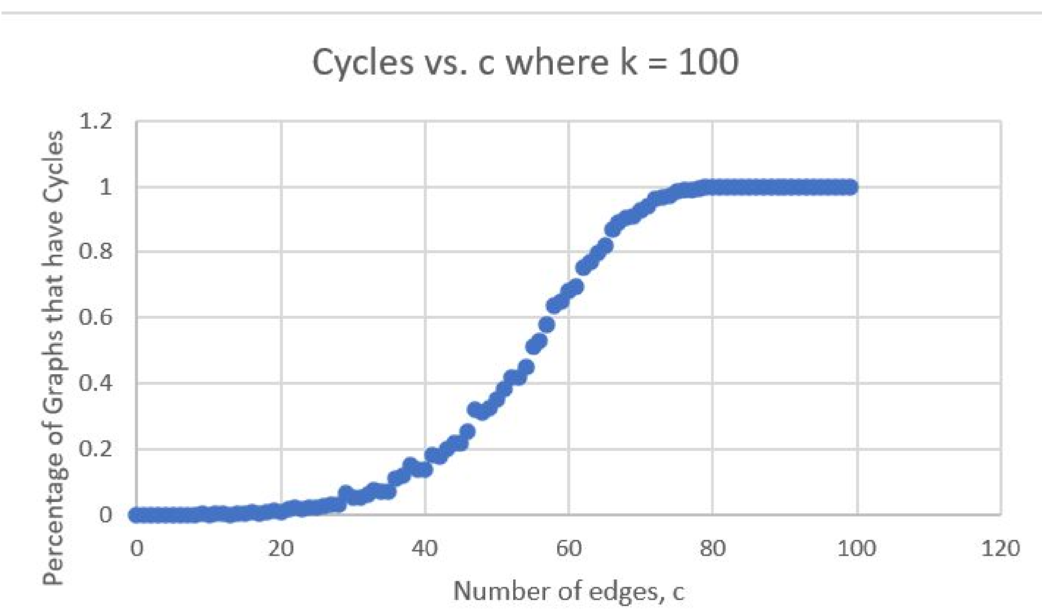
Cycles and Connected Probability:

We initially constructed a function to generate a random graph based on the number of nodes, k, and the number of edges, c. This function generates two random numbers between 0 and k and adds an edge connecting one random number to the other in our graph. We also discovered that there is a limit on the maximum number of edges in a graph based on the number of nodes. We found the formula that determines the maximum number of edges based on the number of nodes which is (k2 - k) / 2. We apply this formula in a conditional to reset the number of edges if a value is given that is greater than the number produced from the formula. There are also a couple of scenarios where the random numbers are not valid and must be picked again. If the two random numbers are equal to each other or one of the random numbers generated already has an edge connecting it to the other random generated node then the random nodes must be reselected until these two conditions are not met. We then created two functions to test and provide us with the percentage of graphs that are connected or have a cycle based on the number of nodes and edges. In each function, we iterate through the c values starting from 0 and generate a graph using the random graph generator function and then we run our has_cycle() or is_connected() functions on the random graph to see whether the graph has a cycle or is connected. For testing, we decided on using k values of 50 and 100 to determine which percentage of the random graphs will hold the properties. For has_cycles() we decided on going up to a c value of 50 when k = 50 and 100 when k = 100. For the is_connected() we decided on going up to a c value of 250 when k = 50 and 500 when k = 100.

For each experiment, what approximately is the value of c which roughly half the graphs to contain a cycle/be connected.

Contain a cycle:

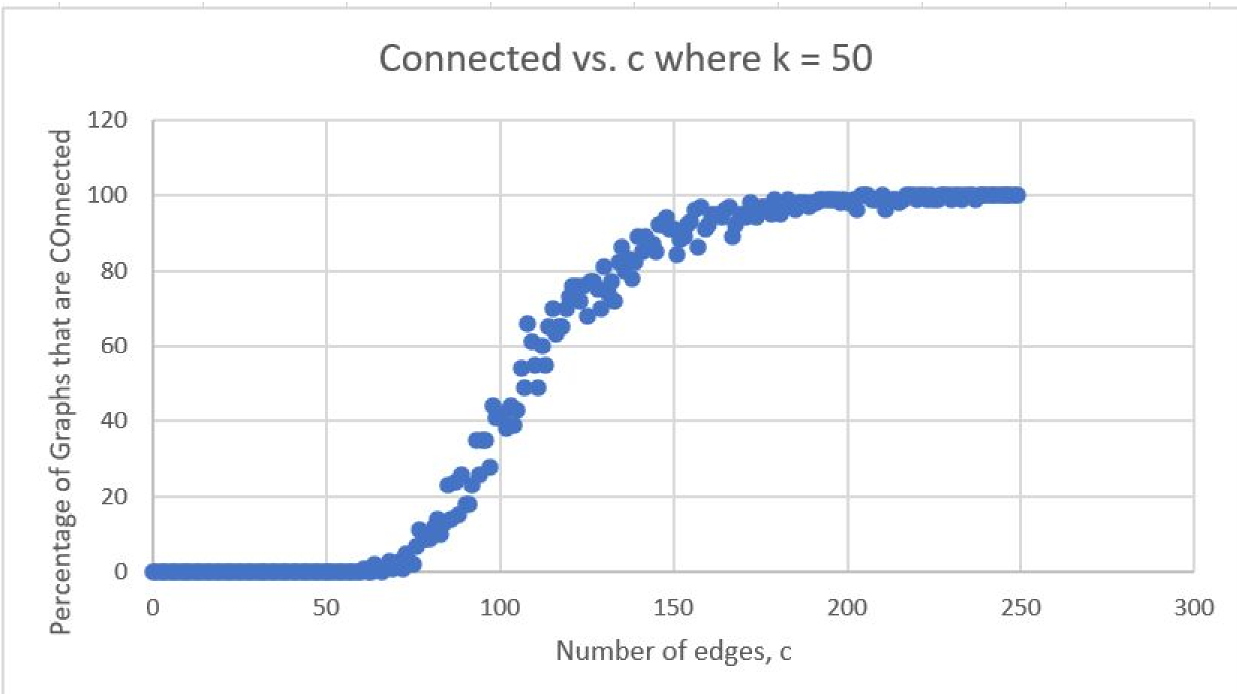


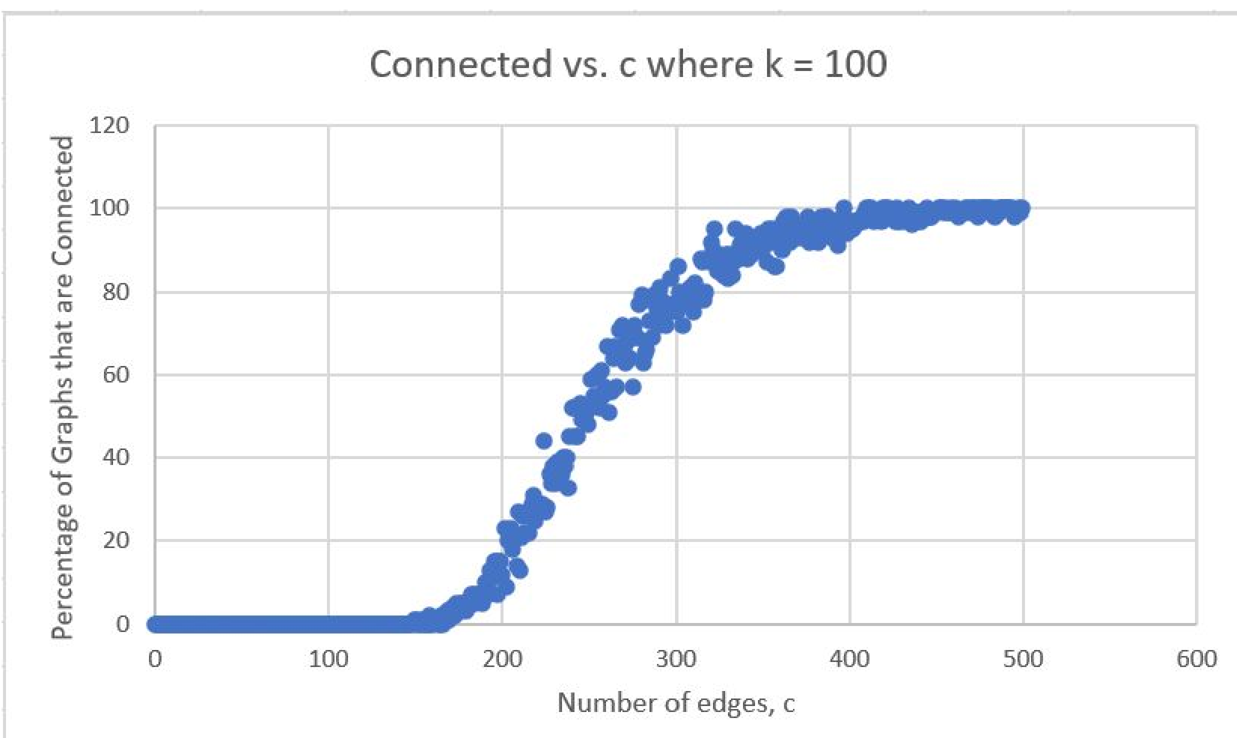


- When we compared the number of graphs that have cycles vs the number of edges, c, when k = 50, our c value at approx 29 showed that about half the graphs will have a cycle.
- When we compared the number of graphs that have cycles vs the number of edges, c, when k = 100, our c value at approx 54 showed that about half of the graphs will have a cycle.

Testing the percentage of graphs that have a cycle vs. the number of edges in a graph showed us that when the number of edges in a graph is at around 50% - 60% of the number of nodes is when approximately half the graphs will have a cycle. When our k value was equal to 50 we found that half of the graphs had cycles when $c \cong 3k/5$ and when our k value was equal to 100 we found that half of the graphs had cycles when $c \cong k/2$. We concluded that when $c = k/2$ to about $c = 3k/5$ is when approximately half of the graphs will contain a cycle.

Be connected:





- The value of c where half of the graphs are connected for k = 50, is roughly 107. We can see that it takes approx. 107 edges for half of the graphs to be connected and 29 for a graph to contain a cycle for k = 50.
- The value of c where half of the graphs are connected for k = 100, is roughly 250. We can see that it takes approx. 250 edges for half of the graphs to be connected and 54 for a graph to contain a cycle for k = 50.

From testing the percentage of graphs that are connected vs. the number of edges in a graph showed us that when the number of edges in a graph is at around 200% - 250% of the number of nodes is when approximately half the graphs will be connected. Looking at both graphs we found the relation between c and k to be approximately $c \cong 2k$ to $c \cong 2.5k$. Comparing the cycles data to the connected data we see that it takes more edges for half the graphs to be connected than for half the graphs to contain a cycle.

Does it make sense why one of these values would be less than the other?

It makes sense that the number of edges in which half the randomly generated graphs have a cycle is lower than that same number for connected graphs. This is because the definition of connected means that a graph must have a path from one node to any other and a cycle only has to involve a subset of nodes. The property of a graph being connected inherently involves every node whereas a graph having a cycle may only need a couple of nodes to be connected. From our tests we discovered that the c value is between k/2 and 3k/5 when half of the random graphs have a cycle. On the other hand, we discovered that the c value is between 2k and 2.5k when half of the random graphs are connected. Intuitively, there’s a higher chance a smaller set of nodes will have the relation between them that makes them a cycle than there is for all of the nodes to share a relation making them connected. So, the number of edges needed for half the graphs to have a cycle is less than the number for the connected graphs.