CX 4010 / CSE 6010
Assignment 3
Phase Transitions and Breadth First Search

Due Dates:

- Due: 11 AM, Friday, September 13, 2019
- No late submissions will be accepted

1. Background

This assignment involves studying phenomena that arise in physics and materials known as phase transitions. A phase transition occurs when a material changes from one state to another causing it to exhibit different physical properties. A phase transition can occur because of a change in some external condition such as temperature or application of a force such as voltage. Phase transitions are important in understanding phenomena such as water freezing and technologies such as liquid crystal displays. Here, we will build some simple computational models to study phase transitions.

Consider a two-dimensional *K by K* grid-structure such as that shown in Figure 1(a). Each grid cell is either colored or uncolored. A cell is colored with probability *P* and uncolored with probability *1.0–P* with the same value of *P* used throughout the entire grid. Here, we are primarily concerned with colored cells and how they cluster for different values of *P*. For example, the grid might represent a body of water at a temperature near freezing. Colored cells represent areas of water that have frozen and uncolored cells areas that are still liquid. The value of *P* is related to the air temperature.

A colored cell *A* is said to be *adjacent* to another colored cell *B* if *B* is the cell immediately to the north, south, east, or west of *A* in the grid. Each cell has 4 adjacent cells (neighbors) except those along the edges of the grid which have only 2 or 3 neighbors. A colored grid can be represented as an undirected graph where each colored cell is represented by a vertex and a link between vertices indicates the corresponding cells are adjacent. For example, the graph shown in Figure 1(b) represents the grid structure shown in Figure 1(a). Here, we have labelled vertices with integers to uniquely identify them.

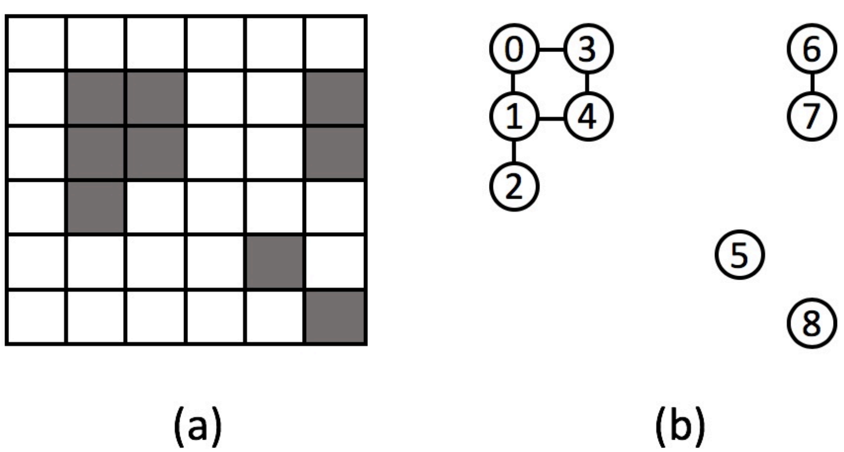

Figure 1. Representations of structures. (a) Grid. (b) Graph.

A connected component (or just *component*) is defined as a subgraph of the original graph where there is a path between every pair of vertices in the subgraph, and there are no edges between a vertex in the component and any other vertex not in the component. The largest component of a graph is the component with the largest number of vertices. For example, the structure shown in Figure 1 consists of four components of size 5, 2, 1 and 1, with the largest component containing 5 cells (vertices).

We are interested in the size of the largest component for different values of P. For example, it is clear that if P is close to 0.0 the graph will consist of mostly isolated, unconnected vertices and perhaps a few small clusters, so most components will have a size of perhaps 1 or 2 cells. On the other hand, if P is close to 1.0, the graph will consist of a single, large component of size N, where N is the number of vertices in the graph. Note that if P is small, the size of the largest component is largely independent of the size of the grid, but if P is large (close to 1.0) the size of the largest component increases with the size of the grid. These two situations might be referred to as "phases." The goal of this assignment is to examine how the size of the largest component varies as P is increased from 0.0 to 1.0, i.e., how this system changes from one "phase" to the other.

This assignment involves two parts. The first creates a grid with colored cells, as discussed above, using P as a parameter. The second determines the size of the largest connected component in the grid that was created. You will then conduct a series of computational experiments to study the relationship between the size of the largest component and the parameter P. You should use *breadth first search* (BFS) to compute the size of the largest component of the graph.

You should write your analysis software to be computationally efficient. For example, you should use a reasonably efficient algorithm for implementing the search and you should minimize the number of search operations; for example, invoking your BFS function on every node of the graph is not very efficient! Also note that you need not create a separate data structure to represent the graph; you should use the grid structure of colored cells you created earlier. Your software must be modular and well documented.

Perform a series of computational experiments to generate a plot of the size of the largest components as P is varied. Use as large a value of N as your programs can process in a reasonable amount of time, e.g., a few minutes per run.

Write a brief report providing evidence that your software executes correctly. Describe how you verified your programs gives correct results. You may use results from small sized grids for this purpose. Include in your report a plot showing the size of the largest component as P is varied. Compare your results concerning average component size with those described in the literature[1]. Discuss the degree to which your experiments agree or disagree and explain why. Be sure to include citations to relevant literature in your report. All code for computing the size of the largest component must be written in C. Once you produce the specified results, however, you may use software packages (e.g., Microsoft Excel) to generate a graph for your report.

[1] For example, see M.E.J. Newman, "Power Laws, Pareto Distributions and Zipf's Law," *Contemporary Physics*, Vol. 46, No. 5. (1 September 2005), pp. 323-351, doi:10.1080/00107510500052444

Grading

Turn in (1) all source code for your program and a README file with instructions how to compile and run your program, (2) example test files demonstrating your program executes correctly, and (3) the report documenting your results. Include some test cases so the TA can test your code to verify correctness.

Your grade will be based on the correct functioning of your software, the quality of the software in terms of functionality (did you check for erroneous inputs?), modularity and documentation, and the quality of your report. The report should be self-contained and explain all the work you did on this assignment as well as your interpretation of the results you obtained.

Collaboration, Citing, and Honor Code

As a reminder, please refer to the course syllabus regarding rules and expectations regarding collaborating with other students and use of other resources, including materials available on the web.