Social and Technological Networks Edinburgh, 2018

Exercises 0. Background Questions - solutions and discusssions.

Rik Sarkar *Exercises*

Below we give answers to some questions. The easier, or direct from book ones are not given. If you had any difficulty with these, reconsider taking the course.

Q 1. *The number of vertices is usually denoted by $|V|$ or n. The number of edges, $|E|$ is often denoted by m. For what value m will the graph have the largest number of connected components?*

Answer. $m = 0$.

Q 2. *Suppose in a graph $G = (V, E)$, $m = n - 5$. What is the minimum number of connected components in G?*

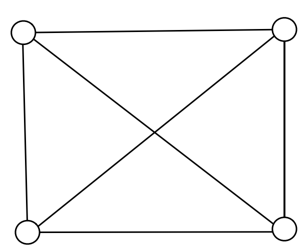
Answer. *5.*

Q 3. *Maximum of how many edges can a graph have? How do you write that in asymptotic notation?*

Q 4. *How many edges does a complete graph have?*

Q 5. *Suppose a graph is known to be sparse. That is, for most of the pairs for vertices, an edge does not exist between them. Let's say that the graph has $O(n)$ edges. How would you encode and store such a graph in python program? (or whichever programming language you want to use.)*

Q 6. *Is the following graph planar?*

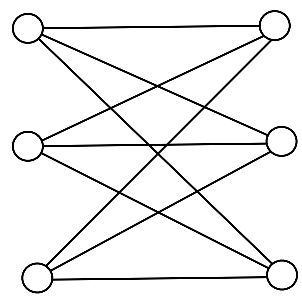


Answer. *Yes. (The graph is intrinsically planar. Though this particular "drawing" is not planar. You can draw a planar embedding of this graph.)*

Q 7. *Are road networks planar?*

Answer. *Strictly speaking, no. Road networks can have bridges and underpasses. In practice though, over large areas, bridges and underpasses are not very frequent and in some ways road networks are similar to planar graphs.*

Q 8. *What is the name of the following bipartite graph? Does it have a planar embedding (drawing)?*



Q 9. *What is the difference between a walk and path?*

Q 10. *Suppose G is an unweighted graph and $u, v \in V$. What algorithm would you use to compute the graph distance (sometimes called number of hops or steps) between u and v?*

Answer. *Breadth first search. (lookup breadth first trees in CLRS algorithms book.)*

Q 11. *Suppose we write the graph distance between u and v as $d(u, v)$. The set of all nodes within distance at most r from u is written as ball $B(u, r) = \{v \in V : d(u, v) \leq r\}$. What algorithm would you use to compute the set $B(u, r)$?*

Answer. *We can use breadth first search starting from u. This gives distances of all nodes to u, and we can select nodes in $B(u, r)$. The search can be restricted to be more efficient by ignoring any node v with $d(u, v) > r$ and not inserting them into the BFS queue.*

Q 12. *If G is a weighted graph, how would you answer the questions above?*

Q 13. *How many edges does a spanning tree have?*

Q 14. *Consider a directed graph G with s strongly connected components. Explain if the following true: "if we add an edge e to G, s can only increase".*

Answer. *This is incorrect. It is possible that adding an edge creates a large strongly connected component which includes multiple of the previous strongly connected components. So the count goes down. (create such an example.)*

Q 15. *Suppose T is the minimum spanning tree of a graph G. Let us write d_G and d_T for distances between nodes as measured in G and T respectively. Give an example of graph G to show that in some cases, there can be nodes u and v such that $d_T(u, v)$ is much larger than $d_G(u, v)$.*

Answer. *Consider vertices placed densely along a narrow U shape. The graph is the complete graph with edge weights equal to the planar distance between end points. The MST will be of a U shape. The distance between the end points of the U are far apart in the graph while they are close in the graph and the plane.*

Q 16. *Suppose every year Mr. X makes double the number of friends he made last year (starting with making 1 friend in first year). In how many years will he make n friends? (asymptotic notation is fine.)*

Answer. *Mr. X makes 1 friend in the first year, 2 in the second year, so he has in total $1 + 2$ friends in the second year. At the end of m-th year he will have $1 + 2 + \ldots 2^{m-1} = 2^m - 1$ friends. Now let us select the smallest m such that $2^m - 1 \geq n$. Observe that by this definition, after year $m - 1$, he had strictly less than n friends, and after year m he can actually have much more than n friends. However, m is still the right answer, because we are counting whole years.*

Expressing m in terms of n, we have $m = \lceil \lg(n + 1) \rceil$ [1]*. We have to use the ceiling function here because $n + 1$ may not be a power of 2, and we need to take the next integer to get a proper count.*

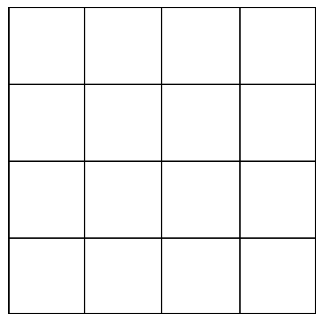
Q 17. *Suppose we throw k balls into n bins randomly, what is the probability that the first bin, bin-1 remains empty?*

Answer. $\Pr[$*bin 1 is empty after 1 throw*$] = 1 - \frac{1}{n}$. *Therefore,* $\Pr[$*bin 1 is empty after k throws*$] = (1 - \frac{1}{n})^k$.

Q 18. *What is union bound? In the problem above, can you put an upper bound the the probability that no bin is empty? [Hint: you know an upper bound on the probability that a particular bin is empty. And you know union bound.]*

Q 19. *The figure below shows a grid graph embedded in the plane. Prove that for an infinite grid (extending in all directions), the number of vertices inside a circle of radius r in the plane is $O(r^2)$.*

[1] The $\lceil \bullet \rceil$ symbol stands for the function *ceiling* implying the integer greater than or equal to its argument.



Answer.

Proof: Let us suppose each grid square has side s, and area s^2. Since the interiors of the grid squares are disjoint, the total area covered by any n distinct grid squares is ns^2. The area of the circle of radius r is πr^2, and the maximum number of possible squares in the circle is $\leq \pi r^2/s^2$. For a given grid s is fixed, so the number of squares in the circle is $O(r^2)$. Each square has 4 vertices, thus number of vertices is also $O(r^2)$. (This argument does the job, but is overcounting vertices; in fact in an infinite grid, we can map vertices to squares 1 to 1. So, the actual number is lower.). $\square$

Q 20. *Show that* $\ln n = \Theta(\lg n)$*, and* $\lg n = \Theta \log(n)$.

$\ln n$, $\lg n$ *and* $\log n$ *are the usual notations for* $\log$ *to base* e, 2 *and* 10 *respectively. This is to show that log functions to different constant bases differ only by constant factors.*