1 There are n galaxies connected by m intergalactic teleport-ways. Each teleport-way joins two galaxies and can be traversed in both directions. However, the company that runs the teleport-ways has established an extremely lucrative cost structure: Anyone can teleport *further* from their home galaxy at no cost whatsoever, but teleporting *toward* their home galaxy is prohibitively expensive.

Judy has decided to take a sabbatical tour of the universe by visiting as many galaxies as possible, starting at her home galaxy. To save on travel expenses, she wants to teleport away from her home galaxy at every step, except for the very last teleport home.

Describe and analyze an algorithm to compute the maximum number of galaxies that Judy can visit. Your input consists of an undirected graph G with n vertices and m edges describing the teleport-way network, an integer $1 \leq s \leq n$ identifying Judy's home galaxy, and an array $D[1 \,..\, n]$ containing the distances of each galaxy from s.

Solution:

We reduce this problem to finding the length of the longest path in a dag $G = (V, E)$ as follows:

- V is the set of n galaxies, plus an artificial target node t. Let s denote Judy's home galaxy.
- E contains a directed edge $u \to v$ if either of the following conditions is satisfied
  - There is a teleport-way between galaxy u and galaxy v, and v is farther from s than u.
  - $v = t$ and there is a teleport-way between galaxy u and s.
- We need to compute the length of the longest path in G from s to t.
- We can compute this length using dynamic programming as described in Tuesday's lecture (and in the lecture notes).
- The algorithm runs in $O(V + E) = \boldsymbol{O(n+m)}$ ***time.***

You probably heard of the phrase "six degrees of separation" and the "small world" phenomenon; see `https://en.wikipedia.org/wiki/Six_degrees_of_separation`. The idea is that in many interesting networks people or objects are within a small distance of each other. At the same time we believe in "locality" in that each person may only a small number of people compared to the total population. The next two problems explore the tradeoffs between diameter and degree in a graph to explore this in a more quantitative fashion.

2 Suppose G is a graph with maximum degree d. The diameter of the graph is $\max_{u,v} \operatorname{dist}(u, v)$. Prove that the diameter of the graph is $\Omega(\log_d n)$ where n is the number of nodes. It is easier to consider $d = 5$ or some other small constant for simplicity. *Hint:* Consider the BFS layers starting at any vertex v.

The point of the problem is to show that if all degrees are small then the diameter must grow with the number of nodes.

Solution:

Suppose the maximum degree is d. Let the diameter be $D = d(s, t)$ for some pair of nodes s and t. Then the number of nodes that can be in layer 1 of a BFS search starting at s (distance 1 away from s) is, by definition of degree, at most d. The number of nodes that can then be at layer 2 is at most d^2, since each node at layer 1 is connected to at most d other nodes. In general, the number of nodes at layer i is at most d^i. Since the diameter is D, the total number of nodes summing across all of the

D layers in a BFS starting from s is at most $\sum_{i=0}^{D} d^i \leq d^{D+1}$. But if there are n nodes, this means $n \leq d^{D+1}$, and taking logs base d, we have $\log_d n \leq D+1$, so diameter $D = \Omega(\log_d n)$ as was to be shown.

3 Suppose the diameter of an undirected simple graph is d. Prove that there is a node with degree at most $3n/d$. *Hint:* Consider the BFS layers for the pair defining the diameter. It is easier to prove a bound such as $9n/d$.

This problem is to show you that if the diameter is small then there must be a large degree node.

Solution:

Note that the graph is connected if it has finite diameter. If the diameter is d then there is a pair of nodes s, t such that $d(s,t) = d$. Consider the BFS layers starting from s. Hence we have $L_0 = \{s\}$, $L_1, L_2, \ldots, L_d$ where $t \in L_d$. Consider any node u in a layer L_i where $1 \leq i \leq d-1$. From the properties of the BFS layers u can be connected only to nodes in L_{i-1}, L_i, L_{i+1}. Thus $deg(u) \leq |L_{i-1}|+|L_i|+|L_{i+1}|$. For $i = 0$ we can say that $deg(s) \leq |L_1|$. For $i = d$ we can say that $deg(t) \leq |L_{d-1}| + |L_d|$. Suppose the claim is false and hence $deg(v) > 3n/d$ for every $v \in V$. For $1 \leq i \leq d-1$ pick an arbitrary node u_i from L_i (each layer is non-empty so we can do this). We have $deg(u_i) \leq |L_{i-1}| + |L_i| + |L_{i+1}|$ and since we assumed that the claim is false, $3n/d < |L_{i-1}| + |L_i| + |L_{i+1}|$ for each $1 \leq i \leq d-1$. For $i = 0$ and $i = d$ we have $|L_1| > 3n/d$ and $|L_{d-1}| + |L_d| > 3n/d$. Summing these inequalities from $i = 0$ to d we obtain

$$|L_0| + 3(|L_1| + |L_2| + \ldots + |L_{d-1}|) + 2|L_d| > (d+1) \cdot 3n/d > 3n.$$

The above implies that

$$|L_0| + |L_1| + \ldots + |L_d| > n$$

which is a contradiction since the every node is in exactly one of the layers. Thus there must be a node with degree at most $3n/d$.

A more elegant solution.

Solution:

Let $P = v_0, v_1, \ldots, v_d$ be the diameter-defining path of the graph. Here, unless $|i-j| = 1$, there cannot be an edge joining v_i and v_j, for otherwise the distance between v_0 and v_d would be smaller than d. In addition, any vertex u not in P can have at most three neighbors in P for the same reason.

Therefore, the sum of degrees of the vertices in P is

$$\sum_{i=0}^{d} \deg(v_i) \leq 2d + 3(n-d-1) = 3n - d - 3$$

By the Pigeonhole Principle, there is a node in P with degree at most

$$\frac{1}{d+1} \sum_{i=0}^{d} \deg(v_i) \leq \frac{3n-d-3}{d+1} \leq \frac{3n}{d}.$$