Assignment 3 – Part 2

Question 1.

In this question, we focus on the independent cascade model for influence maximization.

The greedy algorithm for solving the influence maximization problem adds nodes to the current set one at a time. Let $s(A)$ denote the number of nodes influenced by the initially active set A. At step 0, the algorithms starts with an empty set A_0. At step $i > 0$, it picks the node u which maximizes the marginal gain:

$$u = \arg\max s(A_{i-1} \cup u) - s(A_{i-1}),$$

and adds it to the set A_{i-1} to create the set A_i.
In general, the greedy algorithm does not guarantee an optimal solution. That is, there may be a set A^* with $|A^*| = i$ and $s(A_i) < s(A^*)$.

Assume that (a) a node influences itself, that the is the count of total influence $s(A)$ includes the nodes in A, and (b) in the case of ties, we choose one of the nodes at random.

a. Construct an example such that for $i = 2, s(A_i) < s(A^*)$.
b. Construct an example such that for $i = 3, s(A_i) \leq 0.8\, s(A^*)$, that is the greedy algorithm finds a solution that is at most 80% of the optimal one.
c. Give example of (a family of) graphs, where the greedy algorithm always produces an optimal solution.

Question 2.

The goal of this question is to understand the proof that decentralized greedy search is efficient for small world networks when the probability of long-range links is selected appropriately. You will consider the case that the underlying structure is a grid, and each node on the grid has a long range link. The endpoint of the link is selected with probability inversely proportional to the square of the distance from starting point of the link. Prove that in this case the expected number of steps of a path found between two nodes s, t by the greedy algorithm is $O(\log^2 n)$.