2012 SMALL WORLD

In this question we will take a look something that has attracted attention in recent years — *social networks*. In particular, we will consider these as being a group of people, some pairs of whom are friends with one another.

An interesting fact about most real-life social networks is the much publicized *small world hypothesis* or *six degrees of separation*. This states that you can usually get between any two people by following a short chain of friendships. This is mainly due to the existence of a comparatively small *core* of people, who have a very large number of friends and who are usually friends with each other. The short chains of friendship usually go through the core and back out again. The people who are not in the core are called the *periphery*.

Split Networks

Consider the situation where the network can be partitioned into a core C and a periphery P such that *every* pair of people in C are friends and *no* pair of people in P are friends. (A pair of people, one in C and one if P, might or might not be friends.) Such a network is called a *split network* and the partition into C and P is called a *splitting*.

Question 1

A social network consists of Alphonse, Beryl, Cassandra and Dave. The only pairs of friends are: Alphonse and Beryl, Beryl and Cassandra, and Cassandra and Dave. Is this a split network? Justify your answer.

Question 2

Give an example of a split network with more than one possible splitting.

Question 3

Suppose Edgar is friends with everyone in a split network. When is it possible to put him in P and when in C? How about if Edgar is friends with nobody?

Question 4

Suppose, in a split network, Fenchurch is friends with Geronimo but not with Hildrada. What can you state about C if Geronimo and Hildrada are friends? How about if they are not friends?

Given a network known to be a split network, consisting of n people, we may be interested in an algorithm to produce a valid splitting.

Question 5

Suppose Indigo is a person in the split network. By considering the cases where Indigo is friends with (i) everyone, (ii) no-one or (iii) has at least one friend and one non-friend, explain how to assign one person to either C or P.

Question 6

Using your answer to Question 5 (or otherwise) briefly outline a reasonably efficient algorithm for finding a valid splitting of a split network.

Question 7

Explain how you could use an algorithm that solves Question 6 to determine if the network is a split network.

General networks, badness and best cores

In practice most networks are not split networks. For these networks we need to change the definition of the core and periphery. There are several ways this might be done.

Given a possible partition of a network in to *C* and *P* (without necessarily the *every pair* and *no pair* conditions we had earlier) we can define the *badness* of the partition as:

badness(C) = number of pairs of friends in *P*
+ number of pairs of non-friends in *C*

since we expect people in *C* to be friends with other people in *C,* and people in *P* not to have friends in *P*. So, *badness(C) = 0* if and only if *C & P* are a splitting of a split network.

A *best core* is a partition of the network with the smallest possible badness. We are interested in finding an algorithm to determine a best core.

Question 8
Suppose (for this question only) our social network consists of 26 people Alphonse, Beryl, …, Ysobel and Zebedee. Alphonse and Beryl are friends with everyone apart from each other and there are no other friendships. What is the best core and what is its badness? Justify your answer.

Question 9
Suppose a possible core *C*, which contains *k* people, does not contain Oswald. If Oswald is moved from *P* to *C* by how much does the badness increase or decrease?

One use of social networks is to spread information. Suppose that the firm of *Widget, Whatsit & Doodah* (est. 1862) wish to promote an exciting new range of gadgets by targeting a small group of people and then letting the message spread from friend to friend.

Question 10
Widget, Whatsit & Doodah have decided to target exactly *k* people and have resolved to pick a *best core of size k*; i.e. a group of *k* people with smallest possible badness. Give a simple outline of which *k* people will be targeted which does not rely on the definition of badness, cores or peripheries.

Question 11
Outline a reasonably *efficient* algorithm for finding a best core in a social network. An *efficient* algorithm would rapidly solve the problem for a large number of people.

Question 12
Suppose that *C* is a best core of a social network with 10 people. What is the large possible value for *badness(C)*? Give an example network that gives rise to this value and indicate a possible best core.

Question 13
Suppose that *C* is a best core of a social network and subsequently given pairs of people either become or cease to be friends. Discuss how you might efficiently generate the new best core *C′*, given that you have access to any data used to generate *C*, and how the number of changing pairs affects your approach.