# Proteins and Graph Theory [50]
*Scientific Mathematics Round*

**Introduction.** Graph theory is the study of vertices and the edges connecting them. Though abstract and innocent at first glance, the subject underlies our analysis of systems as diverse as media networks, road layouts, astronomy, meeting schedules, map colorings, computer science, and archaeology (just to name a few). The problems given here provide a flavor of *bioinformatics*, a new subject that is applying theoretical graph theory to discover the secrets of biological systems.

**Modeling Time Relationships with Interval Graphs.** In a cell, proteins are large molecules that interact with one another to achieve cellular processes. While some proteins hang around indefinitely, others are actively produced when they are needed and actively destroyed when they are no longer of use.

Suppose scientists have just discovered 7 new proteins via genetic analysis that are of the actively produced/destroyed variety. They would like to know if these proteins interact while they are in the cell, and *a good indicator of whether they do is whether they exist in the cell during overlapping time intervals*. A bright scientist suggests using *interval graphs* to help model the time relationships of the proteins.

Let's build an example interval graph. First we draw 7 vertices, one for each protein:
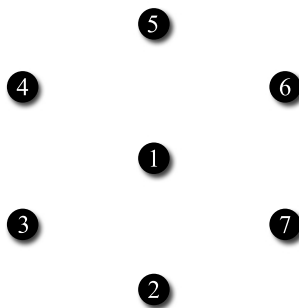


Next, we use experimental means to check if the proteins exist in the cell during overlapping time intervals; if they do, we connect them with an edge. This completes the interval graph.

For example, our experiments could yield the following interval graph:
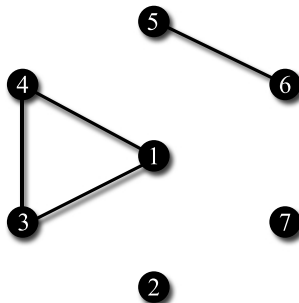


This interval graph implies that

- there is a time interval during which proteins 1 and 3 are both present,

- there is a time interval during which proteins 1 and 4 are both present,

- there is a time interval during which proteins 3 and 4 are both present,

- there is a time interval during which proteins 5 and 6 are both present, and

- for every pair of proteins not connected by an edge (2 and 7, 1 and 6, and so on), there is no time interval during which they are both present.

One protein-existence schedule we could hypothesize from the data that shows the time intervals of existence for the different proteins is
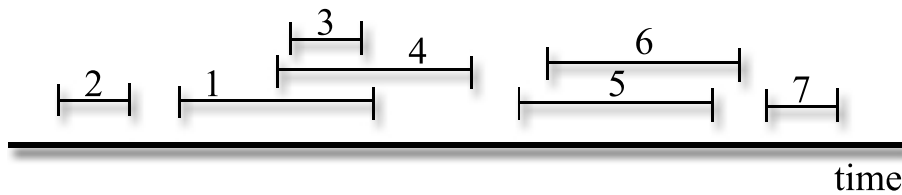


Make sure you understand how this is consistent with the interval graph and that you could conceivably come up with many more possible schedules of this sort that are still consistent with the interval graph that the experiments provide.

(a) [5] **Draw another protein-existence schedule that is consistent with the same interval graph above.**

Suppose that scientists are struggling to find the interval graph for our 7 proteins, and Dr. Sherbet suddenly declares what he believes to be the interval graph:
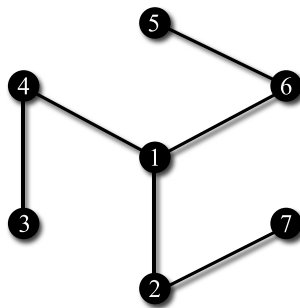


(b) [20] **Prove that Dr. Sherbet is wrong. You can assume that after a protein is made and destroyed, it is not remade by the cell in any reasonable amount of time.**

**Modeling Physical Interactions with Graphs.** As shown above, graph theory provides a great way to organize information about complex systems and make nontrivial conclusions. Another example is as follows.

2

Suppose our 7 proteins interact by passing "activity markers" to one another. One protein cannot have two markers at the same time, *so a protein can only pass a marker to a protein that doesn't have one at the time.* A biochemist gives you the following data:

- 1 can give a marker to 3 and 5.

- 2 can give a marker to 6 and 4.

- 3 can give a marker to 1 and 4.

- 4 can give a marker to 2 and 3.

- 5 can give a marker to 1 and 6.

- 6 can give a marker to 2 and 5.

- 7 can't give markers to any proteins.

(c) [25] **Say a cell is observed to be in a state where $1$ has marker $A$, $3$ has marker $B$, and $5$ has marker $C$. Based on some of his data, Dr. Sherbet hypothesizes that the proteins will keep passing markers around until $2$ has marker $A$, $4$ has marker $B$, and $6$ has marker $C$. Using the biochemist's data, prove that Dr. Sherbet is wrong.**

# Radioactive Decay [50]
*Scientific Mathematics Round*

**Introduction.** Atoms are the fundamental particles that comprise matter. We focus on the nucleus of an atom, a densely packed region which contains protons and neutrons. Some atoms are unstable because the numbers of their protons and neutrons are not energetically optimal. Unstable atoms undergo *radioactive decay,* a process in which protons and neutrons (and other parts of atoms that we won't mention here) are ejected out of the atom to help the atom reach a stable state.

At any given time, we define the *state* of an atom by the pair of integers $(A, Z)$. The *atomic weight* $A$ is the sum of the number of protons and neutrons the atom contains at that time, and the *atomic number* $Z$ is just the number of protons it contains.

(a) [2] **What is the number of neutrons in terms of A and Z?**

**Types of Radioactive Decay.** We consider three types of radioactive decay here:

- *Alpha decay,* in which the nucleus releases a cluster of 2 protons and 2 neutrons.

- *Proton emission,* in which the nucleus releases 1 proton.

- *Neutron emission,* in which the nucleus releases 1 neutron.

(b) [3] **If the nucleus is in state $(A, Z)$, what is its state after one alpha decay? What about after one proton emission? After one neutron emission?**

(c) [10] **Uranium $(238, 92)$ is known to undergo a chain of radioactive decay steps to reach a stable form of lead $(206, 82)$. Assuming all conceivable paths using the three above modes are allowed, find the number of decay steps *in the shortest chain.***

(d) [15] **Suppose the uranium atom was not allowed to undergo alpha decay and could perform only proton and neutron emission. How many different decay paths could uranium $(238, 92)$ take to reach lead $(206, 82)$? You do not have to compute any well-known mathematical quantities.**

**Poisson Model for Atomic Decay.** The time it takes for an unstable atom to undergo one decay is not fixed. Thus in some given time interval the number of decay steps that the nucleus undergoes varies. A good model for the probability $P(k)$ that we get $k$ decay steps in time $T$ is given by the *Poisson distribution:*

$$P(k) = \frac{e^{-\lambda T}(\lambda T)^k}{k!},$$

where $\lambda$ is the *decay rate constant* giving the number of decays we expect for that atom on average per unit time.

(e) [10] **Say we have an unstable atom with $\lambda = 0.1$ decays per day. What is the probability that after 10 days we see more than 2 decays? Assume that once the nucleus decays it can decay again at an undiminished rate.**

**Bulk Decay and the Half-Life.** We can analyze the decay rates for individual atoms as above, but practically if we have a bulk sample of some substance it helps us to know how long it takes for some portion of it to decay. We use the *exponential decay model* for this purpose:

$$N(t) = N_0 e^{-\lambda t},$$

where $N_0$ is the initial number of radioactive nuclei in our sample before decay starts, $N(t)$ is the number of radioactive nuclei left at time $t$ (assume for this problem that *nuclei only decay once and after they do they are no longer radioactive*), and $\lambda$ is the *decay constant* as described above.

(f) [10] **We define the *half-life* $T_{1/2}$ to be the time it takes for half of the radioactive nuclei to decay. Prove that:**

$$T_{1/2} = \frac{\ln 2}{\lambda}.$$