CSCI 0220 Discrete Structures and Probability B. Raphael

Homework 7

Due: Tuesday, 1 April 2008 at 10 am

Suggested Reading: 11.4, 11.5

Problem 7.1

Draw the following undirected, connected graphs, or if they cannot exist, explain why:

a. A graph with five vertices of degrees 1, 2, 3, 3, and 5.

b. A graph with four vertices of degrees 1, 1, 1, and 4.

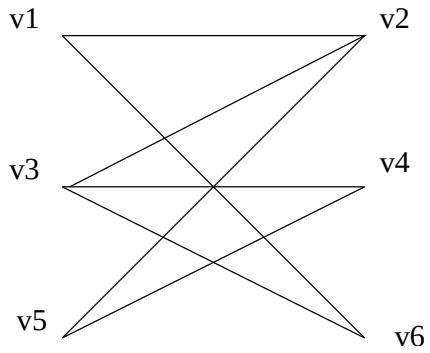
c. A simple graph with more than one vertex in which no two vertices have the same degree.

d. The graph K_6.

Problem 7.2

A *2-colored graph* G is a simple graph whose vertices can be assigned into two colors so that no edge connects two vertices of the same color.

For example, the graph below is 2-colored:



Prove that an undirected 2-colored graph cannot contain a cycle that has an odd number of vertices. A *cycle* is a closed walk.

Problem 7.3

Prove that if G is an undirected connected graph with an Eulerian circuit, and G' is isomorphic to G, then G' has an Eulerian circuit.

Note: You may not use other isomorphic invariants in your proof, unless you prove these invariants as well.

Problem 7.4

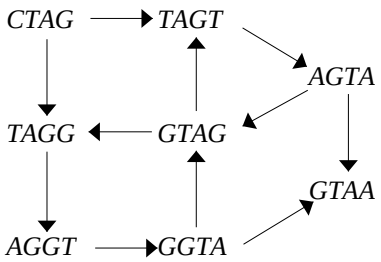
Current technology allows a DNA[1] sequence to be read in short, overlapping pieces called 'fragments'. These fragments must then be put together to reconstruct the original DNA string.

For example, let us say you are given following fragments of length 4:

AGGT, AGTA, CTAG, GGTA, GTAA, GTAG, TAGG, TAGT

One way to reconstruct the original DNA string is to find the shortest string that contains each of these fragments as substrings. This can be accomplished by identifying fragments that "overlap" such that the last three letters of one fragment (since the fragment length is 4) are equal to the first three letters of another fragment.

We can use a directed graph to represent these overlaps. Each vertex represents a single fragment, and each edge represents an overlap. For example, we would draw an arrow from vertex *AGGT* to *GGTA*. The full graph for this example is the following:



To construct the full sequence, we find a Hamiltonian path through the graph. In this case, one possible Hamiltonian gives an ordering of fragments

[1] DNA (Deoxyribose nucleic acid) is our genetic code. For the purposes of this problem, all you need to know is that it can be represented as a string over the alphabet *A, C, G,* and *T*.

that reconstructs the sequence *CTAGTAGGTAA*.

Given the following fragments of length 3, construct a directed graph representing the overlaps. Can this set of fragments reconstruct a full sequence, containing every fragment exactly once? If so, what is that sequence? Is this sequence reconstruction unique; that is, is there no other sequence that you can reconstruct from this set of fragments?

AAG, AGC, ATG, CAT, CGT, CTG, GCG, GCT, GTC, GTT, TCA, TGC, TGT