Introduction to Algorithms	**Problem Set 6**
CS 482 Spring 2000	**Due April 3, 2000**

(1) The *Hitting Set* problem can be described as follows. Given a set A of n elements, a collection $B_1, \ldots, B_m$ of subsets of U, and a number k, does there exist a set $H \subseteq A$ so that H has size at most k, and H contains at least one element from each of $B_1, \ldots, B_m$? (Such an H is a small set that "hits" every B_i.)

Prove that *Hitting Set* is NP-complete.

(2) As some people remember, and many have been told, the idea of hypertext predates the World Wide Web by decades. Even hypertext fiction is a relatively old idea — rather than being constrained by the linearity of the printed page, you can plot a story that consists of a collection of interlocked virtual "places" joined by virtual "passages."[1] So a piece of hypertext fiction is really riding on an underlying directed graph; to be concrete (though narrowing the full range of what the domain can do) we'll model this as follows.

Let's view the structure of a piece of hypertext fiction as a directed graph $G = (V, E)$. Each node $u \in V$ contains some text; when the reader is currently at u, they can choose to follow any edge out of u; and if they choose $e = (u, v)$, they arrive next at the node v. There is a start node $s \in V$ where the reader begins, and an end node $t \in V$; when the reader first reaches t, the story ends. Thus, any path from s to t is a valid *plot* of the story. Note that, unlike a Web browser, there is not necessarily a way to go back; once you've gone from u to v, you might not be able to ever return to u.

In this way, the hypertext structure defines a huge number of different plots on the same underlying content; and the relationships among all these possibilities can grow very intricate. Here's a type of problem one encounters when reasoning about a structure like this. Consider a piece of hypertext fiction built on a graph $G = (V, E)$ in which there are certain crucial *thematic elements* — love; death; war; an intense desire to major in computer science; and so forth. Each thematic element i is represented by a set $T_i \subseteq V$ consisting of the nodes in G at which this theme appears. Now, given a particular set of thematic elements, we may ask: is there a valid plot of the story in which each of these elements is encountered? More concretely, given a directed graph G, with start node s and end node t, and thematic elements represented by sets $T_1, T_2, \ldots, T_k$, the *Plot Fulfillment* problem asks: is there a path from s to t that contains at least one node from each of the sets T_i?

Prove that *Plot Fulfillment* is NP-complete.

(3) *(Adapted from a problem proposed by Maverick Woo.)* There are those who insist that the initial working title for Episode XXVII of the Star Wars series was "P = NP" — but this is surely apocryphal. In any case, if you're so inclined, it's easy to find NP-complete problems lurking just below the surface of the original Star Wars movies.

Consider the problem faced by Luke, Leia, and friends as they tried to make their way from the Death Star back to the hidden Rebel base. We can view the galaxy as an undirected graph $G = (V, E)$, where each node is a star system and an edge (u, v) indicates that one can travel directly from u to v. The Death Star is represented by a node s, the hidden Rebel base by a node t. Certain edges in this graph represent longer distances than others; thus, each edge e has an

[1] See e.g. http://www.eastgate.com

integer *length* $\ell_e \geq 0$. Also, certain edges represent routes that are more heavily patrolled by evil Imperial spacecraft; so each edge e also has an integer *risk* $r_e \geq 0$, indicating the expected amount of damage incurred from special-effects-intensive space battles if one traverses this edge.

It would be safest to travel through the outer rim of the galaxy, from one quiet upstate star system to another; but then one's ship would run out of fuel long before getting to its destination. Alternately, it would be quickest to plunge through the cosmopolitan core of the galaxy; but then there would be far too many Imperial spacecraft to deal with. In general, for any path P from s to t, we can define its *total length* to be the sum of the lengths of all its edges; and we can define its *total risk* to be the sum of the risks of all its edges.

So Luke, Leia, and company are looking at a complex type of shortest-path problem in this graph: they need to get from s to t along a path whose total length and total risk are *both* reasonably small. In concrete terms, we can phrase the *Galactic Shortest Path* problem as follows: given a set-up as above, and integer bounds L and R, is there a path from s to t whose total length is at most L, *and* whose total risk is at most R?

Prove that *Galactic Shortest Path* is NP-complete.