CS3100: Algorithms
Homework 7

1. Consider a set $A = \{a_1, \ldots, a_n\}$ and a collection $B_1, \ldots, B_m$ of subsets of A (so that each $B_i \subseteq A$).

   We say that H is a *hitting set* for the collection $B_1, \ldots, B_m$ if H contains at least one element from each B_i, so that $H \cap B_i$ is not empty for each i. In other words, H "hits" each of the sets B_i.

   We now ask the following: Given a set A and a collection of subsets B_i as described above, and a number k, is there a hitting set $H \subseteq A$ of size at most k? Prove that this problem is NP-Complete.

2. A store trying to analyze the behavior of its customers will often maintain a 2-dimensional array A, where the rows correspond to its customers and the columns corespond to the products it sells. The entry $A[i, j]$ specifies the quantity of product j that has been purchased by customer i.

   Here's a tiny example of such an array A:

	detergent	beer	diapers	cat litter
Alice	0	6	0	3
Bob	2	3	0	0
Eve	0	0	1	3

   One thing the store might want to do with this data is the following. We say that a subset S of the customers is *diverse* if no two of the customers have ever bought the same produce (i.e. for each product, at most one person in S has ever bought it). A diverse set of customers can be useful, for example, as a target pool for market research.

   We can now define the diverse subset problem as follows: Given an $m \times n$ array A defined as above, and a number $k \leq m$, is there a subset of at least k customers which is diverse?

   Show that the diverse subset problem is NP-Hard.

3. In my excitement over the upcoming release of Episode VIII, I have been hunting for algorithms problems while re-watching episodes I-VII. To my delight, I have found several!

   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 start system and each edge $\{u, v\}$ indicates you can travel between u and v. The Death Star is represented by a node s, the hidden Rebel base by a node t. Certain edges have longer distances than others; thus we will give each edge an integer length $l_e \geq 0$. Also, certain edges represent routes that are more heavily patrolled by evil Imperial spacecraft; so each edge e also gets an integer risk $r_e \leq 0$, indicating the expended amount of damage incurred from the special-effects-intensive space battles if you use this edge.

   There is a tradeoff here: it would be safest to travel through the outer rim of the galaxy, from one quiet far away star system to another, but then the ship would likely run out of fuel long before getting to its destination. (After all, they are on the run, so stopping to refuel should be avoided!) Alternatively, it would be fasted to dive 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 from s to t, we get both a length (the sum of all the lengths of its edges), and a total risk (the sum of all the risks of its edges).

   So Luke, Leia, and company are looking at a complex shortest path type problem in this graph: they want to get from s to t along a path whose total length and risk are *both* reasonably small. In concrete terms, we will phrase this as the *Galactic Shortest-Path Problem* as follows: Given a setup 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?

   Show that Galactic Shortest Paths is NP-Complete.

   (Hint: You might have an easier time with a reduction that takes a numeric problem and builds a graph out of it, as opposed to starting with a graph problem.)