CMPSCI 611 FALL '09: FINAL

Rules:

- Do not turn over the page until you are told to do so.
- Write your name at the top of the answer sheets.
- No communicating with other students, or looking at notes, or using electronic devices. You can ask the TA to clarify the meaning of a question but do so in a way that causes minimal disruption.
- If you finish early, you may leave early but do so as quietly as possible. Your exam script should be given to the TA.
- There are five questions. All carry the same number of marks but some questions are easier than others. Don't spend too long on a problem if you're stuck – you may find there are other easier questions.
- Answers should be written on the answer sheets on the appropriate page. There are also 3 blank pages at the end that can be used. If you are using these pages, clearly indicate which question you're answering. Further paper can be requested from the TA if required.
- Take a deep breath...

Question 1. This problem concerns finding independent sets. Recall that an independent set in an undirected graph $G = (V, E)$ is a subset $V' \subset V$ such that for each $(u, v) \in E$, at least one of u or v are not contained in V'.

(1) Suppose the degree of each node in G is at most d. Design an efficient approximation algorithm for finding the largest independent set in G. Your algorithm should have an approximation ratio of $(d + 1)$ and you should prove this.
(2) Suppose G is a tree but the degree of each node can be arbitrary. Design an efficient exact algorithm to find the maximum independent set.

Question 2. The CLIQUE COVER problem is, given a graph $G = (V, E)$ and an integer k, is there a partition of V into k disjoint sets $V_1, \ldots, V_k$ such that for each i, for every pair of vertices $u, v \in V_i$, there exists an edge $(u, v) \in E$. Prove that CLIQUE COVER is NP-complete. [Hint: Consider a reduction from graph coloring.]

Question 3. You are given a string of n characters $s[1 \ldots n]$, which you believe to be a corrupted text document in which all punctuation has vanished (so it looks something like "myfavoritealgorithmsarerandomized…"). You wish to reconstruct the document using a dictionary, which is available in the form of a Boolean function:

$$\texttt{dict}(w) = \begin{cases} \texttt{true} & \text{if } w \text{ is a valid word} \\ \texttt{false} & \text{otherwise} \end{cases}$$

Design a dynamic programming algorithm that determines whether the string $s[1 \ldots n]$ can be reconstituted as a sequence of valid words. The running time should be at most $O(n^2)$ assuming calls to `dict` take unit time. Note that some words are prefixes of others, e.g., "the" and "there".

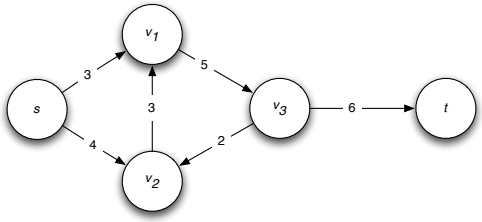
Question 4. 1) Draw the feasible region, find the optimal solution, and write the dual of the LP:

$$\max \; x_1 + x_2$$

subject to

$$\begin{aligned} x_1/2 + x_2 &\leq 2 \\ x_1 &\leq 2 \\ x_1, x_2 &\geq 0 \end{aligned}$$

2) Consider the following network flow problem on 5 nodes where the edges are directed and the numbers are the capacity of the edges.



Write a linear program that would maximize the flow from s to t.

Question 5. Let ϕ be a 3-SAT formula with m clauses and n variables. You set each variable x_i to TRUE or FALSE at random and each x_i is determined independently.

(1) What is the expected number of clauses that are satisfied? Include a proof.
(2) Prove that the probability that less than 5/8 of the clauses are satisfied is at most 1/3.
(3) Suppose $n = 3m$ and each variable appears exactly once in the formula. What's the probability that all the clauses are satisfied?