Exercises for the *Approximation Algorithms* course given by Prof. Maurice Queyranne at the Facultad de Ingeniería, Universidad de la República, Montevideo (August 3–7, 2009)

Do **<u>six</u>** (6) of the twelve exercises below, at your choice.
You must work on these exercise on your own, without any external help.
Please contact Maurice.Queyranne@sauder.ubc.ca if you have any questions.

Please email your work, in *pdf* format, to this address by **<u>Monday September 14, 2009</u>**.

1. (Adapted from Exercise 3.2 in [V][1]) Consider a graph $G = (V, E)$ with positive edge costs, and two disjoint subsets of V : S, the *senders*, and R, the *receivers*. The problem is to find a minimum cost subgraph $G' = (V, E')$ of G such that every receiver is connected by a path in G' to at least one of the senders. Distinguish two cases: (i) $E \cup D = V$; and (ii) $E \cup D \neq V$. Prove that case (i) is in **P** (i.e., solvable in polynomial time), and case (ii) is **NP**-hard. Give a 2-approximation for case (ii). (*Hint*: Add a new vertex connected to every sender by a zero-cost edge. The new vertex and all receivers are the required vertices; all other vertices are Steiner vertices. Find a minimum cost Steiner tree.)

2. Consider a graph $G = (V, E)$ with positive edge costs. For every edge subset $S \subseteq E$ let $f(S)$ denote the maximum cardinality of a subset $F \subseteq S$ such that the subgraph (V, F) does not contain a cycle. Show that the set function $f : 2^E \rightarrow \Re$ is nondecreasing and submodular. Characterize the possible values of the marginal costs $f_S(j) = f(S \cup \{j\}) - f(S)$. Show that the Submodular Set Cover (SSC) problem associated this set function f reduces to a well-known combinatorial optimization problem. (*Hint*: you may want to distinguish whether G is connected or not.) Show that the first Greedy Algorithm for SSC studied in class reduces to a well-known combinatorial algorithm. What does the approximability theorem for the Greedy Algorithm imply for this well-known combinatorial algorithm?

3. (Adapted from Exercise 14.2 in [V]) Let C denote the collection of given subsets selected by the Randomized Rounding algorithm for Weighted Set Cover. Find a constant $\beta > 0$ such that, with probability at least β, C covers at least half of the elements of the ground set and has cost at most $4 \cdot \text{OPT}$.

4. (Adapted from Exercise 16.2 in [V]) Show that the following algorithm is a 2-approximation for MAX-SAT : Let X be a random instantiation of the variables, as produced by the Large Clauses Randomized algorithm; let X' be the inverse instantiation (where each X'_i is True if and only if X_i is False); compute the total weight of all clauses satisfied by X, and then that of all clauses satisfied by X'; return the better of these two instantiations.

---

[1] [V] refers to the textbook: Vijay V. Vazirani, *Approximation Algorithms*, Springer, 2001. (Corr. 2nd printing 2003: ISBN : 3-540-65367-8.)

5. Let S be a set of vertices obtained after applying the Modified Local Search (MLS) algorithm, with parameter $\varepsilon > 0$, to a MAX-CUT instance. How good is this solution relative to the optimum objective value (OPT) for this MAX-CUT instance? How should one choose ε for MLS to be a $(2+\delta)$-approximation for MAX-CUT, for a given $\delta > 0$?

6. Consider the Submodular Function Maximization (SFMAX) problem $\max\{ f(S) : S \subseteq N \}$ where N is a given finite set and $f : 2^E \rightarrow \Re$ a submodular set function. We may assume, w.l.o.g.[2], that $f(\emptyset) = 0$ (but <u>not</u> that f is nonincreasing or nondecreasing, for the problem would then be trivial). Show that this problem is **NP**-hard. Let S be a local maximum for the neighborhood structure defined (as for MAX-CUT) by moving one element, from S to $V \setminus S$ or in the reverse direction. Prove: (i) if $R \subseteq S$ then $f(R) \leq f(S)$; and (ii) if $S \subseteq T$ then $f(T) \leq f(S)$. Let $S' \in \operatorname{argmax}\{ f(S), f(V \setminus S) \}$ denote the better of S and its complement. Show that S' is a 3-approximate solution to SFMAX. Show that it is 2-approximate if f also satisfies the following property: $f(T) = f(V \setminus T)$ for all $T \subseteq N$.

7. (Adapted from Exercise 26.12 in [V]) Consider the constrained MAX-CUT problem where certain given pairs of vertices should be separated by the cut, and other pairs must be on the same side of the cut. Formally, two (possibly empty) subsets of pairs of vertices are given: each pair in P_1 must be separated by the cut, and each pair in P_2 must be on a same side of the cut. Assume that these constraints are consistent. We seek a cut with maximum weight subject to satisfying all these constraints. Define an SDP relaxation for this problem and show that the Goemans & Williamson algorithm (separation by a random hyperplane) produces a cut satisfying the additional constraints and with expected cost $E[W] \geq \alpha \cdot OPT_{SDP}$ (where $\alpha \approx 0.87856$ is the same constant as in the unconstrained case, and OPT_{SDP} is the optimum value of <u>your</u> SDP relaxation). (*Hint*: any polynomial-sized SDP relaxation will suffice; you are <u>not</u> being asked to find a smallest or most effective such SDP formulation.) Explain how to use this result to define a de-randomized version of the Goemans-Williamson original (i.e., unconstrained) algorithm. How many SDP problems need to be solved in this de-randomized algorithm?

8. (Adapted from Exercise 8.3 in [V]) Produce an FPTAS for the following SUBSET-SUMS RATIO problem.[3] Given n positive integers, $0 < a_1 < \ldots < a_n$, find two disjoint nonempty subsets $S, T \subseteq N = \{1,\ldots,n\}$ such that $\sum_{i\in S} a_i \geq \sum_{i \in T} a_i$ and such that the ratio $\frac{\sum_{i\in S} a_i}{\sum_{i\in T} a_i}$ is minimized. (*Hint*: First produce a pseudo-polynomial algorithm for this problem. Then suitably round the data.)

---

[2] <u>w</u>ithout <u>l</u>oss <u>o</u>f <u>g</u>enerality

[3] C. Bazgan, M. Santha, and Z. Tuza. Efficient approximation algorithms for the subset-sum problem. In *Proc. 25th International Colloquium on Automata, Languages, and Programming*, volume 1443 of *Lecture Notes in Computer Science*, pages 387–396. Springer-Verlag, Berlin, 1998.

9. (Adapted from Exercise 5.1 in [V]) Show that if the edge lengths do not satisfy the triangle inequality, then the k-CENTER problem cannot be approximated within a factor $\alpha(n)$ for any function $\alpha(n)$, computable in polynomial time, of the number n of vertices of the k-CENTER instance. (Hint: combine the ideas of the inapproximability theorems for TSP and Metric k-CENTER.) On the other hand, show that for any fixed k the k-CENTER problem is in **P** (i.e., can be solved to exact optimality in polynomial time).

10. (Adapted from Exercise 29.1 in [V]) Prove that **PCP**(log n, 1) $\subseteq$ **NP**. (*Hint*: Let language $L \in$ **PCP**(log n, 1). The **NP** Turing machine accepting L guesses the proof y, simulates the verifier V for L on all possible strings of O(log n) bits, and accepts if and only if the verifier accepts on all these strings.)

11. (Adapted from Exercise 29.3 in [V]) Prove that if there exists a gap-introducing reduction from SAT to MAX-3SAT then **NP** $\subseteq$ **PCP**(log n, 1). A *gap-introducing reduction* from SAT to MAX-3SAT associates with every instance (logical expression) ϕ of SAT an instance I of MAX-3SAT such that (i) if ϕ is satisfiable then OPT(I) $\leq f(I)$ and (ii) ϕ is not satisfiable then OPT(I) > $\alpha(|I|)\cdot f(I) > 0$, where f is a real-valued function of the instance and α a real-valued function of the instance size satisfying $f(I) > 0$ and $\alpha(|I|) \geq 1$ for all instances I of MAX-3SAT. (*Hint*: Reduce every SAT instance ϕ to a MAX-3SAT instance I. Input to the verifier V an instantiation X which is optimum for I. The error probability is then at most $1 - \varepsilon_M$ for a certain positive constant ε_M . Repeat a polynomial number of times to reduce this error probability below 1/2.)

12. Let $G = (V,E)$ be a graph and $G^C = (V, E^C)$ be its *complement*, where $E^C = \{ \{i,j\} \subseteq V : \{i,j\} \notin E\}$ is the set of all edges of the complete graph over V which are not in E. Let $\alpha(G)$ denote the maximum cardinality of a stable set in G (i.e., of a subset of vertices no two of which are adjacent in G); $\gamma(G)$ the maximum cardinality of a vertex cover in G; and $\omega(G)$ the maximum cardinality of a clique in G (i.e., of a subset of vertices all of which are pairwise adjacent in G). Prove two very simple equations relating the values of $\alpha(G)$ and $\gamma(G)$ to $\omega(G^C)$. Assuming that **P** $\neq$ **NP**, what are the implications, if any, on the approximability of the STABLE SET problem (find a maximum cardinality of a stable set in a given graph G) and on the VERTEX COVER problem, of the fact that there cannot exist an $n^{1-\varepsilon}$ -approximation for the CLIQUE problem for any $\varepsilon > 0$? Explain.