1 MAX-SAT [recommended]

Consider a conjunctive normal form (CNF) formula on Boolean variables $x_1, x_2, \ldots, x_n$, that is, a formula defined as AND of a number $m$ of clauses, each of which is the OR of some literals appearing either positively as $x_i$ or negated as $\overline{x}_i$. Suppose that each clause $c_j$ has some weight $w_j$. The goal is to find an assignment of TRUE/FALSE values to the literals with the objective of maximizing the total weight of the satisfied clauses.

(A) Consider a clause that has $k \geq 1$ literals. Argue that the simple randomized algorithm that sets each variable at random (true or false, each with probability half) satisfied this clause with probability $(1 - 2^{-k})$. Argue that this gives a randomized algorithm that outputs a solution that is in expectation $1/2$-approximation for our problem. Then turn this algorithm into one that gives $0.49$-approximation with probability at least 99%.

(B) Consider the linear program with objective

$$\max \sum_{j=1}^{m} w_j z_j$$

subject to $\forall j \in \{1, 2, \ldots, m\}$:

$$\sum_{i \in S_j^+} y_i + \sum_{i \in S_j^-} (1 - y_i) \geq z_j$$

$\forall j \in \{1, 2, \ldots, m\}$: $z_j \in [0, 1]$

$\forall i \in \{1, 2, \ldots, n\}$: $y_i \in [0, 1]$

Here, $S_j^+$ denotes the set of variables that appear in the $j^{th}$ clause positively, and $S_j^-$ denotes the set of variables that appear in the $j^{th}$ clause in a negated form. Explain how this is a relaxation of an integer linear program for our objective. Moreover, let $(y^*, z^*)$ denote the optimal solution of this LP. Show that the natural randomized rounding algorithm that sets each $x_i$ to be True with probability $y_i^*$ provides the following guarantee: if the $j^{th}$ clause has $k$ literals, it is satisfied with probability at least $(1 - (1 - \frac{1}{k})^k)z_j^*$.

As in the previous exercise, argue that this gives an algorithm that in expectation provides $1 - 1/e$-approximation for our problem.

(C) Notice that the first algorithm handles well large clauses and the second algorithm handles well the smaller clauses. Put the two together to get a $3/4$ approximation algorithm for the MAX-SAT problem.

(D) Find a CNF such that there is a $3/4$ gap between the value of the solution of LP described in part (B) and the optimal Boolean assignment to the variables. Hint: find a CNF with 4 clauses, each of weight 1, such that the LP has value 4 but any assignment satisfies at most 3 clauses. This implies that the $3/4$ factor is the integrability gap of this LP formulation and no rounding technique for it will give an approximation better than $3/4$. 

1
(E) (optional) In part (B), we considered a linear randomized rounding process. Consider a non-linear rounding which rounds variable $x_i$ to be true with probability $f(y_i^*)$ where $f : [0, 1] \to [0, 1]$ is an arbitrary function such that $f(y) \in [1 - 4^{-y}, 4^{y-1}]$. Prove that this non-linear rounding directly gives a $3/4$ approximation algorithm.