Algorithms and Complexity. Exercise session 8

Approximation Algorithms

Approximation of Independent Set Let INDEPENDENT SET$-B$ be the problem of finding a maximum number of independent vertices in a graph whose degree (for each vertex) is limited by the constant B. Show that this problem is in APX, ie. it can be approximated with a constant in polynomial time.

Solution to Approximation of Independent Set
Given a graph $G = (V, E)$ whose degrees are bounded by B, construct an independent set of vertices as follows.

$V' \leftarrow \emptyset$
$W \leftarrow V$
for $v \in W$ **do**
  $V' \leftarrow V' \cup \{v\}$
  $W \leftarrow W - \{w \in W : (w, v) \in E\} - \{v\}$

The algorithm yields an independent set of vertices V'. It is also easy to see that V' is a dominating set. The optimal independent set V'_{opt} can not be more than B times larger than $|V'|$. To see this, you only need to look at one of the vertices of V' and its neighbors in V. If all its neighbors are in V'_{opt}, they can't be more than B. Since V' is a dominating set, every vertex of V'_{opt} is either in V' or a neighbor to a vertex of V'. This means that V'_{opt} is at most B times larger than V', namely the approximation factor is B, which is a constant. □

Probabilistic not-all-equal-satisfying The NP-hard problem MAX NOT-ALL-EQUAL 3-CNF SAT is defined as follows.

> INPUT: A CNF-formula consisting of clauses $c_1, c_2, \dots, c_m$ where each clause is a disjunction of exactly three literals (variables or negated variables). The variables are $x_1, x_2, \dots, x_n$.
> SOLUTION: A variable assignment.
> OBJECTIVE FUNCTION: The number of clauses that contain at least one true literal and at least one false literal.
> PROBLEM: Maximize the objective function.

Since this problem is NP-hard, we want an algorithm that approximates the input within a constant in polynomial time. Constructing a probabilistic approximation algorithms for the problem to give an expected approximation factor of 4/3. Analyze your algorithm time complexity and expected approximation guarantee. You will need a randomized algorithm.

Solution to Probabilistic not-all-equal-satisfying
The algorithm is quite simple: set each variable to a random value. It should be equally likely that a variable is set to true or false. If we now look at an arbitrary clause, there are only two cases out of eight (ie false-false-false-and true-true-true) that will not be counted. The expected value of the number of clauses that counts is

E[number of clauses that have both true and false literals] =
$= m \cdot Pr$[an arbitrary clause is neither completely true nor completely false] =
$= m \cdot (1 - Pr$[an arbitrary clause is either completely true or completely false]) =
$= m \cdot (1 - 2/8) = 3m/4$.

Since at most m clauses can be chosen, the expected approximation factor is

$$\frac{OPT}{APPROX} \geq \frac{m}{3m/4} = \frac{4}{3}.$$

The algorithm runs in linear time and requires a linear number of random variables. □

Approximation of linear inequalities MAX SAT LR$^{\geq}$ (maximum satisfiable linear subsystem) is the problem that, given a set of linear inequalities of type $\geq$, find a variable assignment that satisfies as many inequalities as possible. Construct an approximation algorithms that approximates MAX SAT LR$^{\geq}$ in factor 2.

Solution to Approximation of linear inequalities
Assume that the input is given as a set X of rational variables and a set of E of linear inequalities over X.

while $E \neq \emptyset$ **do**
  if (there are inequalities in E with a single variable) **then**
    $U \leftarrow \{x \in X : x$ is the only variable in at least one inequality in $E\}$
    Choose arbitrarily $y \in U$.
    $F(y) \leftarrow \{e \in E : e$ contains only variable $y\}$
    Give y a value that satisfies as many inequalities in $F(y)$ as possible.
    $E \leftarrow E - F(y)$
  else
    Choose arbitrarily $y \in X$.
    $y \leftarrow 0$
  Evaluate all inequalities in E which contain y.
  $X \leftarrow X - \{y\}$

The algorithm always assigns to y a value which satisfies at least half of the inequalities in $F(y)$. Thus, it builds a solution that satisfies at least half of the inequalities in input. Therefore, the maximum number of inequalities that can be satisfied gives an approximation factor of 2. The time complexity is polynomial, since each variable and term is processed only once. □

Upper bound for approximation of homogeneous bipolar inequalities
MAX HOM BIPOLAR SAT LR$^{\geq}$ (maximum homogeneous bipolar satisfiable linear subsystem) is the same problem as MAX SAT LR$^{\geq}$ but the variables may only assume the values 1 and -1, and all inequalities are homogeneous, ie no constant terms.

Show that MAX HOM BIPOLAR SAT LR$^{\geq}$ can be approximated with a factor 2 and that MAX HOM BIPOLAR SAT LR$^{>}$ can be approximated with a factor 4.

Solution to Upper bound for approximation of homogeneous bipolar inequalities
We start with an approximation of MAX HOM BIPOLAR SAT LR$^{\geq}$. Take an arbitrary bipolar vector $\mathbf{x}$ and look for numbers satisfying the inequalities on the variables set to $\mathbf{x}$ and $-\mathbf{x}$. If the left hand side of an inequality is positive for $\mathbf{x}$, it is negative for $-\mathbf{x}$ and vice versa. Therefore, one of these two vectors will satisfy at least half of the inequalities, that is, the approximation factor is 2.

This trivial algorithm does not work for MAX HOM BIPOLAR SAT LR$^{>}$since for many relations it can be zero for both vectors. Therefore, we start to look for a solution with many non-zero relations. The approximation algorithm for MAX SAT LR$^{\geq}$ above can be modified so that it finds a solution $\mathbf{x}$ for which at least half of the relations are non-zero. The same is true also for $-\mathbf{x}$, and one of these two vectors must then satisfy at least a quarter of all relations. □

Lower bound for the approximation of binary inequalities MAX BINARY SAT LR$^{\geq}$ (maximum binary satisfiable linear subsystem) is the same problem as MAX SAT LR$^{\geq}$ where the variables may only assume the values 0 and 1.

Show that MAX BINARY SAT LR$^{\geq}\notin$APX.

Solution to Lower bound for the approximation of binary inequalities

We show this by reducing MAX CLIQUE to MAX BINARY SAT LR$^{\geq}$ with an approximation preserving reduction. Since we know that MAX CLIQUE is not in APX it will mean that MAX BINARY SAT LR$^{\geq}\notin$APX.

Let $G = (V, E)$ be the input to MAX CLIQUE. For each vertex $v_i \in V$, we assign a variable x_i and the inequality

$$x_i - \sum_{j \in N(v_i)} x_j \geq 1$$

where j is in $N(v_i)$ if $v_j \neq v_i$ and $(v_i, v_j) \notin E$ (ie. v_j is not a neighbor of v_i). We have a system of $|V|$ inequalities. Note that the i-th inequality is satisfied if and only if $x_i = 1$ and $x_j = 0$ for all $j \in N(v_i)$.

It is easy to verify that if one has a s-click $V' \subseteq V$, he can obtain a binary solution that satisfies the s corresponding inequalities by setting $x_i = 1$ if $v_i \in V'$ and $x_i = 0$ otherwise. On the other hand, if one gets a binary solution $\mathbf{x}$ that satisfies s inequalities, he can get an s-click bu putting in V' all the vertices that satisfy the corresponding inequalities.

This reduction is not only *approximation preserving* but also *cost preserving* as it preserves the objective function value as well.

Note that this reduction works also for $A\mathbf{x} > \mathbf{0}$ and $A\mathbf{x} = \mathbf{1}$, so we can show that MAX BINARY SAT LR$^{>}\notin$APX and MAX BINARY SAT LR$^{=}\notin$APX. □
