Thomas Rothvoß
Location: ELD120 Discussion: 19.05.10

Exercises

Approximation Algorithms

Spring 2010

Sheet 11

Exercise 1
Here, we want to consider NON-METRIC FACILITY LOCATION, where facilities F with open cost f_i for $i \in F$, cities C and connection cost c_{ij} are given. Differently from the variant, studied in the lecture, we do not assume anymore, that c is metric.

i) Model the problem as SET COVER problem and obtain a polynomial time $O(\log n)$-approximation ($n := |C|$) by using the greedy algorithm for SET COVER.
Hint: Even if the defined set system has exponentially many sets, under some conditions the greedy algorithm can still be made to run in polynomial time.

ii) A result of Raz and Safra (1997) says the following:

> There is a constant $c > 0$ such that, given a SET COVER instance $S_1, \ldots, S_m$ and a parameter $k \in \mathbb{N}$ it is **NP**-hard to distinguish
>
> - YES: $OPT_{\text{SETCOVER}} \leq k$
> - NO: $OPT_{\text{SETCOVER}} \geq k \cdot c \cdot \log n$
>
> Here OPT_{SETCOVER} denotes the smallest number of sets that are needed to cover all n elements.
> **Remark:** This result means that there is a polynomial time reduction, taking a SAT clause $\mathscr{C}$ as instance and mapping it to a SET COVER instances $\mathscr{S} = \{S_1, \ldots, S_m\}$ such that: If $\mathscr{C}$ is satisfiable, then $OPT_{\text{SETCOVER}}(\mathscr{S}) \leq k$ and $OPT_{\text{SETCOVER}}(\mathscr{S}) \geq k \cdot c \cdot \log n$ otherwise (for more details on gap reductions, I recommend Chapter 29 of Vazirani's book *Approximation Algorithms*).

Show that it is also **NP**-hard to approximate NON-METRIC FACILITY LOCATION by a factor better than $c \cdot \log n$.

Exercise 2
We consider the FACILITY LOCATION problem, with given facilities F, cities C, opening cost f_i for every facility i. Assume that the cost function c_{ij} is *metric*. In this exercise, we want to show that there is no 1.46-approximation algorithm for the (metric) FACILITY LOCATION problem.

For the sake of contradiction, suppose that we have a polynomial time algorithm $\texttt{algo}(F, C, c_{ij}, f_i)$ that produces a 1.46-approximate solution $F' \subseteq F$ (note that knowing the set of open facility suffices — the cities are then automatically connected to the nearest such facility).

Let $S_1, \ldots, S_m$ be a SET COVER instance (with unit cost per set) on elements $\{1, \ldots, n\}$. We may assume to know the value k of sets that are contained in an optimum solution. We will now show, how to obtain a $0.999 \cdot \ln(n) + O(1)$ approximate SET COVER solution in polynomial time. This would then contradict an inapproximability result of Feige (1998) (given that **NP** is not contained in $\mathbf{DTIME}(n^{O(\log \log n)})$).

We use the following SET COVER algorithm:

(1) Let $C := \{1, \ldots, n\}$, $F := \{1, ..., m\}$ and $c_{ij} := \begin{cases} 1 & j \in S_i \\ 3 & \text{otherwise} \end{cases}$

(2) WHILE $C \neq \emptyset$ DO

  (3) Let $f_i := 0.46 \cdot \frac{|C|}{k}$ be the facility cost $\forall i \in F$

  (4) $F' := \texttt{algo}(F, C, c_{ij}, f_i)$

  (5) Buy the sets in F'

  (6) C' := cities covered at cost 1; set $C := C \backslash C'$

(7) Return the bought sets

Perform the following analysis:

i) Consider any iteration and let APX be the cost of the FACILITY LOCATION solution F'. Show that $APX \leq 1.46^2 \cdot |C|$.

ii) Suppose the algorithm needs T iterations. For iteration $t \in \{1, \ldots, T\}$, define β_t and α_t such that $|F'| = \beta_t k$ is the number of opened facilities and $|C'| = \alpha_t |C|$ is the number of elements that are covered in this iteration. Show that $\beta_t \leq \ln(\frac{1}{1-\alpha_t})$ holds for any $t < T$.
Hint: It is OK if your solution contains the phrase "*By a Maple/Matlab plot we see that..*".

iii) Why is $\prod_{t=1}^{T-1}(1 - \alpha_t) \geq \frac{1}{n}$?

iv) Show that the algorithm needs at most $0.999 \cdot \ln(n) \cdot k$ many sets (plus $O(k)$ for the last iteration).