Fair and Efficient Social Decision Making (CSCI 699)
Assignment 2
Due Date: November 22th, 11am

Notes

- Citation policy
  - It is certainly preferable for you to solve the questions without consulting a peer or a published source. However, if you do consult and obtain useful insights, you must cite the name of the peer or the link of the source you referred.
  - Further, you should write the solution in your own words. The best way to do so is to not take any notes during discussions, spend at least a few hours doing something else, and then construct the solution on your own. to a subproblem).
- No garbage policy: Stating 'I do not know how to approach this' gets you 20% of the points (this also applies
- Typed assignments are highly preferred (LaTeX or Word), especially if your handwriting is possibly illegible or if you do not have access to a good quality scanner. Please submit a single PDF at Brightspace.

Question 1 [35 Points] Facility Location

Recall the facility location setting from the class. There are n agents. Each agent i holds private a point $y_i \in \mathbb{R}$ on the real line, but may report a potentially different point $\bar{y}_i$ when asked. A mechanism $f : \mathbb{R}^n \to \mathbb{R}$ takes as input the reported points $(\bar{y}_1, \dots, \bar{y}_n)$ and outputs $a \in \mathbb{R}$. The cost to agent i is $c_i(a) = |a - y_i|$ and the social cost is $sc(a) = \sum_{i=1}^{n} c_i(a)$. In class, we saw that the median rule is both strategyproof and 1-approximation algorithm for social cost minimization.

Let us consider a slightly different setting. Fix $k \in \mathbb{N}$. Suppose each agent i now holds private a vector of k points $\overrightarrow{y_i} = (y_{i,1}, \dots, y_{i,k}) \in \mathbb{R}^k$ but may report $\overrightarrow{\bar{y}_i} = (\bar{y}_{i,1}, \dots, \bar{y}_{i,k}) \in \mathbb{R}^k$. A mechanism $f : \mathbb{R}^{n \times k} \to \mathbb{R}$ takes as input the reported points $(\overrightarrow{\bar{y}_1}, \dots, \overrightarrow{\bar{y}_n})$ and outputs $a \in \mathbb{R}$. The cost to agent i is now $c_i(a) = \sum_{j=1}^{k} |a - y_{i,j}|$ and the social cost is $sc(a) = \sum_{i=1}^{n} c_i(a)$.

Consider the following mechanism f:

1. For each $i = 1, \dots, n$, let $med(\overrightarrow{\bar{y}_i})$ be the median of $(\bar{y}_{i,1}, \dots, \bar{y}_{i,k})$.
2. Return the median of $\left(med(\overrightarrow{\bar{y}_1}), \dots, med(\overrightarrow{\bar{y}_n})\right)$

That is, the mechanism first takes the median of the reported points of each agent, and then returns the median of these medians. For simplicity, assume that both n and k are odd, so the median is uniquely defined at each step.

(a) [20 Points] Prove that f is strategyproof.

[Hint: Instead of proving this from scratch, see if you can cleverly use the fact that the median rule is strategyproof in the original setting where each agent holds a single private point.]

(b) [15 Points] Prove that f is a 3-approximation algorithm for minimizing the social cost.

[Hint: Let a be the location returned by f and a^* be the optimal location. Without loss of generality, assume $a < a^*$ and denote $d = |a - a^*|$. First, show that out of all nk points of the agents, at least $nk/4$ of them must be at most a. Then, write $sc(a) = \sum_{i,j} |a - y_{i,j}|$ and $sc(a^*) = \sum_{i,j} |a^* - y_{i,j}|$, and break each summation over three sets: $(i, j) : y_{i,j} \leq a, (i, j) : a < y_{i,j} \leq a^*$, and $(i, j) : y_{i,j} > a^*$.]

Question 2 [30 Points] Envy-Cycle Elimination For Bads

Recall that Envy-Cycle Elimination Algorithm for goods preserves an envy graph where each agent is a node and for every agent i that envies an agent j, there exists a directed edge (i, j). In each round, the algorithm either allocates an arbitrary unallocated good to a node with no input edges or resolves an envy cycle by cyclic exchange of bundles. In class, we show that the "natural" modification of this algorithm for bad does not work, i.e. either allocating an arbitrary unallocated bad to a node with no outgoing edges or resolving an envy cycle by cyclic exchange of bundles. Consider the following modification:

- At the envy graph, each agent has an outgoing edge only to the agent whom she envies and whose bundle has maximum utility, i.e even if an agent envies multiple other agents, their node has at most one outgoing edge
- While there are unassigned bads
  - If there is an agent with no outgoing edges, allocate an arbitrary unallocated bad to this agent
  - Else resolve an envy cycle by cyclic exchange of bundles

Note that the only modification is with respect to the envy graph. Show that this modification of Envy-Cycle Elimination Algorithm terminates with an EF1 allocation.

[Hint: Show that either by allocating a new bad or resolving a cycle, the EF1 property remains true for the partial allocation.]

Question 3 [35 Points] Core in Non-Centroid Clustering

Recall the following definition of the α-core with respect to the maximum loss function at a non-centroid clustering setting:

- A non-centroid k-clustering $(C_1, \ldots, C_k)$ is in the α-core, with $\alpha > 1$, if there is no group of agents $S \subseteq N$ with $|S| \geq n/k$ such that $\alpha \cdot max_{i' \in S} d(i, i') < max_{i' \in C(i)} d(i, i')$ for all $i \in S$, where $C(i)$ denotes the cluster that i is assigned to.

Consider the following modification of Greedy Capture that we discussed in class:

- Grow balls around the points that have not been disregarded, expanding at the same rate
- When a ball captures at least n/k points, assign them to a cluster and disregard them
- Continue until fewer than n/k points remain. Assign the remaining points to a single cluste.

(a) [20 points] Show that this modification of Greedy Capture returns a clustering solution that is in the 2-core with respect to the maximum loss function.

[Hint: Assume for contradiction that this is not true for some $S \subseteq N$ with $|S| \geq n/k$. Consider the first agent in S that was assigned to some cluster and do similar arguments like the ones we did in class.]

(b) [15 points] Show that the above analysis is tight by showing that there exists an instance at which this modification of Greedy Capture does not provide an approximation of $2 - \epsilon$ for arbitrary small ϵ.

[Hint: Consider an instance on the line with $n = 6$ and $k = 2$, such that the first point is located at -10, the second and the third points are located at -1, the fourth point is located at 0, the fifth is located at $1 - \epsilon$ and the last one is located at $2 - 2\epsilon$, for arbitrary small ϵ.]