Theory of Mechanism Design - Assignment 2

1. We consider the house allocation model with existing tenants. Consider the following form of manipulation by a coalition of agents in the TTC mechanism. A coalition of agents S exchange their houses before the start of the mechanism (i.e., they end up with an endowment which is different from their actual endowment). Now, the TTC mechanism is executed. Do you think each agent in S will now get a house which is either the same house he gets if he had not done the manipulation or a house which is higher ranked than the house he gets if he does not do manipulation?

   Solution: Such a manipulation is possible. We give an example with four agents with four houses. Let $N = \{1, 2, 3, 4\}$ and the set of houses be $\{a_1, a_2, a_3, a_4\}$. The initial endowments of houses are given by a^*: $a^*(i) = a_i$ for all $i \in N$. The preferences of agents are shown in Table 1 - some of the preferences of agents are not shown completely, implying that it can be anything in the parts not shown.

$\succ_1$	$\succ_2$	$\succ_3$	$\succ_4$
a_2	a_1	a_4	a_4
a_3		a_2	
		a_3	

Table 1: An example for housing model

   If we run the TTC mechanism on this problem, the outcome will be a: $a(1) = a_2, a(2) = a_1, a(3) = a_3, a(4) = a_4$.

   Now, suppose agents 2 and 3 swap their endowments. So, the initial endowments of agents look as a': $a'(1) = a_1, a'(2) = a_3, a'(3) = a_2, a'(4) = a_4$. If we run the TTC mechanism on this problem, the outcome will be $\hat{a}$: $\hat{a}(1) = a_3, \hat{a}(2) = a_1, \hat{a}(3) = a_2, \hat{a}(4) = a_4$. Note that $\hat{a}(2) = a(2)$ and $\hat{a}(3) \succ_3 a(3)$. Hence, agents 2 and 3 successfully manipulated their initial endowments.

2. Consider the house allocation model with three agents $N = \{1, 2, 3\}$ and three objects $M = \{a, b, c\}$. Let f be a social choice function defined as follows. At any preference profile $\succ \equiv (\succ_1, \ldots, \succ_n)$, if $\succ_2 (1) = a$, then agent 1 gets the best element in $\{b, c\}$ according to his preference ordering $\succ_1$, agent 2 gets a, and agent 3 gets the remaining object (i.e., a serial dictatorship with the highest priority to agent 2, followed by agent 1, and finally to agent 3). In all other cases, agent 1 gets the best object in M, agent 2 gets the best remaining object according to $\succ_2$, and agent 3 gets the remaining object (i.e., a serial dictatorship with the highest priority to agent 1, followed by agent 2, and

finally to agent 3).

Is f strategy-proof? Is f non-bossy, i.e., can an agent change the outcome at a profile without changing the object assigned to him?

Answer: f is strategy-proof. Agents 1 and 3 cannot change the priority. So, they have no incentive to manipulate. Agent 2 can change the priority. But he will not manipulate if he gets the top priority. When agent 2 gets the second priority, he can change the priority by saying that his top is a, and in this case he gets a. But a is not his top according to his true preference. So he gets an object which is at least his second preferred object. But that he could have got even if he did not change the priority. So, he does not gain by manipulation.

f is also non-bossy. Note that if the serial dictatorship with a given priority is non-bossy. So, if an agent does not change his own allocation in f, it does not change the priority in f. So, by the same reasoning, it is non-bossy - other agents will continue to choose the best from same set of available objects to them.

3. Consider a two-sided matching model with men and women. Let $\succ$ be a profile of preference orderings as shown in Table 2.

$\succ_{m_1}$	$\succ_{m_2}$	$\succ_{m_3}$	$\succ_{w_1}$	$\succ_{w_2}$	$\succ_{w_3}$
w_1	w_2	w_2	m_2	m_1	m_1
w_3	w_1	w_1	m_1	m_2	m_2
w_2	w_3	w_3	m_3	m_3	m_3

Table 2: Preference orderings of men and women

Suppose μ is the outcome of the women-proposing deferred acceptance algorithm for the preference profile $\succ$. Let μ' be the outcome of the fixed-priority TTC mechanism where the priorities of men are fixed according to their preference orderings in $\succ$ in Table 2, and then each woman points to the woman with her favorite man in every stage of the TTC.

(a) Verify that $\mu \neq \mu'$.

(b) Verify that μ' is not stable by identifying a blocking pair.

(c) Verify that μ' women-dominates μ.

Answer. This can be verified in a straightforward manner.

4. Prove that if m and w are matched to each other in the men-proposing and women-proposing DAA, then they are matched to each other in *every* stable matching.

   Answer. This follows from the fact that a man gets the best possible woman among all stable matchings in the men-proposing DAA and gets the worst possible woman in the women-proposing DAA. Suppose m is matched to w' in some stable matching μ, where $w \neq w'$. Then, by construction, $w \succ_m w'$ and $w' \succ w$, a contradiction.

5. A stage in a men-proposing DAA involves proposals from rejected men in the previous stage and tentative acceptances and rejections by women based on these proposals. What is the maximum possible number of stages in a DAA with n men and n women.

   Answer. Consider a stage t of men-proposing DAA. If man m is rejected for the $(n-1)$-th time, then his next proposal must be accepted. So, every man can be rejected at most $(n-1)$ times. In every stage, at least one man is rejected. So, after $(n-1)^2$ stages, at least $(n-1)$ men must be rejected $(n-1)$ times. Clearly, everyone will be matched in the next stage. So, the maximum number of stages is $(n-1)^2 + 1$.

   It is possible to construct preference profiles where this bound is achieved. For instance, with $n = 3$, it is possible to construct profiles where it takes 5 stages.

6. Let μ and μ' be two stable matchings. Define $\mu \wedge^m \mu'$ as follows: for every $m \in M$,

$$(\mu \wedge^m \mu')(m) := \min_{\succ_m}(\mu(m), \mu'(m)).$$

   Either show that $(\mu \wedge^m \mu') = (\mu \vee^w \mu')$ or provide a counterexample.

   Answer. We will show that $\mu \wedge^m \mu'$ is a stable matching - in fact it is equal to $\mu \vee^w \mu'$. We do that in two steps.

   Step 1. First, we show that $\mu \wedge^m \mu'$ is a matching. Suppose not. Then, there is some w, m, m' such that $\min_{\succ_m}(\mu(m), \mu'(m)) = \min_{\succ_{m'}}(\mu(m'), \mu'(m')) = w$. This also implies that there exists some $w' \in W$ such that for every man $m'' \in M$, $\min_{\succ_{m''}}(\mu(m''), \mu'(m'')) \neq w'$. In particular, let $\mu(m_1) = w'$ and $\mu'(m_2) = w'$. First, note that $m_1 \neq m_2$. This is because if $m_1 = m_2$, then $\min_{\succ_{m_1}}(\mu(m_1), \mu'(m_1)) = w'$, contradicting our assumption about w'. Next, $\max_{\succ_{m_1}}(w = \mu(m_1), \mu'(m_1)) = \max_{\succ_{m_2}}(\mu(m_2), \mu'(m_2) = w) = w$. But, then the matching $\mu \vee^m \mu'$ will assign w to both m_1 and m_2, contradicting the fact that it is a matching.

   Step 2. Now, we show that $\mu \wedge^m \mu'$ is a stable matching. For this we show that $\mu \wedge^m \mu' = \mu \vee^w \mu'$. Pick $m \in M$ and let $w = \min_{\succ_m}(\mu(m), \mu'(m)) = \mu(m)$. We show that $\max_{\succ_w}(\mu^{-1}(w), \mu'^{-1}(w)) = m$, and this will establish the claim. Suppose

$\mu'(m') = w$. Since $\min_{\succ_{m'}}(\mu(m'), \mu'(m')) \neq w$, we know that $w \succ_{m'} \mu(m')$. Since μ is stable, we get that $m = \mu^{-1}(w) \succ_w m'$. Hence, $\max_{\succ_w}(\mu^{-1}(w), \mu'^{-1}(w)) = m$.