Discrete Optimization 2015. Exam Questions

Please observe the instructions given in a separate document.
Good luck!

Problem 1.
We managed to write an LP for fitting a line $y = ax + b$ to a given set of data points (x_i, y_i), $i = 1, \ldots, n$, so as to minimize the sum of differences of y-values (vertical distances between the the data points and the line). Wriiten as a formula, the objective was: $\min \sum_{i=1}^{n} |ax_i + b - y_i|$. One may be wondering if the applied method still works if the "target function" is not linear. Specifically:

Suppose you have to fit a quadratic function $y = ax^2 + bx + c$ to the data points, with the same goal as above: $\min \sum_{i=1}^{n} |ax_i^2 + bx_i + c - y_i|$. Either write this problem as an LP, or give a clear argument why this is not possible. (Of course, only one of these two alternatives can be correct.)

Problem 2.
A number of jobs with known processing times p_i need to be scheduled without preemption. That is, we have to decide on their start times t_i.
But we bought yet another machine, and now two machines are available. Two jobs may be processed at the same time, but not three jobs.

How can we write linear constraints to exclude that three jobs are done simultaneously? You don't have to write down all formulas in detail.
A high-level description is enough, however, say clearly and unambiguously how you proceed.

Hint (to avoid too much guesswork): We may consider any three jobs, say, with indices i, j, k, and express the following condition: "If job k gets the latest start time of these three, then job k can only start after the first completion time of jobs i and j." But also observe that the order of jobs is not given.

See reverse page.

Problem 3.
Recall the matrix game where a player chooses a row i and an adversary chooses a column j of a matrix, and then the player has to pay c_{ij} units of money to the adversary. Now suppose that, for the player, some row is consistently better than another row, for instance: we have $c_{1j} \leq c_{2j}$ for all columns j. One may conjecture that in such a case the player has no reason to ever choose row 2. In other words: There exists an optimal mixed strategy with $x_2 = 0$ (where x_i denotes the probability to choose row i).

Is this conjecture true or false? Give either a proof or a counterexample.

Problem 4.
Within a branch-and-bound heuristic for ILP one may insert cutting planes in the ILP formulations of subproblems, prior to further branching steps. (This leads to the branch-and-cut paradigm.)

Briefly explain in your own words: Why is the insertion of cutting planes correct in this context, and what is the potential benefit?

Problem 5.
In an undirected graph, an induced matching is a matching M with the additional property that no further edges exist between the $2|M|$ nodes in M, besides the $|M|$ matching edges. Remember that an independent set is a set of nodes without any edges between them.

Consider the following parameters of a given graph:
a: minimum number of edges that cover all nodes
b: maximum size of an induced matching
c: maximum size of an independent set of nodes

We claim that these numbers form the same chain of inequalities in every graph. Bring a, b, c in the correct order, from smallest to largest, and explain why these inequalities hold in every graph.

Answers attached after the deadline:

1. It is possible. Our objective is $\min \sum_{i=1}^{n} |ax_i^2+bx_i+c-y_i|$. Our variables are a, b, c, and the coefficients are the x_i and y_i.
We create new variables z_i and linear constraints
$z_i \geq (ax_i^2 + bx_i + c - y_i)$ and $z_i \geq -(ax_i^2 + bx_i + c - y_i)$,
and we write the objective as $\min \sum_{i=1}^{n} z_i$. By the same argument as before, the constraints together with the objective function ensure that
$z_i = |ax_i + b - y_i|$ holds in any optimal solution. Everything works precisely as before. It is not important at all that the regression function is linear, only our objective and constraints need to be linear. We could do the same for any class of functions equipped with some parameters.

2. For any two time variables r and s we create a binary variable x such that $x = 1$ if $r \leq s$, and $x = 0$ else. To enforce this relationship we create the constraints $r \leq s + M(1 - x)$ and $s \leq r + Mx$ with a large enough constant M. For convenience we define variables for the completion times by $u_i = t_i + p_i$, and so on. Now we can describe all valid permutations of any times $t_i, t_j, t_k, u_i, u_j, u_k$ by a Boolean formula in the binary variables introduced above (where r, s are any two of these start and completion times). Finally, the Boolean formula can be translated into linear constraints, this was already subject of an exercise.

Some more details (not expected): Constraints $t_i \leq u_i$, and similarly for j and k, are obvious. As indicated in the hint, we also need:
if $t_i \leq t_k$ and $t_j \leq t_k$ then $u_i \leq t_k$ or $u_j \leq t_k$.
Boolean algebra turns this into: not $t_i \leq t_k$ or not $t_j \leq t_k$ or $u_i \leq t_k$ or $u_j \leq t_k$. The four corresponding binary variables (the first two negated) must have a sum at least 1.

3. It is true. An optimal mixed strategy minimizes z under the constraints $\sum_i x_i c_{ij} \leq z$ for all i, and $\sum_i x_i = 1$, with non-negative variables x_i. Assume $x_2 > 0$. Then we can change the strategy into $x'_1 = x_1 + x_2$, $x'_2 = 0$, and $x'_i = x_i$ for all other i. Since $c_{1j} \leq c_{2j}$ for all j, the constraints are still satisfied for the same z (and the minimial z may even become smaller).

4. We assume the case of a minimization problem. Cutting planes do not remove feasible integer solutions, thus we cannot wrongly discard an optimal solution. But they are beneficial when the minimal values of LP relaxations are used as lower bounds of subproblems: Cutting planes can increase their minimum values, and these improved lower bounds can prune larger subtrees of the search tree.

5. We have $b \leq c \leq a$. Motivation: Every induced matching also contains an independent set of the same size, obtained by taking one node from every matching edge. (This holds, in particular, for a maximum matching, but there could be even larger independent sets.) Covering nodes by few edges and the maximum independent set problem are (weakly) dual problems. This can be seen as follows. The dual packing problem is to select nodes such that no two of them belong to the same edge. But this is only the maximum independent set problem rephrased.