ISyE 3833 - HW 5 - DO NOT SUBMIT - PRACTICE ONLY

1. *ISyE* is organizing a field trip for n of its students. There are m cabins available, and the jth cabin has capacity K_j, for each $j = 1, \ldots, m$. Assume that there is enough capacity for everyone, i.e., $\sum_{j=1}^{m} K_j \geq n$. If at least one person stays in cabin j, its cost is c_j, independent of how many people stay in the cabin, for each $j = 1, \ldots, m$. (There is no cost associated with a cabin in which no-one stays.) Each student must be assigned to exactly one cabin.

   (a) Formulate an *IP* model that minimizes the total cost of renting the necessary cabins.

   There are many possible, correct, ways to model this problem as an integer linear program. Here is one. Let

$$y_j = \begin{cases} 1, & \text{if the } j\text{th cabin is occupied} \\ 0, & \text{otherwise} \end{cases} \qquad \forall j = 1, \ldots, m.$$

   Now in order to have a feasible solution, *ISyE* merely needs to ensure it has paid for enough cabin capacity, in total, to house all n students. Thus the model is

$$\begin{aligned} \min \quad & \sum_{j=1}^{m} c_j y_j \\ \text{s.t.} \quad & \sum_{j=1}^{m} K_j y_j \geq n, \\ & y \in \{0,1\}^m \end{aligned}$$

   (b) Now, assume that each student has a car and that there is a cost for each student, $i \in \{1, \ldots, n\}$, to use his/her car to drive to any cabin: d_i. Each car can take up to 5 people. Each student must be assigned to exactly one car. If a student i doesn't drive then he/she doesn't have to pay the cost d_i, but does incur a cost for getting to the home of the student that is driving them. Suppose that the cost for student $i \in \{1, \ldots, n\}$ to get to the home of student $k \in \{1, \ldots, n\}$ is given by g_{ik}. Formulate an *IP* model that minimizes the total cost of renting the necessary cabins plus the total transportation cost (the cost to students who are not driving to get to the home of the student that is giving them a ride plus the cost to the students who are driving).

   Again, there are many possible, correct, ways to model this. Here is one. In addition to the y variables from part (a), we introduce the variables

$$x_{ik} = \begin{cases} 1, & \text{if student } i \text{ travels in the car driven by student } k \\ 0, & \text{otherwise} \end{cases} \qquad \forall i, k = 1, \ldots, n.$$

   Note that $x_{kk} = 1$ for some k indicates that student k is a driver. Each student has to be driven by someone, so we will need the constraint

$$\sum_{k=1}^{n} x_{ik} = 1, \qquad \forall i = 1, \ldots, n.$$

   Also, if at most 5 students can travel with any driver (including the driver), so we need the constraint

$$\sum_{i=1}^{n} x_{ik} \leq 5 x_{kk}, \qquad \forall k = 1, \ldots, n.$$

This can be interpreted as saying that if student k is not a driver, then $x_{kk} = 0$, so the capacity of student k's car to carry students is $5x_{kk} = 5(0) = 0$, whereas if student k is a driver, then $x_{kk} = 1$, so the capacity of student k's care to carry students i $5x_{kk} = 5(1) = 5$. Now we can put it all together as the model

$$\begin{array}{llll} \min & \sum_{j=1}^{m} c_j y_j + \sum_{k=1}^{n} d_k x_{kk} + \sum_{i=1}^{n}\sum_{k=1}^{n} g_{ik} x_{ik} & & \\ \text{s.t.} & \sum_{j=1}^{m} K_j y_j & \geq n, & \\ & \sum_{k=1}^{n} x_{ik} & = 1, & \forall i = 1, \ldots, n \\ & \sum_{i=1}^{n} x_{ik} & \leq 5x_{kk}, & \forall k = 1, \ldots, n \\ & & y \in \{0,1\}^m & \\ & & x \in \{0,1\}^{n \times n} & . \end{array}$$

Note that this assumes that if student k is a driver, the student's total cost is $d_k + g_{kk}$. Note also that this model is actually separable: it is equivalent to the sum of two LPs, one in the y variables and one in the x variables.

2. A company is considering opening warehouses in four cities: New York, Los Angeles, Chicago, and Atlanta. Each warehouse can ship 100 units per week. The weekly fixed cost of keeping each warehouse open is \$400 for NY, \$500 for LA, \$300 for Chicago, and \$150 for Atlanta. Region 1 of the country requires 80 units per week, region 2 requires 70 units per week, and region 3 requires 40 units per week. The costs (including production and shipping costs) of sending one unit from plant to a region are shown in the table.

From	To region 1	To region 2	To region 3
New York	20	40	50
Los Angeles	48	15	26
Chicago	26	35	18
Atlanta	24	50	35

We want to meet weekly demands at minimum cost, subject to the preceding information and the following restrictions:

- If the New York warehouse is opened, then the Los Angeles warehouse must be opened.
- At most two warehouses can be opened.
- Either the Atlanta or the Los Angeles warehouse must be opened.

(a) Formulate an IP that can be used to minimize the weekly costs of meeting demand.
Define the index sets

$$W = \{NY, LA, CHI, ATL\} \qquad \text{and} \qquad R = \{1, 2, 3\}$$

to represent the four cities and the three regions respectively. Now, define the parameter d_j to represent the demand of region $j \in R$, so $d = (80, 70, 40)$. Also, define the parameter f_i to be the cost, in dollars per week, of keeping a warehouse open in city i, for each $i \in W$, so $f = (400, 500, 300, 150)$. Finally, define parameter c_{ij} to indicate the cost, in dollars per unit, of sending product from warehouse i to region j, for each $i \in W$ and $j \in R$. The values for c are given in the table.

Now define the variables

$$y_i = \begin{cases} 1, & \text{if a warehouse is opened in city } i \\ 0, & \text{otherwise} \end{cases} \qquad \forall i \in W$$

and

$$x_{ij} = \text{units shipped from a warehouse at city } i \text{ to region } j, \qquad \forall i \in W, j \in R.$$

Now the total, weekly, cost, in dollars, to be minimized, can be expressed as

$$\sum_{i \in W} f_i y_i + \sum_{i \in W} \sum_{j \in R} c_{ij} x_{ij}.$$

To ensure the demand of each region is met, we use the constraint

$$\sum_{i \in W} x_{ij} \geq d_j, \qquad \forall j \in R.$$

Now the capacity of the warehouse in city i to supply product can be expressed as $100 y_i$, so to ensure each city supplies a quantity no greater than its capacity, we use the constraint

$$\sum_{j \in R} x_{ij} \leq 100 y_i, \qquad \forall i \in W.$$

To model the requirement that if the NY warehouse is open, then LA must also be open, we use

$$y_{NY} \leq y_{LA}.$$

To model the requirement that at most two warehouses can be opened, we use

$$\sum_{i \in W} y_i \leq 2.$$

Finally, to model the requirement that either the ATL or LA warehouses must be open, we use

$$y_{ATL} + y_{LA} \geq 1.$$

Putting it all together, we obtain the model

$$\begin{array}{lrll} \min & \displaystyle\sum_{i \in W} f_i y_i + \sum_{i \in W} \sum_{j \in R} c_{ij} x_{ij} & & \\ \text{s.t.} & \displaystyle\sum_{i \in W} x_{ij} & \geq d_j, & \forall j \in R \\ & \displaystyle\sum_{j \in R} x_{ij} & \leq 100 y_i, & \forall i \in W \\ & y_{NY} & \leq y_{LA} & \\ & \displaystyle\sum_{i \in W} y_i & \leq 2 & \\ & y_{ATL} + y_{LA} & \geq 1 & \\ & & x_{ij} \geq 0, & \forall i \in W, j \in R \\ & & y_i \in \{0, 1\} & \forall i \in W. \end{array}$$

(b) Now suppose that if any units are shipped on a transport link (from a warehouse to a region), then at least 20 units must be shipped; anything less is not economic. Modify your integer linear programming model accordingly. [Hint: you may need to introduce additional variables.]

To model this, we add new, binary, variables

$$u_{ij} = \begin{cases} 1, & \text{if the transport link from city } i \text{ to region } j \text{ is used} \\ 0, & \text{otherwise} \end{cases} \qquad \forall i \in W, j \in R.$$

Now we can ensure that the quantity transported on this link is at least 20 if the link is used, and is zero otherwise, via the constraints

$$20u_{ij} \le x_{ij} \le \min\{100, d_j\}u_{ij}, \qquad \forall i \in W, j \in R.$$

Here, the coefficient on the right-hand side, $\min\{100, d_j\}$, is justified by the assumption that all costs are non-negative, so there is no need for more than d_j units to be transported on a link to region $j \in R$.

3. A local radio station is going to schedule commercials within 60 second blocks. There are m comercials to schedule. The duration of each commercial is t_i seconds, for $i = 1 \dots m$. You can assume $t_i \le 60$ for any i. Each commercial must be played exactly once a day. Formulate an IP for which the optimal value corresponds to the minimum number of blocks necessary to run all ads in a day.

First, let n be a parameter set to the maximum number of blocks that might be run during the day. For example, n could be chosen to be m or $24 \times 60 = 1440$, whichever is smaller. Define variables

$$y_j = \begin{cases} 1, & \text{if block } j \text{ is open for ads} \\ 0, & \text{otherwise,} \end{cases} \qquad \forall j = 1, \dots, n,$$

and

$$x_{ij} = \begin{cases} 1, & \text{if ad } i \text{ runs in block } j \\ 0, & \text{otherwise,} \end{cases} \qquad \forall i = 1, \dots, m, \ \forall j = 1, \dots, n,$$

To ensure every ad is run, we use the constraint

$$\sum_{j=1}^{n} x_{ij} = 1, \qquad \forall i = 1, \dots, m.$$

Now the time-capacity of a block is 60 seconds if the block is open for ads, and is zero if the block is not open for ads. So this capacity can be represented by the expression $60y_j$ for each block j. The total time of ads assigned to block j is $\sum_{i=1}^{m} t_i x_{ij}$. So the time-capacity of blocks is respected by requiring

$$\sum_{i=1}^{m} t_i x_{ij} \le 60y_j, \qquad \forall j = 1, \dots, n.$$

An alternative way to model this is to simply ask that the total time of ads assigned to a block is no more than 60, via the constraint

$$\sum_{i=1}^{m} t_i x_{ij} \le 60, \qquad \forall j = 1, \dots, n,$$

and then add a separate constraint to model the logical requirement that if a block is not open for ads, then an add cannot be assigned to it, i.e., model the logic "if $y_j = 0$ then $x_{ij} = 0$", for all j and all i:

$$x_{ij} \leq y_j, \qquad \forall i = 1, \ldots, m, \ \forall j = 1, \ldots, n.$$

We want to minimize the total number of blocks open for ads, which is given by $\sum_{j=1}^{n} y_j$. So one possible integer linear programming model is:

$$\begin{aligned}
\min & \sum_{j=1}^{n} y_j \\
\text{s.t.} & \sum_{j=1}^{n} x_{ij} = 1, & \forall i = 1, \ldots, m, \\
& \sum_{i=1}^{m} t_i x_{ij} \leq 60 y_j, & \forall j = 1, \ldots, n, \\
& x_{ij} \in \{0, 1\}, & \forall i = 1, \ldots, m, \ \forall j = 1, \ldots, n, \text{and} \\
& y_j \in \{0, 1\}, & \forall j = 1, \ldots, n.
\end{aligned}$$

An alternative correct model is given by replacing the second set of constraints by

$$\begin{aligned}
& \sum_{i=1}^{m} t_i x_{ij} \leq 60, & \forall j = 1, \ldots, n, \\
& x_{ij} \leq y_j, & \forall i = 1, \ldots, m, \ \forall j = 1, \ldots, n.
\end{aligned}$$

4. Suppose x and y are binary variables, w is an integer variable, and z is a continuous variable, with $0 \leq w \leq 6$ and $0 \leq z \leq 100$. Write linear constraints to model the following logical relationships. In same cases, you may need to define additional variables.

   (a) If $x = 1$ then $y = 1$. $x \leq y$

   (b) x is not equal to y. $x + y = 1$

   (c) If $x = 1$ then $y = 0$. $y \leq 1 - x$

   (d) If $y = 1$ then $z = 0$. $z \leq 100(1 - y)$

   (e) Either $x = 1$ or $y = 0$ or both. $x + (1 - y) \geq 1$, equivalently $x \geq y$

   (f) If $x = 1$ and $y = 0$ then $z = 0$. $z \leq 100(1 - x + y)$

   (g) If $x = 1$ then $w \geq 3$. $w \geq 3x$

   (h) Either $z = 0$ or $z \geq 20$. Introduce new binary variable u, and require $20u \leq z \leq 100u$.

   (i) If $x = 0$ then $z \geq 50$ and $w \leq 3$. Add the two constraints $z \geq 50(1 - x)$ and $w \leq 3(1 - x) + 6x$, which is equivalent to $w \leq 3 + 3x$.

   (j) If $x = 1$ then $w = 3$. Add the two constraints $w \geq 3x$ and $w \leq 6 - 3x$.

   (k) If $z > 25$ then $x = 1$. $x \geq \frac{1}{75}(z - 25)$, equivalently $z \leq 25 + 75x$

   (l) If $y = 1$ then $w + x \geq 2$. $w + x \geq 2y$

(m) If $w+3x \leq 4$ then $y=0$. $5(1-y) \geq 5-(w+3x)$, since $w+3x$ must be integer, as both x and w are, and since the maximum value of $5-(w+3x)$ is $5-0=5$. Equivalently $5y \leq w+3x$.

(n) If $w+3x \leq 4$ then $w+x \geq 2$. Add a new binary variable, say u. Enforce that if $w+3x \leq 4$ then $u=1$, and that if $u=1$ then $w+x \geq 2$. Thus add constraints $5(1-u) \leq w+3x$ and $w+x \geq 2u$.

(o) Either $w+3x \leq 4$ or $w+x \geq 2$, or both. This is equivalent to asking that if $w+3x > 4$ then $w+x \geq 2$. Since w and x are integer, this is equivalent to if $w+3x \geq 5$ then $w+x \geq 2$. This can be modeled by adding a new binary variable, say u, and asking that if $w+3x \geq 5$ then $u=1$, and if $u=1$ then $w+x \geq 2$. This can be modeled by adding constraints $5u \geq w+3x-4$, (since the maximum of $w+3x-4$ is $6+3-4=5$), and $w+x \geq 2u$.

5. Answer Extra Exercise 3 from the Tutorial 14 handout.
(a) There are integer feasible solutions at Node 5 and Node 8. Since this is minimization, the best is at Node 5, giving a best upper bound of 15.9.
(b) The only active nodes are Node 4 and Node 7. From these we see there is some hope of getting an IP solution with value as low as 14.6, so 14.6 is the best (global) lower bound on the IP that can be deduced from the tree.
(c) The LP relaxation at Node 6 must have been solved before that at Node 5, since otherwise Node 6 would have been pruned by the solution at Node 5; its lower bound is higher than (worse than, for a minimization problem) the value of the feasible IP solution found at Node 5.
In addition, answer the following questions about the partial branch-and-bound tree given in that exercise.

   (a) To create the LP solved at Node 5, what constraint(s) were added to the (original) LP relaxation of the IP? $x_1 \geq 1$ and $x_2 \leq 3$

   (b) For each of Nodes 4, 5, 7, 8 and 9, indicate whether you need to branch on the node. In each case, briefly explain your answer. If branching is required, describe the branches.
   Branching is required at Node 4, since it has a fractional solution, and its lower bound (14.6) is better than the best upper bound found so far (15.9). The two branches would represent the dichotomy $x_3 \leq 1$ or $x_3 \geq 2$.
   Branching is not required at Node 5: it is a leaf node, yielding an integer feasible solution.
   Branching is required at Node 7, since it has a fractional solution, and its lower bound (14.7) is better than the best upper bound found so far (15.9). The two branches would represent the dichotomy $x_3 \leq 4$ or $x_3 \geq 5$.
   Branching is not required at Node 8: it is a leaf node, yielding an integer feasible solution.
   Branching is not required at Node 9: it is a leaf node, with an infeasible LP relaxation, proving that the IP represented by this node is also infeasible.

Important note: this is a tree for a **minimization** problem.

6. Consider an IP (maximization) problem P_{IP} and its linear relaxation P. Let Z_{IP} and Z_P denote their optimal values. Note that P has a dual problem: (D). Let Z_D denote the value of this dual. Assume that all problems have an optimal solution. Currently you have one feasible solution x^* to P_{IP} with a value equal to Z^*. You know that Z^* and the values of P and D lay in the set $\{50, 51, 52\}$:

- What are all possible combination of values for Z^*, Z_P and Z_D? The possible triples of the form (Z^*, Z_P, Z_D) are: $(50, 50, 50)$, $(50, 51, 51)$, $(50, 52, 52)$, $(51, 51, 51)$, $(51, 52, 52)$, and $(52, 52, 52)$. These are all cases for which $Z_P = Z_D$, which holds by strong LP duality, and $Z^* \leq Z_P$, since P is a relaxation of the P_{IP}, so Z_P gives an upper bound on the value of the IP, while Z^*, as the value of a feasible solution for the IP, gives a lower bound.
- In which cases can we be sure that x^* is optimal? The only cases where we can be sure that x^* is optimal for the IP are those with $Z^* = Z_P$: $(50, 50, 50)$, $(51, 51, 51)$, and $(52, 52, 52)$. These are all cases for which $Z^* = Z_P = Z_D$.
- In which cases can we be sure that P has an optimal solution that is integer? Similarly, the only cases we can be sure of are those with $Z^* = Z_P$, since in these cases, x^* must also be an optimal solution of P (if it is feasible for the IP, it is also feasible for its LP relaxation, and its value equals the optimal value of the LP).