Answer the questions in the boxes provided on the question sheets. If you run out of room for an answer, add a page to the end of the document.

Name: **Carson Miller**   Wisc id: **9081473028**

Randomization

1. *Kleinberg, Jon. Algorithm Design (p. 782, q. 1).*

   3-Coloring is a yes/no question, but we can phrase it as an optimization problem as follows.

   Suppose we are given a graph $G = (V, E)$, and we want to color each node with one of three colors, even if we aren't necessarily able to give different colors to every pair of adjacent nodes. Rather, we say that an edge (u, v) is *satisfied* if the colors assigned to u and v are different. Consider a 3-coloring that maximizes the number of satisfied edges, and let c^* denote this number. Give a polynomial-time algorithm that produces a 3-coloring that satisfies at least $\frac{2}{3}c^*$ edges. If you want, your algorithm can be randomized; in this case, the expected number of edges it satisfies should be at least $\frac{2}{3}c^*$.

   > We can design a randomized procedure to make the approximation factor at least 2c/3. For each node, set the probability that it is each of three colors to 1/3 each. Let k represent the number of edges in the graph.
   >
   > $$E(E) = E\big(edge\ satisfied\big) = \text{P} * (\text{Cxi}) + (1 - \text{P}) * (\text{Cnotxi})$$
   > $$E(E) = P * (Cxi + Cnotxi) - P * Cnotxi = k - P * Cnotxi$$
   > $$= k - \left(\frac{Cxi}{Cxi + Cnotxi}\right) * Cnotxi = k - Cxi + \frac{Cxi * Cnotxi}{Cxi + Cnotxi} \geq k - Cxi + \frac{Cxi}{2} = \frac{k + Cxi}{2}$$
   > $$Cxi \geq \left(\frac{k}{2}\right) + 1,\ Thus\ E(E) \geq \frac{k + \left(\frac{k}{2} + 1\right)}{2} = \frac{\frac{3k}{2} + 1}{2} = \frac{3k}{4} + 0.5 = \frac{2}{3}c$$

2. *Kleinberg, Jon. Algorithm Design (p. 787, q. 7).*

   In lecture, we designed an approximation algorithm to within a factor of 7/8 for the MAX 3-SAT Problem, where we assumed that each clause has terms associated with three different variables. In this problem, we will consider the analogous MAX SAT Problem: Given a set of clauses $C_1, \cdots, C_k$ over a set of variables $X = \{x_1, \cdots, x_n\}$, find a truth assignment satisfying as many of the clauses as possible. Each clause has at least one term in it, and all the variables in a single clause are distinct, but otherwise we do not make any assumptions on the length of the clauses: There may be clauses that have a lot of variables, and others may have just a single variable.

   (a) First consider the randomized approximation algorithm we used for MAX 3-SAT, setting each variable independently to true or false with probability 1/2 each. Show that in the MAX SAT, the expected number of clauses satisfied by this random assignment is at least $k/2$, that is, at least half of the clauses are satisfied in expectation.

   > The expected number of satisfied clauses can be calculated as follows:
   >
   > $$E(S) = \sum_{i=1}^{k} P\big(C_i \text{ is satisfied}\big) = \sum_{i=1}^{k} (1 - P(C_i \text{ is not satisfied}))$$
   >
   > $$= \sum_{i=1}^{k} \left(1 - \left(\frac{1}{2}\right)^{l_i}\right), \text{where } l_i \text{ is the number of literals in } C_i$$
   >
   > $$\geq \sum_{i=1}^{k} \left(1 - \frac{1}{l_i}\right) \geq k - \sum_{i=1}^{k} \frac{1}{l_i} \geq k - \frac{k}{2} = \frac{k}{2}$$
   >
   > Thus, the expected number of satisfied clauses is greater than k/2

   (b) Give an example to show that there are MAX SAT instances such that no assignment satisfies more than half of the clauses.

   > MAX SAT Instance: {C1 = [x1, x2], C2 = [x1, not x2], C3 = [not x1, x2], C4 = [not x1, not x2]}
   >
   > All 4 possible assignments (00, 01, 10, 11) yield exactly 2/4 clauses satisfied.

(c) If we have a clause that consists only of a single term (e.g., a clause consisting just of x_1, or just of $\overline{x_2}$), then there is only a single way to satisfy it: We need to set the corresponding variable in the appropriate way. If we have two clauses such that one consists of just the term x_i, and the other consists of just the negated term $\overline{x_i}$, then this is a pretty direct contradiction. Assume that our instance has no such pair of "conflicting clauses"; that is, for no variable x_i do we have both a clause $C = \{x_i\}$ and a clause $C' = \{\overline{x_i}\}$. Modify the randomized procedure above to improve the approximation factor from $1/2$ to at least 0.6. That is, change the algorithm so that the expected number of clauses satisfied by the process is at least $0.6k$.

We can change the randomized procedure to improve the approximation factor from 1/2 to 0.6 by altering the probabilities for variables to be 0 or 1. For each variable xi, count the occurrences of xi and not xi in all clauses. Set the probability for xi to be a 1 to: (# of xi/(# of xi + # of not xi)). Similarly, set the probability for xi to be a 0 to: (# of not xi/(# of xi + # of not xi)). Let's call the probability that xi is set to 1 P. Also, assume that "xi" represents the more common of xi and not xi, thus P(xi) >= 0.5

$$E(S) = E\big(Clauses\ w/\ xi\ satisfied\big) + E\big(Clauses\ w/\ not\ xi\ satisfied\big)$$
$$E(S) = P * (Cxi) + (1 - P) * (Cnotxi)$$
$$E(S) = P * (Cxi + Cnotxi) - P * Cnotxi = k - P * Cnotxi$$
$$= k - \left(\frac{Cxi}{Cxi + Cnotxi}\right) * Cnotxi = k - Cxi + \frac{Cxi * Cnotxi}{Cxi + Cnotxi} \geq k - Cxi + \frac{Cxi}{2} = \frac{k + Cxi}{2}$$
$$Cxi \geq \left(\frac{k}{2}\right) + 1, Thus\ E(S) \geq \frac{k + \left(\frac{k}{2} + 1\right)}{2} = \frac{\frac{3k}{2} + 1}{2} = \frac{3k}{4} + 0.5 = 0.75k \geq 0.6k$$

(d) Give a randomized polynomial-time algorithm for the general MAX SAT Problem, so that the expected number of clauses satisfied by the algorithm is at least a 0.6 fraction of the maximum possible. (Note that, by the example in part (a), there are instances where one cannot satisfy more than $k/2$ clauses; the point here is that we'd still like an efficient algorithm that, in expectation, can satisfy a 0.6 fraction of the maximum that can be satisfied by an optimal assignment.)

This is the same as problem c. We can change the randomized procedure to improve the approximation factor from 1/2 to 0.6 by altering the probabilities for variables to be 0 or 1. For each variable xi, count the occurrences of xi and not xi in all clauses. Set the probability for xi to be a 1 to: (# of xi/(# of xi + # of not xi)). Similarly, set the probability for xi to be a 0 to: (# of not xi/(# of xi + # of not xi)). Let's call the probability that xi is set to 1 P. Also, assume that "xi" represents the more common of xi and not xi, thus P(xi) >= 0.5

$$E(S) = E\big(Clauses \; w/ \; xi \; satisfied\big) + E\big(Clauses \; w/ \; not \; xi \; satisfied\big)$$
$$E(S) = P * (Cxi) + (1 - P) * (Cnotxi)$$
$$E(S) = P * (Cxi + Cnotxi) - P * Cnotxi = k - P * Cnotxi$$
$$= k - \left(\frac{Cxi}{Cxi + Cnotxi}\right) * Cnotxi = k - Cxi + \frac{Cxi * Cnotxi}{Cxi + Cnotxi} \geq k - Cxi + \frac{Cxi}{2} = \frac{k + Cxi}{2}$$
$$Cxi \geq \left(\frac{k}{2}\right) + 1, \; Thus \; E(S) \geq \frac{k + \left(\frac{k}{2} + 1\right)}{2} = \frac{\frac{3k}{2} + 1}{2} = \frac{3k}{4} + 0.5 = 0.75k \geq 0.6k$$

3. *Kleinberg, Jon. Algorithm Design (p. 789, q. 10).*

   Consider a very simple online auction system that works as follows. There are n *bidding agents*; agent i has a bid b_i, which is a positive natural number. We will assume that all bids b_i are distinct from one another. The bidding agents appear in an order chosen uniformly at random, each proposes its bid b_i in turn, and at all times the system maintains a variable b^* equal to the highest bid seen so far. (Initially b^* is set to 0.) What is the expected number of times that b^* is updated when this process is executed, as a function of the parameters in the problem?

   > When agent i appears, there are two cases to consider:
   >   bi < b*: Agent i's bid is lower than the highest bid seen already, so b* does not change.
   >   bi > b*: Agent i's bid is higher than the highest bid seen already, so we need to update b*.
   >
   > To count the number of times b* is updated, we need to determine the probability that bi > b*. It is expected that the number of times it will be updated can be represented by E[n-1]+1 if bi > b* and E[n-1] is bi < b*.
   >
   > $$P(bi > b*) = \frac{1}{n-i+1}$$
   >
   > Therefore,
   >
   > $$E[n] = E[n-1] + \frac{1}{n-i+1} * (E[n-1] + 1 - E[i-1])$$

4. Recall that in an undirected and unweighted graph $G = (V, E)$, a cut is a partition of the vertices $(S, V\backslash S)$ (where $S \subseteq V$). The size of a cut is the number of edges which cross the cut (the number of edges (u, v) such that $u \in S$ and $v \in V\backslash S$). In the MAXCUT problem, we try to find the cut which has the largest value. (The decision version of MAXCUT is NP-complete, but we will not prove that here.) Give a randomized algorithm to find a cut which, in expectation, has value at least $1/2$ of the maximum value cut.

5. Implement an algorithm which, given a MAX 3-SAT instance, produces an assignment which satisfies at least 7/8 of the clauses, in either C, C++, C#, Java, or Python.

   The input will start with a positive integer n giving the number of variables, then a positive integer m giving the number of clauses, and then m lines describing each clause. The description of the clause will have three integers $x\ y\ z$, where $|x|$ encodes the variable number appearing in the first literal in the clause, the sign of x will be negative if and only if the literal is negated, and likewise for y and z to describe the two remaining literals in the clause. For example, $3\ -1\ -4$ corresponds to the clause $x_3 \vee \overline{x_1} \vee \overline{x_4}$. A sample input is the following:

   ```
   10
   5
   -1 -2 -5
   6 9 4
   -9 -7 -8
   2 -7 10
   -1 3 -6
   ```

   Your program should output an assignment which satisfies at least $\lfloor 7/8 \rfloor m$ clauses. Return n numbers in a line, using a ± 1 encoding for each variable (the ith number should be 1 if x_i is assigned TRUE, and -1 otherwise). The maximum possible number of satisfied clauses is 5, so your assignment should satisfy at least $\lfloor \frac{7}{8} \times 5 \rfloor = 4$ clauses. One possible correct output to the sample input would be:

   ```
   -1 1 1 1 1 1 -1 1 1 1
   ```