CECS 528, Learning Outcome Assessment 11, 4/30/2025, Dr. Ebert

Problem

LO7. An instance of the `Load Balancing` decision problem is an array a of size n, a positive integer $p \geq 2$, as well as a positive integer k. The problem is to decide if the members of a can be partitioned into p different subarrays $b_1, \ldots, b_p$ so that the sum of the members in any one of the arrays does not exceed k. Our goal is to prove that `Load Balancing` is an `NP` problem.

(a) For a given instance (a, p, k) of `Load Balancing`, describe a certificate in relation to (a, p, k). Hint: a single array will suffice.

(b) Provide a semi-formal verifier algorithm that takes as input i) an instance (a, p, k) of `Load Balancing`, ii) a certificate for (a, p, k) as defined in part a, and decides if the certificate is valid for (a, p, k).

(c) Clearly define size parameters for the `Load Balancing` decision problem.

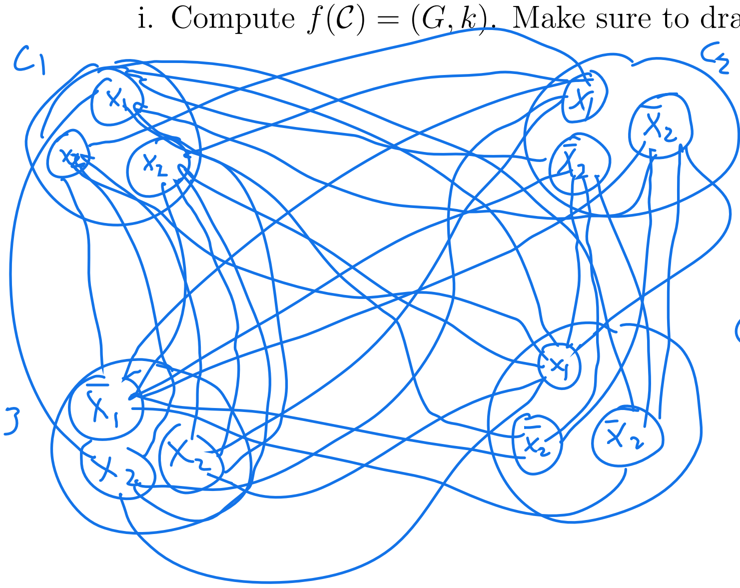
(d) Use the size parameters from part c to describe the running time of your verifier. Defend your answer.

LO8. Answer the following.

(a) Consider the mapping reduction f : `3SAT` $\to$ `Clique` from `3SAT` to `Clique` that was presented in lecture and the `3SAT` instance

$$\mathcal{C} = \{c_1 = (x_1, x_2, x_2), c_2 = (\overline{x}_1, \overline{x}_2, \overline{x}_2), c_3 = (\overline{x}_1, x_2, x_2), c_4 = (x_1, \overline{x}_2, \overline{x}_2)\}.$$

i. Compute $f(\mathcal{C}) = (G, k)$. Make sure to draw the entire graph.



, $k = 4$

ii. Verify that f is valid for input $\mathcal{C}$ from part i) in the sense that both $\mathcal{C}$ and $f(\mathcal{C})$ are either both positive instances or both negative instances. Make sure your answer is specific to the particular instances from part i).

f is valid since C is unsatisfiable (none of the four possible assignments satisfies all four clauses). Thus, G has no 4-Clique.

(b) Consider the mapping reduction f : `3SAT` $\to$ `Subset Sum` from `3SAT` to `Subset Sum` that was presented in lecture. Draw the complete table associated with $f(\mathcal{C}) = (S, t)$ and either provide a subset $A \subseteq S$ that sums to t or explain why one does not exist.

	x_1	x_2	c_1	c_2	c_3	c_4
y_1	1	0	1	0	0	1
z_1	1	0	0	1	1	0
y_2		1	1	0	1	0
z_2		1	0	1	0	1
g_1			1	0	0	0
h_1			1	0	0	0
g_2				1	0	0
h_2				1	0	0
g_3					1	0
h_3					1	0
g_4						1
h_4						1
t	1	1	3	3	3	3

There is no subset that sums to t since C is unsatisfiable (see aii).

LO9. Do the following.

(a) Demonstrate the `Vertex Cover` approximation algorithm described in lecture when applied to the graph $G = (V, E)$ whose edges are

$$\begin{array}{l} \quad e_1 \quad\ e_2 \quad e_3 \quad\ e_4 \quad\ e_5 \quad\ e_6 \quad\ e_7 \quad\ e_8 \quad\ e_9 \quad e_{10} \quad e_{11} \\ E = \{(0,5),(0,9)(1,6),(2,5),(2,7),(2,3),(3,4),(3,6),(3,9),(1,4),(1,5), \end{array}$$

$$\begin{array}{l} (5,8),(3,7),(4,8),(8,9),(0,10),(7,10)\} \\ \ e_{12} \quad e_{13} \quad e_{14} \quad e_{15} \quad e_{16} \quad e_{17} \end{array}$$

Clearly indicate how each vertex was selected for the cover.

$$C = \{ \underset{e_1}{0}, \underset{e_1}{5}, \underset{e_3}{1}, \underset{e_3}{6}, \underset{e_5}{2}, \underset{e_5}{7}, \underset{e_7}{3}, \underset{e_7}{4}, \underset{e_{15}}{8}, \underset{e_{15}}{9} \}$$

$|C| = 10. \therefore$ min cover size ≥ 5.

(b) Answer the following regarding the analysis of the `k-Clustering` approximation algorithm.

i. In the analysis of the k-`Clustering` approximation algorithm recall that point x was introduced as what would have been the next center (assuming $k+1$ or more desired clusters). How is r defined in relation to x and how do we know that, for any two centers c_1 and c_2, $d(c_1, c_2) \geq r$?

Let C = set of centers.
$r = \max(d(x, C))$. If some center c satisfied $d(c, C) < r$, $x \notin C$ then x would have been chosen over c, since x is further from the closest center than is c.

ii. For any non-center y that is in some cluster with center c, how do we know that $d(y, c) \leq r$?

Since $d(y,c) = d(y, C)$, we must have $d(y,c) \leq r$, since otherwise $d(y,C) > d(x,C) \Rightarrow y$ should have been selected as the $k+1$ center instead of x.

iii. How do we know that the optimal solution must have at least one cluster with a diameter of at least r?

This is true since x must belong in one of the k clusters and thus $d(x,c) \geq r$ where c is the center of that cluster. True since $d(x,c) = r = d(x, C)$

LO10. Solve the following and show all work.

(a) A bag consists of four coins. Two of the coins are two-headed, one of the coins is one-headed, and the fourth coin is zero-headed. If two of the coins are randomly selected from the bag and are randomly tossed on to a table, determine the probability that both will show a head. Show work.

Each of the following pairs has an equal chance of being selected: $\{(2H_1, 2H_2), (2H_1, 1H), (2H_2, 1H), (2H_1, 0H), (2H_2, 0H), (1H, 0H)\}$

Thus, $P(HH) = \frac{1}{6}\left(1 + \frac{1}{2} + \frac{1}{2}\right) = \frac{1}{3}$

(b) For the experiment described in part a) determine the domain and probability distribution for the random variable $E[H|S]$, where H is the number of heads that appear on the table, and S is the number of two-headed coins that were selected.

$\text{dom}(E[H|S]) = \{E[H|S=0], E[H|S=1], E[H|S=2]\}$

$E[H|S=0] = 0.5$, $E[H|S=1] = 1 + \frac{1}{4} = 1.25$, $E[H|S=2] = 2$

$P(0.5) = \frac{1}{6}$ $\quad P(1.25) = \frac{2}{3}$ $\quad P(2) = \frac{1}{6}$

$$E[E[H|S]] = E[H] = \frac{1}{12} + \left(\frac{5}{4}\right)\left(\frac{2}{3}\right) + \frac{1}{3} = \frac{15}{12} = 1.25$$

(c) For the experiment described in part a), how many times should we expect to have to repeat the experiment before we witness no heads appearing? Show work and explain.

For no heads to appear, first pick the 1H and 0H coins with 1/6 prob. Then have the 1H coin land with no head showing with 1/2 prob. ∴ 1/12 prob. of no heads ⟹ 12 trials are expected before witnessing it.

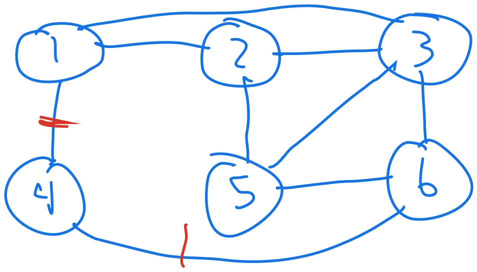
LO11. Answer the following questions about Karger's algorithm.

(a) A graph G has edges

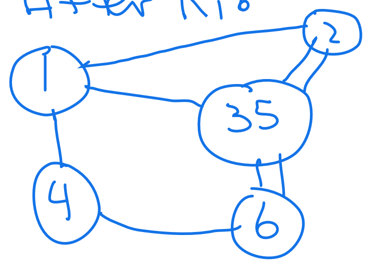
$$\{e_1 = (1,4), e_2 = (4,6), e_3 = (1,2), e_4 = (2,5), e_5 = (2,3),$$

$$e_6 = (3,5), e_7 = (3,6), e_8 = (5,6), e_9 = (1,3)\}.$$

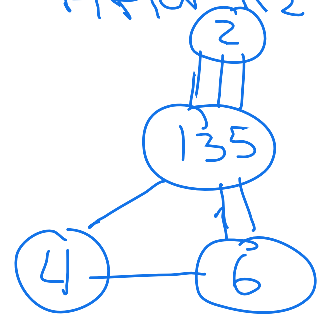
Draw the graph that remains after edges e_6, e_9, and e_4 have been selected in rounds 1, 2, and 3 of Karger's algorithm, respectively.



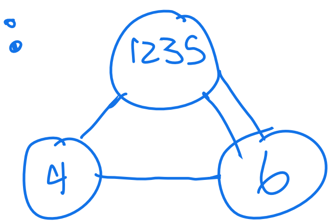
After R_1:



After R_2:



After R_3:



(b) Regarding part a), provide the probability of not selecting a min-cut edge in Round 1. Do the same for Round 2, conditioned on e_6 being selected in Round 1. Do the same for Round 3, conditioned on e_6 and e_9 being selected in Rounds 1 and 2, respectively, and finally for Round 4 assuming that a non-min-cut edge is selected in this final round. Compare the product of these four probabilities with the lower-bound probability p (as stated in Karger's Theorem) that a single execution of Karger's algorithm will return G's min cut.

Probabilities: $\frac{7}{9}, \frac{3}{4}, \frac{5}{7}, \frac{1}{2}$

$$\frac{7 \cdot 3 \cdot 5}{3 \cdot 3 \cdot 2 \cdot 2 \cdot 7 \cdot 2} = \frac{5}{24}$$

versus $p = \frac{2}{n(n-1)} = \frac{2}{(5)(6)} = \frac{1}{15}$

(c) When proving the lower-bound probability that Karger's algorithm finds a minimum cut, the proof heavily relies on a property that is posssessed by every vertex of the input graph during each round of the algorithm, assuming the algorithm still has a chance of returning the minimum cut. State this property.

$\forall v \in V$, $\deg(v) \geq K$, where K is the minimum cut size.