CECS 528, Quiz 4, Spring 2026, Dr. Ebert

IMPORTANT: READ THE FOLLOWING DIRECTIONS.

- For each LO, write your solutions to all parts using **ONE SHEET OF PAPER ONLY (BOTH FRONT AND BACK)**. Write **NAME** and **PROBLEM NUMBER** on each sheet. Use a moderate font size and section off the page in order to maximize solution space.
- Use **SEPARATE SHEETS** for different LO problems. Please raise your hand if you need more paper.
- **Papers that do not adhere to the above rules will not be graded**.
- The quiz ends at 8:45 am

LO10. Solve the following problems.

(a) In the analysis of the k-`Clustering` approximation algorithm recall that point x is introduced as what would have been the next center (assuming $k+1$ or more desired clusters). What makes x special is that its distance r from the center of its cluster is maximum amongst all data points.

i. Because of this, how do we know that any two points y and z in some cluster are within a distance of $2r$ from each other?
Solution. Since x's distance from its center is maximum of length r, then we must have $d(y, c) \leq r$ and $d(z, c) \leq r$, where c is the center of the cluster in which y and z belong. Thus, by the triangle inequality,

$$d(y, z) \leq d(y, c) + d(c, z) = d(y, c) + d(z, c) \leq r + r = 2r.$$

ii. How do we know that the optimal solution must have at least one cluster with a diameter of at least r?
Solution. It is also the case that $\mathrm{diam}(c_1, ..., c_k, x) \geq r$, where $\{c_1, \ldots, c_k\}$ are the k centers. Indeed, we already know that x is at least r away from every center. In addition, every center c must be at least r away from every other center. Otherwise, based on the criterion for selecting a center in each round, x would have been selected as the next center instead of c.

(b) A delivery truck starts at location $\mathcal{O} = (0,0)$, and must deliver packages to locations

$$A = (0,6), B = (2,2), C = (6,2), D = (7,1), E = (7,6), F = (7,9), G = (8,4), H = (9,7),$$

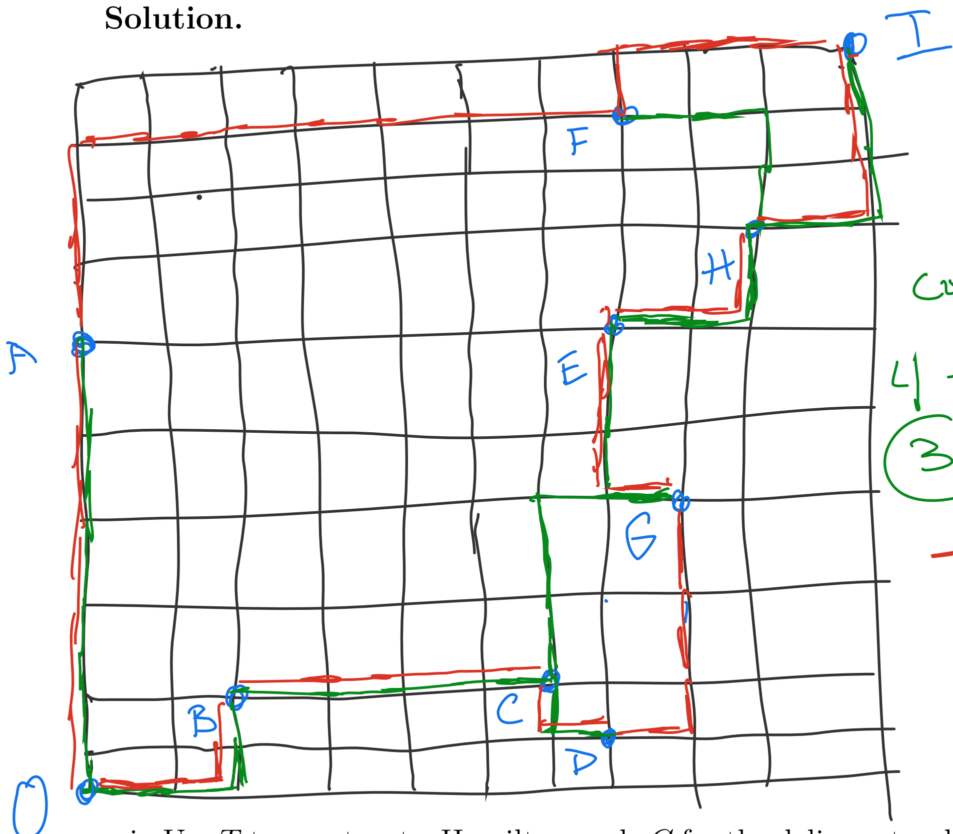
$$I = (10,10)$$

before returning back to $\mathcal{O}$. The distance from one location to another is measured by adding both the number of horizontal and vertical units that separate them. This way of measuring distance is commonly referred to as the **taxi-cab metric**. For example, location D is a distance

$$|7-8| + |1-4| = 4$$

units from location G. Plot the points on grid paper and connect them together with a minimum spanning tree T whose edges are in line with the grid lines.

Solution.



– mst T

cost(T) = 2+3+3+3 4+4+4+4+6 = (33)

– Hamilton Cycle

i. Use T to construct a Hamilton cycle C for the delivery truck using the 2-approximation algorithm for Δ-TSP. When encountering any forks in the tree, give precedence to the rightmost branch, followed by the middle branch, etc..

Solution.

C = O, B, C, D, G, E, H, I, F, A, O

cost(C) = 44 ≤ (2)(33) = 66

ii. Verify that $\text{cost}(C) \le 2\text{cost}(T)$, and hence C has a cost that is no more than twice the cost of the optimal cycle.

Solution.

cost(C) = 44 ≤ (2)(33) = 66

Ratio = 44/33 = 1.3̄

LO9. Answer the following.

(a) Recall the mapping reduction $f(\mathcal{C}) = (S, t)$, where f maps an instance of `3SAT` to an instance of the `Subset Sum (SS)` decision problem and hence establishes the NP-completeness of `SS`. Suppose $\mathcal{C}$ is a satisfiable instance of `3SAT` and has 100 variables and 300 clauses.

i. What is the cardinality of S? Explain.
Solution. $|S| = 2n + 2m = 2(100) + 2(300) = 800$.

ii. One of the numbers in S is y_{72}. How many digits does this number have? Hint: it is fewer than 700.
Solution. $(100 - 71) + 300 = 329$.

iii. If you wanted to write the digits of y_{72} what information would you need to have about $\mathcal{C}$?
Solution. The first digit equals 1, followed by 28 0's. To complete the rest of the digits, for each $1 \leq i \leq 300$, if clause c_i has x_{72} as one of its literals, then digit $29 + i$ equals 1. Otherwise, digit $29 + i$ equals 0.

(b) Recall the Tseytin transformation f : `SAT` $\rightarrow$ `3SAT` that is used to map reduce an instance F of `SAT` to an instance $f(F)$ of `3SAT`.

i. If $F(x_1, x_2, x_3, x_4) = \overline{x}_4 \wedge (x_3 \vee (\overline{x}_1 \wedge x_2))$, then draw its parse tree and provide the Boolean formula that is satisfiability equivalent to F. and is the starting point of the transformation. Hint: the desired formula has double-arrow subformulas.

ii. Show the steps of the Tseytin transformation that converts the Boolean formula

$$y_1 \leftrightarrow (x_3 \vee y_2)$$

to an intance of `3SAT`. Hint: this formula is unrelated to part i.

LO8. Do the following.

(a) Recall that an instance of the `Hamilton Cycle (HC)` decision problem is a simple graph $G = (V, E)$, and the problem is to decide if G has a cycle of length n, where $n = |V|$ and each vertex in V appears in the cycle. Thus, a reasonable certificate for G is a sequence of vertices $C = v_1, v_2, \ldots, v_n$, where each $v_i \in V$.

i. Provide the pseudocode for a verifier program for `HC` that takes as inputs (G, k) and C and returns 1 iff the C is a Hamilton cycle.

ii. Given that the size parameters for `HC` are $m = |E|$ and $n = |V|$, determine the big-O number steps required by your verifier. Defend your answer.

(b) Classify each of the following problems as being in `P`, `NP`, or co-`NP`. Note: at least three correct answers is necessary for passing this part of LO8.

i. An instance of `Set Packing` is a positve integer $k > 0$ and a collection of sets $\mathcal{S}$, where each set $C \in S$ is a subset of $\{1, \ldots, n\}$ for some $n \geq 1$. The problem is to decide if, for any k sets $C_1, \ldots, C_k$ in $\mathcal{S}$, their intersection is nonempty.

ii. An instance of the 50-`Independent Set` decision problem is a simple graph $G = (V, E)$, and the problem is to decide if G has an independent set of size 50.

iii. An instance of the k-`Clustering` decision problem is a set of n points in some metric space (X, d) and an integer $r \geq 0$. The problem is to decide if there is a k-partition of the points so that every set in the partition has a diameter no greater than r.

iv. An instance of the `Bounded Cycles` decision problem is a simple graph $G = (V, E)$ and an integer $k \geq 0$ and the problem is to decide if every cycle of G has a length that does not exceed k.

LO7. Answer the following.

(a) Provide the definition of what it means to be a mapping reduction from decision problem A to decision problem B.

(b) For the mapping reduction f : `Subset Sum` $\rightarrow$ `Set Partition`, provide $f(S, t)$ for the `Subset Sum` instance with $S = \{6, 11, 12, 16, 20, 28\}$ and $t = 35$.

(c) Verify that the reduction is valid for (S, t) in that both (S, t) and $f(S, t)$ are either both positive or negative instances. **Justify your answer**.

LO6. Do the following.

(a) Consider the `2SAT` instance

$$\mathcal{C} = \{(x_1, \overline{x}_2), (x_2, x_3), (\overline{x}_1, x_3), (\overline{x}_4, x_5), (\overline{x}_4, \overline{x}_5), (\overline{x}_3, x_4), (\overline{x}_3, x_6), (\overline{x}_1, \overline{x}_6)\}.$$

i. Draw the implication graph $G_{\mathcal{C}}$.

ii. Perform the `Improved 2SAT` algorithm by computing the necessary reachability sets. Use numerical order (in terms of the variable index) and positive literal before negative literal when choosing the reachability set to compute next. Draw the resulting reduced `2SAT` instance whenever a consistent reachability set is computed. Either provide a final satisfying assignment for $\mathcal{C}$ or indicate why $\mathcal{C}$ is unsatisfiable.

(b) Given `2SAT` instance $\mathcal{C}$, suppose its implication graph has the path $x_2, \overline{x}_3, x_1, \overline{x}_2$. What clauses in $\mathcal{C}$ cause this path, and show that there is no satisfying assignment α that satisfies all of them and for which $\alpha(x_2) = 1$.

LO5. Do the following.

(a) The dynamic-programming algorithm that solves the single-source distances problem in an acyclic graph $G = (V, E)$ defines a recurrence for the function $d(s, v)$, where $s, v \in V$ and s is the source vertex. Provide the recurrence.

(b) Use the recurrence from part a) to determine all the distances from source $s = a$ to each of the other vertices in the graph G for which

$$E = (a, c, 1), (a, b, 4), (a, d, 8), (b, c, 5), (b, e, 5), (b, f, 3), (c, d, 7), (c, e, 8), (d, f, 6), (e, f, 7).$$

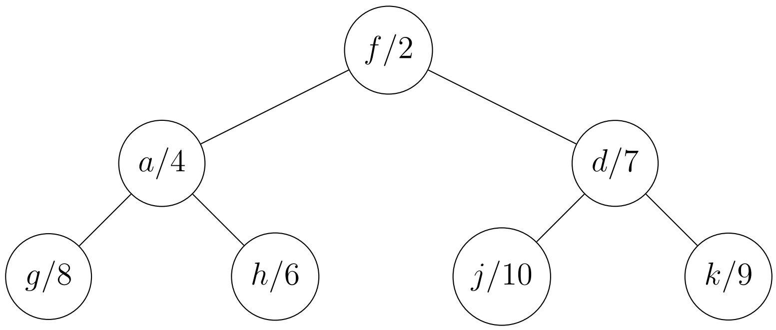
Show all work and make sure to explicitly use the recurrence in your calculations.

LO4. Do the following.

(a) The tree below shows the state of the binary min-heap at the beginning of some round of Prim's algorithm, applied to some weighted undirected graph G. If G has edges

$$(f, d, 3), (a, f, 6), (d, g, 4), (f, g, 3), (f, h, 8), (f, k, 6),$$

then draw a plausible state of the heap at the end of the round.



(b) State the greedy choice that is being made in each round of Prim's algorithm. The weighted edges of a graph $G = (V, E)$ are

$$E = \{(1, 2, 11), (1, 3, 19), (1, 4, 15), (1, 5, 13), (1, 6, 8), (2, 3, 14), (2, 4, 17),$$
$$(2, 5, 16), (2, 6, 22), (3, 4, 10), (3, 5, 12), (3, 6, 19), (4, 5, 15), (5, 6, 20)\}.$$

For each round of Prim's algorithm applied to G, indicate the selection for that round and provide a drawing of the final output of the algorithm. Hint: you do *not* need to use a heap data structure. Also, break any ties by choosing the vertex having least value. For example, if there is a tie beetween vertices 2 and 4, then choose vertex 2.

LO3. Do the following.

(a) The divide-and-conquer FFT algorithm provides an efficient way to simultaneously evaluate a degree-$(n-1)$ polynomial at n different inputs, where n is a power of two. i) What do we call these n inputs? ii) Provide the equation that relates $A(x)$ to the two subproblem polynomials $A_e(x)$ and $A_o(x)$. iii) Based on your answers to i) and ii), what special property must each of these n inputs possess?

(b) If $p(x) = 5 + 3x - 4x^2 - 2x^3$, then compute DFT($p$) using the FFT algorithm. Show the entire recursion tree as was done in the lecture notes.

LO2. Solve each of the following problems.

(a) Provide the arithmetic series that gives the big-O number of steps required by Hoare's `Quicksort` algorithm when the pivot selected is always the second-least element in the array.

(b) Recall that the `Minimum Positive Subsequence Sum (MPSS)` problem admits a divide-and-conquer algorithm that, on input integer array a, requires computing the mpss with respect to all subarrays of a that contains both $a[n/2-1]$ and $a[n/2]$ (the end of a_{left} and the beginning of a_{right}. For

$$a = 48, -37, 29, -33, 51, -64, 46, -34, 45, -36$$

provide the two sorted arrays LeftSorted and RightSorted that are used to compute the values of these subsequence sumps and demonstrate how the minimum positive subsequence sum can be computed via a linear scan of LeftSorted and RightSorted.

LO1. Solve the following problems.

(a) Use the Master Theorem to determine the growth of $T(n)$ if it satisfies the recurrence $T(n) = 70T(n/3) + n^2$. Defend your answer.

(b) Use the substitution method to prove that, if $T(n)$ satisfies

$$T(n) = 2T(n/4) + 5\sqrt[3]{n},$$

then $T(n) = \mathrm{O}(\sqrt{n})$.