Part 1: ACO Comprehensive Exam, Fall 2021

Graduate Algorithms

Professor G likes to relax sitting in the Clough roof garden while watching the Midtown skyline. From his point of view, the buildings are abstracted as rectangles with one side touching the ground (the x axis), and the skyline is the upper contour of the building, or the boundary of the union of these rectangles, minus the ground. Given the locations and heights of all the buildings in Midtown, help professor G compute the skyline formed by these buildings collectively.

Your input is the array `buildings` where $\texttt{buildings}[i] = [left_i, right_i, height_i]$:

- $left_i$ is the x-coordinate of the left edge of the i^{th} building.
- $right_i$ is the x-coordinate of the right edge of the i^{th} building.
- $height_i$ is the height of the i^{th} building.

For example, consider three buildings, $[1, 10]$ with height 4, $[2, 7]$ with height 6, and $[3, 5]$ with height 7:

```
          +-------+
          |   +---|------+
          |   |   |      |
    +-----+---|---+------|--------+
    |     |   |   |      |        |
    |     |   |   |      |        |
    |     |   |   |      |        |
____+_____+___+___+______+________+_
```

The skyline consist of the line segments passing through the points $(1, 0) - (1, 4) - (3, 4) - (3, 7) - (5, 7) - (5, 6) - (7, 6) - (7, 4) - (10, 4) - (10, 0)$

You may assume all buildings are perfect rectangles grounded on an absolutely flat surface at height 0.

1. Show that the skyline consists of $O(n)$ segments.
2. Given the description of the buildings as above, output a description of the skyline.

To earn full credit, you must describe your design in detail and justify its correctness. You should also state and analyse its runtime, aiming for the most efficient algorithm.

Your output must be a list of points $\texttt{contour}= [x_1, y_1], [x_2, y_2], \ldots, [x_k, y_k]$ sorted by their $x-$coordinate. Each point is the left endpoint of some horizontal segment in the skyline except the last point in the list, which always has a y-coordinate 0 and is used to mark the skyline's termination where the rightmost building ends. There must be no consecutive horizontal lines of equal height in the output skyline.

Example:
Input: $\texttt{buildings}= [[0, 2, 3], [2, 5, 3]]$.
Output: $[[0, 3], [5, 0]]$.

Example:
Input: $\texttt{buildings}= [[2, 9, 10], [3, 7, 15], [5, 12, 12], [15, 20, 10], [19, 24, 8]]$.
Output: $[[2, 10], [3, 15], [7, 12], [12, 0], [15, 10], [20, 8], [24, 0]]$.

Solution:

1. [2points] Consider the $2n$ coordinates on the $x-$axis limiting the buildings: $\{x_1, x_2, \ldots, x_{2n}\}$. The claim follows if one can show that only $O(1)$ many intervals in the skyline have $x-$coordinate at x_i for $1 \leq i \leq 2n$. Clearly there can be only one vertical segment at x_i. For the horizontal direction, note that distinct intervals have distinct $y-$coordinate (otherwise they would merge into one) and hence, only two such intervals can "share" an $x-$coordinate. This concludes the proof.

2. [8points] Sort the $x-$coordinates from left to right. We loop through these values keeping track of the visible horizon in the skyline at each x_i. When at x_i, the horizon will result from the tallest building with left edge $\leq x_i$ and right edge $> x_i$. This can be implemented efficiently using a priority queue with the following update: label the top corners of each building as $(left_i, height_i, \texttt{up})$ and $(right_i, height_i, \texttt{down})$. When processing x_i, remove all (x_i, $height_i$, `down`) and include all (x_i, $height_i$, `up`). The horizon at x_i is given by the highest value on the queue. Finally, one must merge consecutive segments with the same height.

   This design runs in time $O(n \log(n))$. The sorting steps achieves this time using `MergeSort`. The pre-processing (`up/down` events) takes linear time. Each building is added and removed once, thus the final loop also runs in linear time.

Note: a suboptimal solution using Dynamic Programming will receive partial credit. The student must realize that the worst case performance of DP is $O(n^2)$.

Graph Theory

Problem: Let T be a tree on $t \geq 1$ vertices. Prove the following statements.

(1) If G is a graph with minimum degree at least t then, for any $v \in V(G)$, G has a subgraph containing v and isomorphic to T.

(2) If G is graph with at least $t|V(G)|$ edges then G has a subgraph isomorphic to T.

Solution. To prove (1), we apply induction on t. The assertion clearly holds when $t = 1$. Now assume that $t \geq 2$ and that (1) holds for trees on fewer than t vertices. Let w be a leaf of T. Consider the tree $T - w$ on $t - 1$ vertices. Note that $\delta(G - v) \geq t - 1$, and let u be a neighbor of v in G. By induction hypothesis, we see that $G - v$ contains a subgraph H such that $u \in V(H)$ and H is isomorphic to $T - w$. Now, $H + \{u, uv\}$ gives the desired subgraph of G.

For (2), we note that any graph G with $e(G) \geq t|V(G)|$ contains a subgraph H with $\delta(H) \geq t$. (This can be shown by repeatedly deleting vertices of degree smaller than t, or by taking a minimum counterexample and deriving a contradiction.) Thus, we derive (2) from (1). (Note: (2) implies that the Turán number $ex(n, T) \leq tn$.)

Linear Inequalities

Let $S_n^{\text{even}} := \{x \in \{0,1\}^n : x \text{ has even number of 1's}\}$. In this problem, you will derive two different formulations for convex hull of S_n^{even}.

1. (2 points) Let $S_n^k = \{x \in \{0,1\}^n : x \text{ has exactly k 1's}\}$. Show that

$$conv(S_n^k) = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = k, 0 \le x_i \le 1 \;\; \forall 1 \le i \le n\}$$

2. (3 points) Let Q be the polyhedron (in higher dimensions $(x, y, \lambda) \in \mathbb{R}^n \times \mathbb{R}^{n \times \lfloor \frac{n}{2} \rfloor} \times \mathbb{R}^{\lfloor \frac{n}{2} \rfloor}$) defined by the following set of inequalities.

$$\begin{aligned}
x_i &= \sum_{k:k \text{ is even}} y_i^k && \forall 1 \le i \le n \\
\sum_{i=1}^{n} y_i^k &= k\lambda_k && \forall k \text{ even} \\
\sum_{k:k \text{ even}} \lambda_k &= 1 && \\
y_i^k &\le \lambda_k && \forall 1 \le i \le n, k \text{ even} \\
y_i^k &\ge 0 && \forall 1 \le i \le n, k \text{ even} \\
\lambda_k &\ge 0 && \forall k \text{ even}
\end{aligned}$$

Show that $conv(S_n^{\text{even}}) = proj_x(Q)$.

3. (5 points) The above formulation gives the description of $conv(S_n^{\text{even}})$ in higher dimensions. Now we give a description of $conv(S_n^{\text{even}})$ in the x space. Let

$$P = \{x \in \mathbb{R}^n : \sum_{i \in O} x_i - \sum_{i \notin O} x_i \le |O| - 1, \forall O \subseteq [n], |O| \text{ odd}, 0 \le x \le 1\}.$$

To show $P = conv(S_n^{\text{even}})$, prove the following claims.

   (a) (2 points) For any $c \in \mathbb{R}^n$, give a simple characterization of $\max\{c^T x : x \in S_n^{\text{even}}\}$.

   (b) (3 points) Show that $\max\{c^T x : x \in P\} = \max\{c^T x : x \in S_n^{\text{even}}\}$. (Hint: Use Duality).

Solution:

1. Observe that constraint matrix of the single constraint in $\{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = k, 0 \le x_i \le 1 \;\; \forall 1 \le i \le n\}$ is totally unimodular. Thus the polytope is integer. Since integer points are exactly S_n^k, we have the statement.

2. Observe that $S_n^{\text{even}} = \cup_{k \text{ is even}} S_n^k$. Let $x \in S_n^k$ for some k even. Then set $\lambda_k = 1$ and $\lambda_j = 0$ for all $j \ne k$. Moreover set $y^k = x$ and $y^j = 0$ for all other j. It is easy to check that $(x, y, \lambda) \in Q$. Thus, we have $proj_x(Q) \supseteq \cup_{k \text{even}} S_n^k = S_n^{\text{even}}$ and therefore, $proj_x(Q) \supseteq conv(S_n^{\text{even}})$.

For the other direction, let $(x, y, \lambda) \in Q$. For every $\lambda_k > 0$, we have $z^k := \frac{y^k}{\lambda_k} \in conv(S_n^k) \subseteq conv(S_n^{\text{even}})$. But we have $x = \sum_{k:k \text{ even}, \lambda_k > 0} \lambda_k z^k$, i.e. x is a convex combination of points in $conv(S_n^{\text{even}})$ and thus $x \in conv(S_n^{\text{even}})$ as required.

3. (a) Now, for any $c \in \mathbb{R}^n$, we compute $\max\{c^T x : x \in S_n^{\text{even}}\}$. For simplicity, we assume that $c_1 \geq c_2 \geq c_l \geq 0 > c_{l+1} \geq \ldots \geq c_n$ for some l. More generally, the indices can be renumbered to sort them in this order. We have three cases. The optimality in all three cases is trivial to check.

   i. l is even. Then $x_i = 1$ for $1 \leq i \leq l$ and 0 otherwise is an optimal solution.

   ii. l is odd and $c_l + c_{l+1} \geq 0$. Then $x_i = 1$ for $1 \leq i \leq l+1$ and 0 otherwise is an optimal solution.

   iii. l is odd and $c_l + c_{l+1} < 0$. Then $x_i = 1$ for $1 \leq i \leq l-1$ and 0 otherwise is an optimal solution.

   (b) Now we show that in each of the three cases the value of the LP relaxation matches the integral optimum. First check that $P \supseteq S_n^{\text{even}}$. Indeed any x with even number of ones satisfies the constraint for each odd sized O. Thus it is enough to show that $\max\{c^T x : x \in P\} \leq \max\{c^T x : x \in S_n^{\text{even}}\}$ to show equality. To compute the value of LP relaxation, we compute its dual.

   The primal is given by

$$\begin{aligned} &\max c^T x \\ &\quad s.t. \\ &\sum_{i \in O} x_i - \sum_{i \notin O} x_i \leq |O| - 1, && \forall O \subseteq [n], |O| \text{ odd}, \\ &x_i \leq 1 && \forall 1 \leq i \leq n \\ &x_i \geq 0 && \forall 1 \leq i \leq n. \end{aligned}$$

   Let $\mathcal{O}$ denote the set of all odd sized sets of $\{1, \ldots, n\}$. The dual is given by

$$\begin{aligned} &\min \sum_{O \in \mathcal{O}} (|O| - 1) y_O + \sum_{i=1}^n z_i \\ &\quad s.t. \\ &\sum_{O \in \mathcal{O}: i \in O} y_O - \sum_{O \in \mathcal{O}: i \notin O} y_O + z_i \geq c_i && \forall 1 \leq i \leq n \\ &y_O \geq 0 && \forall O \in \mathcal{O} \\ &z_i \geq 0 && \forall 1 \leq i \leq n \end{aligned}$$

   We now give a feasible solution to the dual of the objective that equals $\max\{c^T x : x \in S_n^{\text{even}}\}$. This will show that $\max\{c^T x : x \in P\}$ is at most the dual objective (by weak duality) and therefore at most $\max\{c^T x : x \in S_n^{\text{even}}\}$. The dual solution will vary depending on the case. Recall, we assume $c_1 \geq c_2 \geq c_l \geq 0 > c_{l+1} \geq \ldots \geq c_n$. (To define the dual solution, we can use the complementary slackness condition with the candidate optimal solution. This gives us hint on which dual variables can take non-zero values.)

   i. l is even. Then set $z_i = c_i$ for $1 \leq i \leq l$ and 0 otherwise. We set $y = 0$. It is easy to check feasibility and the objective value.

ii. l is odd and $c_l + c_{l+1} \geq 0$. Then $z_i = c_i + c_{l+1}$ for $1 \leq i \leq l$ and 0 otherwise. Define $y_{\{1,\ldots,l\}} = -c_{l+1}$ and $y_O = 0$ for all other $O \in \mathcal{O}$. Observe that (y, z) is feasible and its objective is exactly the same as objective of the primal $\sum_{i=1}^{l+1} c_i$.

iii. l is odd and $c_l + c_{l+1} < 0$. Then $z_i = c_i - c_l$ for $1 \leq i \leq l-1$ and 0 otherwise. Define $y_{\{1,\ldots,l\}} = c_l$ and $y_O = 0$ for all other $O \in \mathcal{O}$. Observe that (y, z) is feasible and its objective is exactly the same as objective of the primal $\sum_{i=1}^{l-1} c_i$.

Since the equality holds for every $c \in \mathbb{R}^n$, we have $P = conv(S_n^{\text{even}})$ as claimed.

Part 2: ACO Comprehensive Exam, Fall 2021

Design and Analysis of Algorithms

Consider the following game played between two players A and B on $n \geq 2$ integer counters:

1. Initially all the counters start at 0.
2. At each step, A can take two of the counters, increment each by 1.
3. At each step, B can take a **single** counter, and decrease it by at most 10, without making it negative.

For this process, show the following:

1. (4 points) No matter how B plays, A can make some counter's value at least $\lfloor \log_2 n \rfloor$.
2. (6 points) For an unbounded game length, no matter how A plays, B can ensure that no counter's value exceeds $100 \log_2 n$.

Solution:

We first show the first part about the lower bound on the value that A can ensure to occur.

If A starts with $2k$ numbers of value x, A can guarantee getting back k numbers of value $x+1$. The strategy is to pair these numbers arbitrarily, and increase both from a pair by 1. This takes k steps, during which B can decrease at most k of these $2k$ counters. So at least k of the numbers have value at least $x+1$.

Repeatedly applying this argument to the n counters initially at 0 gives that A can guarantee at least one counter at value $\lfloor \log_2 n \rfloor$.

We now give a proof of the second part based on potential functions. Define the potential function

$$\sum_i 1.1^{x_i}.$$

We will show that if this potential is more than $(1.1)^{11}n$, then B can ensure that it does not increase by repeatedly decreasing the largest value. It then follows that any x_i is at most

$$\log_{1.1}\left((1.1)^{11}\, n\right) < 11 + 25 \log n < 50 \log n$$

for $n >= 2$.

There are two cases to consider:

(1) If the potential is less than $(1.1)^{10}n$, then in one step it can increase by at most a factor of 1.1, and hence it is always at most $(1.1)^{11}n$.

(2) If the potentail is more than $(1.1)^{10}n$. Let the entry of largest value among $x_1 \ldots x_n$ be y. Without loss of generality, we have $1.1^y \geq 1.1^{10}$ by the pigeon hole principle. Then A can increase the potential function by at most

$$2 \cdot \left((1.1)^{y+1} - 1.1^y\right) = 0.2(1.1)^y.$$

On the other hand, because $(1.1)^y \geq (1.1)^{10}$, we have $y \geq 10$. So the decrease is at least

$$(1.1)^y - (1.1)^{y-10} = \left(1 - 1.1^{-10}\right) 1.1^y \geq 0.5 \cdot 1.1^y.$$

So in this case, the potential function can only decrease.

Probabilistic Combinatorics

Prove that for any $\epsilon > 0$ there is $\Delta_0 > 0$ such that the following holds: if G is a Δ-regular graph (i.e., in which every vertex has degree exactly Δ) with $\Delta \geq \Delta_0$, then there is a partition $V(G) = V_1 \cup V_2$ of the vertex set of G such that every vertex of G has between $(1-\epsilon)\Delta/2$ and $(1+\epsilon)\Delta/2$ many neighbors in each V_i.

Solution: Pick a random subset $V_1 \subseteq V(G)$ by including each $v \in V(G)$ independently with probability $1/2$, and set $V_2 := V(G) \setminus V_1$. Let $\mathcal{B}_v$ denote the 'bad' event that v has fewer than $(1-\epsilon)\Delta/2$ or more than $(1+\epsilon)\Delta/2$ many neighbors in V_1. Let

$$\mathcal{B} := \bigcup_{v \in V(G)} \mathcal{B}_v.$$

Using standard Chernoff bounds, it is easy to see that for some $c = c(\epsilon) > 0$ we have

$$\Pr(\mathcal{B}_v) \leq 2 \cdot e^{-c\Delta} =: p.$$

Note that $\mathcal{B}_v$ is mutually independent of all other events $\mathcal{B}_w$ except for at most Δ^2 many (corresponding to events $\mathcal{B}_u$ for which u shares a neighbor with v), say. Since $e^{-cx}(x^2+1) \to 0$ as $x \to \infty$, we infer that for sufficiently large $\Delta \geq \Delta_0 = \Delta_0(c)$ we have

$$ep(\Delta^2 + 1) = 2e \cdot e^{-c\Delta}(\Delta^2 + 1) < 1.$$

Hence (the symmetric form of) the Lovász Local Lemma implies that

$$\Pr\big(\neg\mathcal{B}) = \Pr\Big(\bigcap_{v \in V(G)} \neg\mathcal{B}_v\Big) > 0.$$

By the probabilistic method, there thus exists a choice of $V_1 \subseteq V(G)$ for which $\neg\mathcal{B}$ holds, i.e., such that every vertex $v \in V(G)$ has between $(1-\epsilon)\Delta/2$ and $(1+\epsilon)\Delta/2$ many neighbors in V_1. Since G is Δ-regular, using $V_2 = V(G) \setminus V_1$ this also implies that every vertex $v \in V(G)$ has between $(1-\epsilon)\Delta/2$ and $(1+\epsilon)\Delta/2$ many neighbors in V_2, completing the proof.

Combinatorial Optimization

(i) Let f be a set function defined on the subsets of S. Let $T \subseteq S$, and define a new set-function as follows. Construct a new dummy element a_t (so that $a_t \notin S$), and set $S' = S \cup \{a_T\}$. Define an extension of f into a new set function f':

$$f'(X) = f(X), \tag{1}$$
$$f'(X \cup \{a_T\}) = f(X \cup T), \text{ for } X \subseteq S. \tag{2}$$

Here, the function f is said to be extended parallel to T. Construct an example so that f' is not submodular. (4 points)

(ii) Let an element $a \in S$ be called *increasing* with respect to f if $f(X \cup \{a\}) \geq f(X)$ for all $X \subseteq S$. Show that f' defined above is submodular if and only if every element of T is increasing with respect to f. (6 points)

Solution.

(i) There can be many such examples. One such example is: consider $f(1,2) = 0$, $f(2) = 3$, $f(1) = 1.5$, $f(\emptyset) = 0$. Note that f is submodular since $f(1,2) + f(\emptyset) \leq f(2) + f(1)$. Consider $T = \{1\}$, $f(1,2) < f(2)$, and note that T is not increasing with respect to f. Consider the parallel extension of f with respect to $T = \{1\}$ as $f' : 2^{\{1,2,a_T\}} \to \mathbb{R}$, where $f'(1,2) = 0, f'(2) = 3, f'(1) = 1.5, f'(\emptyset) = 0, f'(1,2,a_T) = 0, f'(1,a_T) = 1.5, f'(2,a_T) = 0, f'(a_T) = 1.5$. But f' violates submodularity since $f'(1,2) + f'(2,a_T) = 0 < 3 = f'(1,2,a_T) + f'(2)$.

(ii) Consider the parallel extension of $f : 2^S \to \mathbb{R}$ with respect to a set T so that every element of T is increasing with respect to f. Due to the latter property, we can show using an inductive argument that:

$$f(X \cup T) \geq f(X \cup (T \setminus W)) \geq f(X), \text{ for any } W \subseteq T. \tag{3}$$

To show submodularity of f', the only non-trivial case is to consider two subsets $X \cup \{a_T\}$ and Y so that $X, Y \subseteq S$:

$$f'(X \cup a_T) + f'(Y) - f'(X \cup Y \cup a_T) - f'(X \cap Y) \tag{4}$$
$$= f(X \cup T) + f(Y) - f(X \cup Y \cup T) - f(X \cap Y) \tag{5}$$
$$\geq f((X \cup T) \cap Y) - f(X \cap Y) \text{ (using submodularity)} \tag{6}$$
$$= f((X \cap Y) \cup (T \cap Y)) - f(X \cap Y) \geq 0. \text{ (using (3))} \tag{7}$$