Computational Complexity - Homework I (Solutions)

K. Subramani
LCSEE,
West Virginia University,
Morgantown, WV
{ksmani@csee.wvu.edu}

1 Problems

1. Let f denote a convex function. It follows that for any x_1, x_2 in the domain and $\lambda \in [0, 1]$,

$$f(\lambda \cdot x_1 + (1-\lambda) \cdot x_2) \leq \lambda \cdot f(x_1) + (1-\lambda) \cdot f(x_2).$$

Let X denote a random variable. Jensen's inequality states that

$$\mathbf{E}[f(X)] \geq f(\mathbf{E}[X])$$

Prove the above inequality for the case when X takes on precisely two values x_1 and x_2 with probabilities p and $(1-p)$ respectively. Argue that Jensen's inequality becomes an equality, when $f(X) = a \cdot X + b$, where a and b are constants.

Solution: Since X takes precisely two values x_1 and x_2 with probabilities p and $(1-p)$ respectively, we have that $f(X)$ takes two values $f(x_1)$ and $f(x_2)$ with the same probabilities. Thus,

$$f(\mathbf{E}[X]) = f(p \cdot x_1 + (1-p) \cdot x_2) \leq p \cdot f(x_1) + (1-p) \cdot f(x_2) = \mathbf{E}[f(X)].$$

The first and the third equalities follow from the definitions of $f(\mathbf{E}[X])$ and $\mathbf{E}[f(X)]$, respectively, and the second inequality follows from the convexity of f. Now, assume that $f(X) = a \cdot X + b$, then

$$\begin{aligned} f(\mathbf{E}[X]) &= f(p \cdot x_1 + (1-p) \cdot x_2) \\ &= a \cdot (p \cdot x_1 + (1-p) \cdot x_2) + b \\ &= a \cdot p \cdot x_1 + a \cdot (1-p) \cdot x_2 + b \\ &= (a \cdot x_1 + b) \cdot p + (a \cdot x_2 + b) \cdot (1-p) \\ &= f(x_1) \cdot p + f(x_2) \cdot (1-p) \\ &= \mathbf{E}[f(X)]. \end{aligned}$$

$\square$

2. Let $E_1, E_2, \ldots E_n$ denote a set of events. Argue that

$$\mathbf{P}[\overline{\cup_{i=1}^n E_i}] = \sum_{T \subseteq \{1,2,\ldots,n\}} (-1)^{|T|} \mathbf{P}[\cap_{i \in T} E_i]$$

Solution:

We assume that $\mathbf{P}[\cap_{i \in \emptyset} E_i] = 1$.

Accordingly, we can rewrite the given inequality as:

$$\mathbf{P}[\cup_{i=1}^n E_i] = \sum_{\emptyset \neq T \subseteq \{1,2,\ldots,n\}} (-1)^{|T|-1} \cdot \mathbf{P}[\cap_{i \in T} E_i].$$

In our proof we will use the following equality:

$$\mathbf{P}[A \cup B] = \mathbf{P}[A] + \mathbf{P}[B] - \mathbf{P}[A \cap B],$$

where A and B are arbitrary events.

Now, let us prove the main equality by induction on n.

Clearly, it holds for $n = 1$, since in this case, the inequality reduces to the identity $\mathbf{P}[E_1] = \mathbf{P}[E_1]$.

Assume that it is true for $n = (k-1)$.

Consider an arbitrary collection of k events $E_1, ..., E_k$.

Observe that,

$$\begin{aligned}
\mathbf{P}[\cup_{i=1}^k E_i] &= \mathbf{P}[(\cup_{i=1}^{k-1} E_i) \cup E_k] \\
&= \mathbf{P}[\cup_{i=1}^{k-1} E_i] + \mathbf{P}[E_k] - \mathbf{P}[(\cup_{i=1}^{k-1} E_i) \cap E_k] \\
&= \mathbf{P}[\cup_{i=1}^{k-1} E_i] + \mathbf{P}[E_k] - \mathbf{P}[(\cup_{i=1}^{k-1} (E_i \cap E_k)] \\
&= \sum_{\emptyset \neq T} (-1)^{|T|-1} \cdot \mathbf{P}[\cap_{i \in T} E_i] + \mathbf{P}[E_k] - \sum_{\emptyset \neq T} (-1)^{|T|-1} \cdot \mathbf{P}[\cap_{i \in T} (E_i \cap E_k)]
\end{aligned}$$

The above step follows from the inductive hypothesis.

Also, note that in the above equation T varies over all subsets of $\{1, 2, \ldots (k-1)\}$.

$$= \sum_{\emptyset \neq S} (-1)^{|S|-1} \cdot \mathbf{P}[\cap_{i \in S} E_i].$$

In the final equation S varies over all subsets of $\{1, 2, \ldots k\}$.

Let us justify the last step. Partition $\sum_{\emptyset \neq S} (-1)^{|S|-1} \cdot \mathbf{P}[\cap_{i \in S} E_i]$ into two sums by considering those subsets S of $\{1, 2, \ldots k\}$ which contain k and those which do not.

Clearly, if these subsets do not contain k, then the corresponding terms are going to appear in

$$\sum_{\emptyset \neq T} (-1)^{|T|-1} \cdot \mathbf{P}[\cap_{i \in T} E_i].$$

If $S = \{k\}$, then

$$\sum_{\emptyset \neq S} (-1)^{|S|-1} \cdot \mathbf{P}[\cap_{i \in S} E_i] = \mathbf{P}[E_k].$$

Finally, if S is any subset that contains k, and is different from $\{k\}$, then $S - \{k\}$ is non-empty, hence the corresponding term will appear in

$$\sum_{\emptyset \neq S} (-1)^{|S|-1} \cdot \mathbf{P}[\cap_{i \in S} (E_i \cap E_k)] = -\sum_{\emptyset \neq T} (-1)^{|T|-1} \cdot \mathbf{P}[\cap_{i \in T} (E_i \cap E_k)]$$

where $T = S \backslash \{k\}$. $\square$

3. Prove that in any finite graph, the number of vertices with odd degrees is even,

   Solution: Let G be a graph, and for a vertex v of G, let $d(v)$ be the degree of the vertex v in G. Consider the subsets V_E and V_O of $V(G)$ defined as follows:

$$V_E = \{v \in V(G) : d(v) \text{ is even}\}, V_O = \{v \in V(G) : d(v) \text{ is odd}\}.$$

Clearly, $V(G) = V_E \cup V_O$ and $V_E \cap V_O = \emptyset$, and we need to show that $|V_O|$ is odd. Observe that:

$$\sum_{v \in V_O} d(v) + \sum_{v \in V_E} d(v) = \sum_{v \in V(G)} d(v) = 2 \cdot |E(G)|.$$

The last equality follows from the observation that each edge has two end-vertices, hence it is counted twice in the sum $\sum_{v \in V(G)} d(v)$. We have that:

$$\sum_{v \in V_O} d(v) = 2 \cdot |E(G)| - \sum_{v \in V_E} d(v).$$

Observe that the right side of this equation is even, hence $\sum_{v \in V_O} d(v)$ is even as well. Taking into account that for each $v \in V_O$, $d(v)$ is odd, we conclude that $|V_O|$ is even (the sum of odd numbers can be even if and only if the number of summands is even). □

4. Consider two variants of the Hamilton cycle problem: In Variant I, you are required to provide a "yes/no" answer to the question: Does the graph G have a Hamilton cycle? In Variant II, you are required to actually provide the Hamilton cycle in G, if one exists. Assume that an oracle for Variant I exists. Argue that by querying this oracle at most a polynomial number of times (polynomial in the size of G), we can solve Variant II.

   Solution: Let $E(G) = \{e_1, ..., e_q\}$. We will show that with at most $(q+1)$ queries of the oracle for Variant I, we can construct the actual Hamiltonian cycle of G, if it exists. Since $q \leq \frac{n \cdot (n+1)}{2}$, this will establish the polynomiality of our approach.

   Consider the following algorithm:

Function HAMILTON-CYCLE-SEARCH($G = \langle V, E \rangle, H$)
1: **if** (HAMILTON-CYCLE-DECISION(G) = "no") **then**
2:     **return**("No Hamilton cycle exists").
3: **else**
4:     HAMILTON-CYCLE-AUX($G = \langle V, E \rangle, H$).
5: **end if**

Algorithm 1.1: Turing reduction of the Hamilton cycle search problem to the Hamilton cycle decision problem

Function HAMILTON-CYCLE-AUX($G = \langle V, E \rangle, H$)
1: **if** ($|V| = 1$) **then**
2:     **return**(H).
3: **else**
4:     Pick some edge $e \in E$.
5:     **if** (HAMILTON-CYCLE-DECISION($G = \langle V, E - \{e\} \rangle$) = "yes") **then**
6:         **return**(HAMILTON-CYCLE-AUX($G = \langle V, E - \{e\} \rangle, H$)).
7:     **else**
8:         $H = H \cup \{e\}$.
9:         Contract edge e reducing the number of vertices in V by 1. If parallel edges are created, remove all except one. Modify V and E accordingly.
10:         **return**(HAMILTON-CYCLE-AUX($G = \langle V, E \rangle, H$)).
11:     **end if**
12: **end if**

Algorithm 1.2: The Auxiliary procedure

It is important to note that f Algorithm 1.2 does not return a unique Hamilton cycle. In particular, if a graph has more than one Hamilton cycle, the cycle that is returned depends upon the implementation of Step 4 of this algorithm.

□

5. Consider the Fibonacci series, which is defined as follows:

$$\begin{aligned} F(1) &= 1 \\ F(2) &= 1 \\ F(n) &= F(n-1) + F(n-2), n \geq 3 \end{aligned}$$

Argue that $F(n) = \Theta(\psi^n)$, where $\psi = \frac{1+\sqrt{5}}{2}$.

What is the complexity of checking whether a given number n is a Fibonacci number?

Solution: In order to show that $F(n) = \Theta(\psi^n)$, we need to show that there are $C_1 > 0$, $C_2 > 0$, such that for each $n \geq 1$, one has

$$C_1 \cdot \psi^n \leq F(n) \leq C_2 \cdot \psi^n.$$

Choose $C_1 = (\frac{2}{1+\sqrt{5}})^2$ and $C_2 = \frac{2}{1+\sqrt{5}}$. It can checked directly that these inequalities are satisfied for $n = 1, 2$. Now, we will show by induction that they are going to hold for all $n \geq 3$. By the choice of C_1 and C_2, the inequalities hold for $n = 1, 2$ so we have the basis of the induction. Consider $F(n)$, $n \geq 3$. By induction, we have that

$$\begin{aligned} F(n) &= F(n-1) + F(n-2) \\ &\geq C_1 \cdot (\frac{1+\sqrt{5}}{2})^{n-1} + C_1 \cdot (\frac{1+\sqrt{5}}{2})^{n-2} \\ &= C_1 \cdot (\frac{1+\sqrt{5}}{2})^n \cdot (\frac{2}{1+\sqrt{5}} + (\frac{2}{1+\sqrt{5}})^2) \\ &= C_1 \cdot (\frac{1+\sqrt{5}}{2})^n \cdot \frac{6+2\cdot\sqrt{5}}{(1+\sqrt{5})^2} \\ &= C_1 \cdot (\frac{1+\sqrt{5}}{2})^n, \end{aligned}$$

and similarly

$$\begin{aligned} F(n) &= F(n-1) + F(n-2) \\ &\leq C_2 \cdot (\frac{1+\sqrt{5}}{2})^{n-1} + C_2 \cdot (\frac{1+\sqrt{5}}{2})^{n-2} \\ &= C_2 \cdot (\frac{1+\sqrt{5}}{2})^n \cdot (\frac{2}{1+\sqrt{5}} + (\frac{2}{1+\sqrt{5}})^2) \\ &= C_2 \cdot (\frac{1+\sqrt{5}}{2})^n \cdot \frac{6+2\cdot\sqrt{5}}{(1+\sqrt{5})^2} \\ &= C_2 \cdot (\frac{1+\sqrt{5}}{2})^n. \end{aligned}$$

In order to check whether a given number n is a Fibonacci number, we need to test whether there exists $k \geq 1$, such that $F(k) = n$. We can do this by the following algorithm: we start with $F(1) = F(2) = 1$, and compute $F(3), F(4), ..., F(k)$, such that $F(k) \leq n$ and $F(k+1) > n$. Clearly, If $F(k) = n$, n is a Fibonacci number, otherwise it is not. The number of additions is

$$k - 2 \leq \Theta(\log(n)).$$

This follows from the following reasoning: as $F(k) = \Theta(\psi^k) \leq n$, we have that $k \leq \Theta(\log(n))$.

□