Jared Yu

Homework 2

1. Problem 1

CLRS 34.3-2: Show that the $\leq_P$ relation is a transitive relation on languages. That is, show that if $L_1 \leq_P L_2$ and $L_2 \leq_P L_3$, then $L_1 \leq_P L_3$.

<u>Ans:</u>

If $L_1 \leq_P L_2$, then there exists a polynomial-time computable function $f: \{0,1\}^* \to \{0,1\}^*$ such that for all $x \in \{0,1\}^*$,

$$x \in L_1 \text{ if and only if } f(x) \in L_2,$$

where the function f is the reduction function, and a polynomial-time algorithm F that computes f is a reduction algorithm. Moreover, if $x \notin L_1$, then $f(x) \notin L_2$.

Similarly, if $L_2 \leq_P L_3$, then there exists a polynomial-time computable function $f: \{0,1\}^* \to \{0,1\}^*$ such that for all $x \in \{0,1\}^*$,

$$x \in L_2 \text{ if and only if } f(x) \in L_3$$

(the function f does not need to be the same in both cases). Also, if $x \notin L_2$, then $f(x) \notin L_3$.

Let $L_1 \leq_P L_2$ and $L_2 \leq_P L_3$. Then $L_1 \leq_P L_2$ and $L_2 \leq_P L_3$ implies that for all $x \in \{0,1\}^*$, $x \in L_1$ iff (if and only if) $f(x) \in L_2$. Using the notation that $y = f(x)$, it is also implied that for all $y \in \{0,1\}^*$, $y \in L_2$ iff $g(y) \in L_3$, where $z = g(y)$ is a reduction function similar to $f(x)$.

Since $\forall\, x \in L_1\ \exists\, y \in L_2$ and $\forall\, y \in L_2\ \exists\, z \in L_3$, then it must be that $\forall\, x \in L_1\ \exists\, w \in L_3$, where $w = g\big(f(x)\big) = z$. This means that $L_1 \leq_P L_3$, so $\leq_P$ is a transitive relation on languages.

2. Problem 2

Recall the definition of a complete graph K_n is a graph with n vertices such that every vertex is connected to every other vertex. Recall also that a clique is a complete subset of some graph. The *graph coloring problem* consists of assigning a color to each of the vertices of a graph such that adjacent vertices have different colors and the total number of colors used is minimized. We define the *chromatic number* of a graph G to be this minimum number of colors required to color graph G. Prove that the chromatic number of a graph G is no less than the size of the maximal clique of G.

References:

https://en.wikipedia.org/wiki/Clique_(graph_theory)

https://math.stackexchange.com/questions/140170/proof-about-clique-number-adjacency-number-and-chromatic-number

http://mathworld.wolfram.com/ChromaticNumber.html

https://www.youtube.com/watch?v=gjRcTH2p65c

https://pwp.gatech.edu/math3012openresources/lecture-videos/lecture-9/

Ans:

The goal is to prove that the chromatic number of a graph G is no less than the size of a maximal clique within G. Denote the chromatic number of a graph G as $\chi(G)$. A maximal clique, $C_{maximal}$, of graph G is a clique such that all the vertices are connected to every other vertex similar to a complete graph K_n, and that it cannot be extended by adding an additional vertex.

Let there be a graph G, which itself contains at least one maximal clique, $C_{maximal}$. The chromatic number is derived by coloring the adjacent vertices within G different colors. Within just $C_{maximal}$, the chromatic number would already be equal to the number of vertices within $C_{maximal}$. With the possibility of additional maximal cliques of different sizes within G, then the chromatic number can increase even more.

Base case: $k = 2$

Let there be a graph G with a maximal clique of size $k = 2$. Then the chromatic number of the graph must be at least equal to the size of the maximal clique, or $\chi(G) \geq 2$.

Induction step: Assume $k = m$ is true.

Then a graph G with a maximal clique of size $k = m$ has a chromatic number at least equal to the size of the maximal clique, or $\chi(G) \geq m$.

(Induction Hypothesis)

Now it must be shown that the same logic holds true for $k = m + 1$.

If a graph G has a maximal clique of size $k = m + 1$, it has been shown that a maximal clique of size $k = m$ required m-colors for each of the adjacent vertices within the maximal clique. Then for a maximal clique of size $k = m + 1$, an additional color, or a total of $(m + 1)$-colors would be required to color each of the adjacent vertices within the maximal clique a different color. So, for the graph G with a maximal clique of size $k = m + 1$, $\chi(G) \geq (m + 1)$.

Therefore, the reasoning for the chromatic number of a graph G with a maximal clique of size $m + 1$ holds true. ■

3. Problem 3

Suppose you're helping to organize a summer sports camp, and the following problem comes up. The camp is supposed to have at least one counselor who's skilled at each of the n sports covered by the camp (baseball, volleyball, and so on). They have received job applications from m potential counselors. For each of the n sports, there is some subset of the m applicants qualified in that sport. The question is "For a given number $k < m$, is it possible to hire at most k of the counselors and have at least one counselor qualified in each of the n-sports?" We call this the *Efficient Recruiting Problem*. Prove that the Efficient Recruiting Problem is NP-complete.

References:

https://studylib.net/doc/6833222/a--a--page-121-of-the-textbook-states-that-the-search-ver...

https://en.wikipedia.org/wiki/Vertex_cover

Ans:

Using $Lemma$ 34.8 from the textbook:

If L is a language such that $L' \leq_P L$ for some $L' \in NPC$, then L is NP-hard. If, in addition, $L \in NP$, then $L \in NPC$.

In problem 3, it will be said that the language L is the *Efficient Recruiting Problem* (ERP). Also, according to the professor during office hours it could be assumed that $L \in NP$ for this case. Therefore, to show that ERP is NP-complete, it only needs to be shown that $L' \leq_P L$ for some $L' \in NPC$. In this proof, the language L' that will be used is $VERTEX\text{-}COVER$ as described below. Rather than proving directly that $VERTEX\text{-}COVER$ is in NPC, this proof will assume this is the case based on the result of $Theorem$ 34.12 on p.1090 in the textbook.

As a language, define (directly from the textbook, p.1090)

$$VERTEX\text{-}COVER = \{< G, k >: \text{graph } G \text{ has a vertex cover of size } k\ \}.$$

We now prove that $VERTEX\text{-}COVER \leq_P ERP$, which shows that ERP is NP-complete (following from what the above said regarding $Lemma$ 34.8). Given an undirected graph $G = (V, E)$ and an integer k, we can show a vertex-cover problem on this graph is analogous to solving the ERP. Let each vertex be a counselor and each edge be a sport in G, and if an edge is connected to a vertex it symbolizes that a counselor is qualified for a sport. In the ERP we are looking to hire the minimum number of k counselors to be qualified in each of the n-sports. This is akin to finding the minimum number of vertices such that each edge of the graph is incident to at least one vertex in the subset of vertices, which precisely is the goal accomplished by finding the vertex cover in G. Therefore, we can say that $VERTEX\text{-}COVER \leq_P ERP$ and conclude that ERP is NP-complete by $Lemma$ 34.8.

4. Problem 4

We start by defining the *Independent Set Problem (IS)*. Given a graph $G = (V, E)$, we say a set of nodes $S \subseteq V$ is *independent* if no two nodes in S are joined by an edge. The Independent Set Problem, which we denote IS, is the following. Given G, find an independent set that is as large as possible. Stated as a decision problem, IS answers the question: "Does there exist a set $S \subseteq V$

such that $|S| \geq k$?" Then set k as large as possible. For this problem, you may take as given that IS is NP-complete.

A store is trying to analyze the behavior of its customers and will often maintain a table A where the rows of the tables correspond to the customers and the columns (or fields) correspond to products the store sells. The entry $A[i, j]$ specifies the quantity of product j that has been purchased by customer i. For example, Table 1 shows one such table.

One thing that a store might want to do with this data is the following. Let's say that a subset S of the customers is *diverse* if no two of the customers in S have ever bought the same product (i.e., for each product, at most one of the customers in S has ever bought it). A diverse set of customers can be useful, for example, as a target pool for market research.

We can now define the *Diverse Set Problem (DS)* as follows: Given an $m \times n$ array A as defined above and a number $k \leq m$, is there a subset of at least k customers that is diverse?

Prove that DS is NP-complete.

<u>Ans:</u>

Using $Lemma$ 34.8 from the textbook:

If L is a language such that $L' \leq_P L$ for some $L' \in NPC$, then L is NP-hard. If, in addition, $L \in NP$, then $L \in NPC$.

In this problem L is DS and L' is IS. The problem already states that IS is NP-complete. All that is needed to show that DS is NP-complete is to show that $IS \leq_P DS$.

Let there be a graph G, such that $G = (V, E)$. In the DS, let the m-customers represent the $|V|$ vertices of the graph. For any of the m-customers, if their n columns are nonzero, connect them with an edge. In the table A given, $A[1,2]$ and $A[2,2]$ are both nonzero in the $j = 2$ column. Therefore, there should be an edge connecting them. This indicates that a group consisting of the two customers are not diverse, because they have both bought the same product (i.e., they both have purchased beer). To find a diverse set of customers, we are looking for an independent set of nodes S, where $S \subseteq V$. Then to find a subset of at least k customers that are diverse, where $k \leq m$, we are asking the question asked by IS that asks, "Does there exist a set $S \subseteq V$ such that $|S| \geq k$?" Therefore, we can say that $IS \leq_P DS$ and conclude that DS is NP-complete by $Lemma$ 34.8.