CS401 - Problem Set 2

1. (a) Prove that the following language is in P:

$$\mathsf{UNARYFACTORING} = \{\langle n_{\text{unary}}, l_{\text{unary}}, k_{\text{unary}} \rangle : \exists j \in \mathbb{N} \text{ s.t. } l \leq j \leq k \text{ and } j \text{ divides } n\}. \tag{1}$$

In this case, n_{unary} means the number n represented in unary. So for example, 2 is represented in unary as 11, and 5 is represented in unary as 11111.
Then $\langle 1111111111, 1111, 1111111 \rangle \in \mathsf{UNARYFACTORING}$ because in base 10, this corresponds to the sequence $< 10, 4, 7 >$. Since 5 divides 10 and $4 \leq 5 \leq 7$, this tuple satisfies the conditions, and so is in the language. On the other hand, $\langle 1111111111, 1111, 1111111 \rangle \notin \mathsf{UNARYFACTORING}$ because there is no number between 4 and 7 that divides 9.

   (b) In class, we only discussed languages associated with functions $f : \{0,1\}^* \to \{0,1\}$. However, we can also create a language associated with a function
$f : \{0, 1, 2, \ldots b-1\}^* \to \{0,1\}$, where the inputs to the function are written in base b. Then the language associated with this function is $L_f \subseteq \{0, 1, 2, \ldots, b-1\}^*$, where $L_f = \{x : f(x) = 1\}$. Since the definition of P depends on the length of the input, and changing the base doesn't change the fundamental problem, but does change the length of the input, it seems like the base we choose to represent the language might be important important. In this problem, you show that the choice of base is not important, because you can always convert a language that is not in base 2 to base 2, without affecting the complexity.

   For S, a subset of the natural numbers, let $L_b^S \subseteq \{0, 1, \ldots, b-1\}^*$, where $L_b^S = \{n_b : n \in S\}$, where n_b is the number n written in base b. Prove that for $b > 2$, if $L_b^S \in \mathsf{P}$, then $L_2^S \in \mathsf{P}$.

   (c) We don't know whether the problem of factoring in base 2 is in P. However, in part (a) you prove that $\mathsf{UNARYFACTORING} \in \mathsf{P}$. Why can't we use the same argument as in part (b) to convert $\mathsf{UNARYFACTORING}$ into base 2, and thus prove factoring is in P?

2. (Extra Practice Problem) Let $L_\triangle = \{G : G \text{ contains a triangle}\}$. Prove $L_\triangle$ is in P.

3. A common function problem is to find the shortest path between two vertices s and t in a graph G. This is not a decision problem, so it is not in P. Please define a language that represents the decision version of this problem. (You do not have to use only mathematical notation - use English or math based on which is clearer!)

4. (a) * A unary language L is a subset of $\{1\}^*$. Let NP_U be the set of unary languages that are also in NP. Prove that if $\mathsf{NP}_U \subseteq \mathsf{P}$, then $\mathsf{EXP} = \mathsf{NEXP}$, where NEXP is defined similarly to NP except now the TM can run for exponential time in the size of the input, and the witness u can be of exponential size in the size of the input.

(b) **EXP** vs **NEXP** is the exponential time version of the **P** vs **NP** question. Based on the result in part (a), how does the **EXP** vs **NEXP** question relate to the **P** vs **NP** question? Is this relationship surprising?

5. (Extra Practice Problem) Prove if $L \in \mathbf{NP}$ and $L' \leq_p L$, then $L' \in \mathbf{NP}$.