Problem 7.21. Let G represent an undirected graph. Also let

$$SPATH = \{\langle G, a, b, k \rangle \mid G \text{ contains a simple path of length at most } k \text{ from } a \text{ to } b\},$$

and

$$LPATH = \{\langle G, a, b, k \rangle \mid G \text{ contains a simple path of length at least } k \text{ from } a \text{ to } b\}.$$

Part a. Show that $SPATH \in P$.

Proof. We show that $SPATH \in P$ by presenting a polynomial time algorithm that decides $SPATH$. A polynomial time algorithm M for $SPATH$ operates as follows.

$M = $ "On input $\langle G, a, b, k \rangle$:

1. Unmark all nodes.

2. Assign each node a value of ∞.

3. Mark a, and set it's value 0.

4. Let integer $d = 0$.

5. Repeat until no additional nodes are marked:

6. $d = d + 1$.

7.   Scan all edges of G. If an edge (u, v) exists between a marked node u and an unmarked node v, then mark v and set value of v to d.

8. If b is marked and value of b is at most k, then *Accept.* Otherwise *reject.*"

Now we analyze this algorithm to show that it runs in polynomial time. Obviously, stages 1 to 4 and stage 8 are executed only once. e. Stage 7 runs at most m times because each time except the last it marks an additional node in G, where m is the number of nodes in G. Stage 6 also runs at most m times. Thus, the total number of stages used is at most $1 + 1 + m + m$, giving a polynomial in the size of G.

Stages 1 to 4 along with stages 6 and 8 are easily implemented in polynomial time on any reasonable deterministic model. Stage 7 involves a scan of the input and a test of whether certain nodes are marked and an update of node values, which also is easily implemented in polynomial time. Hence M is a polynomial time algorithm for $SPATH$. □

Part b. Show that $LPATH$ is NP-complete.

Proof. First, we need to show that $LPATH \in P$, which is easy as certificate is the path. Next, to show that all problems in NP are polynomial time reducible to $LPATH$, we show $3SAT \leq_p$ $LPATH$. We show how to construct an integer k and an undirected graph G with two nodes, s and t, where a simple path of length at least k exists between s and t , iff ϕ is satisfiable.

Let ϕ be any Boolean formula in 3CNF containing m clauses:

$$\phi = (a_1 \vee b_1 \vee c_1) \wedge (a_2 \vee b_2 \vee c_2) \wedge \cdots \wedge (a_m \vee b_m \vee c_m).$$

where each a, b and c is a literal x_i or $\overline{x_i}$, and $x_1, x_2 \cdots x_l$ are the l variables of ϕ. Now we show how to convert ϕ to G. $\qquad\square$