Proof for Dijkstra's Algorithm

Recall that Dijkstra's algorithm finds the length of all the shortest paths in a directed graph with non-negative weights on the edges, from a source vertex s to every other vertex v_i in the graph.

We define the distance d_i to be the length of the shortest path from s to vertex v_i. Dijkstra's algorithm maintains a set of vertices S, with two properties. First, S is a set of vertices in the graph nearest to s; that is:

$$\forall v_i \in S,\, \forall v_j \in V - S,\;\; d_i \leq d_j$$

And second, for all vertices $v_j \in S$, there is a shortest path from s to v_j using only vertices of S as intermediates. There might be several different choices for S; Dijkstra's algorithm chooses one arbitrarily.

For each "outside" vertex $d_j \in V - S_k$, we define an estimated distance:

$$d_j^{\text{est}} = \min_{v_i \in S} d_i + w(e_{i,j})$$

Since the estimated distance is the length of some path to d_j, it is an upper bound on the length of the true shortest path from s.

Lemma 1 *The estimated distance to d_j^{est} is the length of the shortest path from s to d_j, using only vertices of S as intermediates.*

Proof: Any path from s to d_j, using only vertices of S as intermediates, consists of a shortest path from from s to some $v_i \in S$, and then one more edge from v_i to v_j. We defined the estimated distance to v_j to be the shortest path of this form.

The basic step of Dijkstra's algorithm adds one more vertex to S. It selects the vertex to add to be one of the $v_m \in V - S$ such that d_j^{est} is minimum; that is, $d_m^{\text{est}} \leq d_j^{\text{est}},\; \forall v_j \in V - S$. The new set $S' = S + v_m$.

We want to prove that this is a correct choice, that is, that S' will have the two properties that S had. This is done using the following essential lemma about S.

Lemma 2 *Let v_m be an outside vertex in $V - S$ such that d_m^{est} is minimum. Then $d_m^{est} \leq d_j$, for all $j \in V - S$. That is, the estimated distance to v_m is a lower bound on the length of the shortest path from s to any vertex in $V - S$.*

Proof: Assume for the purpose of contradiction that there is some vertex $v_j \in V - S$, with $d_j < d_m^{\text{est}}$. Since $d_m^{\text{est}} \leq d_j^{\text{est}}$, we have $d_j < d_j^{\text{est}}$. So any true shortest path P from s to v_j is shorter than the length of a shortest path using only vertices from S as intermediates. Then P must use at least one vertex from $V - S$ as an intermediate.

Let v_x be the first vertex from $V - S$ along P, as we go from s to v_j, so that the predecessor of v_x along the path belongs to S. Since v_x comes before v_j, $d_x^{\text{est}} \leq d_j < d_m^{\text{est}}$. But v_m was defined to be a vertex of $V - S$ such that d_m^{est} is minimum. This is a contradiction. So our assumption that there is some v_j such that $d_j < d_m^{\text{est}}$ has to be wrong.

Now we use the essential Lemma to prove that our new set S' has the two properties of S.

Theorem 3 *S' is a set of vertices nearest to s, that is,*

$$\forall v_i \in S',\ \forall v_j \in V - S',\ \ d_i \leq d_j$$

We use the essential lemma in two ways. First, since the lemma is true for v_m as well as any other vertex in $V - S$, we have $d_m^{\text{est}} \leq d_m \leq d_m^{\text{est}}$, that is, $d_m^{\text{est}} = d_m$. So we have computed the length of the shortest path to v_m. And second, since $d_m^{\text{est}} = d_m \leq d_j$, for all $v_j \in V - S$, then it is also true that $d_m \leq d_j$ for all $v_j \in V - S'$.

Finally, we need to prove that for all vertices $v_i \in S'$, there is a shortest path from s to v_i using only vertices of S' as intermediates. This is your question to think about.