Single Source Shortest Path (Dijkstra's Algorithm)

1 Key Ideas:

Let $G = (V, E)$ be a graph[1] with a weight[2] function $f : V \times V \rightarrow R^+ \cup \{+\infty\}$.

1.1 Set of vertices S for which the shortest paths are known

Dijkstra's algorithm maintains a set of vertices S such that for all $v \in S$ the cost of the shortest path from the starting vertex v_0 to v is known. At each step, we add a new vertex to S.

1.2 Cost array D

Dijkstra's algorithm also maintains an array $D[v]$ of cost estimates known so far. A key to understanding the correctness proof for Dijkstra's algorithm is a precise interpretation of the values contained in $D[v]$:

- If $v \in S$ then $D[v] =$ the cost of the shortest path from v_0 to v.
- If $v \notin S$ then $D[v] =$ the cost of the shortest path from v_0 to v that lies[3] entirely within S, except for the endpoint v. It is possible that no such path to v exists, and in this case $D[v] = \{+\infty\}$.

2 Dijkstra's Algorithm

Input: Graph $G = (V, E)$ and a weight function $f : V \times V \rightarrow R^+ \cup +\infty$
Output: Array D such that $D[v]$ is the cost of the shortest path from v_0 to v.

```
S = {v0}
D[v0] = 0
for all v ∈ (V − {v0})
    D[v] = f(v0, v)
end for all
while ( S ≠ V )
    Select vertex w in V − S such that D[w] is minimal
    S = S ∪ { w }
    for all v in V − S do
        D[v] = min(D[v], D[w] + f(w, v))
    end for all
end while
```

[1] G may be either directed or undirected.
[2] The set R^+ denotes the set of positive real numbers.
[3] A path P lies entirely within a vertex set S if and only if every vertex w along path P is also in S.

3 Correctness Proof of Dijkstra's Algorithm

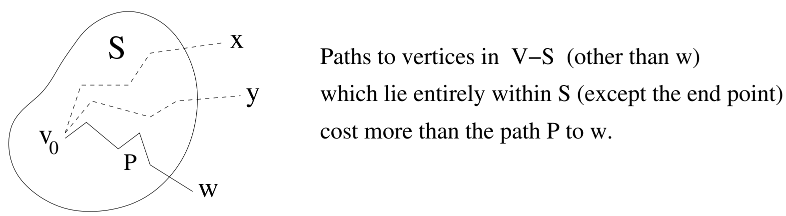
Proof is by induction on $|S|$.

Base Case: $S = \{ v_0 \}$ The cost of the shortest path from v_0 to v_0 is 0 by definition, i.e., $D[v] = 0$. For all other vertices v: a path which lies entirely within $S = \{ v_0 \}$ except for the endpoint, v itself consists only of one edge. Therefore, the cost of that path is:

$$D[v] = f(v_0, v) \quad \text{where } f \text{ is the weight function for the graph.} \tag{1}$$

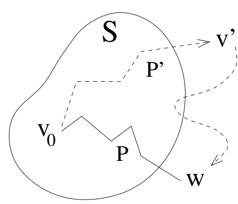
Notice that the definition of D in equation (1) is consistent with the claim stated in subsection §1.2.

General Case: By the inductive hypothesis, the values of $D[v]$ is the cost of the shortest path for all $v \in S$. Let w denote the vertex in $V - S$ such that $D[w]$ is minimal. The situation is illustrated in the diagram below:



Claim: $D[w]$ is the cost of the shortest path from v_0 to w (with no restrictions on the possible path).

Proof of the Claim: (by contradiction) Suppose there is a shorter path P' to w. The path P' must go to some vertex (other than w) outside of S and then proceed to w.[4] The situation is illustrated in the diagram below:



Let v' be the first vertex outside of S along the path P'. Observe that the cost of path P' to w is higher than the partial path along P' to v'. This contradicts the way in which w was chosen. I.e., if the path P' exists, Dijkstra's algorithm would have chosen v' instead of w.

The update step for $D[v]$ considers the fact that in the newly expanded set S, the paths from v_0 to vertices x still outside of S may go to w, and then proceed to x.

∎

[4] If P' stays entirely within S except for the endpoint w, then the inductive hypothesis implies that $\text{cost}(P') = \text{cost}(P)$.