BFS(s) Computes the Shortest Paths from s – Proof Sketch

Recall that during the execution of a BFS started from s (denoted BFS(s)), if a node u discovers a node v, then $d[v]$ is set to $d[u] + 1$. Initially, $d[s]$ is set to 0.

It easy to see that the $d[v]$ computed by BFS(s) is equal to the length of the BFS path that starts from s and "discovers" node v.

Let $\delta(s, v)$ be the length of a *shortest* path from s to v.

Lemma 1: If u is enqueued before v during the execution of BFS(s), then $d[u] \leq d[v]$.

The proof of the above Lemma is in CLRS (Corollary 22.4).

Lemma 2: After the execution of BFS(s), for all nodes v, $d[v] \geq \delta(s, v)$.

The proof of Lemma 2 is obvious: After the execution of BFS(s), $d[v]$ is the length of *some* path from s to v (namely, the path that "discovers" v), while, by definition, $\delta(s, v)$ is the length of a *shortest* path from s to v. Thus, $d[v] \geq \delta(s, v)$.

We now show that BFS(s) correctly computes the distance of every node v from s:

Theorem: After the execution of BFS(s), for all nodes v, $d[v] = \delta(s, v)$.

Proof (Sketch): Suppose, for contradiction, that for some node x, $d[x] \neq \delta(s, x)$. Let v be the *closest* node from s such that $d[v] \neq \delta(s, v)$. By Lemma 2, $d[v] > \delta(s, v)$.

Consider a shortest path from s to v (there may be several ones, chose and fix one of them). Let (u, v) be the last edge on that shortest path.

Note that the length of this path is $\delta(s, v)$, and that $\delta(s, v) = \delta(s, u) + 1$. By our choice of v, since u is closer to s than v, we have $d[u] = \delta(s, u)$.

Putting all this together, we get $d[v] > \delta(s, v) = \delta(s, u) + 1 = d[u] + 1$.

That is, $d[v] > d[u] + 1$ (*).

We now obtain a contradiction to (*). To do so, consider the color of v at the time node u is first explored by the BFS(s). There are three possible cases:

1. *v is* ***white****: v is not yet discovered.*

   In this case, v is discovered during the exploration of u, and so $d[v] = d[u] + 1$ — a contradiction to (*).

2. *v is* ***black****: v was already discovered and explored.*

   In this case, v was enqueued (and removed) before u was enqueued. By Lemma 1, $d[v] \leq d[u]$ — a contradiction to (*).

3. *v is* ***grey****: v was already discovered but it was not yet explored.*

   Let w be the node that discovered v. This discovery occured before the exploration of u. So w was explored before u was explored. Thus, w was enqueued before u was enqueued. So, by Lemma 1, $d[w] \leq d[u]$. This implies $d[w] + 1 \leq d[u] + 1$. Since v was discovered by w, $d[v] = d[w] + 1$. We conclude that $d[v] \leq d[u] + 1$ — a contradiction to (*).

Q.E.D.