Finding a Shortest Path

1. A special characteristic of the BFS style of traversing a graph is that, with very little extra processing, it outputs the shortest path between any two vertices in the graph. (There is no straightforward to do this with DFS.)

2. If p: $v_0 – v_1 – v_2 – \ldots – v_n$ is a path in G, recall that its length is n, the number of edges in the path. BFS can be used to compute the shortest path between any two vertices of the graph.

3. As with DFS, the discovered edges during BFS collectively form a spanning tree (assume G is connected, so there is a spanning tree). A tree obtained in this way, starting with a vertex s (the *starting vertex when performing BFS*) is called the *BFS rooted spanning tree.*

4. Recall that a rooted tree can be given the usual *levels.*

5. Given a connected graph G with vertices s and v, Here is the algorithm to return the shortest length of a path in G from s to v

   A. Perform BFS on G starting with s to obtain the BFS rooted tree T with root s, together with a map recording the levels of T
   B. Return the level of v in T

Theorem. For any simple graph G and any vertices s, v in G, the length of the shortest path from s to v is equal to the level to which v belongs in any BFS rooted tree with root s.

Exercise 1. If $v_1 - v_2 - \ldots - v_k - v_{k+1}$ is a shortest path from v_1 to v_{k+1}, then whenever $1 \leq i \leq k$, the path $v_1 - v_2 - \ldots - v_i$ is a shortest path from v_1 to v_i .

Exercise 2. Suppose v has level i in a BFS rooted tree with root s. Then there is a path from s to v of length i. [Proof is by induction on levels.]

(Optional) **Proof of Theorem**. First build a BFS rooted tree with root s. Proceed by induction to prove the following statement φ(i), for all i:

> For every vertex v in G, the shortest path from s to v has length i
> if and only if the level of v is i.

For i = 0, the only vertex that has shortest path length 0 from s is s itself, which has level 0. Therefore, φ(0) holds.

Assume φ(j) is true for all j < i and let v be a vertex in G. For one direction, assume v has level i. Then there is (by Exercise 2) a path from s to v of length i. Suppose there is also a shorter path, of length j < i, from s to v. By induction hypothesis, this would imply that v is in level j, and this is not the case.

For the other direction, assume there is a shortest path from s to v of length i – denote this path p: s – a – b – . . . – u – v. By Exercise 1, s – a – … – u is a shortest path from s to u, and this path clearly has length i - 1. By the induction hypothesis, lev(u) = i - 1. Since (u,v) is an edge in the tree, there are just two possibilities for the level of v: i – 2 or i (recall that in a BFS rooted tree, there is never an edge between two vertices at the same level). If lev(v) = i – 2, then by the induction hypothesis, there is a shortest path from s to v of length i - 2; but we already know that the shortest path length from s to v is i, so this possibility must be ruled out. It follows therefore that lev(v) = i, as required.

By induction, we have shown that, for every i, φ(i) holds, and this establishes the theorem.