Lecture 16, Nov 1, 2023

Correctness of BFS

- Denote $\delta(s, v)$ as the length of the shortest path from the starting node s to a node v
- Suppose we execute BFS on a graph G; we wish to prove that, for every node $v \in V, d[v] = \delta(s, v)$, i.e. the length of the discovery path by BFS is the length of the shortest path, or that the discovery path is the shortest path
- Lemma 0: $\delta(s, v) \leq d[v]$ (trivially true)
- Lemma 1: if a node u is inserted into Q before v, then $d[u] \leq d[v]$
  - Proof by contradiction: suppose that this lemma is false, and let u, v be the first of such a pair so that $d[u] > d[v]$
  - $v \neq s$, because no vertices enter before s
  - $u \neq s$, because $d[s] = 0$, so if $u = s$, then that would imply $d[v] < 0$
  - Therefore both u and v must have been discovered by some other nodes, u', v'
  - $d[u'] = d[u] - 1, d[v'] = d[v] - 1 \implies d[u'] > d[v']$, so u', v' must be different
  - Since u was inserted into Q first, u' must have been in the queue before v'
  - Hence we have u' in the queue before v' and $d[u'] > d[v']$, and both were inserted before u, v
  - But we assumed that u, v were the first pair that violated the lemma, so since u', v' came before, $d[u'] \leq d[v']$, leading to a contradiction
- Theorem: after BFS, for every $v \in V, d[v] = \delta(s, v)$
  - Proof by contradiction: suppose there exists $x \in V$ such that $d[x] \neq \delta(s, x)$
  - Pick v to be the close node from s that that $d[v] \neq \delta(s, v)$
  - By Lemma 0, $d[v] > \delta(s, v)$
  - Consider an actual shortest path from s to v; let (u, v) be the last edge on that path; then $\delta(s, u) = \delta(s, v) - 1$
  - Since $\delta(s, u) < \delta(s, v)$, u cannot violate the lemma and therefore $d[u] = \delta(s, u)$
  - $d[v] > \delta(s, v) = \delta(s, u) + 1 = d[u] + 1$
  - Now consider the color of v just before u is explored
    - Case 1: v is white
      - When we explore u, we will set $d[v] = d[u] + 1$ since v is still white and connected to u
      - But $d[v] > d[u] + 1$ so this is a contradiction
    - Case 2: v is black:
      - Since v has already been explored before u, it has entered the queue before u
      - By Lemma 1, this means $d[v] \leq d[u]$, leading to a contradiction
    - Case 3: v is grey:
      - Since v is grey before we've explored u, it must have been discovered by a node $w \neq u$
      - Therefore w must have been in the queue before u, so by Lemma 1 $d[w] \leq d[u]$
      - Since w discovered v, $d[v] = d[w] + 1$
      - $d[w] \leq d[u] \implies d[w] + 1 \leq d[u] + 1 \implies d[v] \leq d[u] + 1$ which is a contradiction
- Therefore the BFS discovery path from s to v is a shortest path
- Note we could have proven this by induction