

The invariant of Dijkstra's shortest-path algorithm

David Gries

The shortest-path algorithm is a breadth-first search algorithm. It has a loop, and each iteration of the loop will identify the shortest distance to one more node. We now present the invariant of the loop and prove a simple theorem about it.

The set of all nodes is partitioned into three sets: A *settled* set S (red), a *frontier* set F (blue), and a *far-off* set (black).



The invariant consists of three points:

1. For a *settled* node s , at least one shortest path from v to s contains only settled nodes, and $d[s]$ is the distance of that shortest path from v to s .

Think of the *settled* set as places that we know all about because we have visited often and settled there.

A node in the *frontier* has been visited at least once but not enough to know for certain that its shortest-path distance has been fully calculated. Think of the *frontier* as the moon and close planets that we know something about, but not everything. We have not settled there yet. Here's the second part of the invariant:

2. For a node f in the *frontier*, at least one path from v to f contains only settled nodes, except for the last one, f (as shown below), and $d[f]$ is the minimum distance of all such paths.

That is, over paths that start with settled node v , perhaps contain more red nodes, and finally have one edge to f . There is a degenerate case: when f and v are the same node and it is in the *frontier* set. This degenerate case will make it easy to initialize the invariant, including putting v into the *frontier* set.



Here's the third part of the invariant:

3. All edges leaving the *settled* set end in the *frontier* set.

That's all there is to the invariant! Three simple and easy-to-remember points.

Stop the video at this point and convince yourself that if v is in either the *settled* or the *frontier* set, the invariant implies that $d[v] = 0$.

A theorem based on the invariant. We now prove an important theorem.

Theorem. For a node f in the *frontier* with minimum d value (over nodes in the *frontier*), $d[f]$ is indeed the shortest-path distance from v to f .

For example, suppose the *frontier* contains three nodes f_0 , f_1 , and f_2 , with shortest-distances $d[f_0] = 5$, $d[f_1] = 4$, and $d[f_2] = 6$. Then f_1 is the node in the frontier with minimum d value.



We prove this theorem by showing that any other path from v to f does not have a smaller distance. We consider two cases:

Case 1. v and f are the same, and that node is in the *frontier* set. First, the settled set is empty, for, by invariant P1, if a node s is in the *settled* set then v must be there too, but it isn't; it is in the *frontier*. Second, since v is in the *frontier* set, by P2, the *frontier* set contains only one node, v . Third, we know that $d[v]$ is 0, and 0 is the distance of the shortest path from v to v .

Case 2. v is in the *settled* set. Suppose that the shown path from v to f is the one with distance $d[f]$.

By part 2 of the invariant, $d[f]$ is the shortest-path distance over paths that contain only *settled* (red) nodes except for f . Consider any other v - f path. It starts at v , goes through *settled* nodes, visits f (so that path had distance $\geq d[f]$) or visits another *frontier* node g (say) and then winds its way to f . (The path can't go directly from a red node to a node in the far-off set because part 3 of the invariant says there are no edges from the settled set to the far-off set). Since $d[g] \geq d[f]$, and since edge weights are positive, the distance of this path is $\geq d[f]$.

Q.E.D. (Quit-End-Done)