

CSCI 241

Lecture 20 Dijkstra's Algorithm: Implementation, Proof of Correctness

Happenings

Today: AWC-organized panel discussion with recent female alumnae working in the tech industry. AW 210 at 4:00.

Tuesday, 3/12 – AWC Bake Sale – 10 am to 2 pm in the CF 1st Floor Lobby

Wednesday, 3/13 – Peer Lecture Series: Machine Learning – 5 pm in CF 420

Thursday, 3/14 – [CS Study Break: Bingo and Brownies!](https://nam03.safelinks.protection.outlook.com/?url=https%3A%2F%2Fcse.wwu.edu%2Fcomputer-science%2Fevent%2Fcs-study-break-6&data=02%7C01%7Cwehrwes%40wwu.edu%7Ceca19019c15c4c268aa008d6a35150b7%7Cdc46140ce26f43efb0ae00f257f478ff%7C0%7C0%7C636875967018325587&sdata=OQO49pLiWE6r20Gtfoz8XRbeecXGwm3V4ySt7ehEgUI%3D&reserved=0) – 3 pm in the CF 4th Floor Foyer

COMPUTER SCIENCE G $\overline{\mathbf{B}}$ STUDY BREAK

Join us for brownies and bingo!

 $3 - 4 PM$ **THURSDAY, MARCH 14 CF 4TH FLOOR FOYER**

Ringo called by Aran Clauson

For disability resources, contact 360-650-3083

Announcements

- No quiz today.
- Quiz Monday instead.

Goals

- Be able to run Dijkstra's algorithm on paper.
- Know how to implement Dijkstra efficiently (as in A4)
- Know how to augment the algorithm to keep backpointers in order to reconstruct the sequence of nodes in a shortest path.

Dijkstra's Shortest Paths: Intuition

- Intuition: explore nodes kinda like BFS.
- There are three kinds of nodes:
	- Settled nodes for which we know the actual shortest path.
	- Frontier nodes that have been visited but we don't necessarily have their actual shortest path
	- Unexplored all other nodes.
- Each node n keeps track of $n.d$, the length of the shortest known known path from start.
- We may discover a shorter path to a frontier node than the one we've found already - if so, update $n.d.$

Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move the node f with smallest d from F to S For each neighbor w of f: if we've never seen w before: set its path length add it to frontier else if the path to w via f is shorter: update w's shortest path length

Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move the node f with smallest d from F to S For each neighbor w of f: if we've never seen w before: set its path length add it to frontier else if the path to w via f is shorter: update w's shortest path length $s \longrightarrow \cdots \longrightarrow f \longrightarrow w$ f.d wt(f,w)

 $Initialiro$ $Set+lad to a$

Node d Best known distance

0 ?

1 ?

2 ?

3 ?

4 ?

Settled set:

Frontier set:

Node d 0 ? 1 ? 2 ? 3 ? Best known distances:

4 **0**

 Ω Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move the node f with smallest d from F to S For each neighbor w of f: if we've never seen w before: set its path length to $f.d + wt(f,w)$ add w to the frontier else if the path to w via f is shorter: update w's shortest path length

Settled set: {}

Frontier set: {4}

0 Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move the node f with smallest d from F to S For each neighbor w of f: if we've never seen w before: set its path length to $f.d + wt(f,w)$ add w to the frontier else if the path to w via f is shorter: update w's shortest path length f: 4

Settled set: {4}

Frontier set: {}

0 Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move the node f with smallest d from F to S For each neighbor w of f: if we've never seen w before: set its path length to $f.d + wt(f,w)$ add w to the frontier else if the path to w via f is shorter: update w's shortest path length f: 0

Settled set: {4, 0}

Frontier set: {}

shortest-paths(4)

shortest-paths(4)

0 Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move the node f with smallest d from F to S For each neighbor w of f: if we've never seen w before: set its path length to $f.d + wt(f,w)$ add w to the frontier else if the path to w via f is shorter: update w's shortest path length f: 1

Settled set: {4, 0, 1}

Frontier set: {2}

 $D - 1$

shortest-paths(4)

0 Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move the node f with smallest d from F to S For each neighbor w of f: if we've never seen w before: set its path length to $f.d + wt(f,w)$ add w to the frontier else if the path to w via f is shorter: update w's shortest path length f: 2

Settled set: {4, 0, 1, 2}

Frontier set: {3}

0 $\frac{2}{1}$ $\frac{4}{1}$ Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move the node f with smallest d from F to S For each neighbor w of f: if we've never seen w before: set its path length to $f.d + wt(f,w)$ add w to the frontier else if the path to w via f is shorter: update w's shortest path length f: 3

4

1

3

4

3

1

2

shortest-paths(4)

3

Settled set: {4, 0, 1, 2, 3}

Frontier set: {} Empty => done!

 $S = \{ \}$; $F = \{v\}$; $v.d = 0$; **while** $(F \neq \{\}) \leq$ $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of $f \}$ **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ add w to F; $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$ } } } Initialize Settled to empty Initialize Frontier to the start node

 $S = \{ \}$; $F = \{v\}$; $v.d = 0$; **while** $(F \neq \{\}) \leq$ $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ add w to F; $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$ } } } Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move node f with smallest d from F to S

```
S = \{ \}; F = \{v\}; v.d = 0;
while (F \neq \{\}) \leqf = node in F with min d value;
          Remove f from F, add it to S;
   for each neighbor w of f \}       if (w not in S or F) {
        w.d = f.d + weight(f, w);            add w to F;
    \} else if (f.d+weight(f,w) < w.d) {
        w.d = f.d + weight(f,w);      }
 }
}
                                   Initialize Settled to empty
                                    Initialize Frontier to the start node
                                     While the frontier isn't empty:
                                        move node f with smallest d 
                                         from F to S
                                    For each neighbor w of f: 
                                       if we've never seen w before:
                                         set its path length
                                         add it to frontier
```
 $S = \{ \}$; $F = \{v\}$; $v.d = 0$; **while** $(F \neq \{\}) \leq$ $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ add w to F; $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$ } } } For each neighbor w of f: if we've never seen w before: set its path length add it to frontier else if path to w via f is shorter: update w's shortest path length Initialize Settled to empty Initialize Frontier to the start node While the frontier isn't empty: move node f with smallest d from F to S

What if we want to know the shortest path?

 $S = \{ \}$; $F = \{v\}$; $v.d = 0$; **while** $(F \neq \{\}) \leq$

 $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) {

 $w.d = f.d + weight(f, w);$ add w to F;

 $w.d = f.d + weight(f,w);$

}

}

}

• At termination: for each reachable node n, n.d stores the **length** of the shortest path from v to n.

• We didn't keep track of $\}$ else if $(f.d+weight(f,w) < w.d)$ { how to get from v to n!

What if we want to know the shortest path?

 $\}$ else if $(f.d+weight(f,w) < w.d$ { Strategy: maintain a $S = \{ \}$; $F = \{v\}$; $v.d = 0$; $v.bp = null$; **while** $(F \neq \{\}) \leq$ $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ **w.bp = f;** add w to F; $w.d = f.d + weight(f,w);$ **w.bp = f** }

}

}

Each node could store the full path, but that would be expensive to keep updated.

backpointer at each node pointing to the previous node in the shortest path.

What if we want to know the shortest path? Example

} else if (f.d+weight(f,w) < w.d) { Strategy: maintain a $S = \{ \}$; $F = \{v\}$; $v.d = 0$; $v.bp = null$; **while** $(F \neq \{\}) \leq$ $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of $f \}$ **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ **w.bp = f;** add w to F; $w.d = f.d + weight(f,w);$ $w \cdot bp = f$

}

}

}

backpointer at each node pointing to the previous node in the shortest path.

Let's Dijkstra

- Half get the algorithm, half get the graphs. Pair up with someone with the other thing.
- Run the algorithm on each graph:
	- (first) 5-node graph: start at node S
	- (second) Other graph: start at node D

The next slide is really important.

and a

- $S = \{ \}$; $F = \{v\}$; $v.d = 0$; $v.bp = null$; ¹ **while** $(F \neq \{\}) \leq$
	- $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of $f \}$ **if** (w not in S or F) {
		- $w.d = f.d + weight(f, w);$ $w \cdot bp = f$; add w to F;
		- $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$ $w.bp = f$

}

}

}

Store Frontier in a min-heap priority queue with d-values as priorities.

- 2. To efficiently iterate over neighbors, use an adjacency list graph representation.
- 3. Could store w.d and w.bp in Node class; in A4, we use a HashMap<Node,PathData>
- 4. No need to explicitly store Settled or Unexplored sets: a node is in S or F iff it is in the map.

 $S = \{ \}$; $F = \{v\}$; $v.d = 0$; $v.bp = null$; ¹ **while** $(F \neq \{\})$ {

- **f = node in F with min d value; Remove f from F, add it to S; for** each neighbor w of $f \}$ **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ $w.bp = f;$ **add w to F;** $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$
	- $w.bp = f$

}

}

}

1. Store Frontier in a min-heap priority queue with d-values as priorities.

- 2. To efficiently iterate over neighbors, use an adjacency list graph representation.
- 3. Could store w.d and w.bp in Node class; in A4, we use a HashMap<Node,PathData>
- 4. No need to explicitly store Settled or Unexplored sets: a node is in S or F iff it is in the map.

$$
S = \{\}; F = \{v\}; v.d = 0; v.bp = null;
$$

while $(F \neq \{\}) \{$

 $f = node$ in F with min d value; Remove f from F, add it to S;

for each neighbor w of f {

}

}

}

```
       if (w not in S or F) {
    w.d = f.d + weight(f, w);w \cdot bp = f;
                add w to F;
```
 $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$ $w.bp = f$

1. Store Frontier in a min-heap priority queue with d-values as priorities.

- **2. To efficiently iterate over neighbors, use an adjacency list graph representation.**
- 3. Could store w.d and w.bp in Node class; in A4, we use a HashMap<Node,PathData>
- 4. No need to explicitly store Settled or Unexplored sets: a node is in S or F iff it is in the map.

- $S = \{ \}$; $F = \{v\}$; $v.d = 0$; $v.bp = null$; ¹ **while** $(F \neq \{\}) \leq$
	- $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of $f \}$ **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ **w.bp = f;** add w to F;
		- $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$ $\bf w.bp = f$

}

}

}

Store Frontier in a min-heap priority queue with d-values as priorities.

- 2. To efficiently iterate over neighbors, use an adjacency list graph representation.
- **3. Could store w.d and w.bp in Node class; in A4, we use a HashMap<Node,PathData>**
- 4. No need to explicitly store Settled or Unexplored sets: a node is in S or F iff it is in the map.

- $S = \{ \}$; $F = \{v\}$; $v.d = 0$; $v.bp = null$; ¹ **while** $(F \neq \{\}) \leq$
	- $f = node$ in F with min d value; Remove f from F, add it to S;
	- **for** each neighbor w of $f \}$ **if** (**w not in S or F**) {
		- $w.d = f.d + weight(f, w);$ $w \cdot bp = f$;
			- add w to F;

}

}

}

```
\} else if (f.d+weight(f,w) < w.d) {
   w.d = f.d + weight(f,w);w.bp = f
```
Store Frontier in a min-heap priority queue with d-values as priorities.

- 2. To efficiently iterate over neighbors, use an adjacency list graph representation.
- 3. Could store w.d and w.bp in Node class; in A4, we use a HashMap<Node,PathData>
- **4. No need to explicitly store Settled or Unexplored sets: a node is in S or F iff it is in the map.**

- Dijkstra's algorithm is **greedy**: it makes a sequence of *locally* optimal moves, which results in the *globally* optimal solution.
	- Most algorithms don't work like this need to prove that it results in the global optimum.
- Specifically: It is not obvious that there cannot still be a shorter path to the Frontier node with smallest d-value.

Proof of Correctness: Invariant **Frontier F Unexplored**

Settled

S

f⁻

The while loop in Dijkstra's algorithm maintains a 3 part invariant:

1. For a Settled node s, a shortest path from v to s contains only settled nodes and s.d is length of shortest $v \rightarrow s$ path.

→ **v** for formal parameters and the set of th

- 2. For a Frontier node f, at least one v -> f path contains only settled nodes (except perhaps for f) and f.d is the length of the shortest such path
- 3. All edges leaving S go to F (or: no edges from S to Unexplored)

 $S = \{ \}$; F = $\{v\}$; v.d = 0; **while** $(F \neq \{\}) \leq$

}

}

Theorem

 $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ add w to F; $\}$ **else if** (f.d+weight(f,w) < w.d) {

```
Theorem: For a node f in the Frontier 
with minimum d value (over all nodes in 
the Frontier), f.d is the shortest-path 
distance from v to f.
```
Proof: Show that any other path from v to if has length $>=$ f.d

```
w.d = f.d + weight(f,w);      }
```
Case 1: if v is in F, then S is empty and $v.d = 0$, which is trivially the shortest distance from v to v.

 $S = \{ \}$; $F = \{v\}$; v.d = 0; **while** $(F \neq \{\}) \leq$

}

}

Theorem

 $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) {

 $w.d = f.d + weight(f, w);$ add w to F;

Theorem: For a node f in the Frontier with minimum d value (over all nodes in the Frontier), f.d is the shortest-path distance from v to f.

Proof: Show that any other path from v to if has length $>=$ f.d

- $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$
- } **Case 2:** v is in S. Part 2 of the invariant says:
	- **•** f.d is the length of the shortest path from v to f containing all settled nodes except f, and f.d is the length of such a path.

 $S = \{ \}$; F = $\{v\}$; v.d = 0; **while** $(F \neq \{\}) \leq$

}

}

Theorem

 $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) {

 $w.d = f.d + weight(f, w);$ add w to F;

Theorem: For a node f in the Frontier with minimum d value (over all nodes in the Frontier), f.d is the shortest-path distance from v to f.

Proof: Show that any other path from v to if has length $>=$ f.d

v or de la communication

- $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$
- } **Case 2:** v is in S. Part 2 of the invariant says:
	- **•** f.d is the length of the shortest path from v to f containing all settled nodes except f, and f.d is the length of such a path. Any other v-f path must either be longer or go through another frontier node g then arrive at f:

 $S = \{ \}$; $F = \{v\}$; v.d = 0; **while** $(F \neq \{\}) \leq$

 $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f {

 if (w not in S or F) { $w.d = f.d + weight(f, w);$ add w to F;

Theorem

Theorem: For a node f in the Frontier with minimum d value (over all nodes in the Frontier), f.d is the shortest-path distance from v to f.

Proof: Show that any other path from v to if has length $>=$ f.d

- $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$
- } **Case 2:** v is in S. Part 2 of the invariant says:
	- **•** f.d is the length of the shortest path from v to f containing all settled nodes except f, and f.d is the length of such a path. Any other v-f path must either be longer or go through another frontier node g then arrive at f: $\longrightarrow 9$

 $d.f \leq d.g,$

}

}

v or de la communication so that path cannot be shorter

Proof of Correctness: Invariant Maintenance

 $S = \{ \}$; F = $\{v\}$; v.d = 0; **while** $(F \neq \{\}) \leq$ $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ add w to F; $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$

}

}

}

- 1. For a Settled node s, a shortest path from v to s contains only settled nodes and s.d is length of shortest $v \rightarrow s$ path.
- 2. For a Frontier node f, at least one v -> f path contains only settled nodes (except perhaps for f) and f.d is the length of the shortest such path
- 3. All edges leaving S go to F (or: no edges from S to Unexplored)

Proof of Correctness: Invariant Maintenance

 $S = \{ \}$; $F = \{v\}$; v.d = 0; **while** $(F \neq \{\}) \leq$ $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ add w to F; $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$ }

}

}

- 1. For a Settled node s, a shortest path from v to s contains only settled nodes and s.d is length of shortest $v \rightarrow s$ path.
- 2. For a Frontier node f, at least one v -> f path contains only settled nodes (except perhaps for f) and f.d is the length of the shortest such path
- 3. All edges leaving S go to F (or: no edges from S to Unexplored)

At initialization:

- 1. S is empty; trivially true.
- 2. $v.d = 0$, which is the shortest path.
- 3. S is empty, so no edges leave it.

Proof of Correctness: Invariant Maintenance

 $S = \{ \}$; $F = \{v\}$; $v.d = 0$; **while** $(F \neq \{\}) \leq$ $f = node$ in F with min d value; Remove f from F, add it to S; **for** each neighbor w of f { **if** (w not in S or F) { $w.d = f.d + weight(f, w);$ add w to F;

 $\}$ **else if** (f.d+weight(f,w) < w.d) { $w.d = f.d + weight(f,w);$

- 1. For a Settled node s, a shortest path from v to s contains only settled nodes and s.d is length of shortest $v \rightarrow s$ path.
- 2. For a Frontier node f, at least one v -> f path contains only settled nodes (except perhaps for f) and f.d is the length of the shortest such path
- 3. All edges leaving S go to F (or: no edges from S to Unexplored)

 } At each iteration:

}

}

- Theorem says f.d is the shortest path, so it can safely move to S
- 2. Updating w.d maintains Part 2 of the invariant.
- 3. Each neighbor is either already in F or gets moved there.