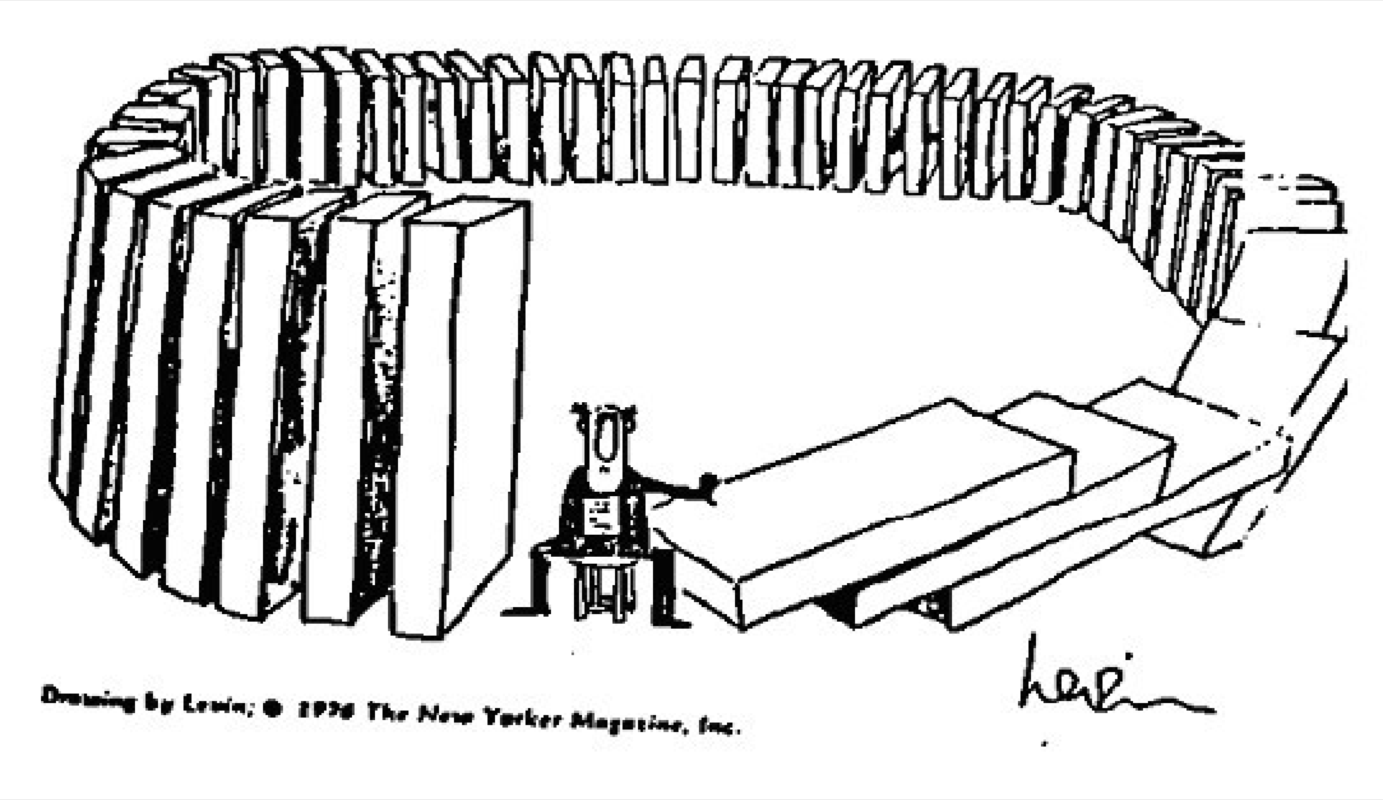
Unweighted directed graphs



Announcements

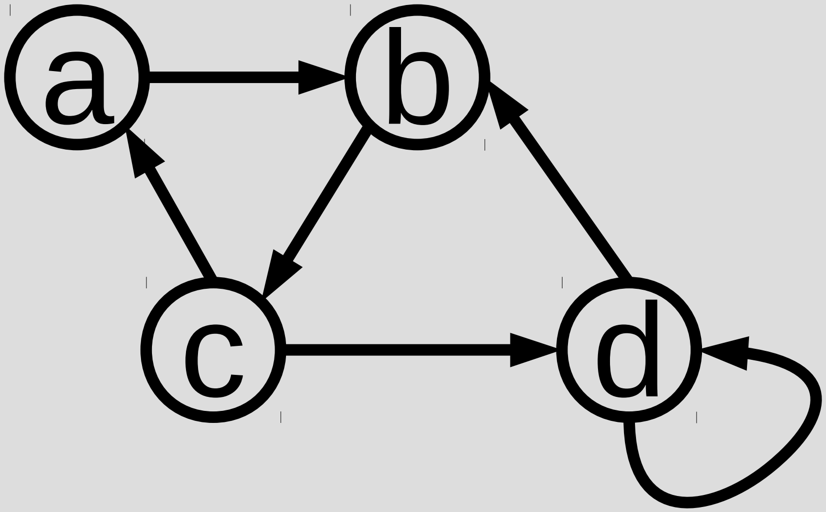
Midterm & gradescope

- will get an email today to register (username name is your email)
- tests should appear by next Monday (nothing there now)

Graph

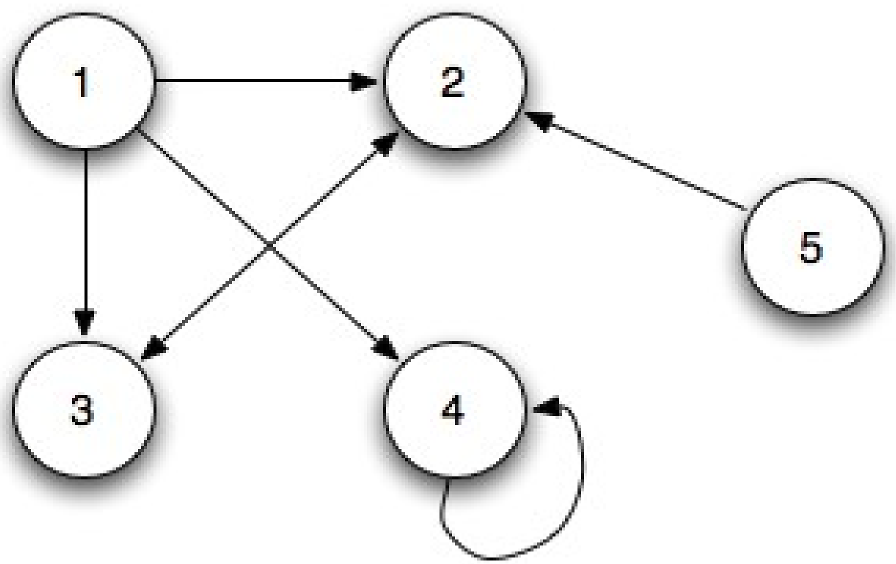
A directed graph G is a set of edges and vertices: G = (V, E)

Two common ways to represent a graph:
-Adjacency matrix
-Adjacency list



Graph

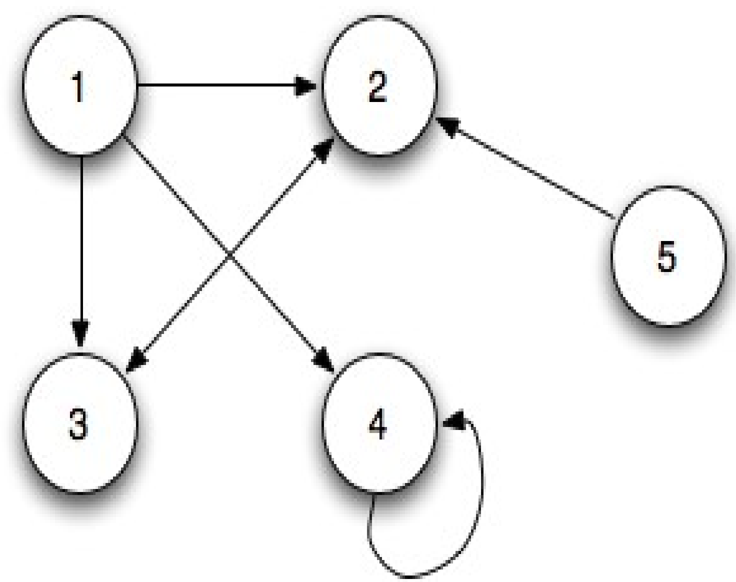
An adjacency matrix has a 1 in row i and column j if you can go from node i to node j



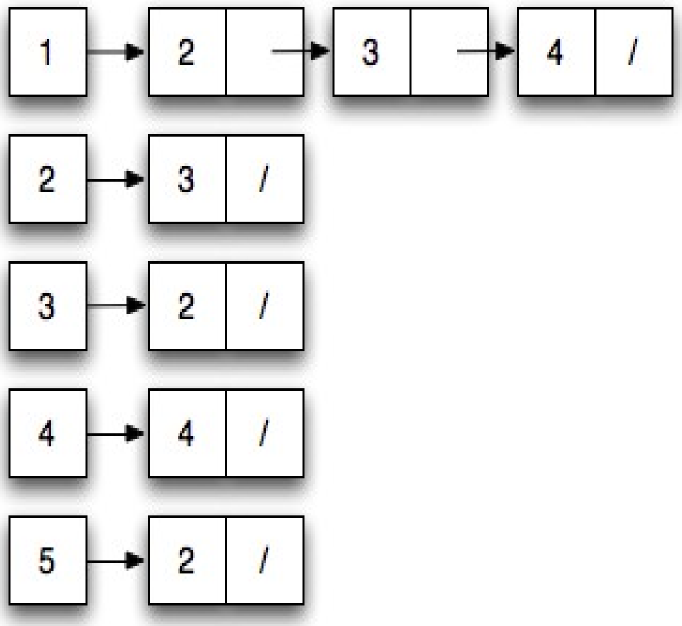
	1	2	3	4	5
1	0	1	1	1	0
2	0	0	1	0	0
3	0	1	0	0	0
4	0	0	0	1	0
5	0	1	0	0	0

Graph

An adjacency list just makes lists out of each row (list of edges out from every vertex)



	1	2	3	4	5
1	0	1	1	1	0
2	0	0	1	0	0
3	0	1	0	0	0
4	0	0	0	1	0
5	0	1	0	0	0



Graph

Difference between adjacency matrix and adjacency list?

Graph

Difference between adjacency matrix and adjacency list?

Matrix is more memory $O(|V|^2)$, less computation: O(1) lookup

List is less memory O(E+V) if sparse, more computation: O(branch factor)

Graph

Adjacency matrix, $A=A^1$, represents the number of paths from row node to column node in 1 step

Prove: A^n is the number of paths from row node to column node in n steps

Graph

Proof: Induction

Base: $A^0 = I$, 0 steps from i is i

Induction: (Assume A^n, show A^{n+1})

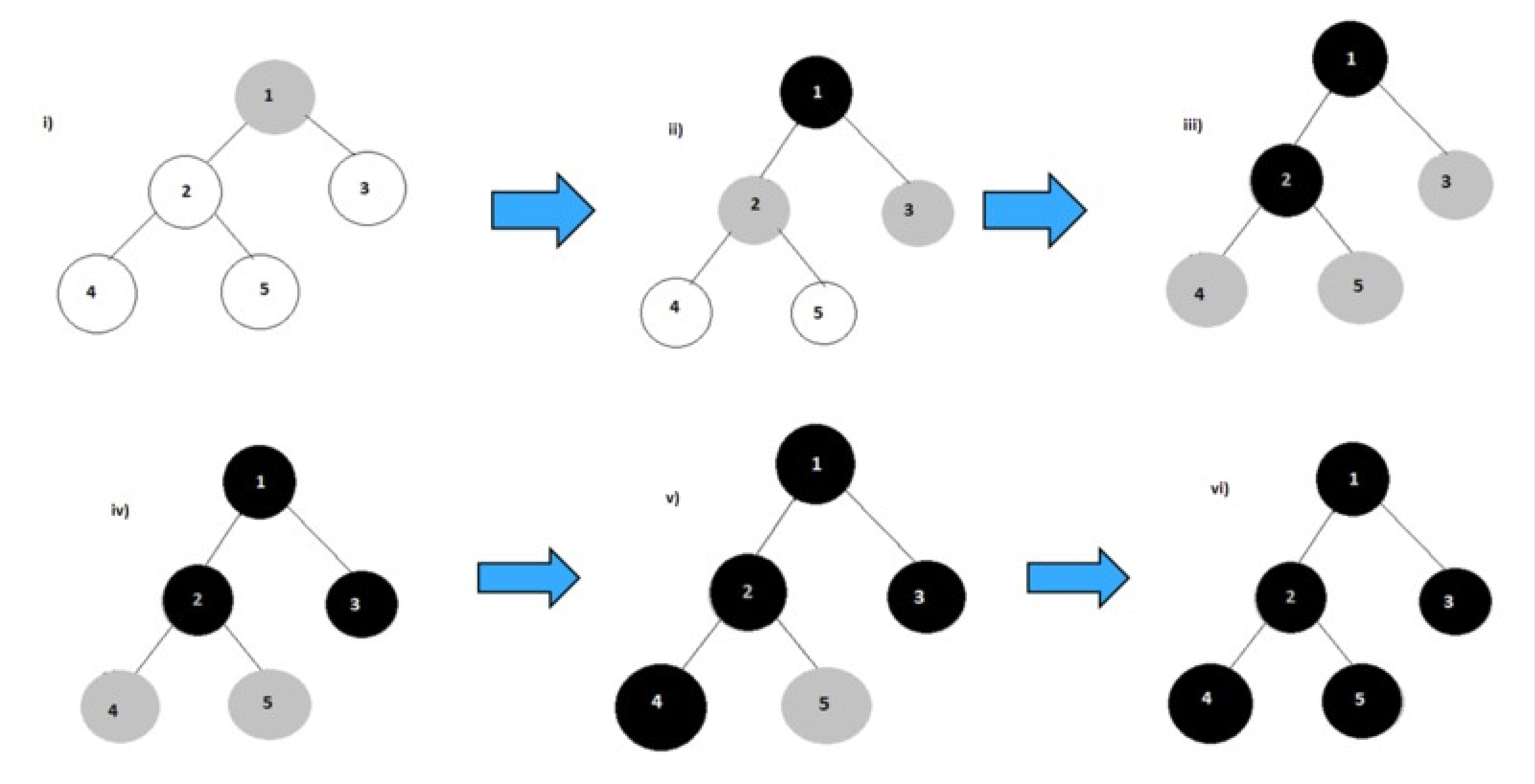
Let $a^n_{i,j}$ = i^{th} row, j^{th} column of A^n

Then $a^{n+1}_{i,j} = \sum_k a^n_{i,k}\, a^1_{k,j}$

This is just matrix multiplication

Breadth First Search Overview

Create first-in-first-out (FIFO) queue to explore unvisited nodes

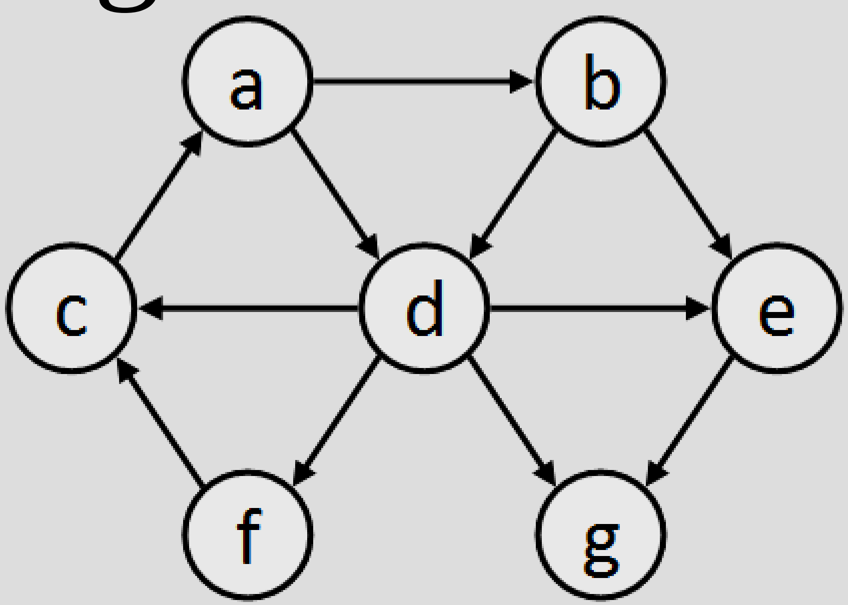


https://www.youtube.com/watch?v=nI0dT288VLs

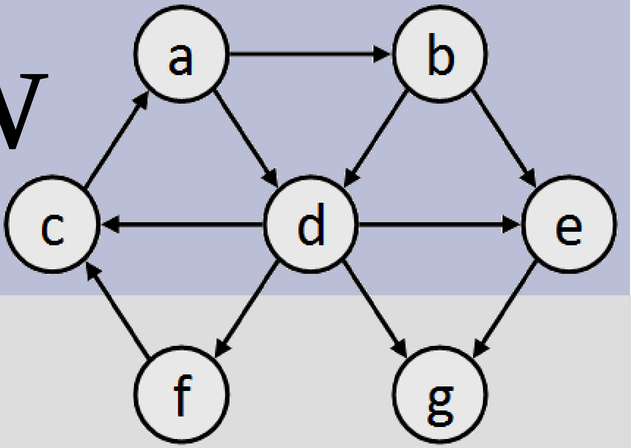
Breadth First Search Overview

Consider the graph below

Suppose we wanted to get from "a" to "c" using breadth first search



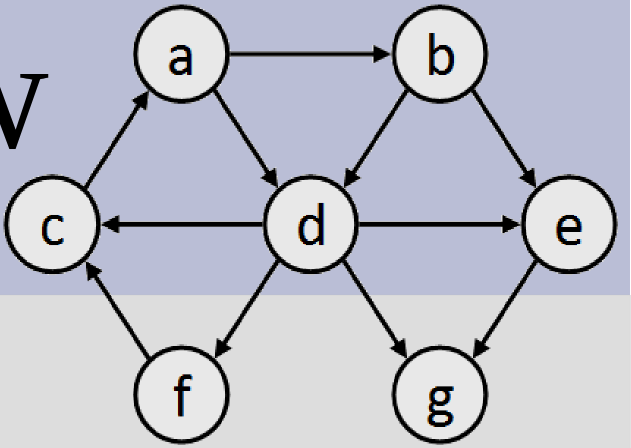
BFS Overview



To keep track of which nodes we have seen, we will do:

White nodes = never seen before
Grey nodes = nodes in Q
Black nodes = nodes that are done
To keep track of who first saw nodes I will make red arrows (π in book)

BFS Overview



First, we add the start to the queue, so Q = {a}

Then we will repeatedly take the left-most item in Q and add all of its neighbors (that we haven't seen yet) to the Q on the right

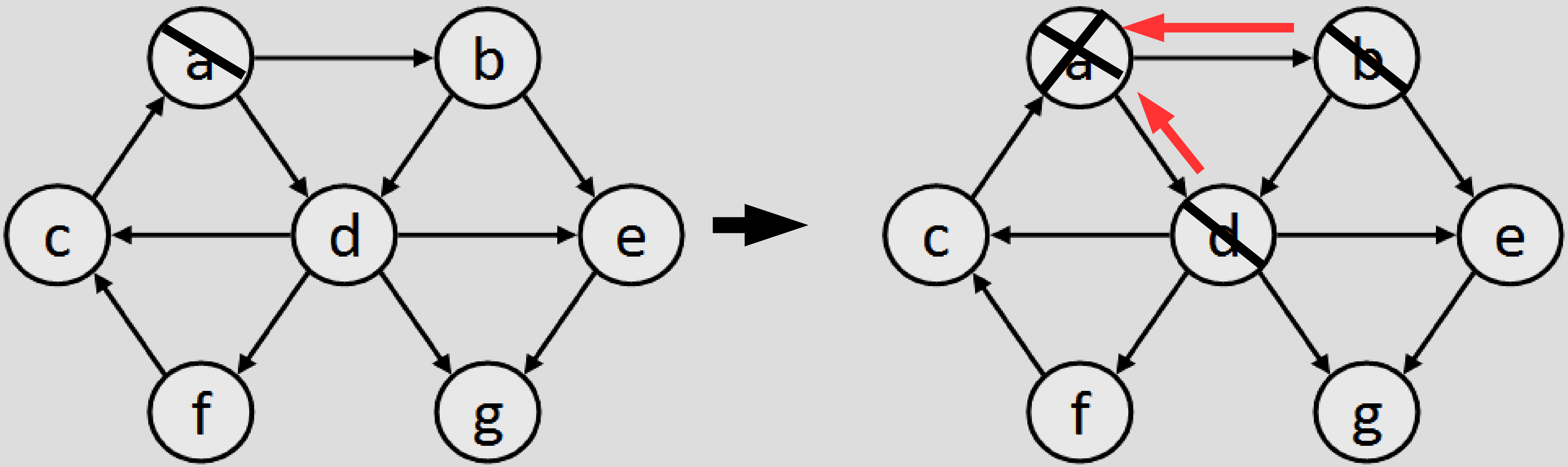
BFS Overview

Q = {a}

Left-most = a

White neighbors = b & d

New Q = {b, d}



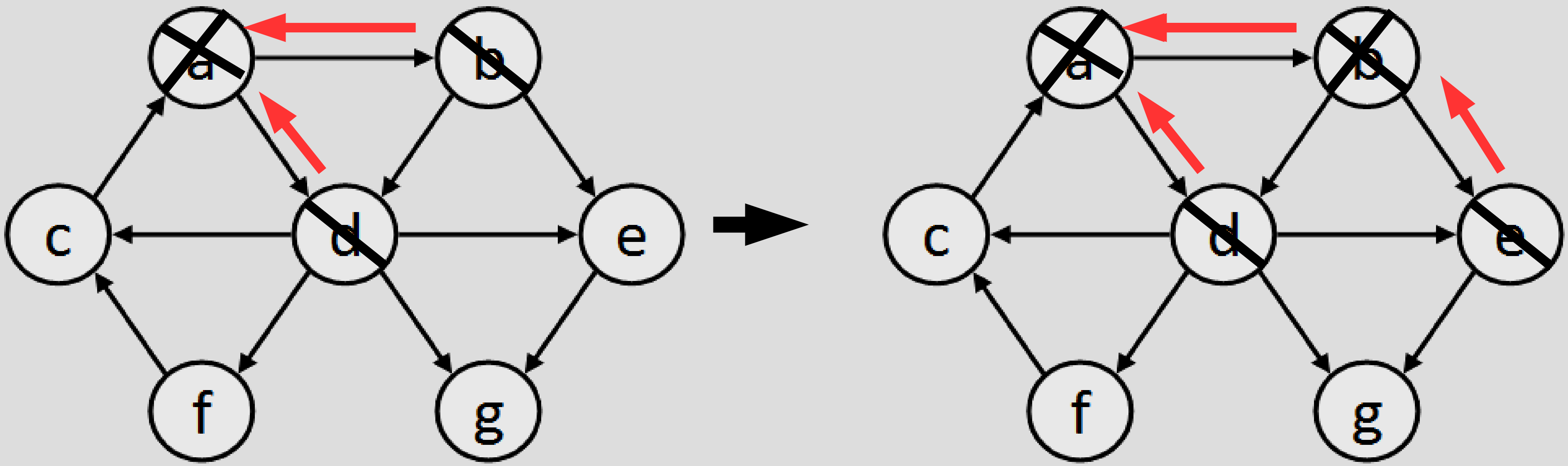
BFS Overview

Q = {b, d}

Left-most = b

White neighbors = e

New Q = {d, e}



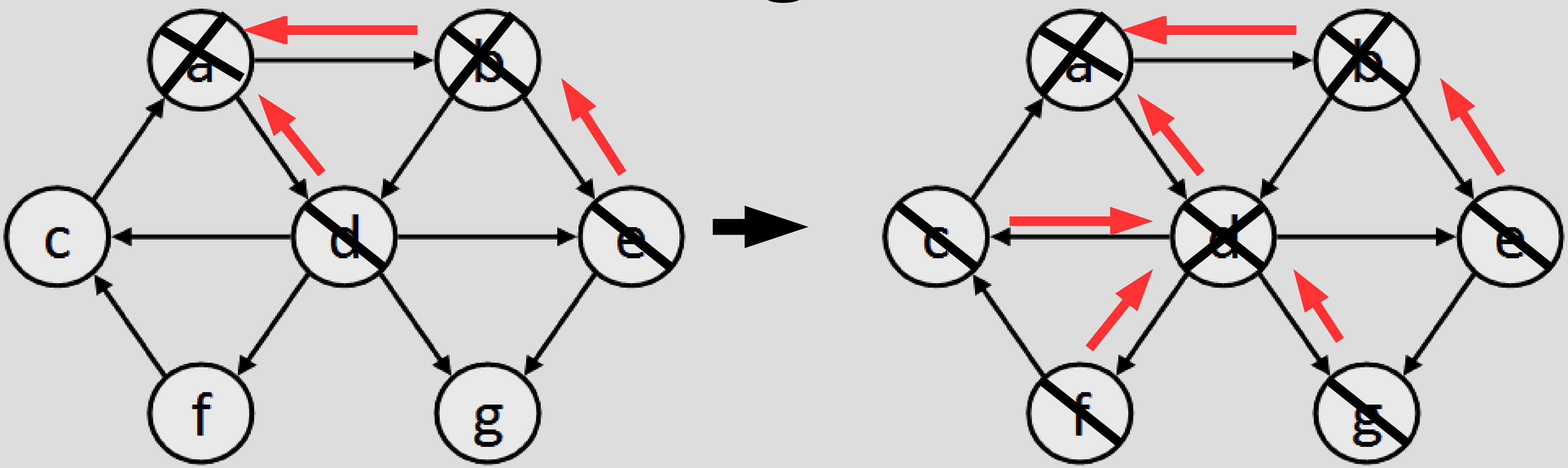
BFS Overview

Q = {d, e}

Left-most = d

White neighbors = c & f & g

New Q = {e, c, f, g}



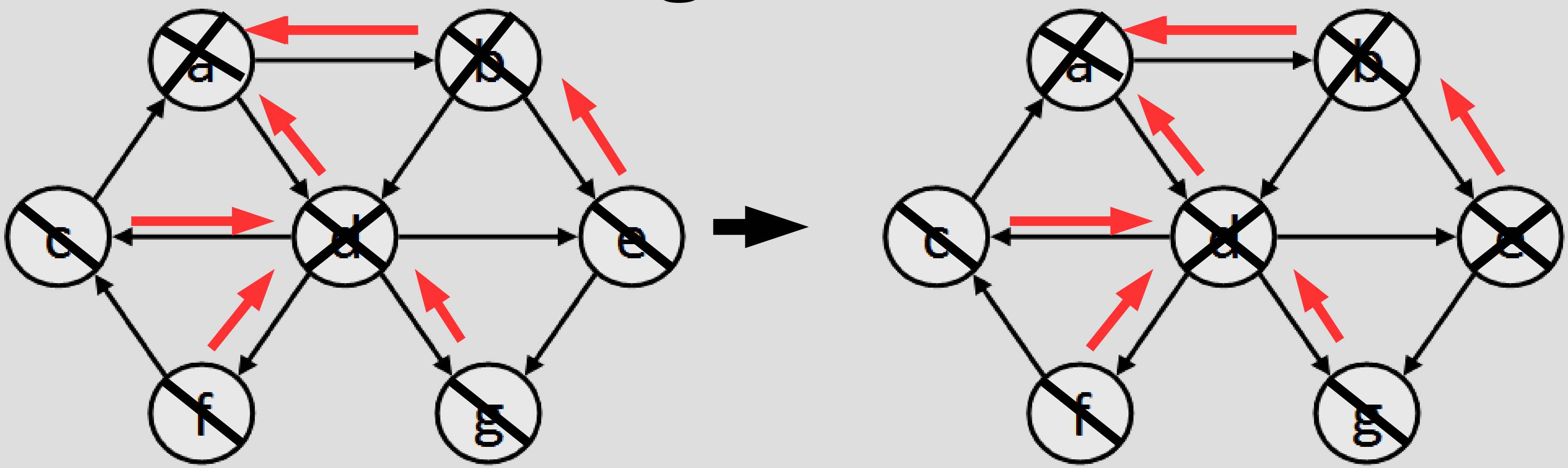
BFS Overview

Q = {e, c, f, g}

Left-most = e

White neighbors = (none)

New Q = {c, f, g}

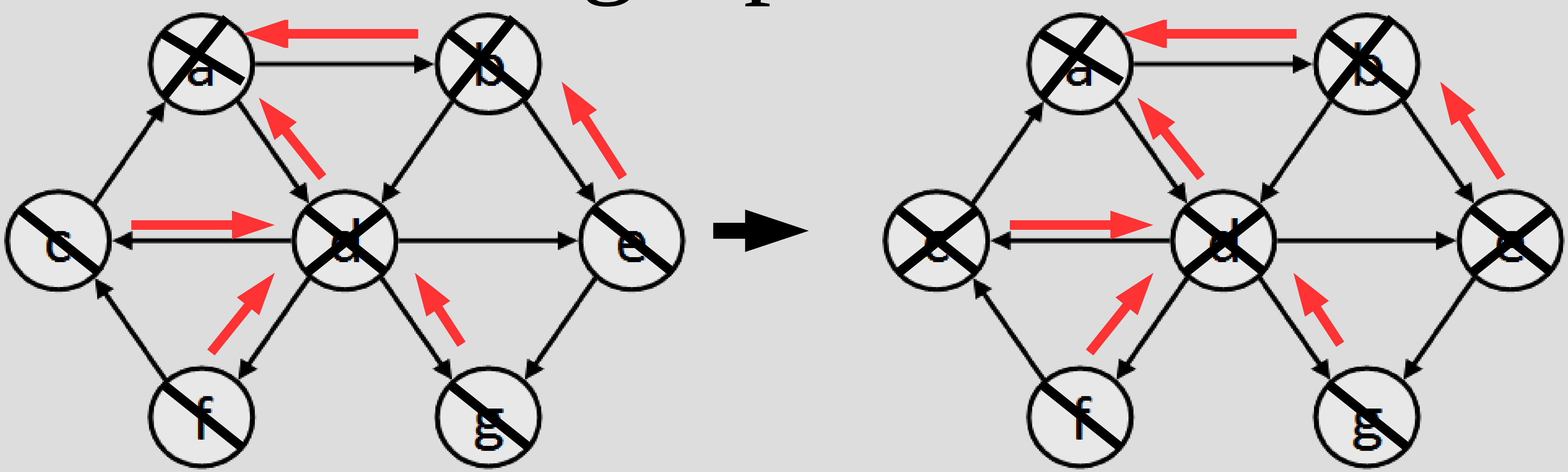


BFS Overview

Q = {c, f, g}

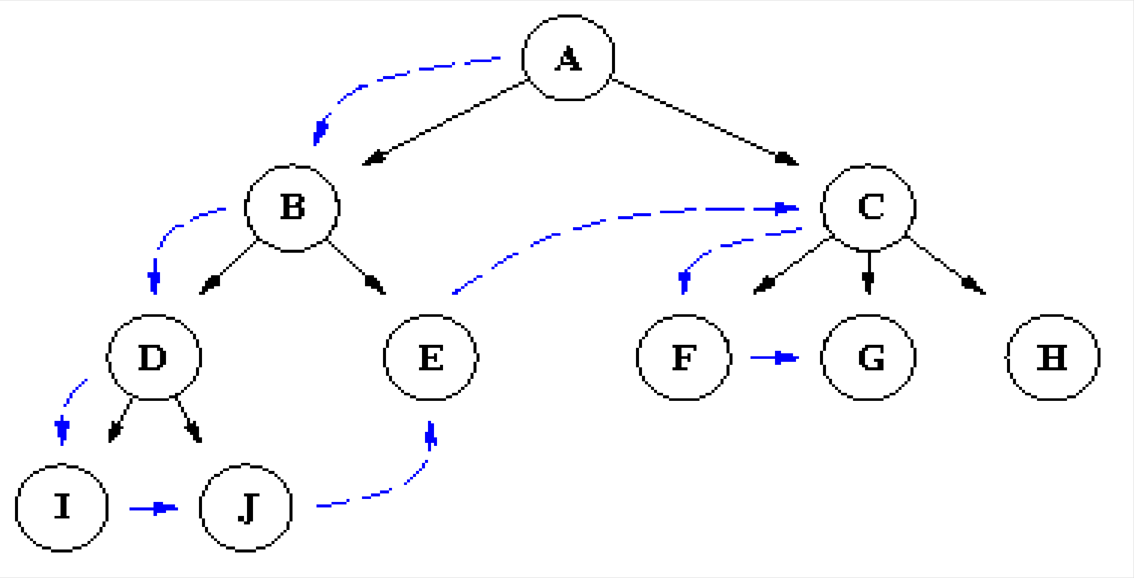
Left-most = c

Done! We found c, backtrack on red arrows to get path from "a"



Depth First Search Overview

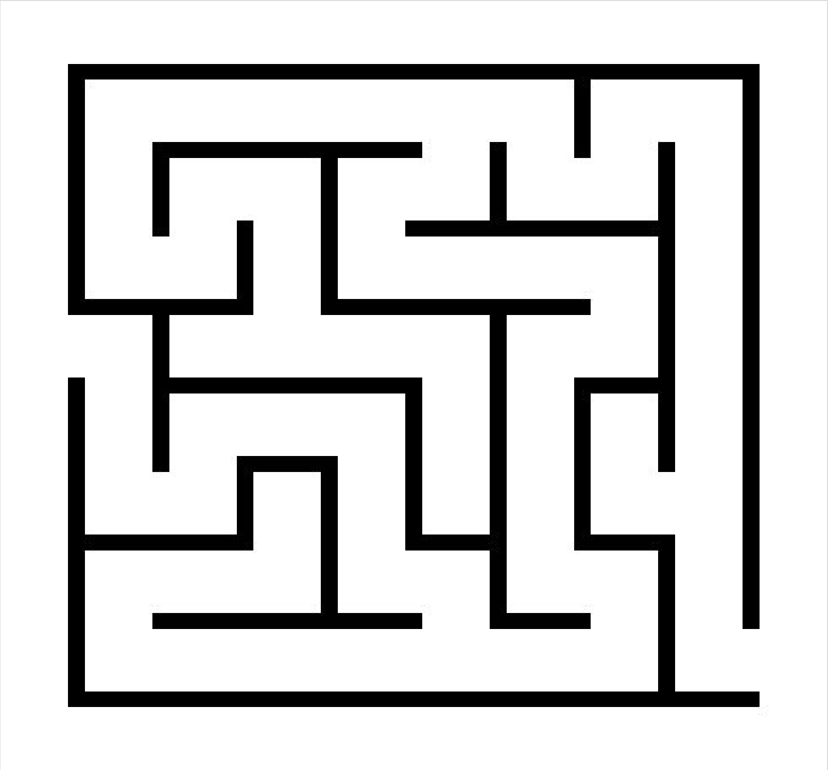
Create first-in-last-out (FILO) queue to explore unvisited nodes



Depth First Search Overview

You can solve mazes by putting your left-hand on the wall and following it

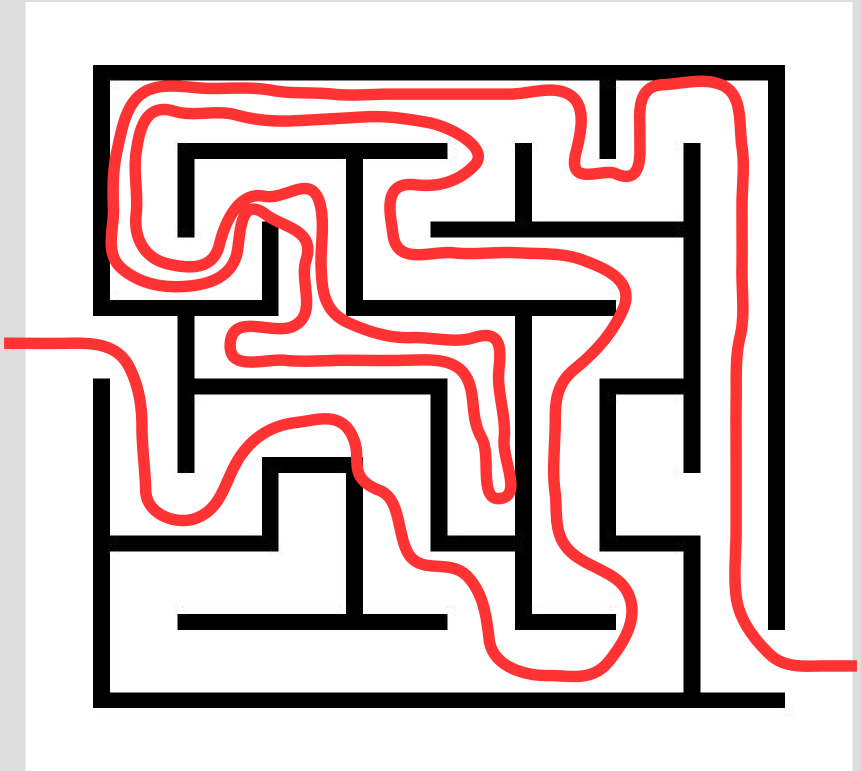
(i.e. left turns at every intersection)



Depth First Search Overview

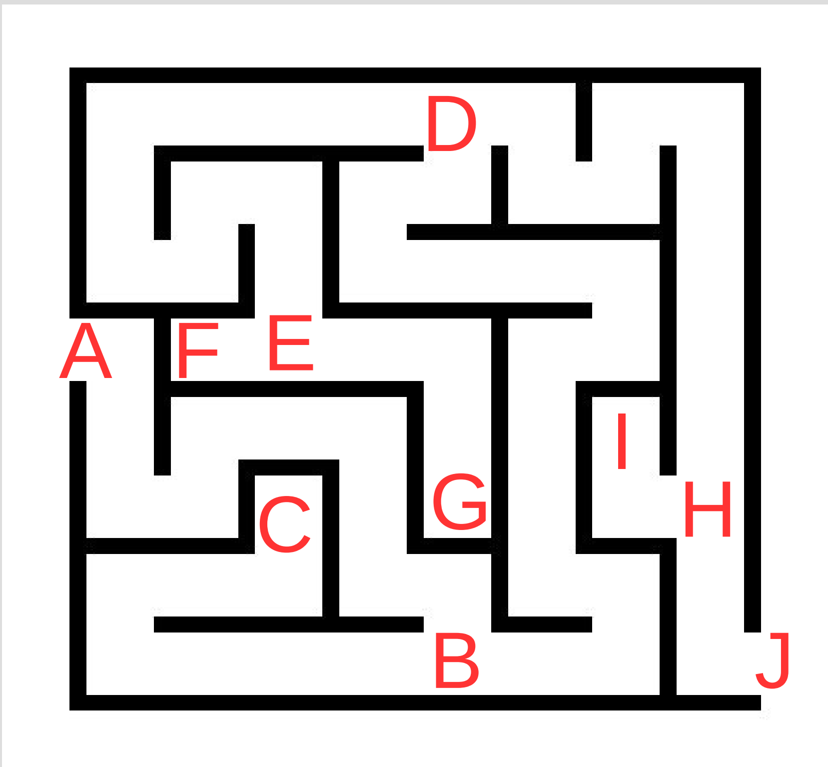
You can solve mazes by putting your left-hand on the wall and following it

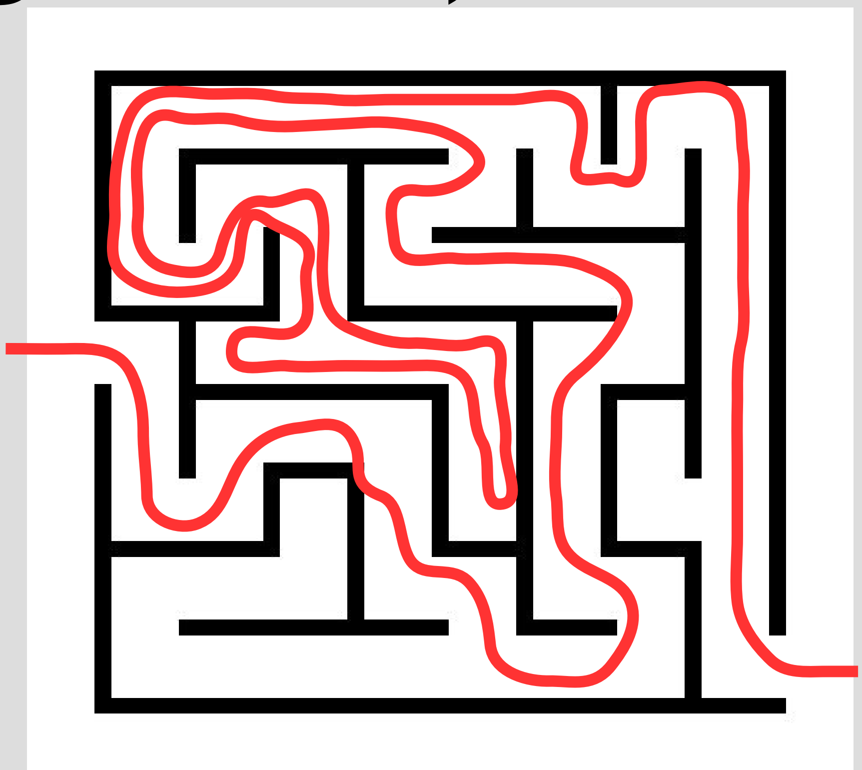
(i.e. left turns at every intersection)



Depth First Search Overview

This is actually just depth first search (add nodes to the "right" first)





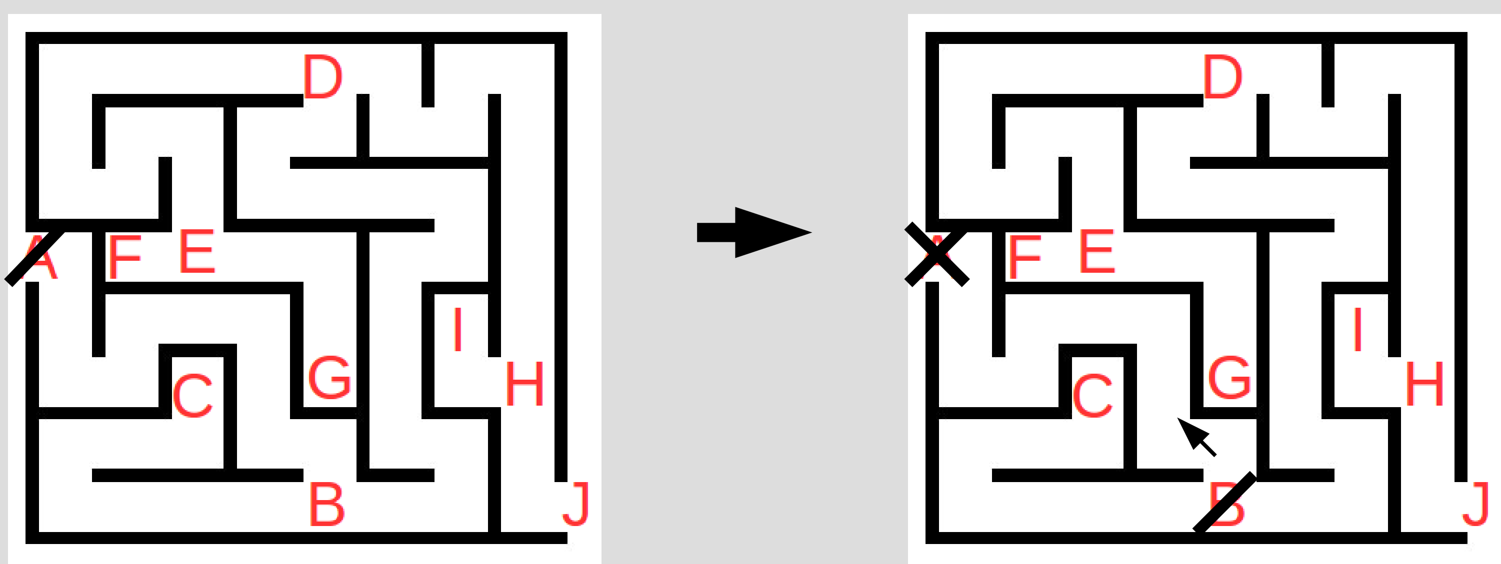
Depth First Search Overview

Q = {A}

Right most = A

White neighbors = {B}

New Q = {B}



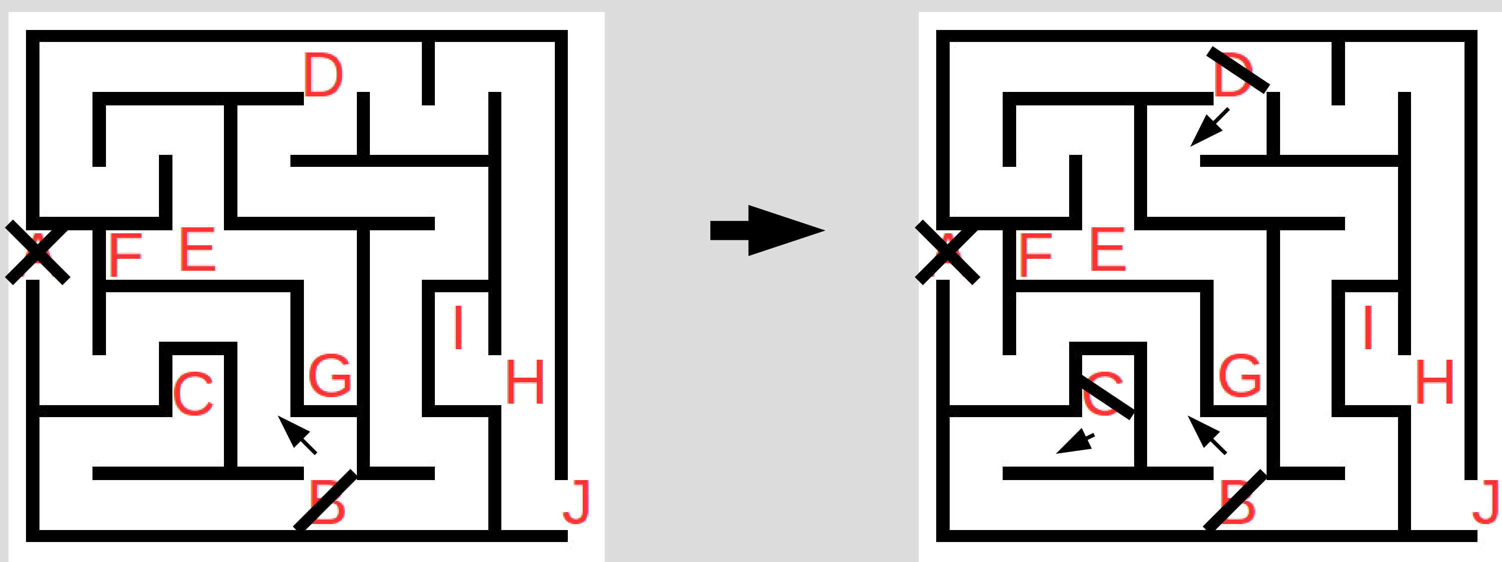
Depth First Search Overview

Q = {B}

Right most = B

White neighbors = {C, D}

New Q = {C, D}



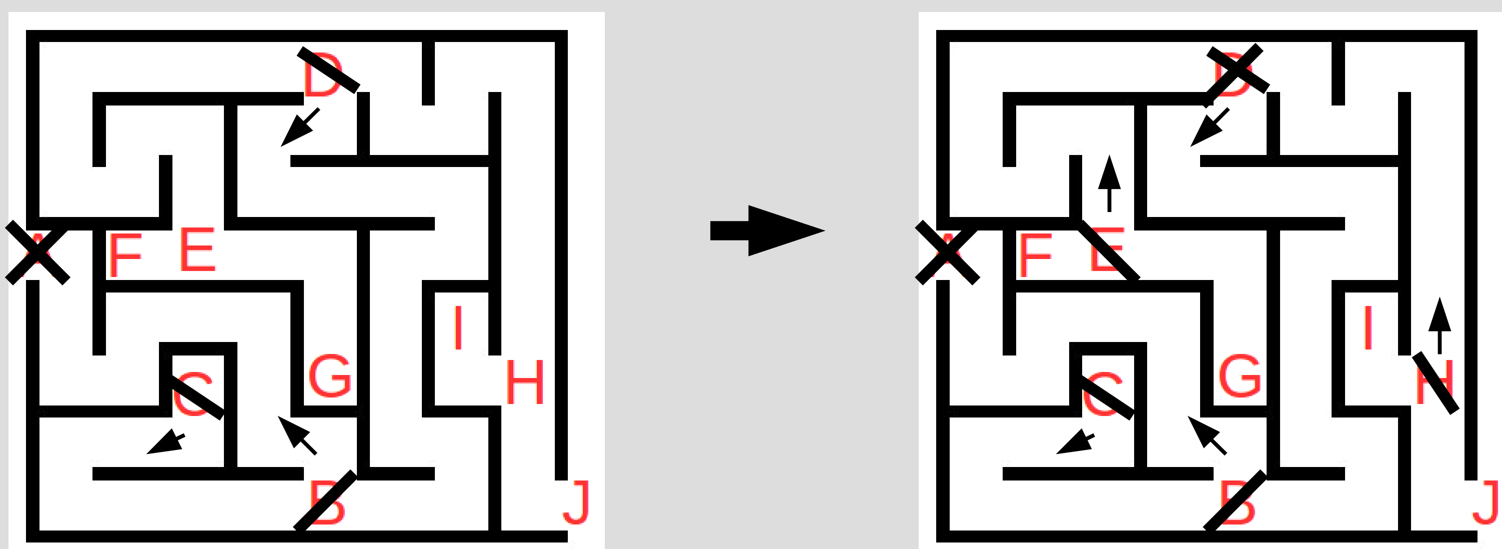
Depth First Search Overview

Q = {C, D}

Right most = D

White neighbors = {H, E}

New Q = {C, H, E}



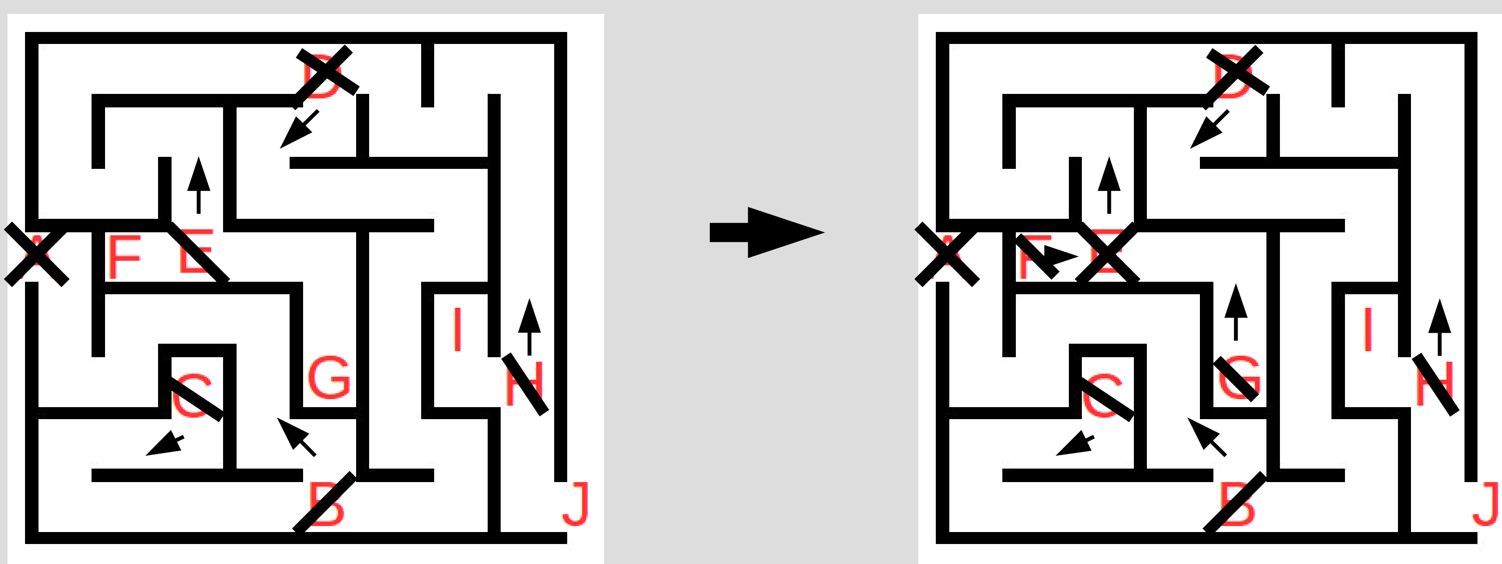
Depth First Search Overview

Q = {C, H, E}

Right most = E

White neighbors = {F, G}

New Q = {C, H, F, G}



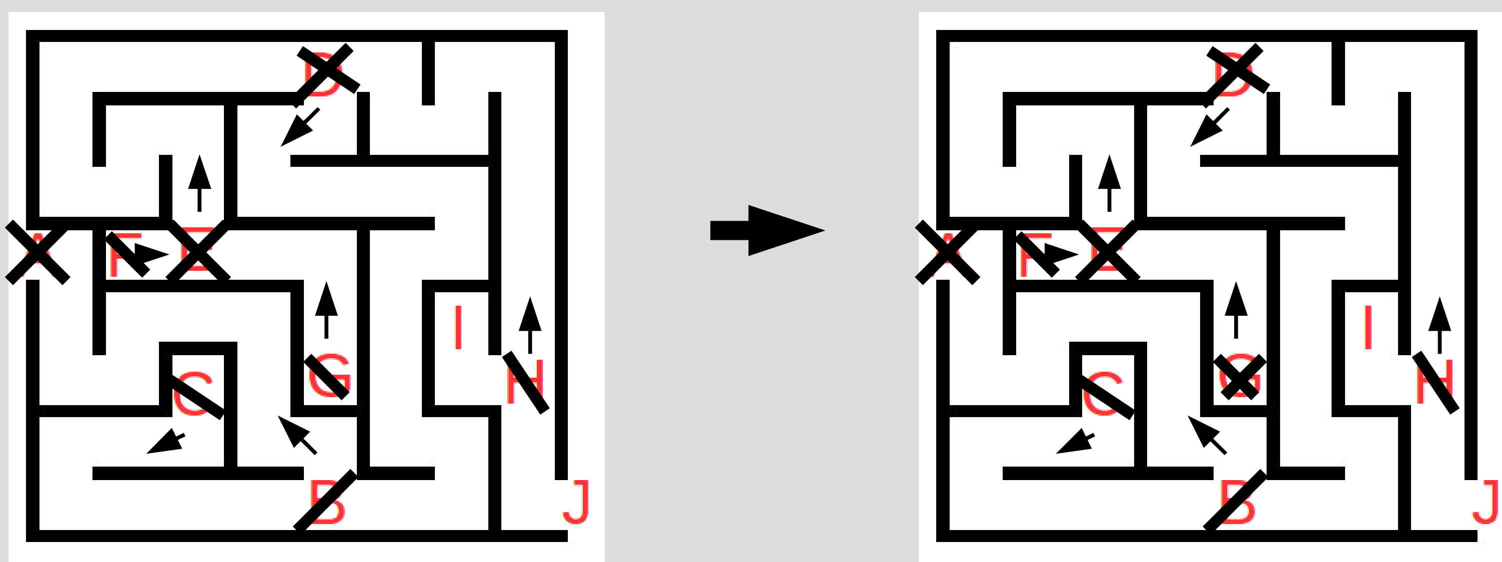
Depth First Search Overview

Q = {C, H, F, G}

Right most = G

White neighbors = {}

New Q = {C, H, F}



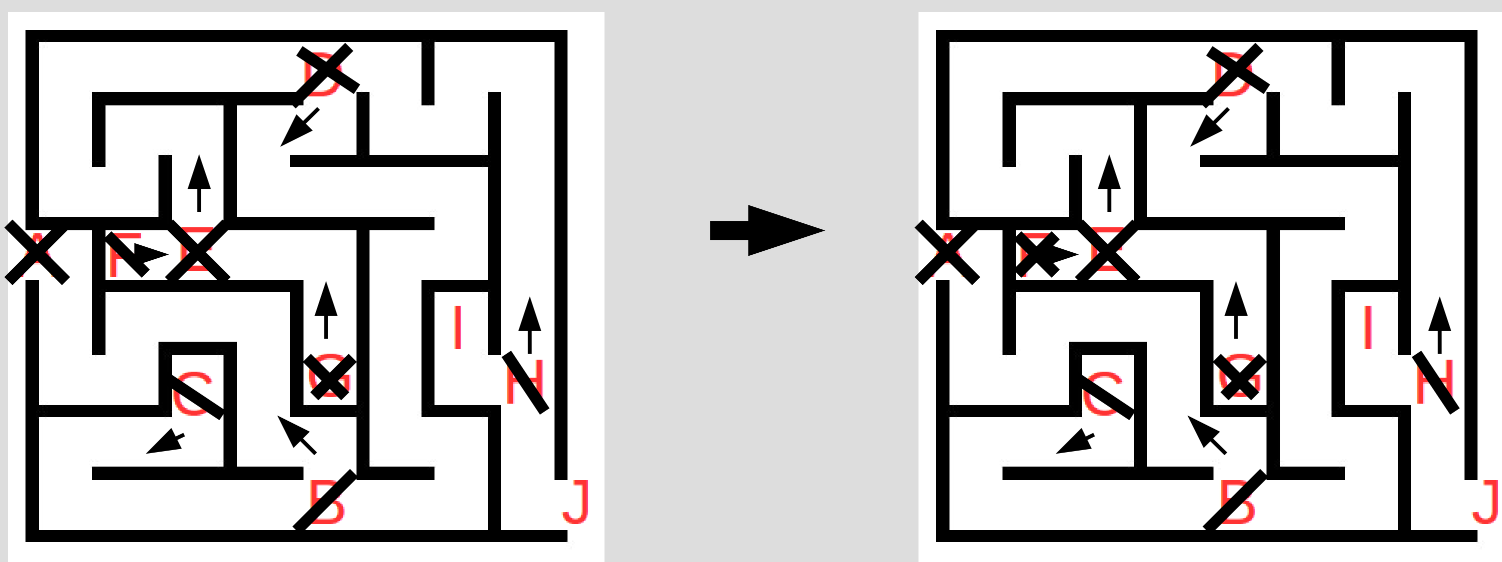
Depth First Search Overview

Q = {C, H, F}

Right most = F

White neighbors = {}

New Q = {C, H}



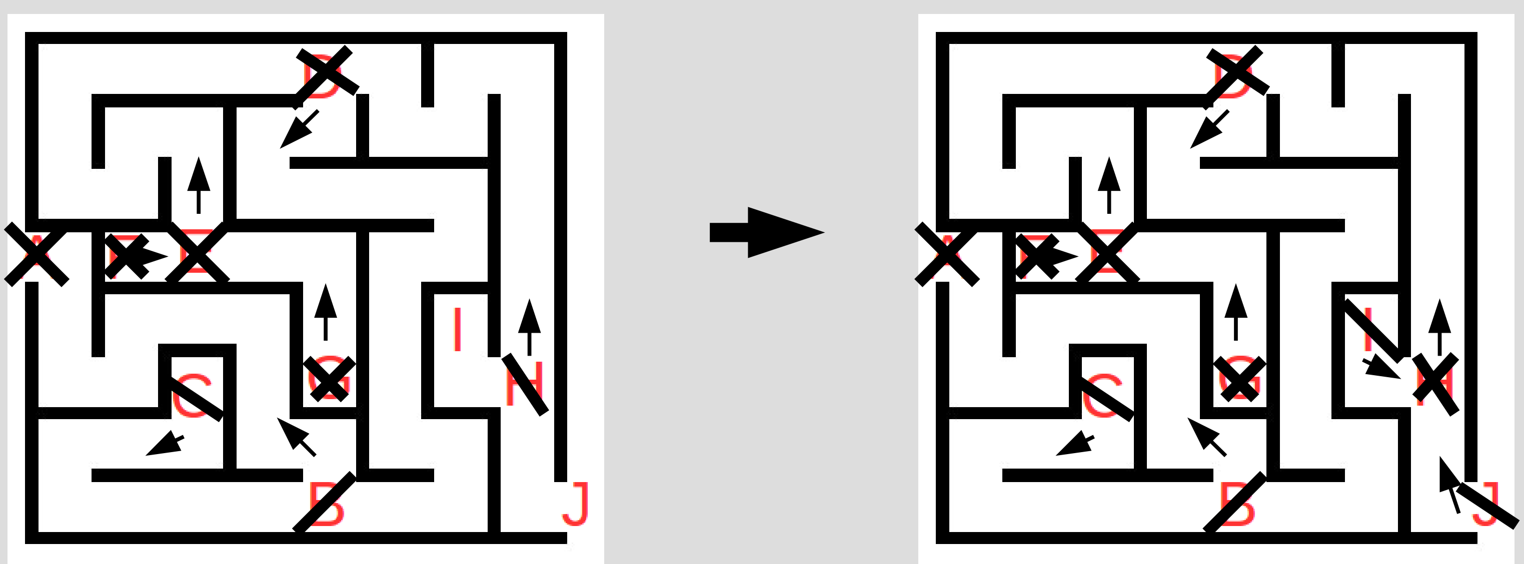
Depth First Search Overview

Q = {C, H}

Right most = H

White neighbors = {I, J}

New Q = {C, I, J}

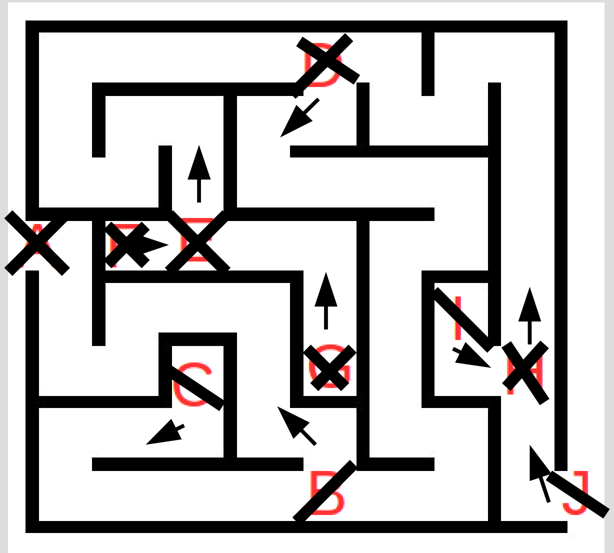


Depth First Search Overview

Q = {C, I, J}

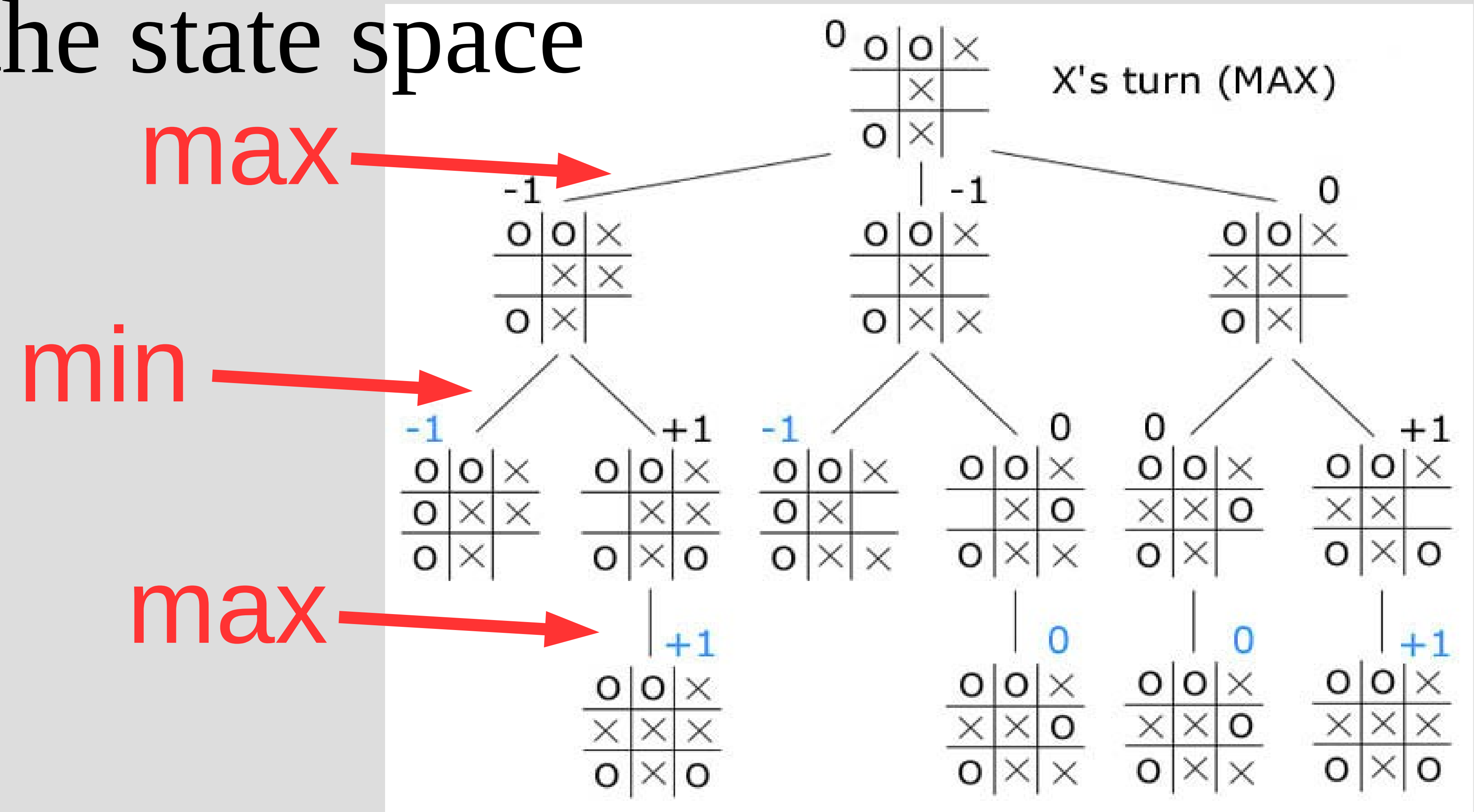
Right most = J

J is exit, we are done



BFS and DFS in trees

Solve problems by making a tree of the state space



BFS and DFS in trees

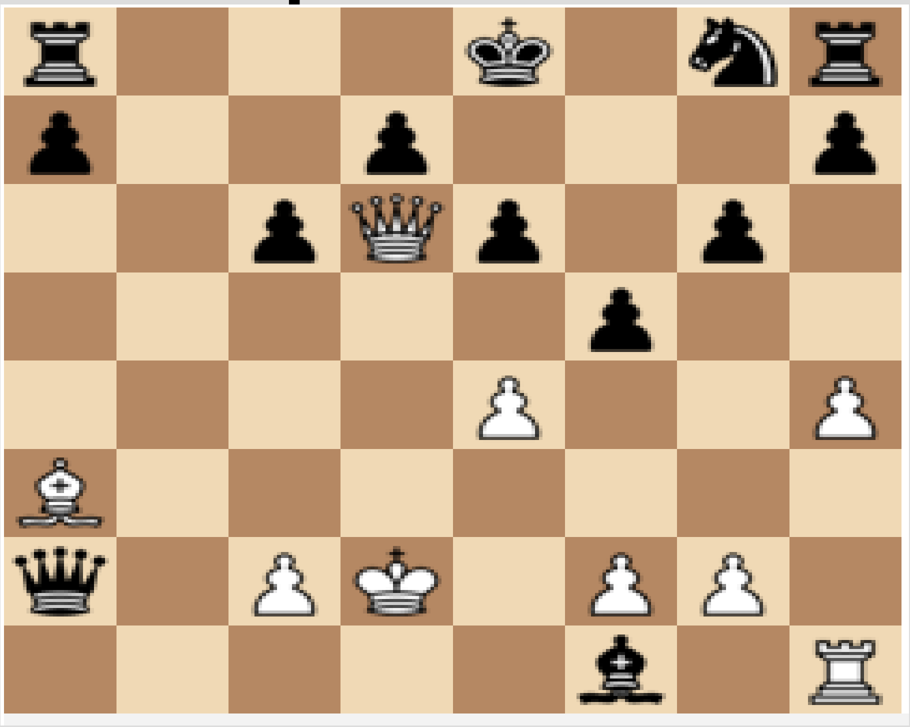
Often times, fully exploring the state space is too costly (takes forever)

Chess: 10^{47} states (tree about 10^{123})
Go: 10^{171} states (tree about 10^{360})
At 1 million states per second...
Chess: 10^{109} years (past heat death
Go: 10^{346} years of universe)

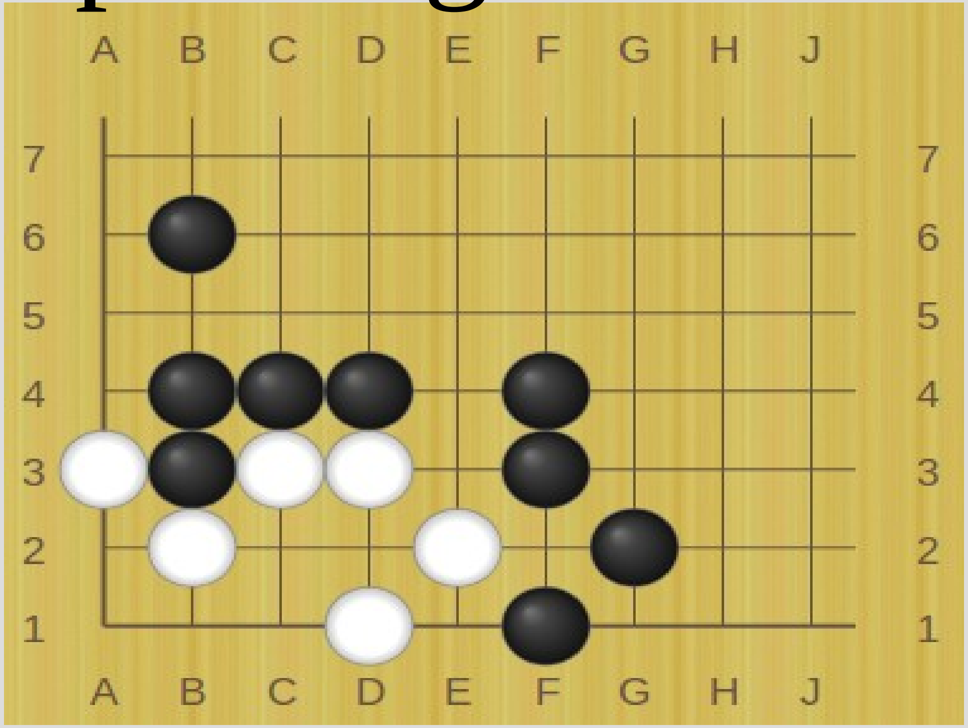
BFS and DFS in trees

BFS prioritizes “exploring”

DFS prioritizes “exploiting”



White to move



Black to move

BFS and DFS in trees

BFS benefits?

DFS benefits?

BFS and DFS in trees

BFS benefits?

-if stopped before full search, can evaluate best found

DFS benefits?

-uses less memory on complete search

BFS and DFS in graphs

BFS: shortest path from origin to any node

DFS: find graph structure

Both running time of O(V+E)

Breadth first search

```
BFS(G,s) // to find shortest path from s
for all v in V
    v.color=white, v.d=∞,v.π=NIL
s.color=grey, v.d=0
Enqueue(Q,s)
while(Q not empty)
    u = Dequeue(Q,s)
    for v in G.adj[u]
        if v.color == white
            v.color=grey, v.d=u.d+1, v.π=u
            Enqueue(Q,v)
    u.color=black
```

Breadth first search

Let $\delta(s,v)$ be the shortest path from s to v

After running BFS you can find this path as: $v.\pi$ to $(v.\pi).\pi$ to ... s

(pseudo code on p. 601, recursion)

BFS correctness

Proof: contradiction

Assume $\delta(s,v) \neq v.d$

$v.d \geq \delta(s,v)$ (Lemma 22.2, induction)

Thus $v.d > \delta(s,v)$

Let u be previous node on $\delta(s,v)$

Thus $\delta(s,v) = \delta(s,u)+1$

and $\delta(s,u) = u.d$

Then $v.d > \delta(s,v) = \delta(s,u)+1 = u.d+1$

BFS correctness

$v.d > \delta(s,v) = \delta(s,u)+1 = u.d+1$

Cases on color of v when u dequeue,
all cases invalidate top equation

Case white: alg sets $v.d = u.d + 1$

Case black: already removed

thus $v.d \leq u.d$ (corollary 22.4)

Case grey: exists w that dequeued v,

$v.d = w.d+1 \leq u.d+1$ (corollary 22.4)

Depth first search

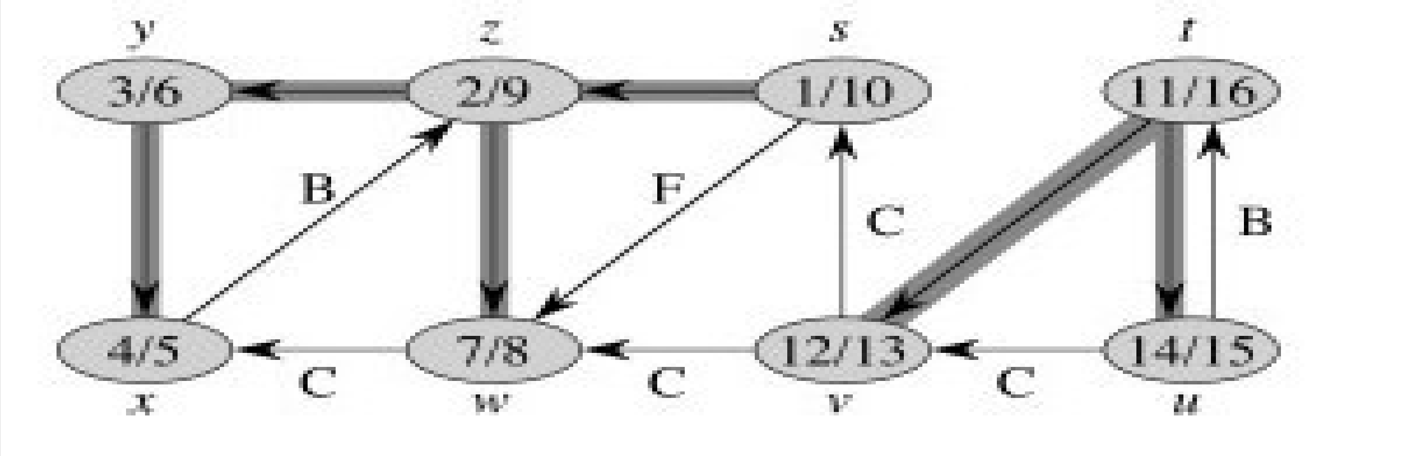
```
DFS(G)
for all v in V
   v.color=white, v.π=NIL
time=0
for each v in V
   if v.color==white
      DFS-Visit(G,v)
```

Depth first search

```
DFS-Visit(G,u)
time=time+1
u.d=time, u.color=grey
for each v in G.adj[u]
   if v.color == white
      v.π=u
      DFS-Visit(G,v)
u.color=black, time=time+1, u.f=time
```

Depth first search

Edge markers:



Consider edge u to v

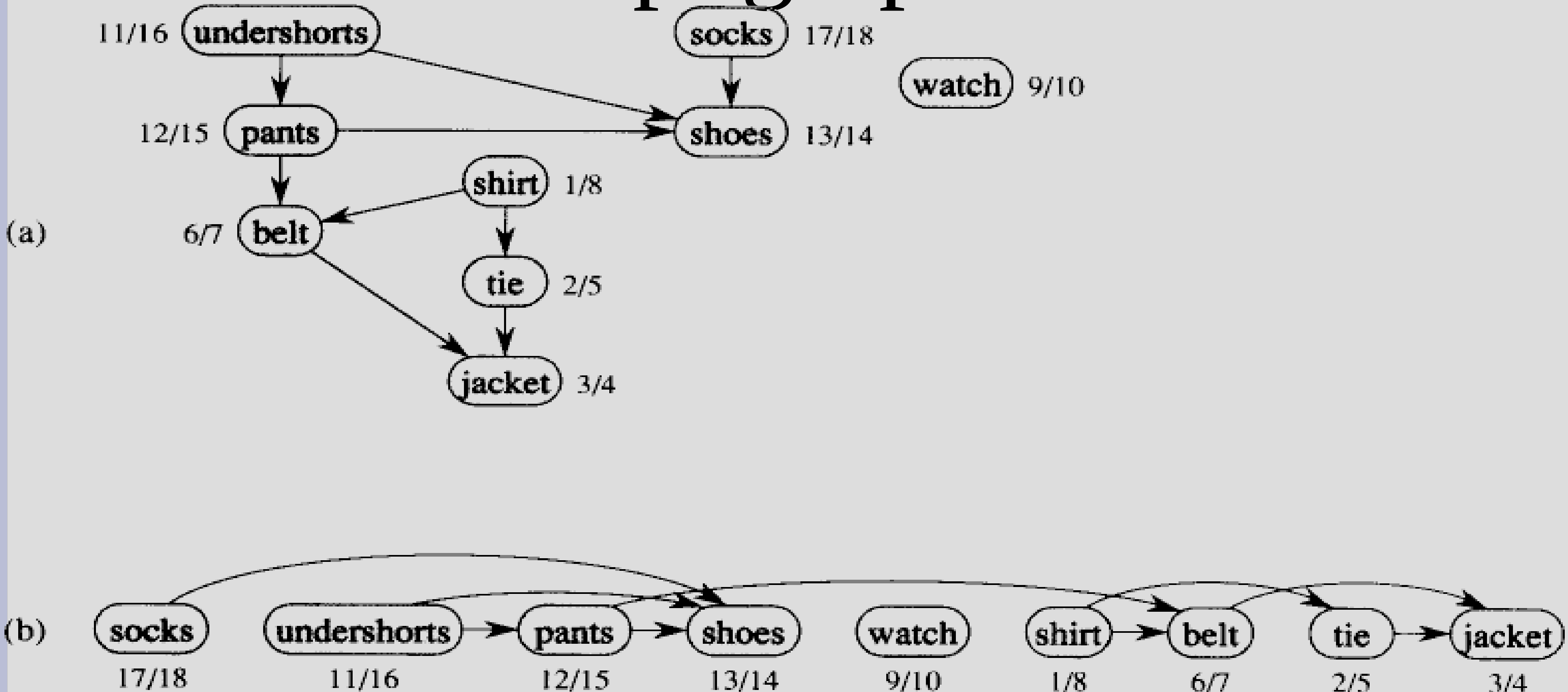
C = Edge to black node (u.d > v.f)

B = Edge to grey node (u.f < v.f)

F = Edge to black node (u.f > v.f)

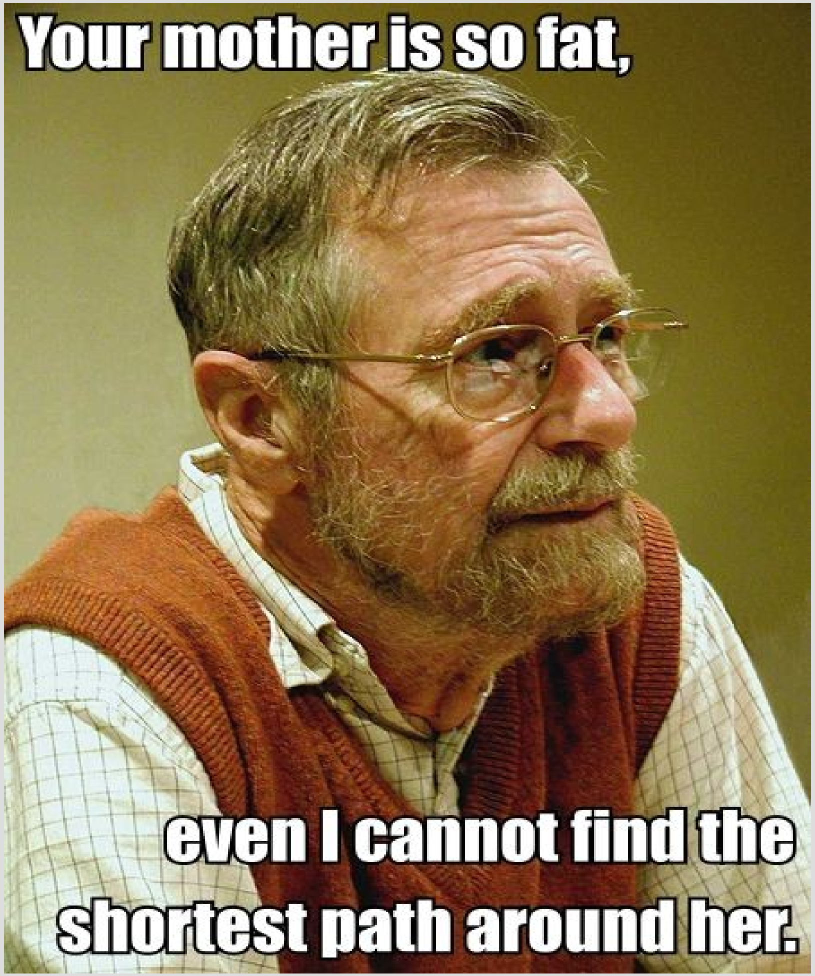
Depth first search

DFS can do topographical sort



Run DFS, sort in decreasing finish time

Weighted graphs



Weighted graph

Edges in weighted graph are assigned a weight: $w(v_1, v_2)$, v_1, v_2 in V

If path $p = <v_0, v_1, \ldots v_k>$ then the weight is: $w(p) = \sum^{k}_{i=0}(v_{i-1},v_i)$

Shortest Path:

$\delta(u,v)$: $\min\{w(p) : v_0=u, v_k=v)\}$

Shortest paths

Today we will look at single-source shorted paths

This finds the shortest path from some starting vertex, s, to any other vertex on the graph (if it exists)

This creates G_π, the shortest path tree

Shortest paths

Optimal substructure: Let $\delta(v_0,v_k)=p$, then for all $0 \leq i \leq j \leq k$, $\delta(v_i,v_j)=p_{i,j}= \langle v_i, v_{i+1}, \ldots v_j \rangle$

Proof?

Where have we seen this before?

Shortest paths

Optimal substructure: Let $\delta(v_0,v_k)=p$, then for all $0 \leq i \leq j \leq k$, $\delta(v_i,v_j)=p_{i,j}=$ $<v_i, v_{i+1}, \ldots v_j>$

Proof? Contradiction!

Suppose $w(p'_{i,j}) < p(_{i,j})$, then let $p'_{0,k} = p_{0,i}\ p'_{i,j}\ p_{j,k}$ then $w(p'_{0,k}) < w(p)$

Relaxation

We will only do relaxation on the values v.d (min weight) for vertex v

```
Relax(u,v,w)
if(v.d > u.d + w(u,v))
   v.d = u.d+w(u,v)
   v.π=u
```

Relaxation

We will assume all vertices start with $v.d=\infty, v.\pi=NIL$ except s, $s.d=0$

This will take $O(|V|)$ time

This will not effect the asymptotic runtime as it will be at least $O(|V|)$ to find single-source shortest path

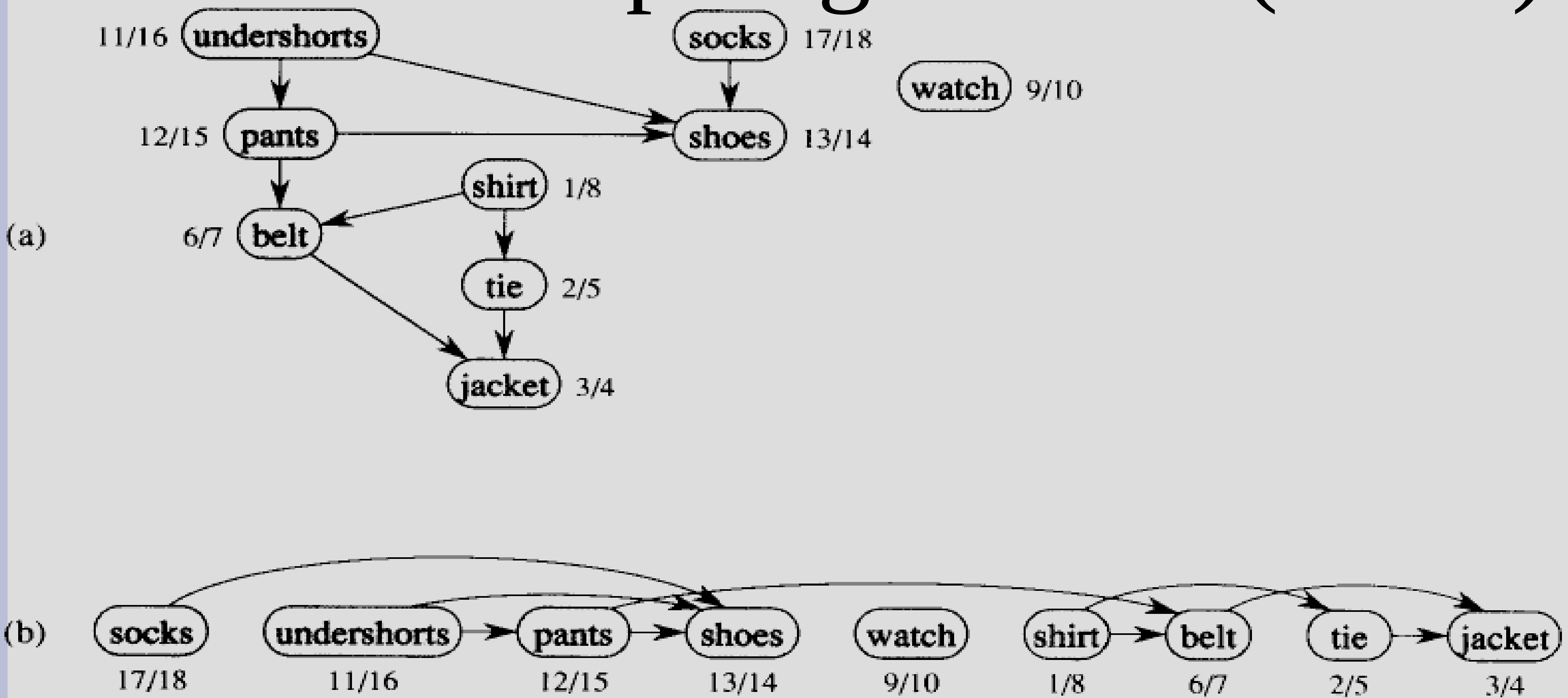
Relaxation

Relaxation properties:

1. $\delta(s,v) \leq \delta(s,u) + \delta(u,v)$ (triangle inequality)
2. $v.d \geq \delta(s,v)$, v.d is monotonically decreasing
3. if no path, $v.d = \delta(s,v) = \infty$
4. if $\delta(s,v)$, when $(v.\pi).d = \delta(s, v.\pi)$ then relax$(v.\pi, v, w)$ causes $v.d = \delta(s,v)$
5. if $\delta(v_0, v_k) = p_{0,k}$, then when relaxed in order $(v_0, v_1), (v_1, v_2), \ldots (v_{k-1}, v_k)$ then $v_{k.d} = \delta(v_0, v_k)$ even if other relax happen
6. when $v.d = \delta(s,v)$ for all v in V, G_π is shortest path tree rooted at s

Directed Acyclic Graphs

DFS can do topological sort (DAG)

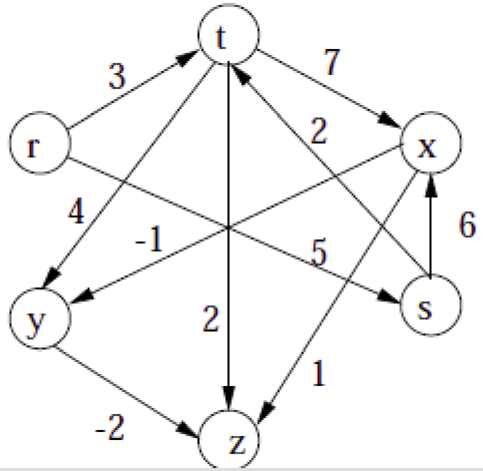


Run DFS, sort in decreasing finish time

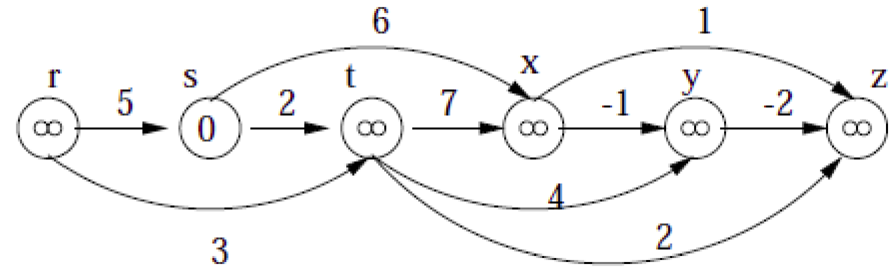
Directed Acyclic Graphs

```
DAG-shortest-paths(G,w,s)
topologically sort G
initialize graph from s
for each u in V in topological order
   for each v in G.Adj[u]
      Relax(u,v,w)
```

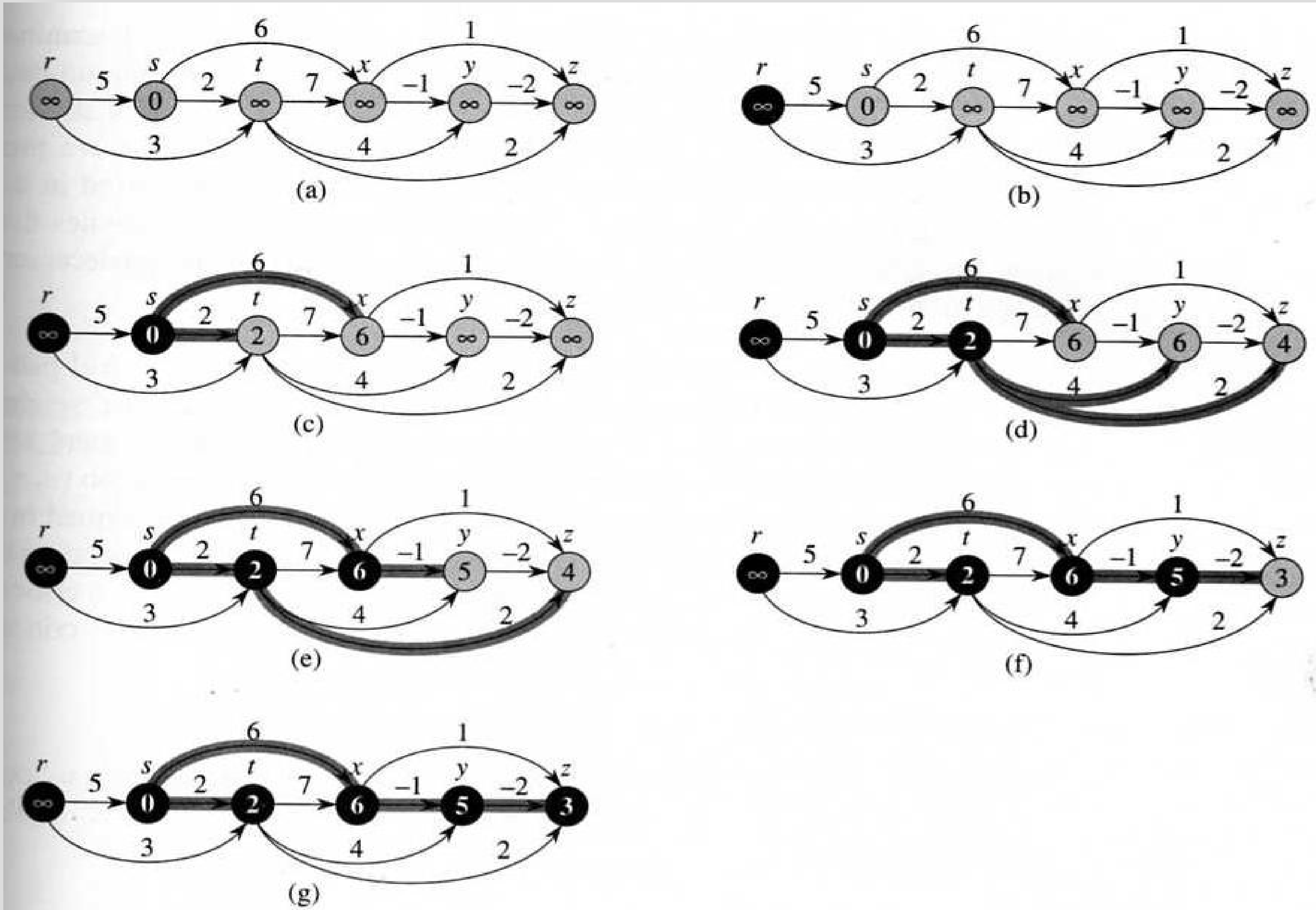
Runtime: $O(|V| + |E|)$



t
r
x
y
s
z
3
7
2
4
-1
5
6
2
1
-2



r
s
t
x
y
z
5
2
7
-1
-2
6
1
3
4
2
0
∞



(a)
(b)
(c)
(d)
(e)
(f)
(g)

Directed Acyclic Graphs

Correctness:

Prove it!

Directed Acyclic Graphs

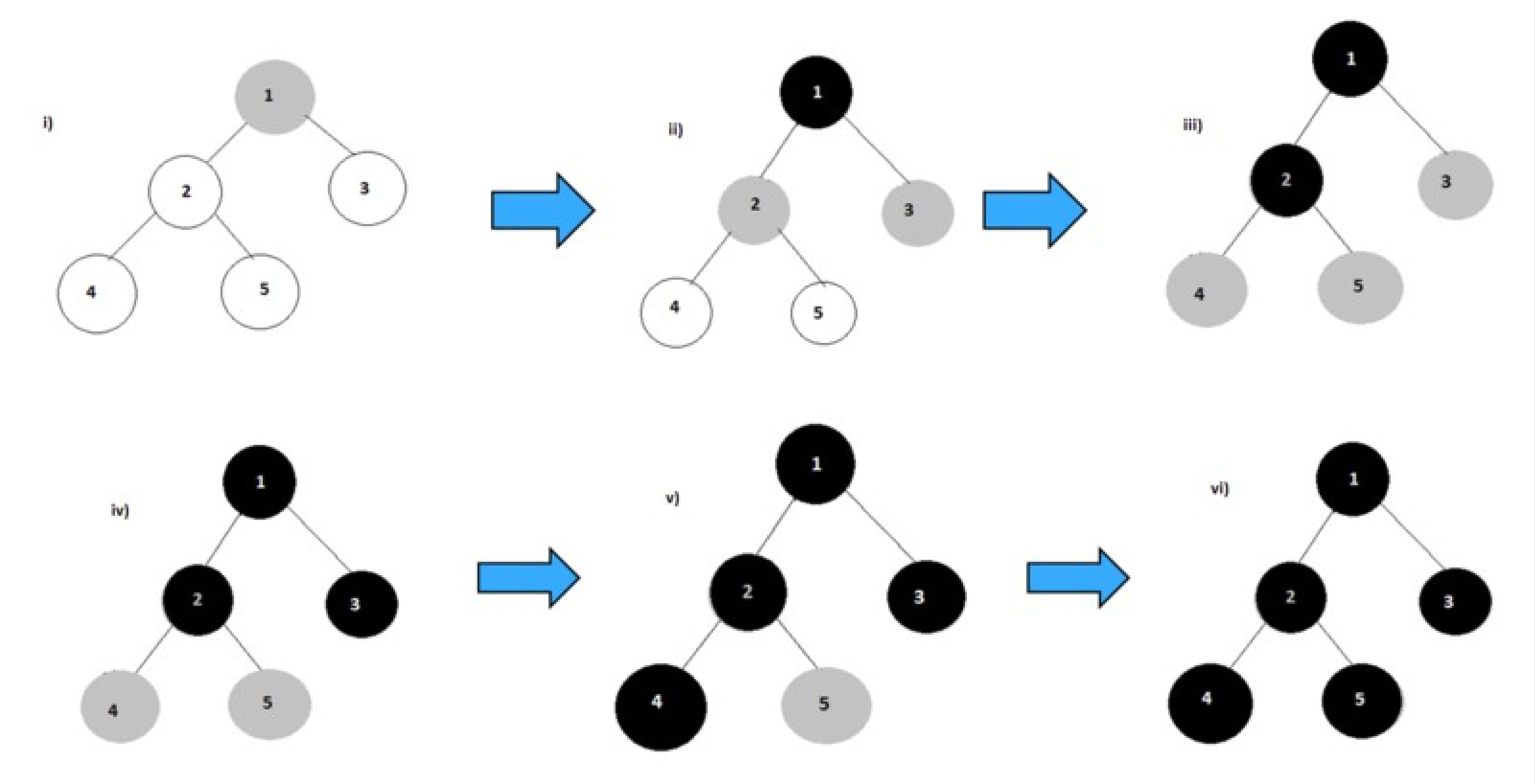
Correctness:

By definition of topological order,
When relaxing vertex v, we have already relaxed any preceding vertices

So by relaxation property 5, we have found the shortest path to all v

BFS (unweighted graphs)

Create FIFO queue to explore unvisited nodes



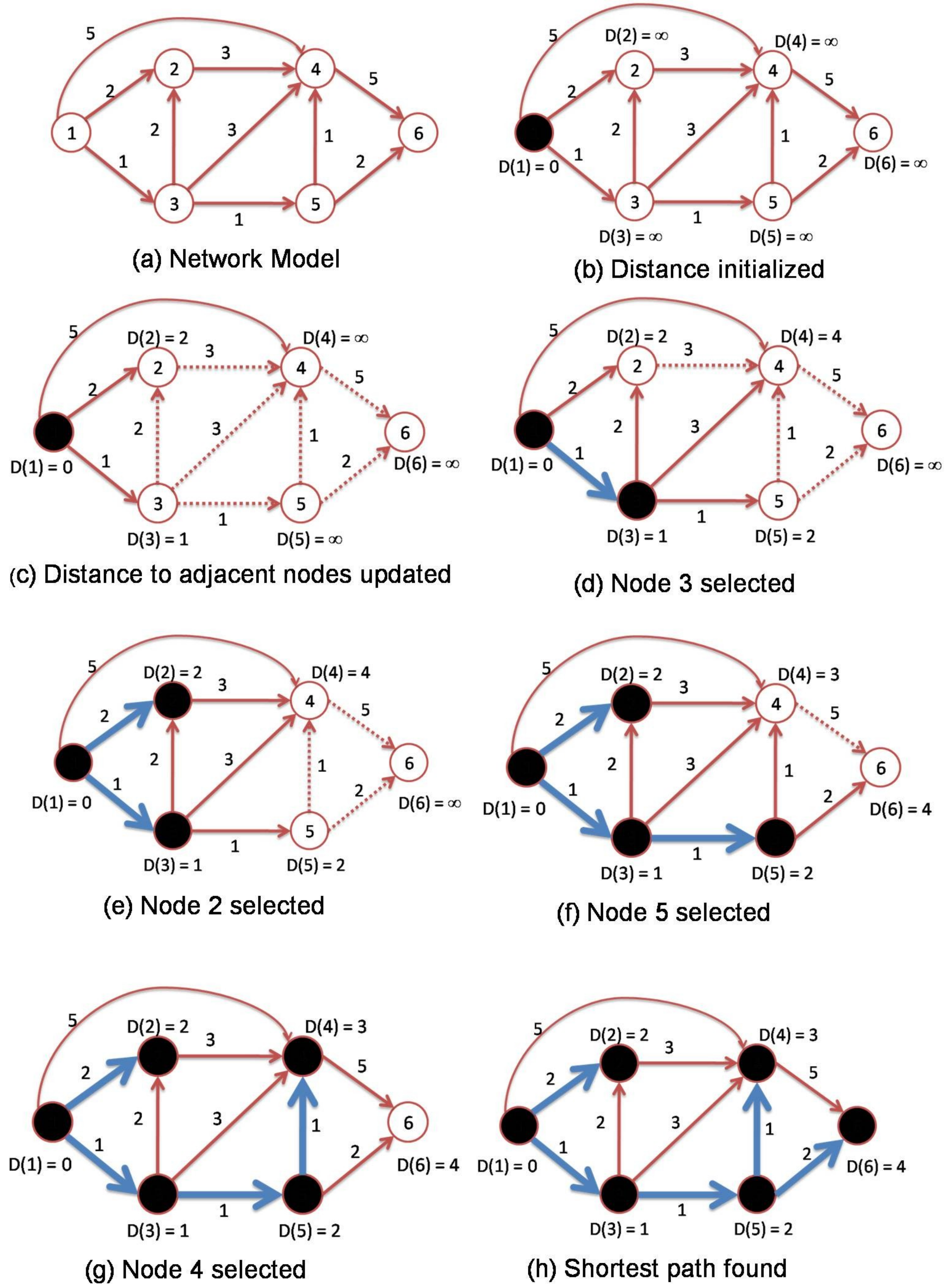
Dijkstra

Dijkstra's algorithm is the BFS equivalent for non-negative weight graphs

Dijkstra

```
Dijkstra(G,w,s)
initialize G from s
Q = G.V, S = empty
while Q not empty
   u = Extract-min(Q)
   S = S U {u}
   for each v int G.Adj[u]
      relax(u,v,w)
```

S optional


(a) Network Model
D(1) = 0
D(2) = ∞
D(3) = ∞
D(4) = ∞
D(5) = ∞
D(6) = ∞
(b) Distance initialized
D(2) = 2
D(3) = 1
(c) Distance to adjacent nodes updated
D(4) = 4
D(5) = 2
(d) Node 3 selected
(e) Node 2 selected
D(4) = 3
D(6) = 4
(f) Node 5 selected
(g) Node 4 selected
(h) Shortest path found

Dijkstra

Runtime?

Dijkstra

Runtime:

Extract-min() run |V| times

Relax runs Decrease-key() |E| times

Both take O(lg n) time

So O((|V| + |E|) lg |V|) time

(can get to O(|V|lg|V| + E) using Fibonacci heaps)

Dijkstra

Runtime note:

If G is almost fully connected,

$|E| \approx |V|^2$

Use a simple array to store v.d

Extract-min() = $O(|V|)$

Decrease-key() = $O(1)$

total: $O(|V|^2 + E)$

Dijkstra

Correctness: (p.660)
Sufficient to prove when u added to S, u.d = δ(s,u)

Base: s added to S first, s.d=0=δ(s,s)

Termination: Loop ends after Q is empty, so V=S and we done

Dijkstra

Step: Assume v in S has $v.d = \delta(s,v)$
Let y be the first vertex outside S
on path of $\delta(s,u)$

We know by relaxation property 4,
that $\delta(s,y)=y.d$ (optimal sub-structure)

$y.d = \delta(s,y) \leq \delta(s,u) \leq u.d$, as $w(p)\geq 0$

Dijkstra

Step: Assume v in S has $v.d = \delta(s,v)$

But as u was picked before y,

$u.d \leq y.d$, combined with $y.d \leq u.d$

$y.d = u.d$

Thus $y.d = \delta(s,y) = \delta(s,u) = u.d$