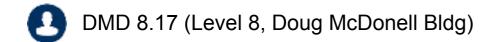


# COMP90038 Algorithms and Complexity

Lecture 8: Graph Traversal (with thanks to Harald Søndergaard)

#### **Toby Murray**





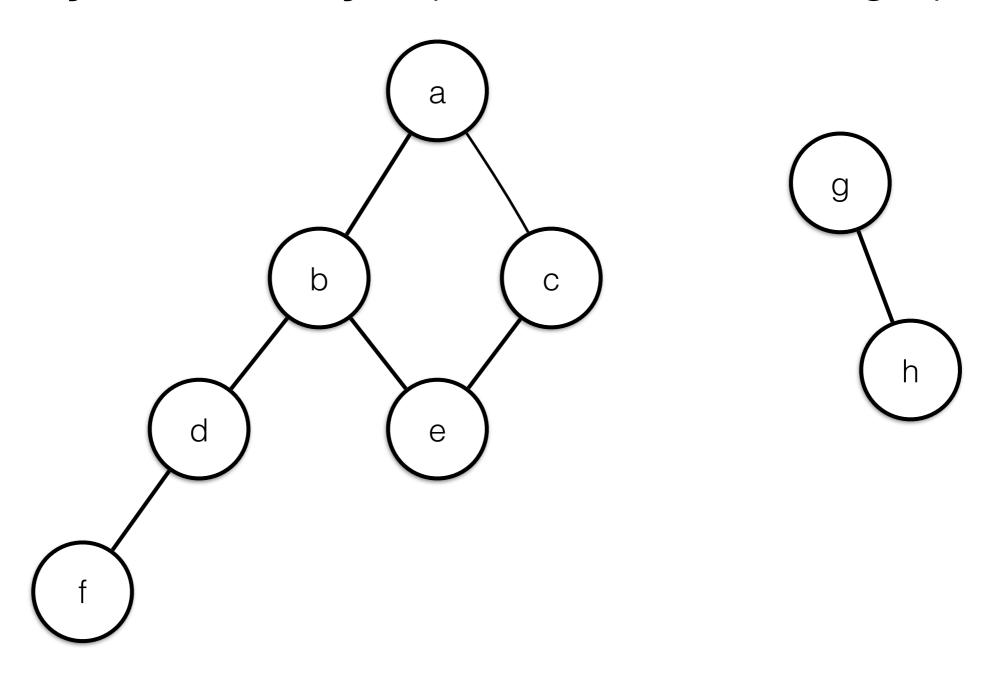


@tobycmurray

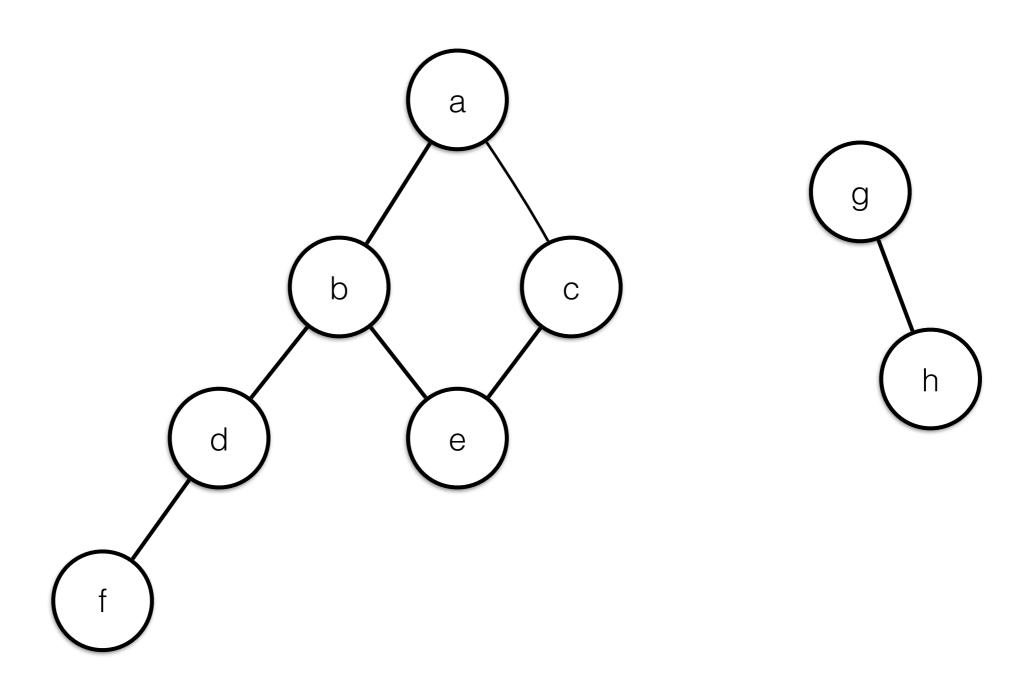
# Breadth-First and Depth-First Traversal



• Used to systematically explore all nodes of a graph

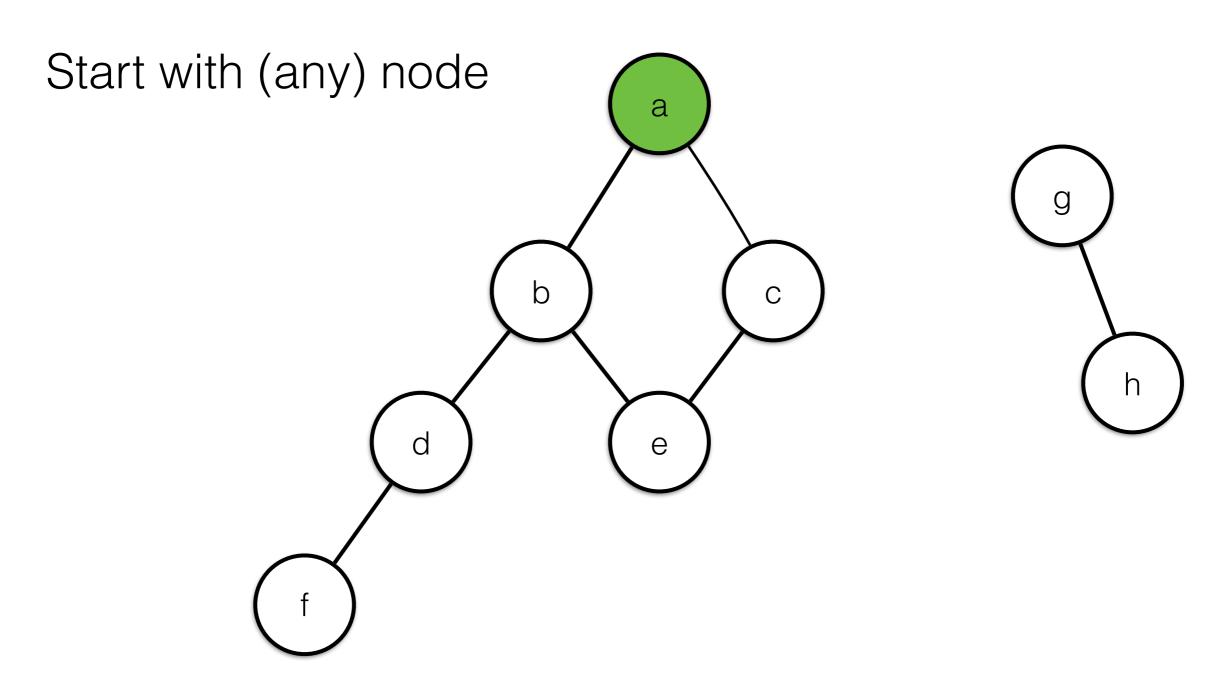






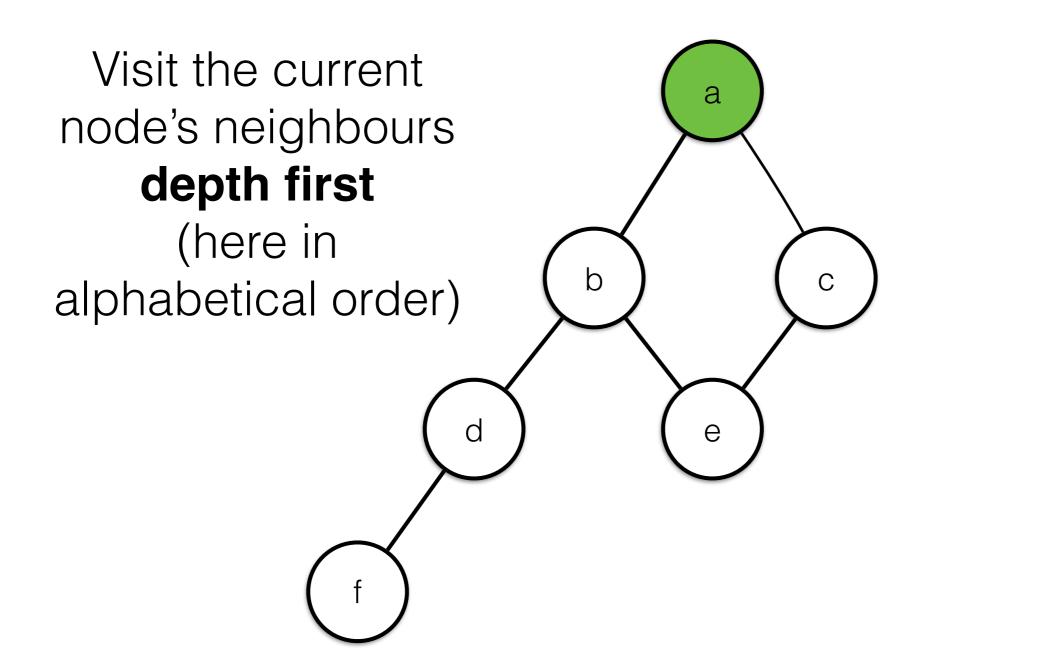


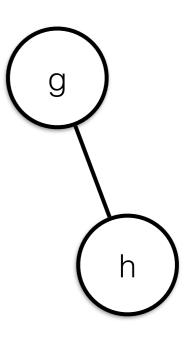
Nodes visited in this order: a





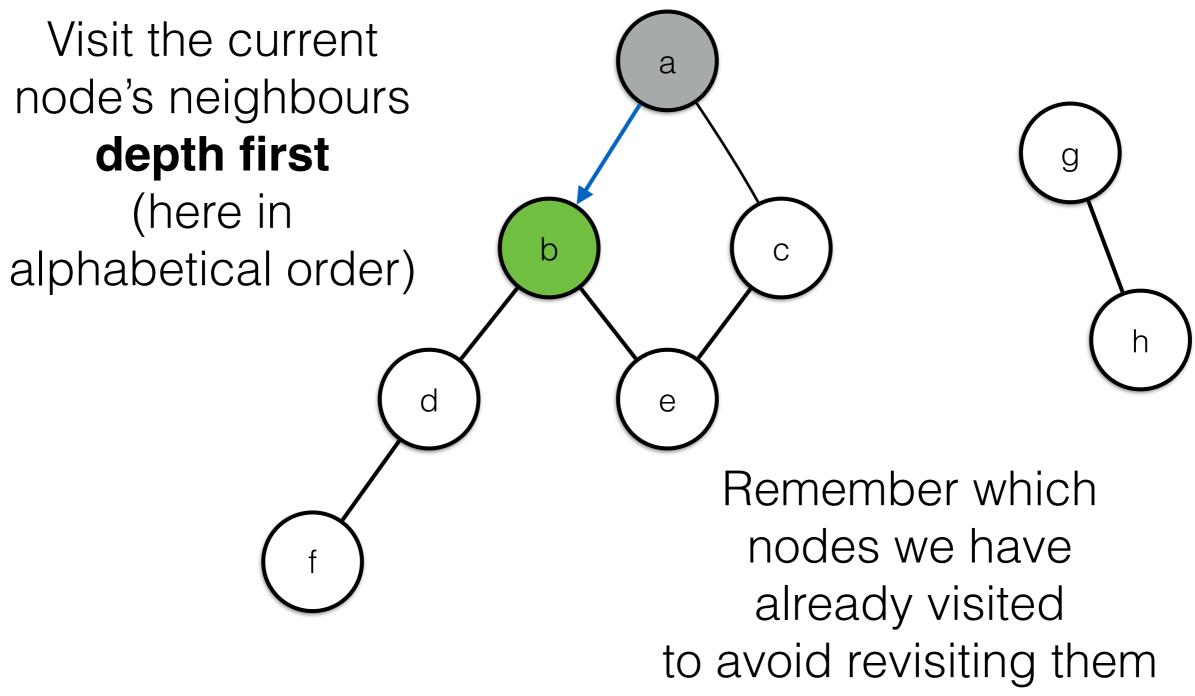
Nodes visited in this order: a





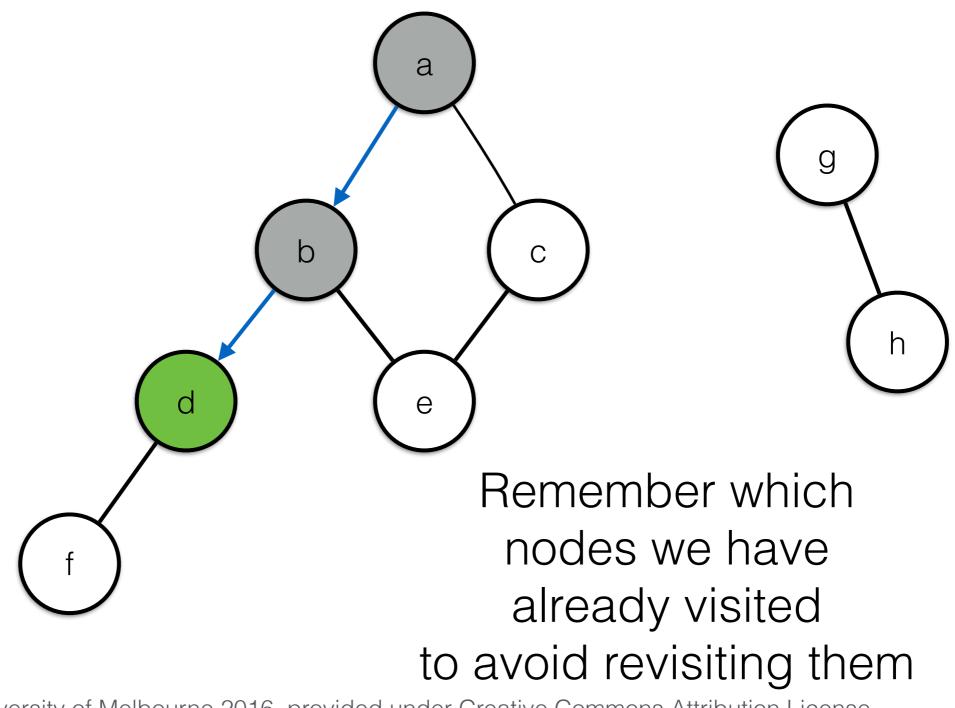


Nodes visited in this order: a b



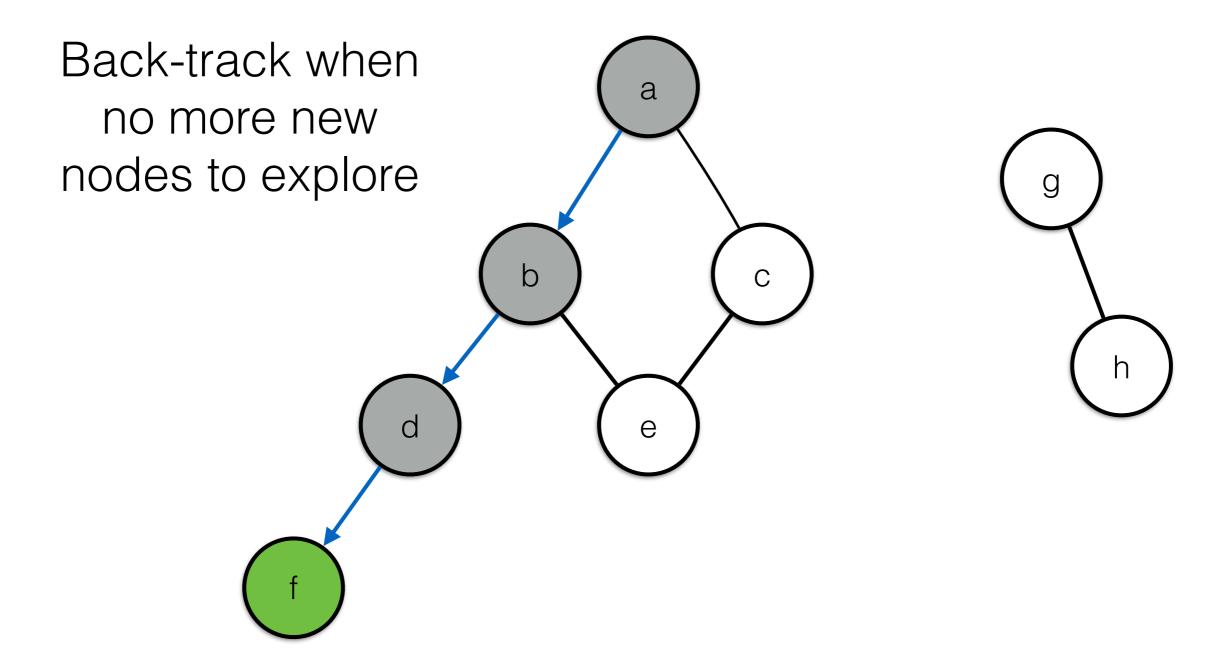


Nodes visited in this order: a b d



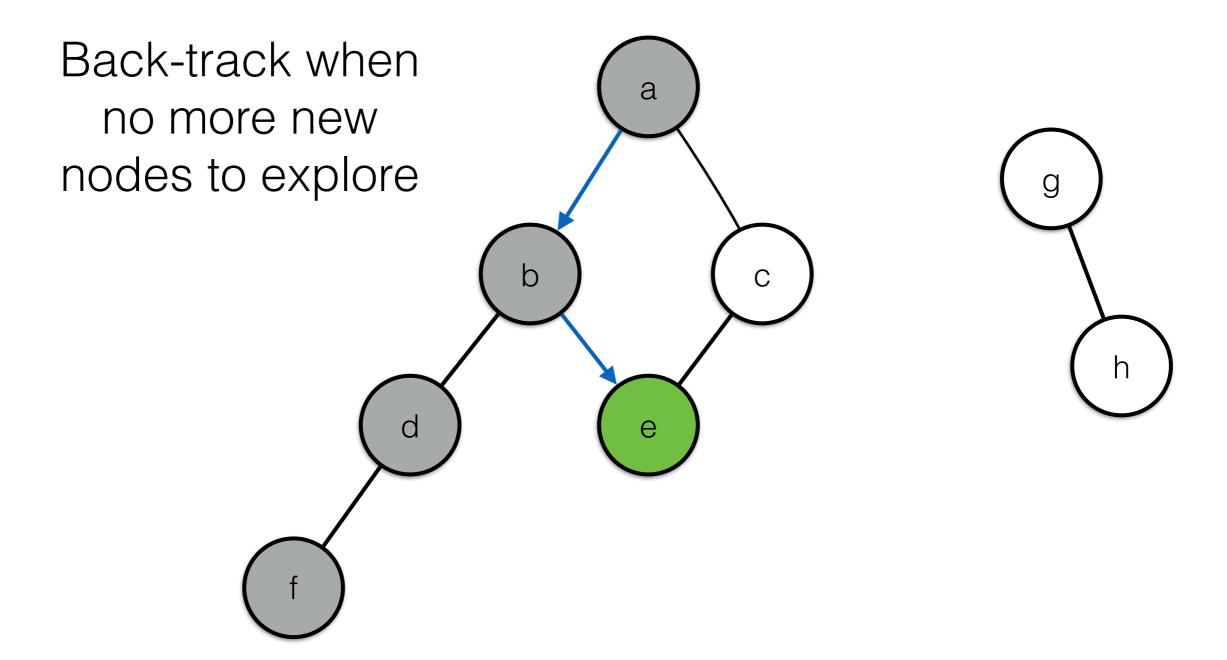


Nodes visited in this order: a b d f



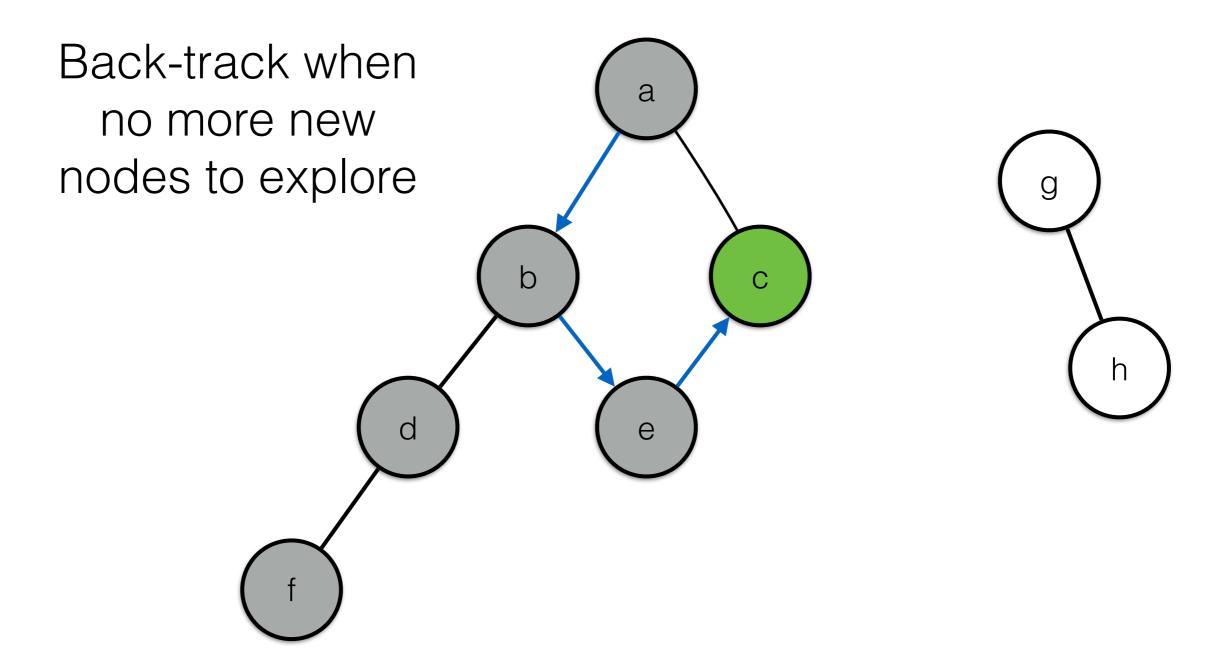


Nodes visited in this order: a b d f e



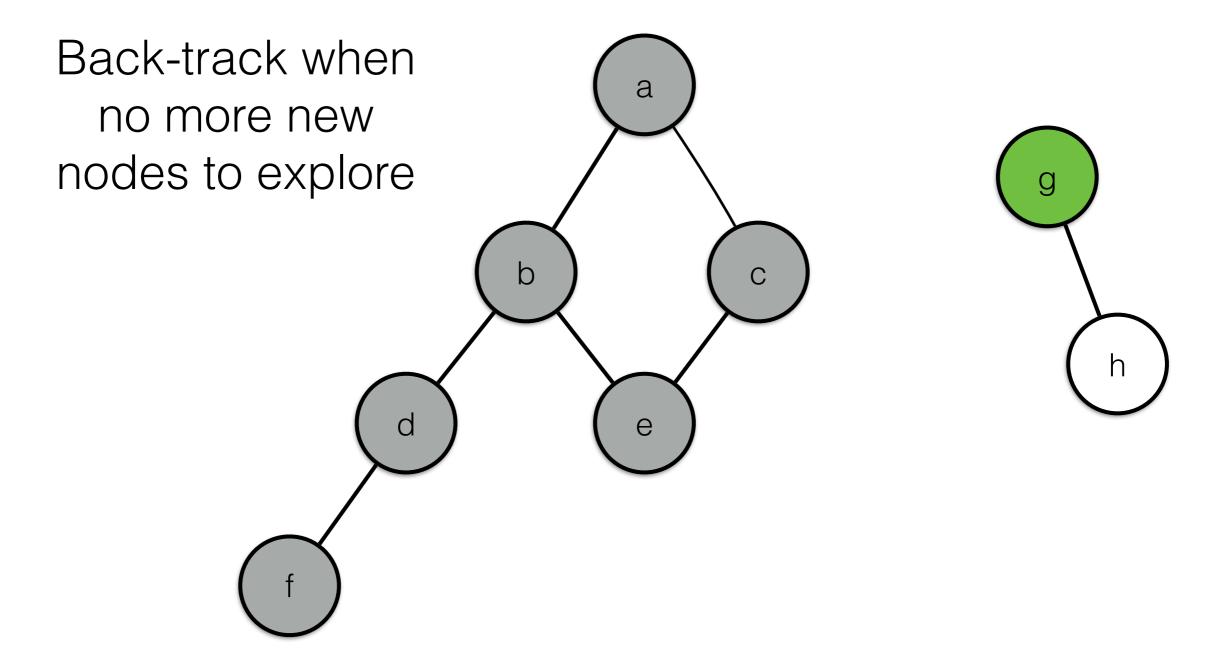


Nodes visited in this order: a b d f e c



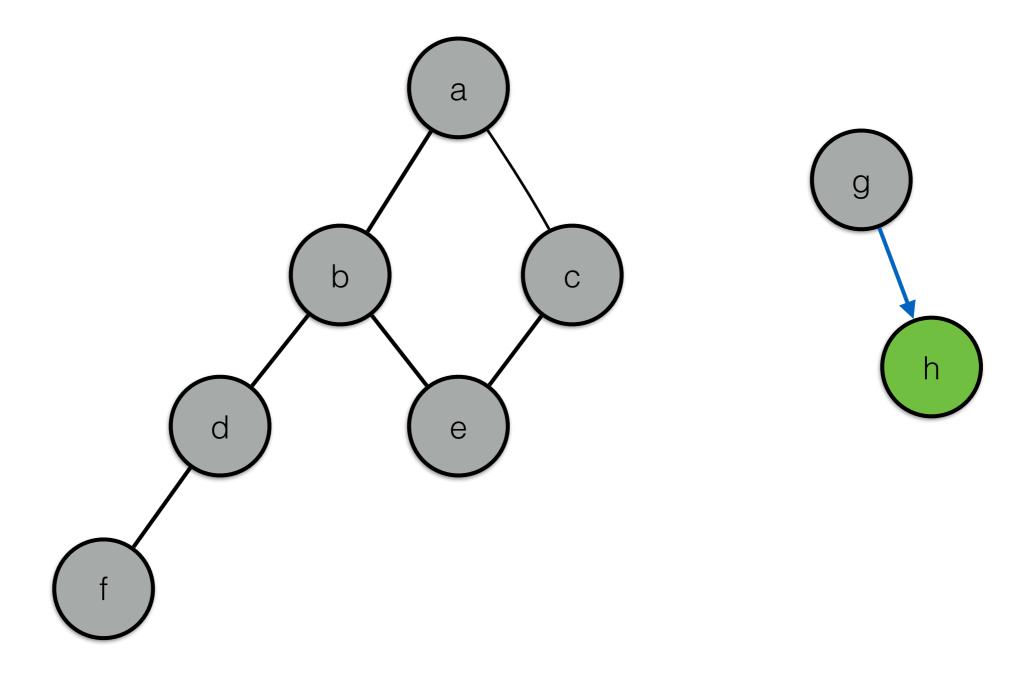


Nodes visited in this order: a b d f e c g





Nodes visited in this order: a b d f e c g h



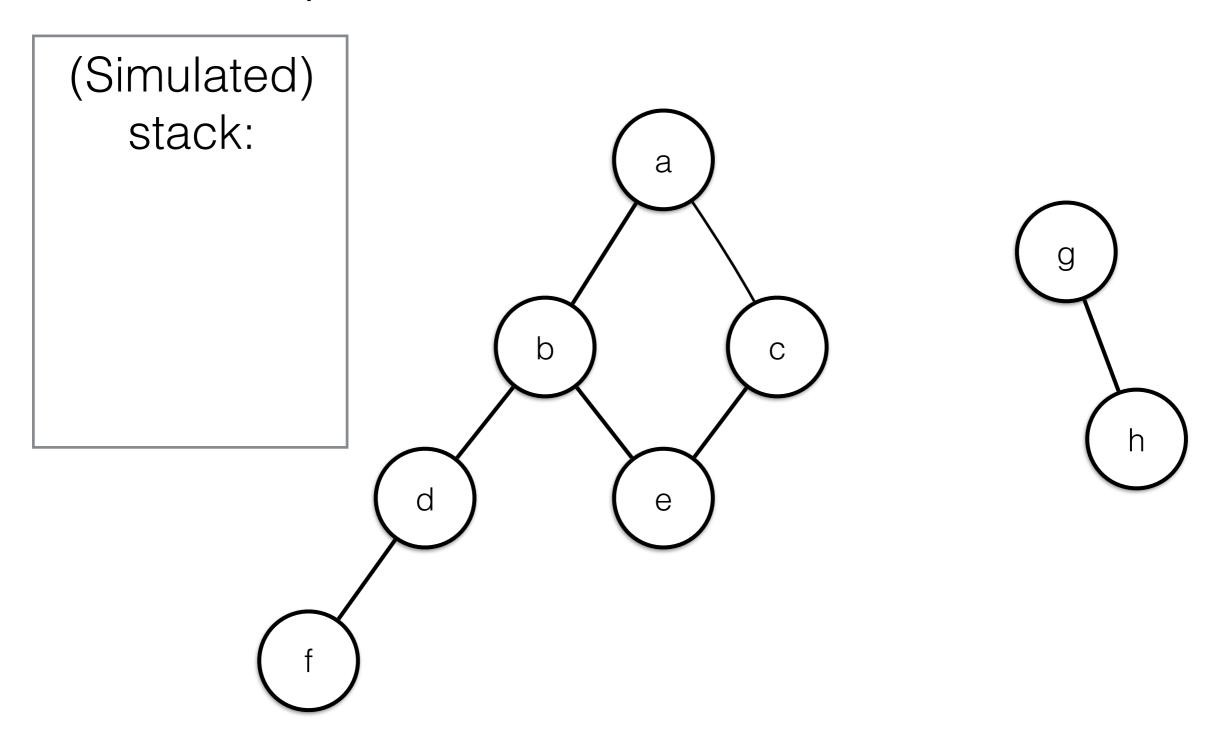


When back-tracking, we go back to the most recently-visited node that still has unvisited neighbours

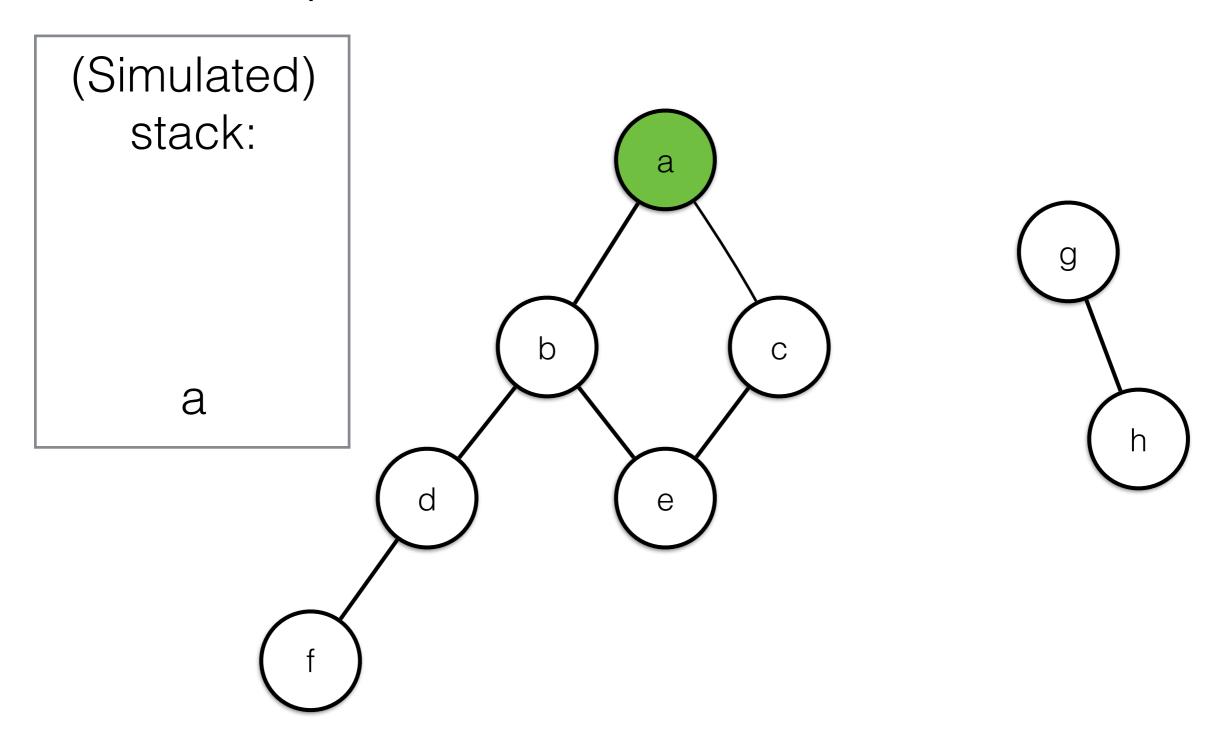
This is simulated by pushing each node onto a stack as it is visited.

Back-tracking then corresponds to popping the stack.

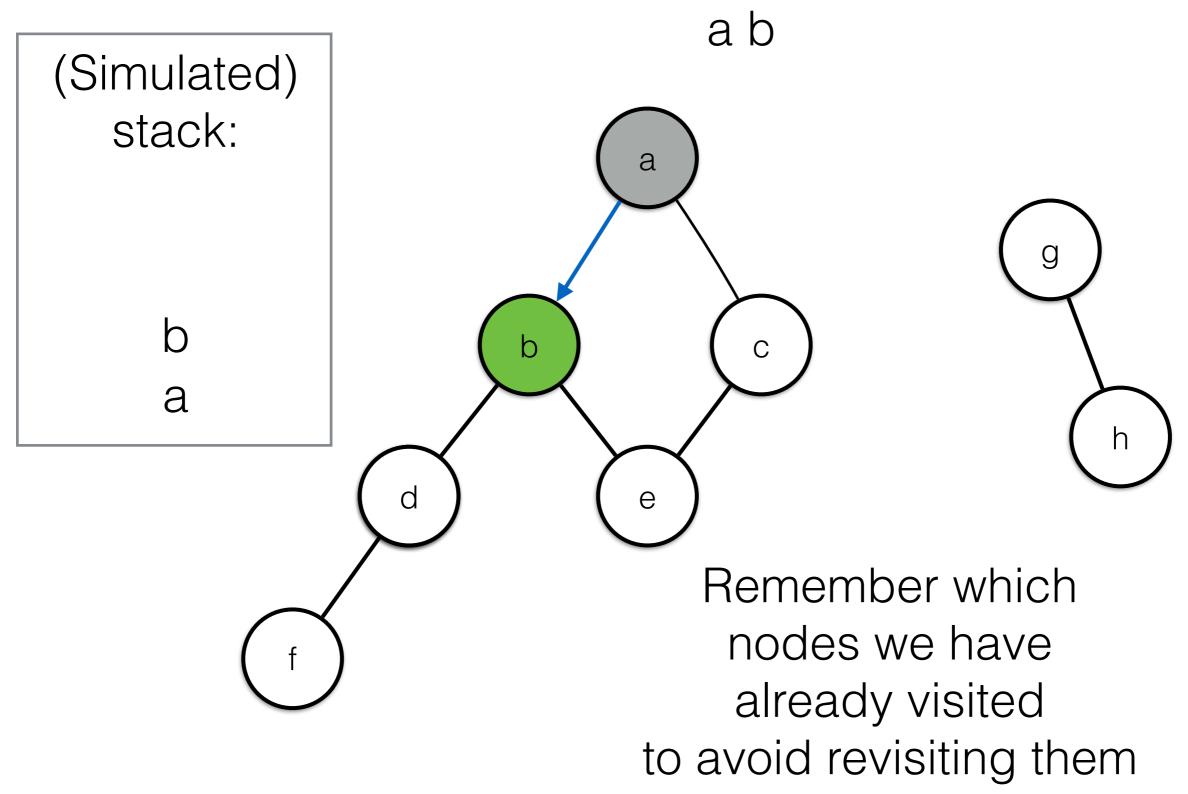




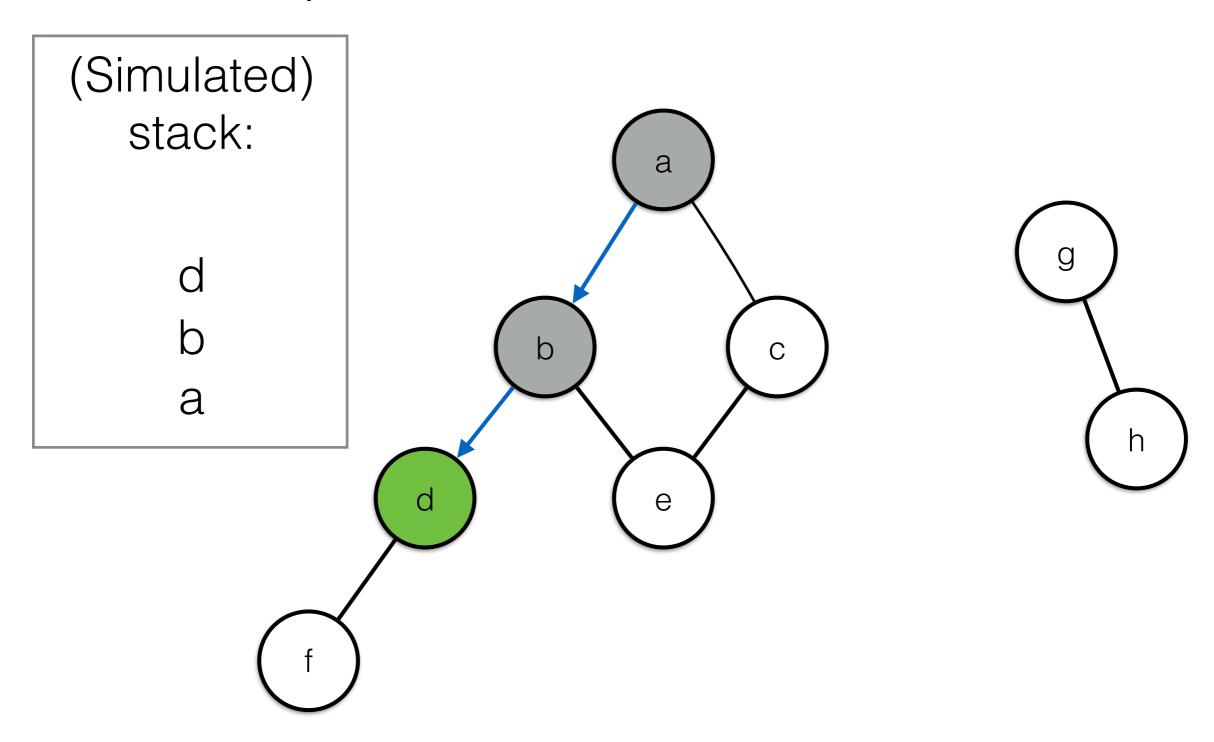




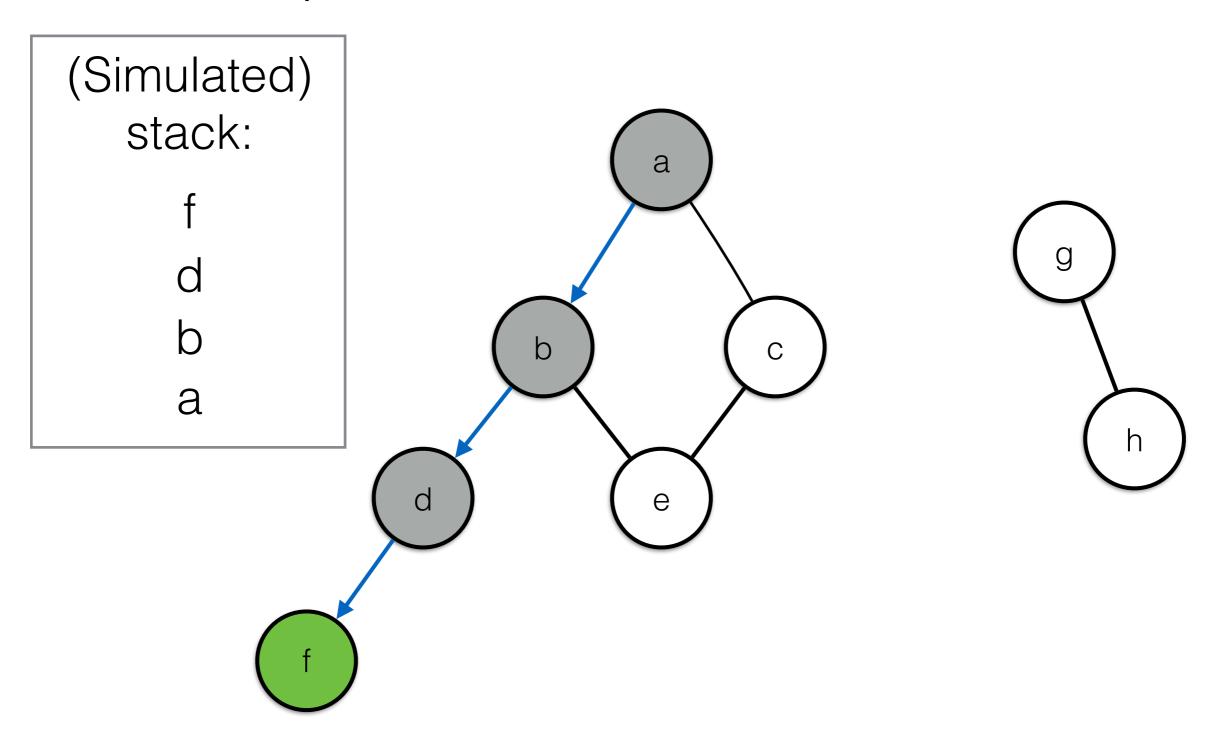




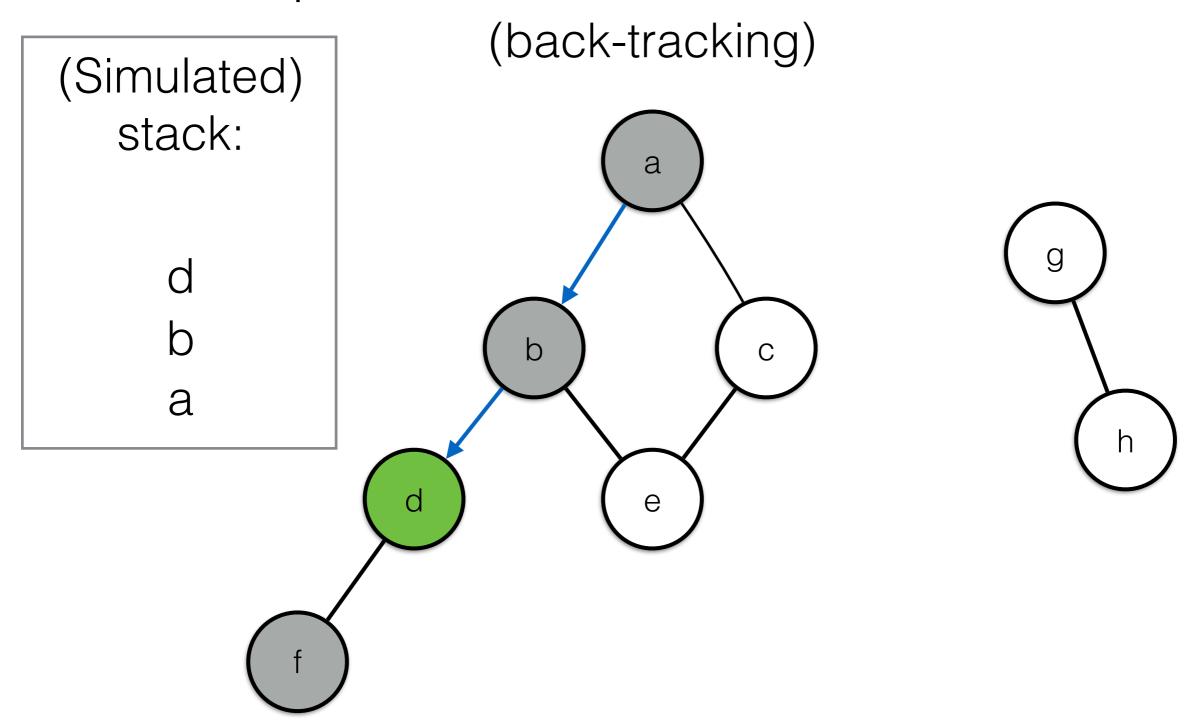




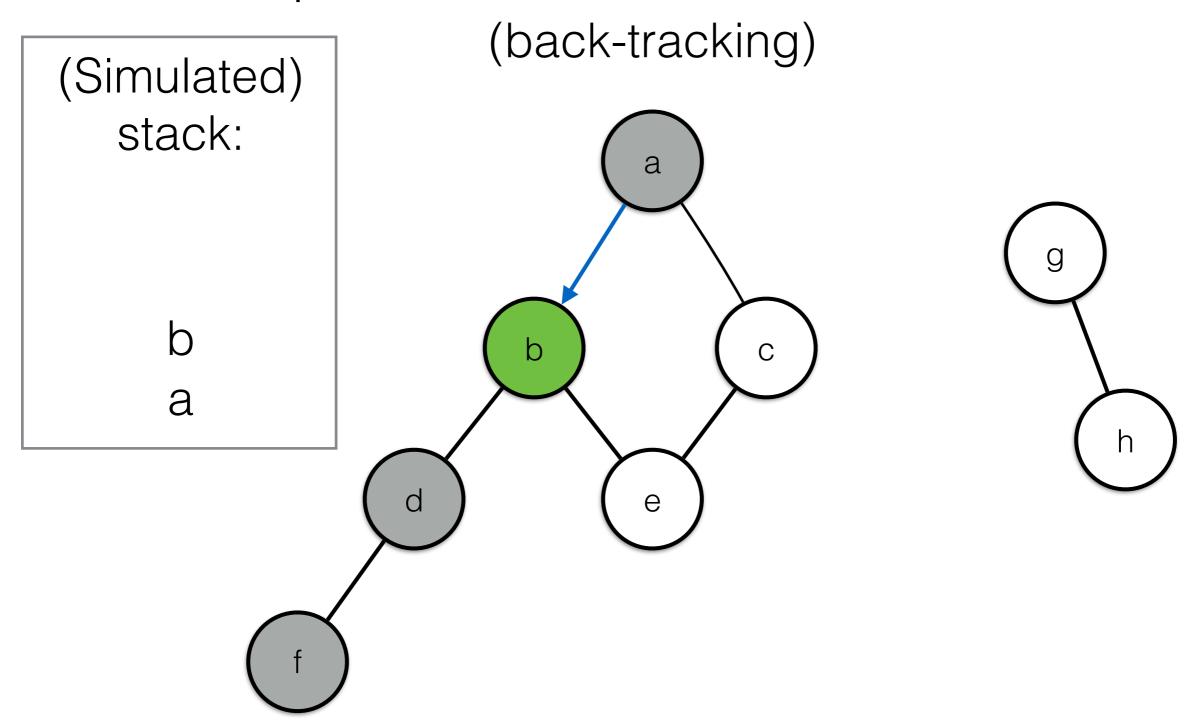




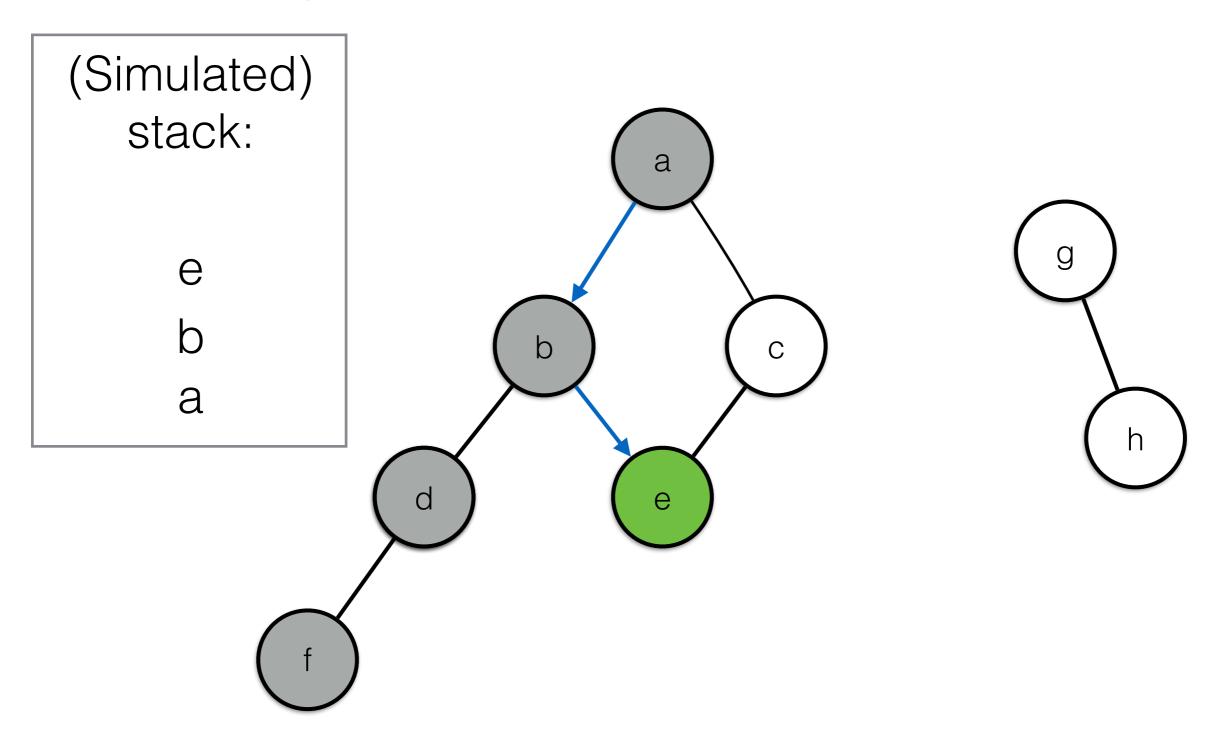




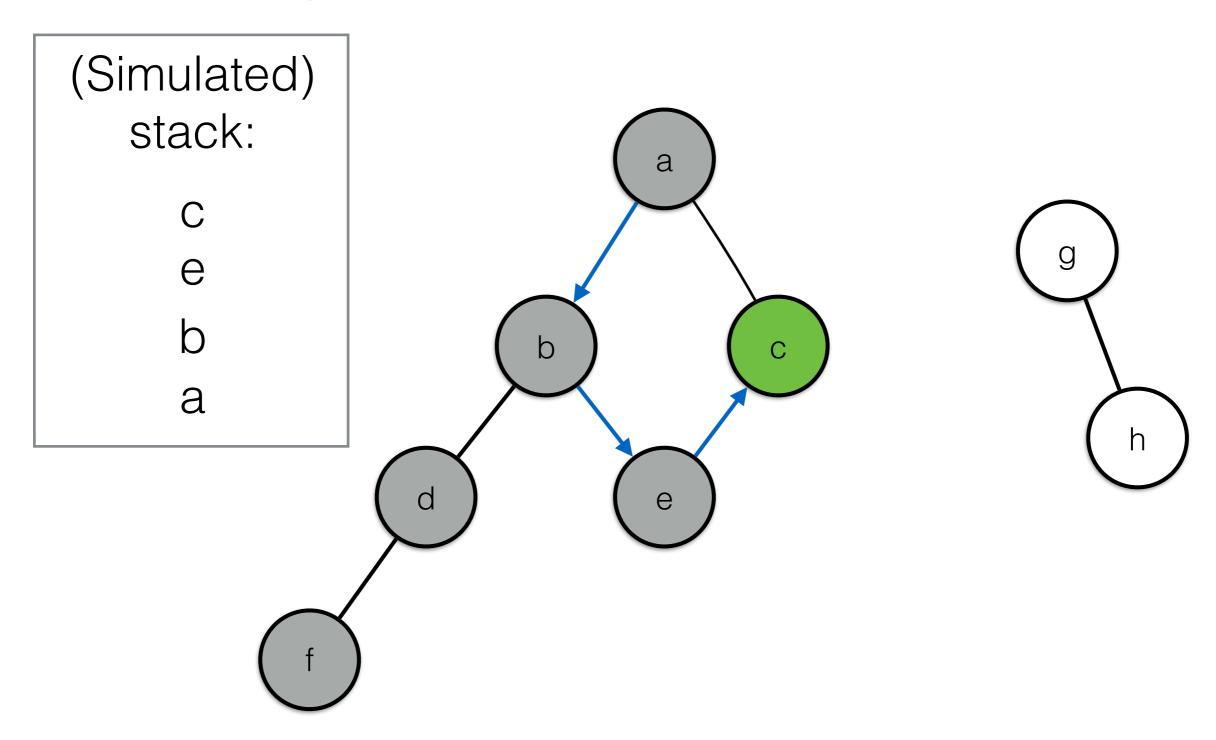








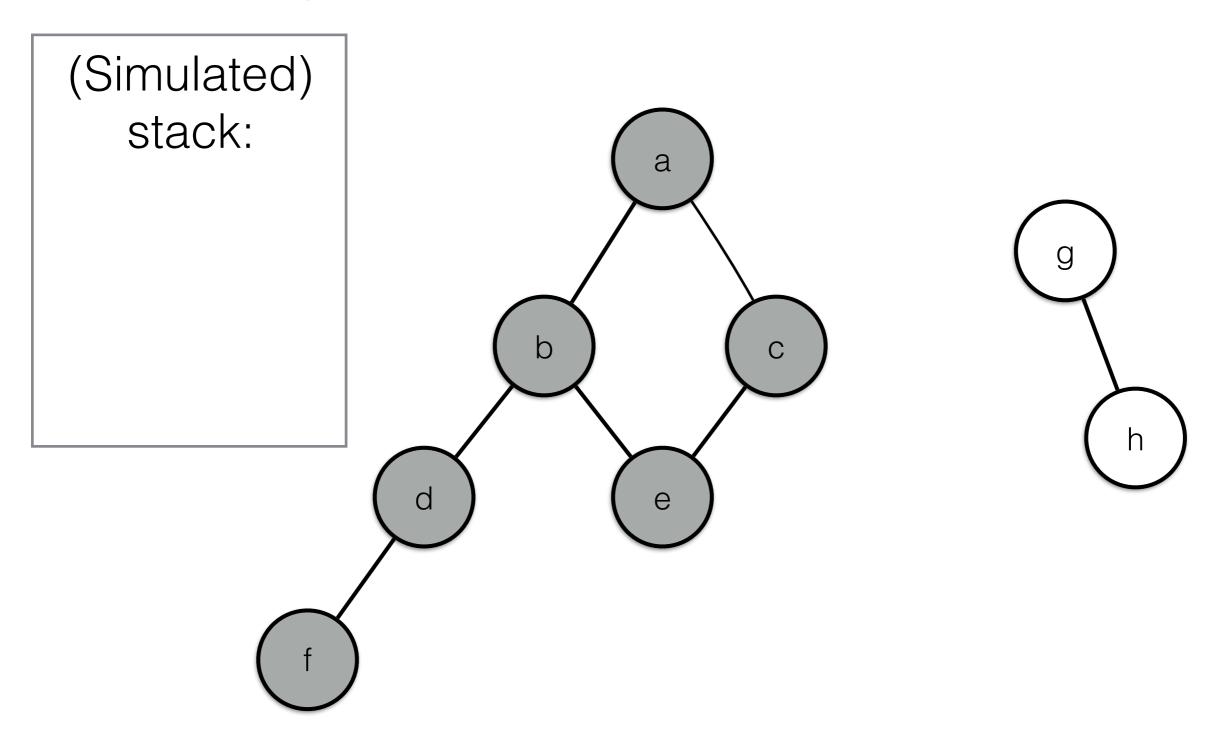




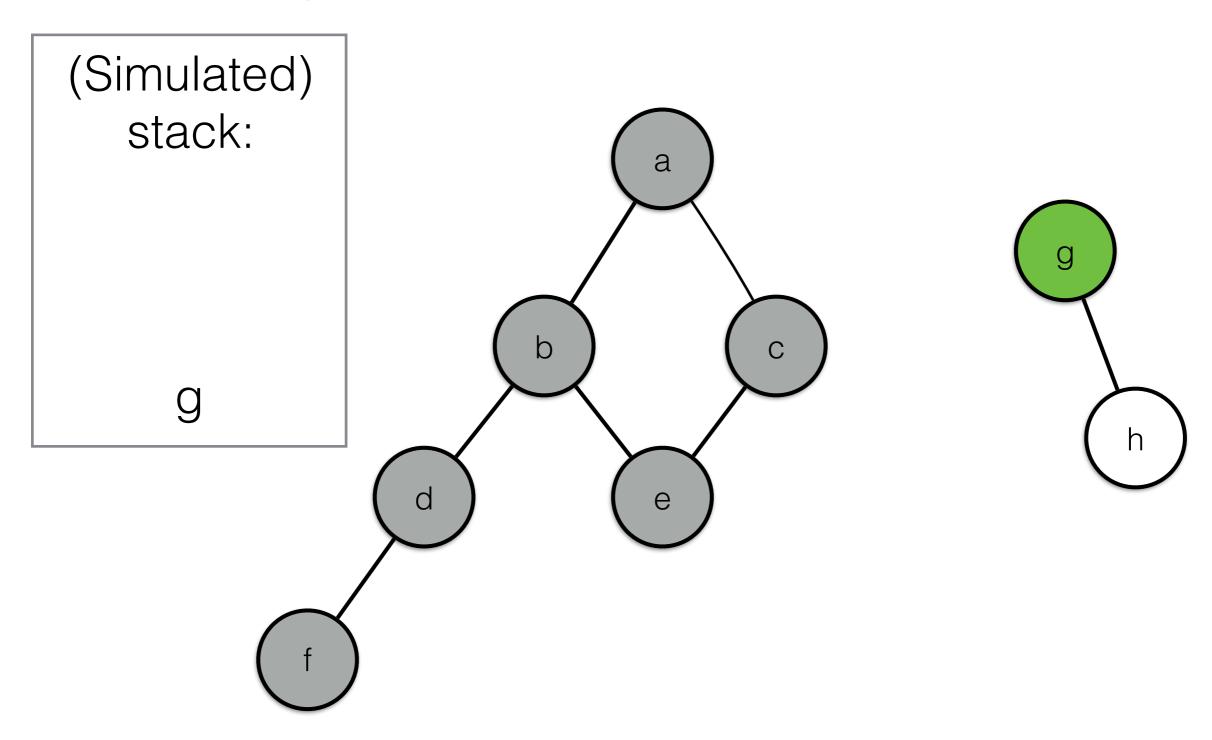


back-tracking ... (Simulated) details omitted ... stack: b

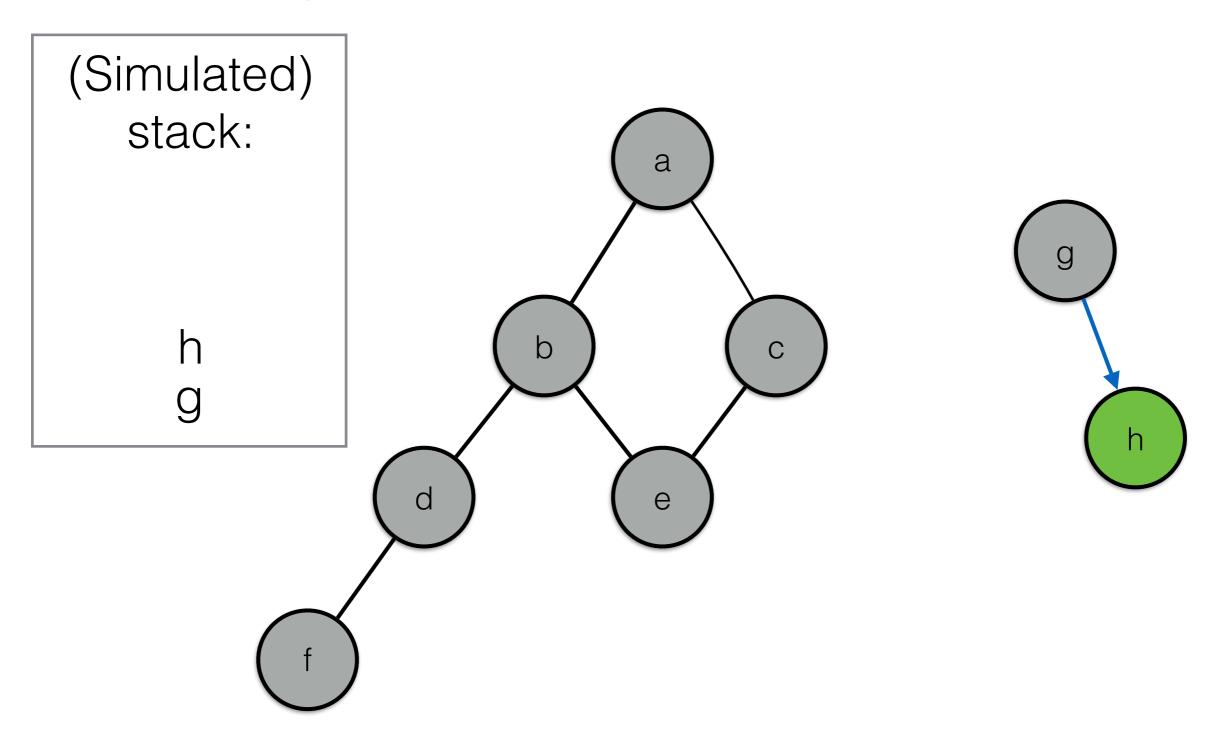












```
function DFS(\langle V, E \rangle)

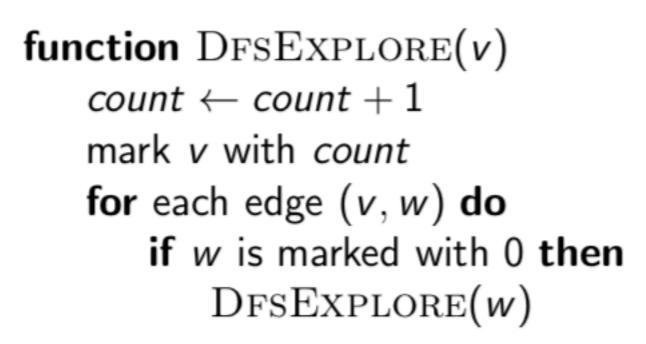
mark each node in V with 0

count \leftarrow 0

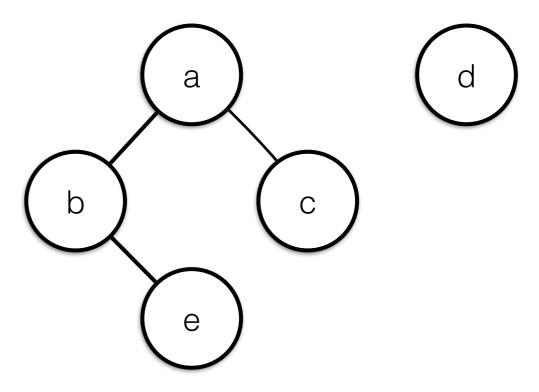
for each v in V do

if v is marked with 0 then

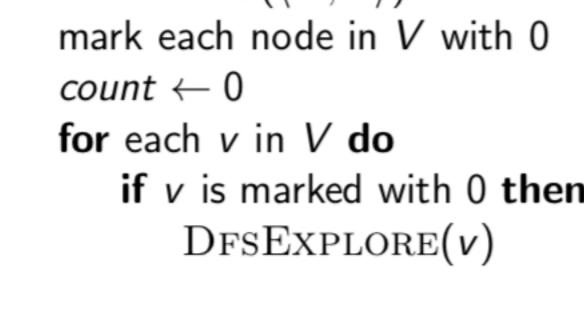
DFSEXPLORE(v)
```





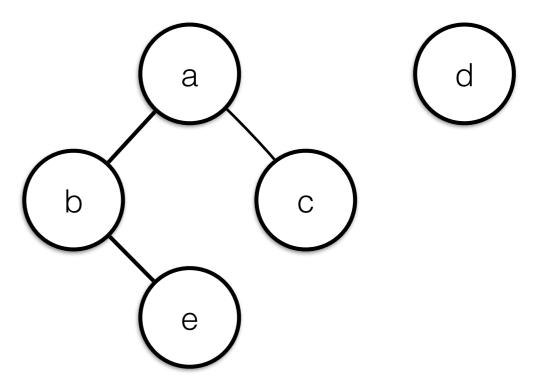


function DFS( $\langle V, E \rangle$ ) mark each node in V with 0 count  $\leftarrow 0$ for each v in V do if v is marked with 0 then



function DfsExplore(v)  $count \leftarrow count + 1$ mark v with count for each edge (v, w) do if w is marked with 0 then DfsExplore(w)





 $DFS(\langle V,E \rangle)$ Call Stack



function DFS( $\langle V, E \rangle$ )

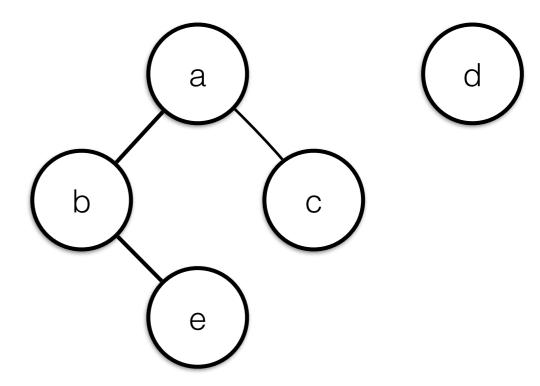
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSEXPLORE(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

 $\mathsf{DFS}(\langle \mathsf{V}, \mathsf{E} \rangle)$ 



function DFS( $\langle V, E \rangle$ )

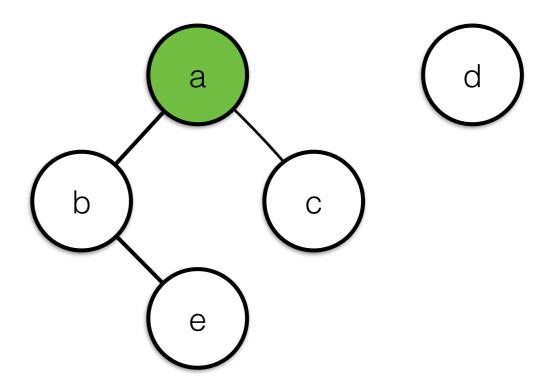
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(a)
DFS(\langle V,E\rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

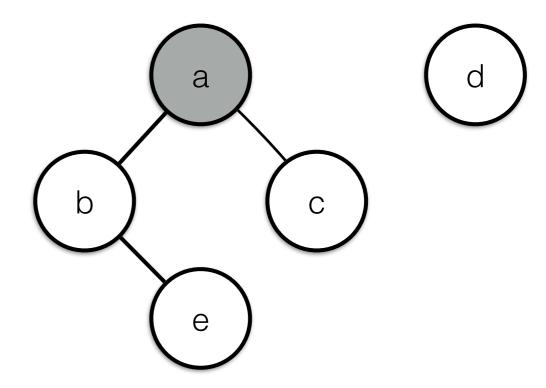
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(a)
DFS(\langle V,E\rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

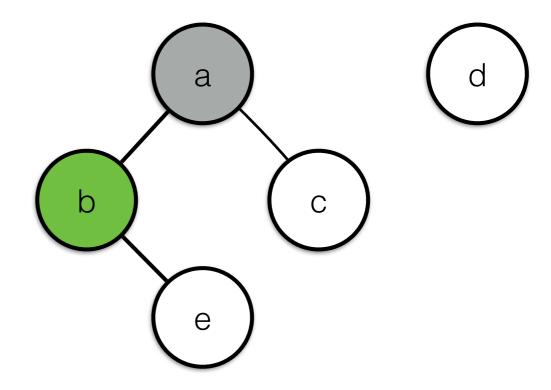
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DfsExplore(v)

count ← count + 1
mark v with count
for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(b)
DFSEXPLORE(a)
DFS(\langle V,E \rangle)



function DFS( $\langle V, E \rangle$ )

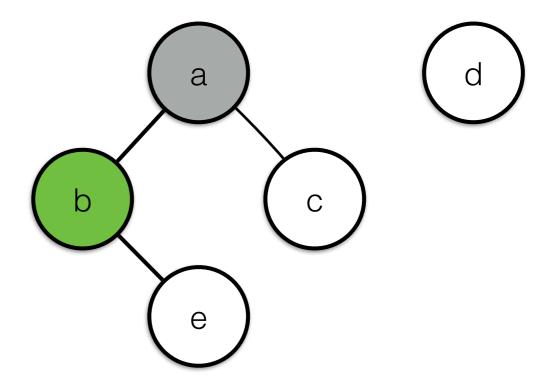
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DfsExplore(v)

 $count \leftarrow count + 1$ mark v with countfor each edge (v, w) do
if w is marked with 0 then DFSEXPLORE(w)

count:

DFSEXPLORE(b)

DfsExplore(a)

 $DFS(\langle V,E \rangle)$ 



function DFS( $\langle V, E \rangle$ )

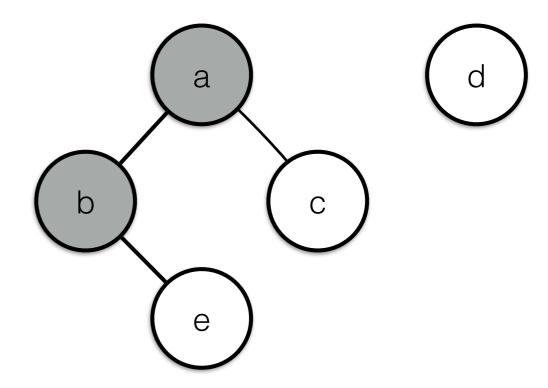
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DfsExplore(v)

 $count \leftarrow count + 1$ mark v with countfor each edge (v, w) do
if w is marked with 0 then DFSEXPLORE(w)

count:

DFSEXPLORE(b)
DFSEXPLORE(a)

 $DFS(\langle V,E \rangle)$ 



function DFS( $\langle V, E \rangle$ )

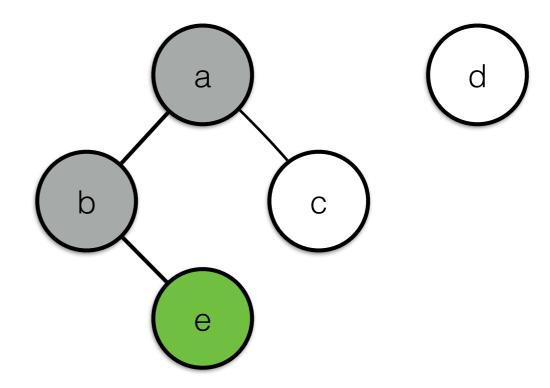
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(e)
DFSEXPLORE(b)
DFSEXPLORE(a)
DFS(\langle V,E \rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

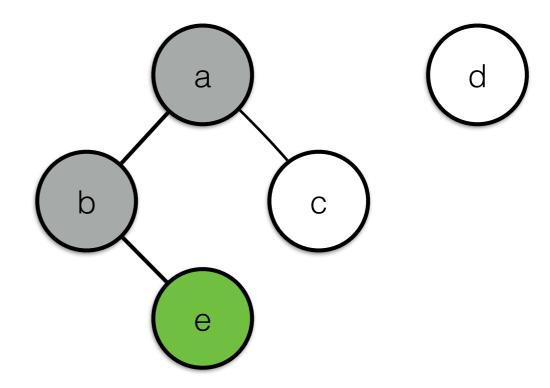
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(e)
DFSEXPLORE(b)
DFSEXPLORE(a)
DFS(\langle V,E \rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

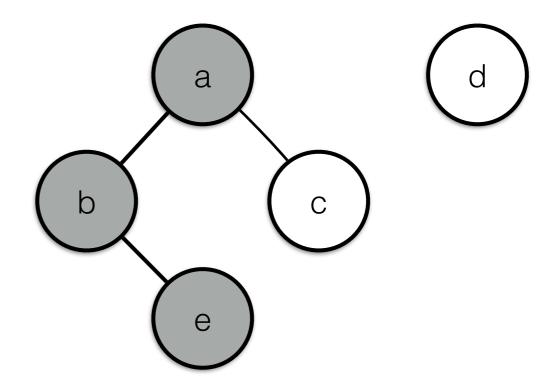
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(e)
DFSEXPLORE(b)
DFSEXPLORE(a)
DFS(\langle V,E \rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

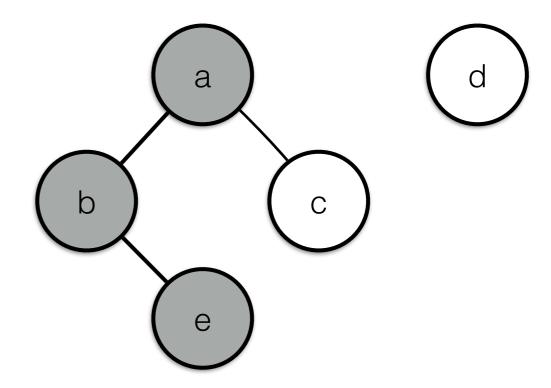
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DfsExplore(v)

count ← count + 1
mark v with count
for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DfsExplore(b)

DfsExplore(a)

 $DFS(\langle V,E \rangle)$ 



function DFS( $\langle V, E \rangle$ )

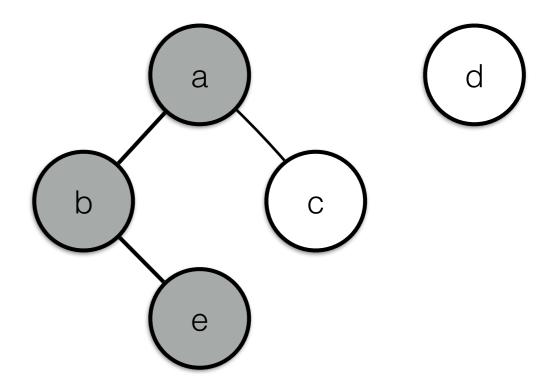
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(a)
DFS(\langle V,E\rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

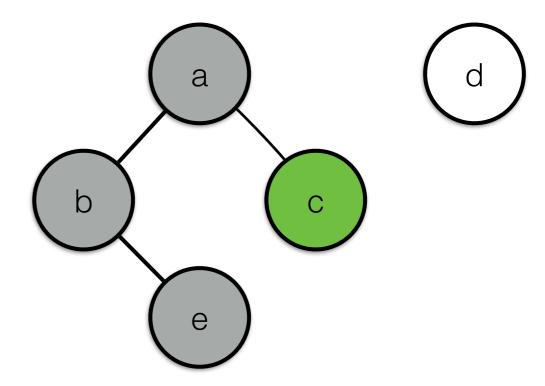
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DfsExplore(v)

 $count \leftarrow count + 1$ mark v with countfor each edge (v, w) do
if w is marked with 0 then DFSEXPLORE(w)

count:

DFSEXPLORE(c)
DFSEXPLORE(a)
DFS((V,E))



function DFS( $\langle V, E \rangle$ )

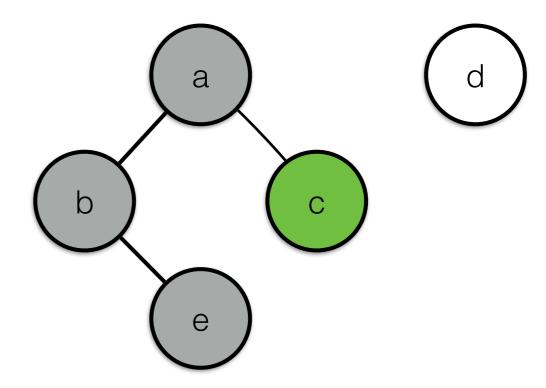
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DfsExplore(v)

 $count \leftarrow count + 1$ mark v with countfor each edge (v, w) do
if w is marked with 0 then DFSEXPLORE(w)

count:

DFSEXPLORE(c)
DFSEXPLORE(a)
DFS((V,E))



function DFS( $\langle V, E \rangle$ )

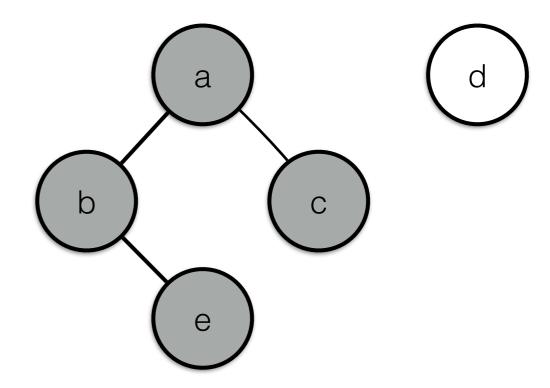
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DfsExplore(v)

count ← count + 1
mark v with count
for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(c)
DFSEXPLORE(a)
DFS((V,E))



function DFS( $\langle V, E \rangle$ )

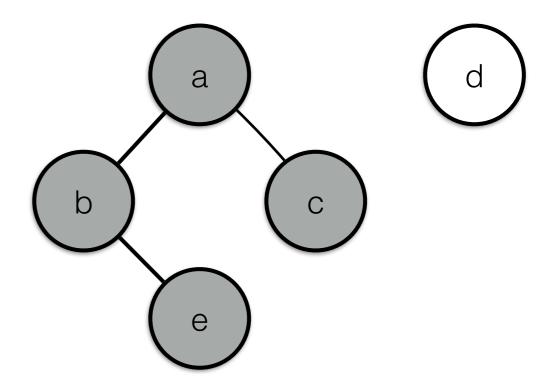
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSEXPLORE(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(a)
DFS(\langle V,E\rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

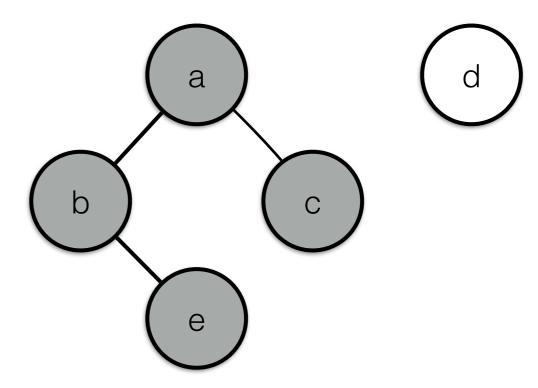
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSEXPLORE(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFS((V,E))
Call Stack



function DFS( $\langle V, E \rangle$ )

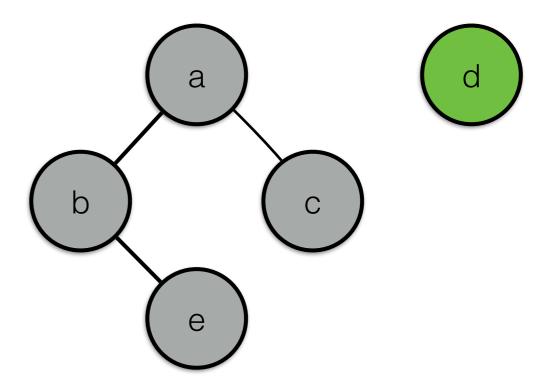
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSExplore(v)



function DfsExplore(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then

DfsExplore(w)

count:

DFSEXPLORE(d)
DFS(\langle V,E\rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

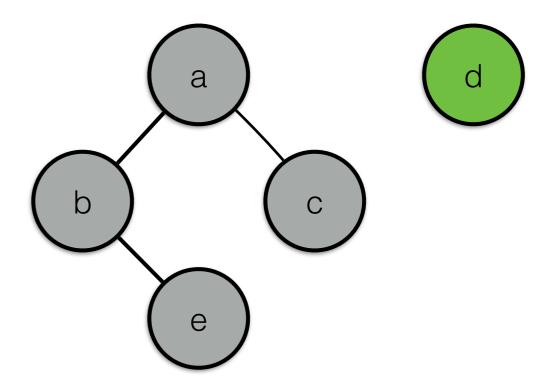
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSEXPLORE(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFSEXPLORE(d)
DFS(\langle V,E\rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

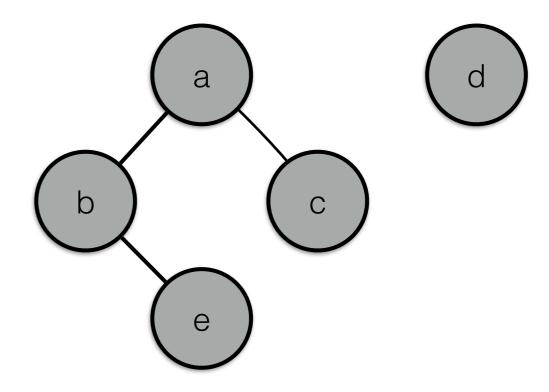
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSEXPLORE(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count: 5 DFSEXPLORE(d)
DFS(\langle V,E\rangle)
Call Stack



function DFS( $\langle V, E \rangle$ )

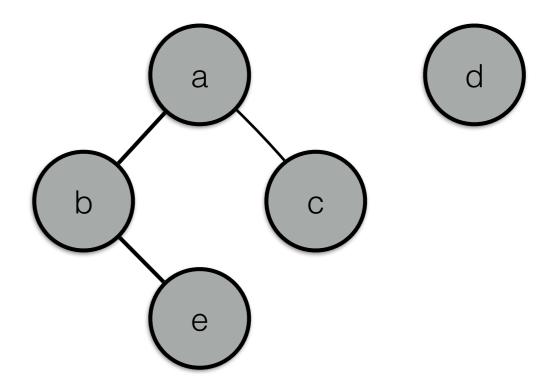
mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

DFSEXPLORE(v)



function DFSEXPLORE(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DFSEXPLORE(w)

count:

DFS((V,E))
Call Stack



```
function DFS(\langle V, E \rangle)

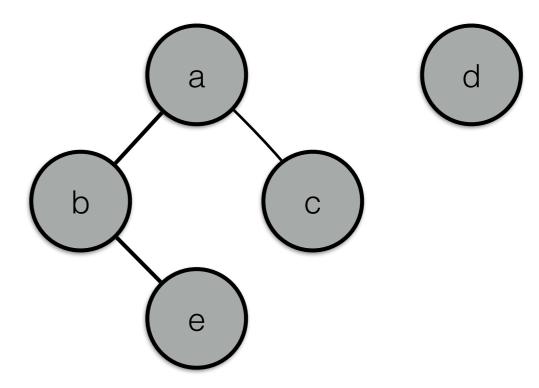
mark each node in V with 0

count \leftarrow 0

for each v in V do

if v is marked with 0 then

DFSEXPLORE(v)
```



```
function DFSEXPLORE(v)
    count ← count + 1
    mark v with count
    for each edge (v, w) do
        if w is marked with 0 then
            DFSEXPLORE(w)
```

count:

5

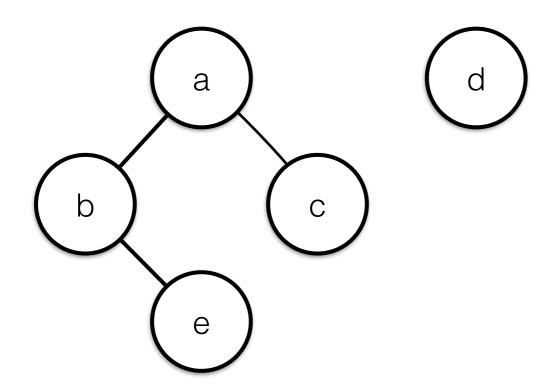
# Depth-First Search: Recursive Algorithm Notes



- Works both for directed and undirected graphs.
- The "marking" of nodes is usually done by maintaining a separate array, mark, indexed by V.
- For example, when we wrote "mark v with count", that would be implemented as "mark[v] := count".
- How to find the nodes adjacent to v depends on the graph representation used.
- Using an adjacency **matrix** adj, we need to consider adj[v,w] for each w in V. Here the complexity of graph traversal is  $\Theta(|V|^2)$ .
- Using adjacency lists, for each v, we traverse the list adj[v].
   In this case, the complexity of traversal is Θ(|V| + |E|).

# Applications of Depth-First Search (DFS)

- Easy to adapt DFS to decide if a graph is connected.
- How?





function DFS( $\langle V, E \rangle$ )

mark each node in V with 0

count  $\leftarrow$  0

for each v in V do

if v is marked with 0 then

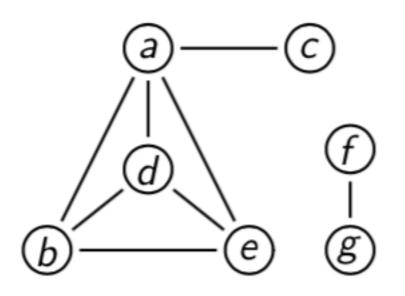
DFSEXPLORE(v)

function DfsExplore(v)
 count ← count + 1
 mark v with count
 for each edge (v, w) do
 if w is marked with 0 then
 DfsExplore(w)

# Depth-First Search: Node Orderings



 We can order nodes either by the order in which they get pushed onto the stack, or by the order in which they are popped from the stack



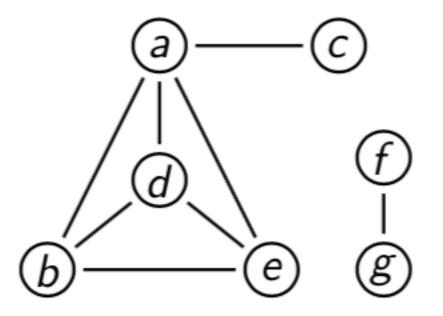
Levitin's compact stack notation

$$e_{4,1}$$
 $d_{3,2}$ 
 $b_{2,3}$   $c_{5,4}$   $g_{7,6}$ 
 $a_{1,5}$   $f_{6,7}$ 

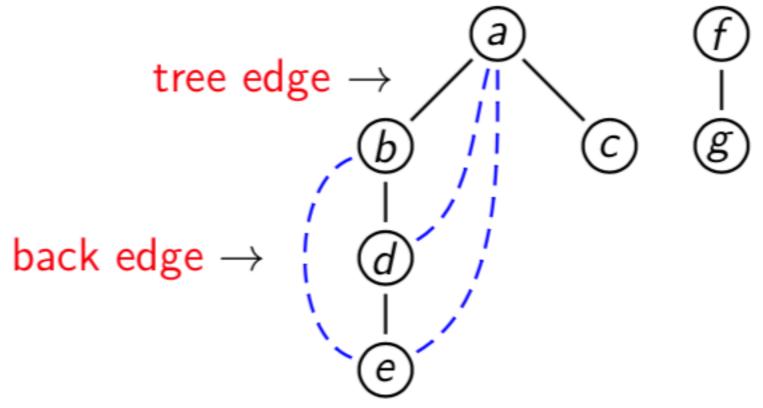
The first subscripts give the order in which nodes are pushed, the second the order in which they are popped off the stack.

# Depth-First Search Forest

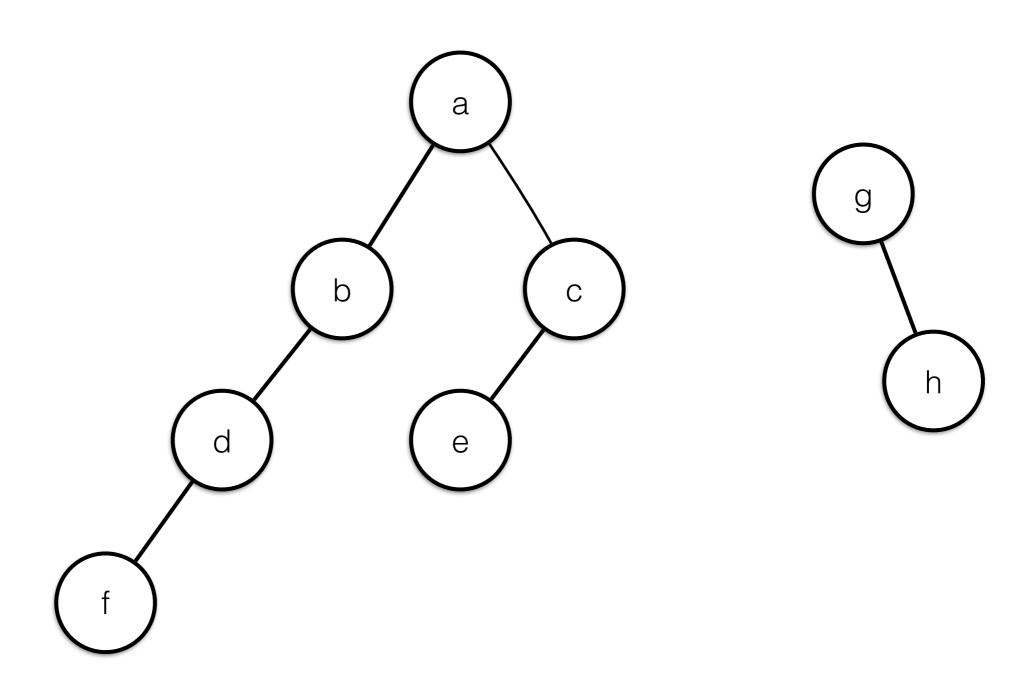




DFS can be depicted by the resulting **DFS Forest** (DFS **Tree** for a connected graph)

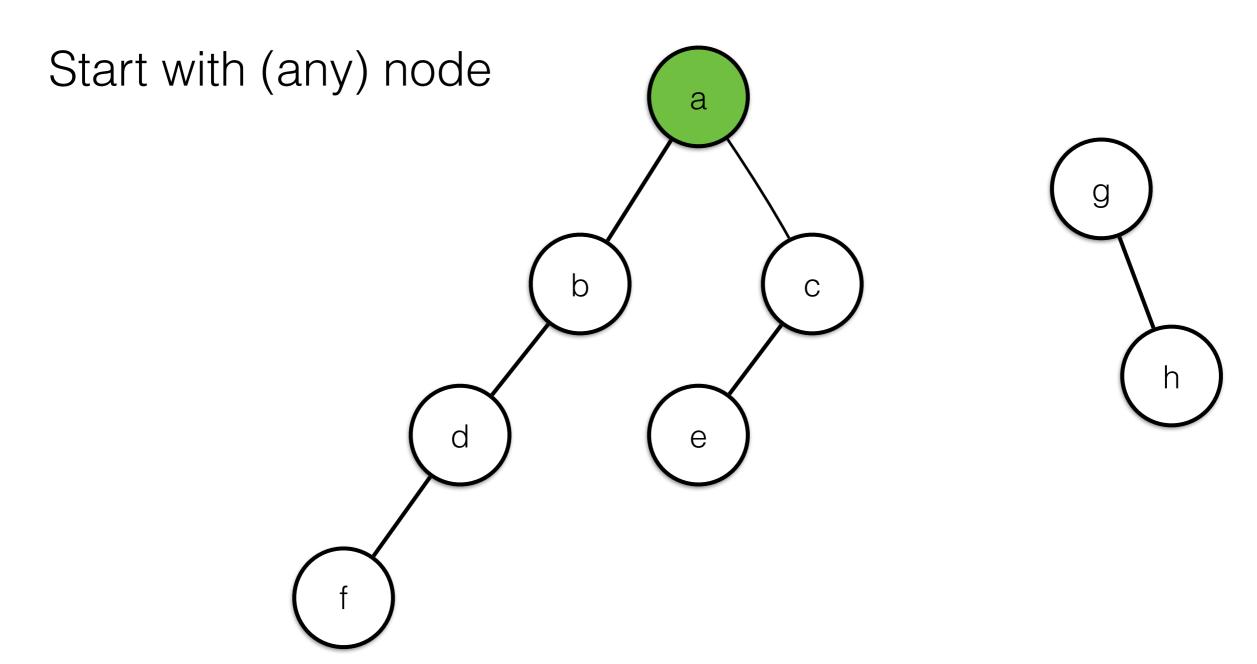






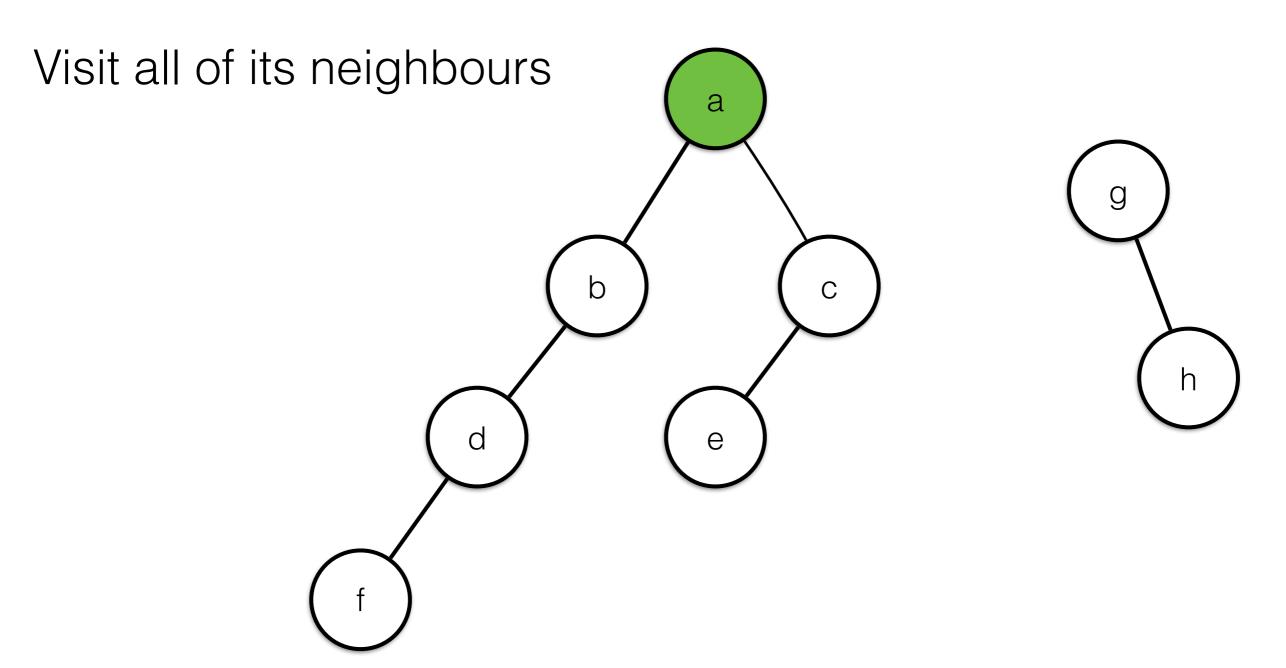


Nodes visited in this order: a



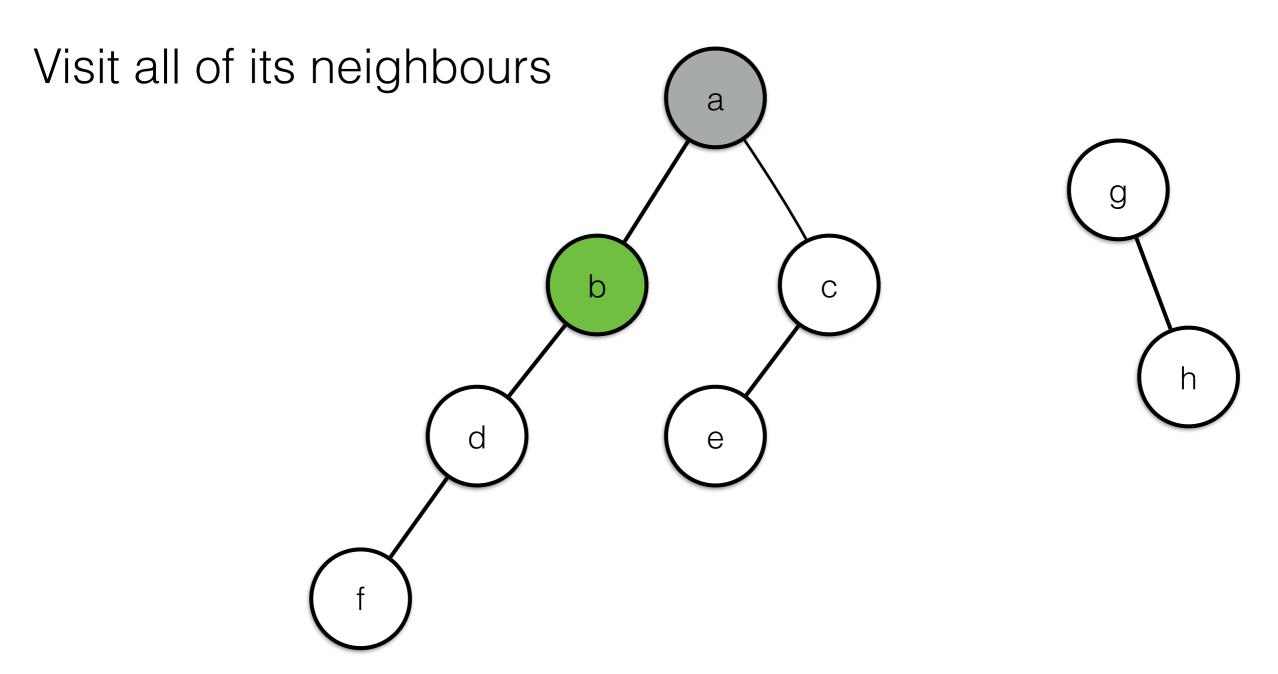


Nodes visited in this order: a



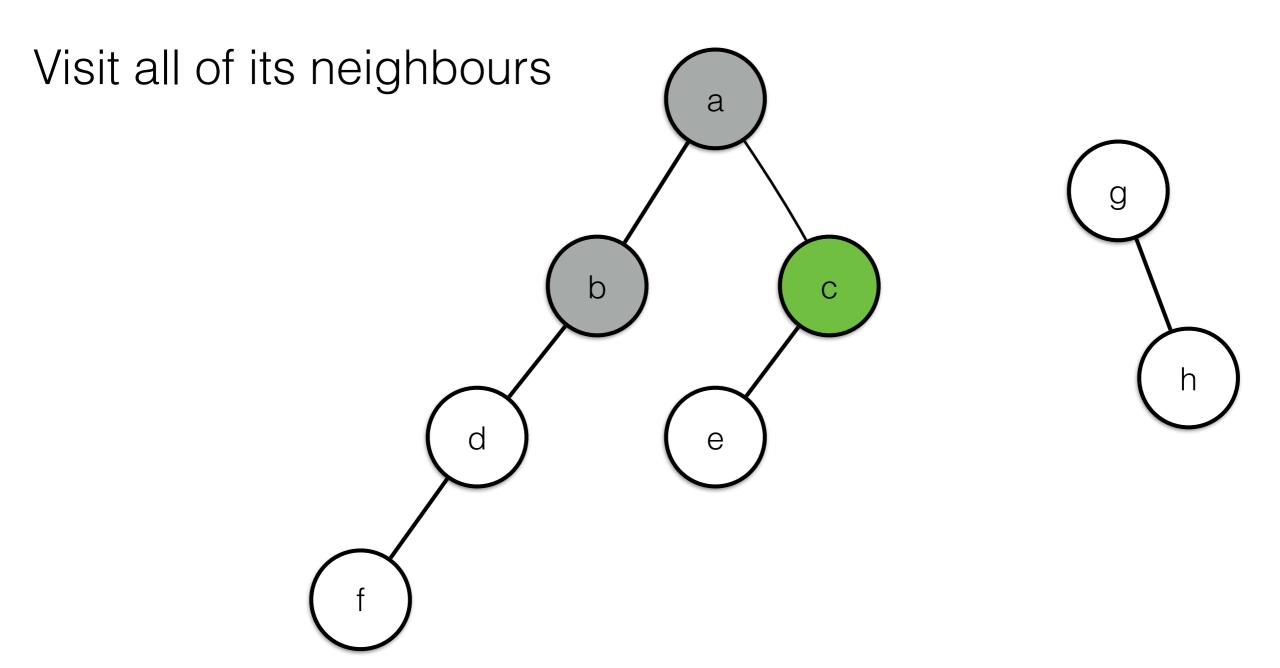


Nodes visited in this order: a b



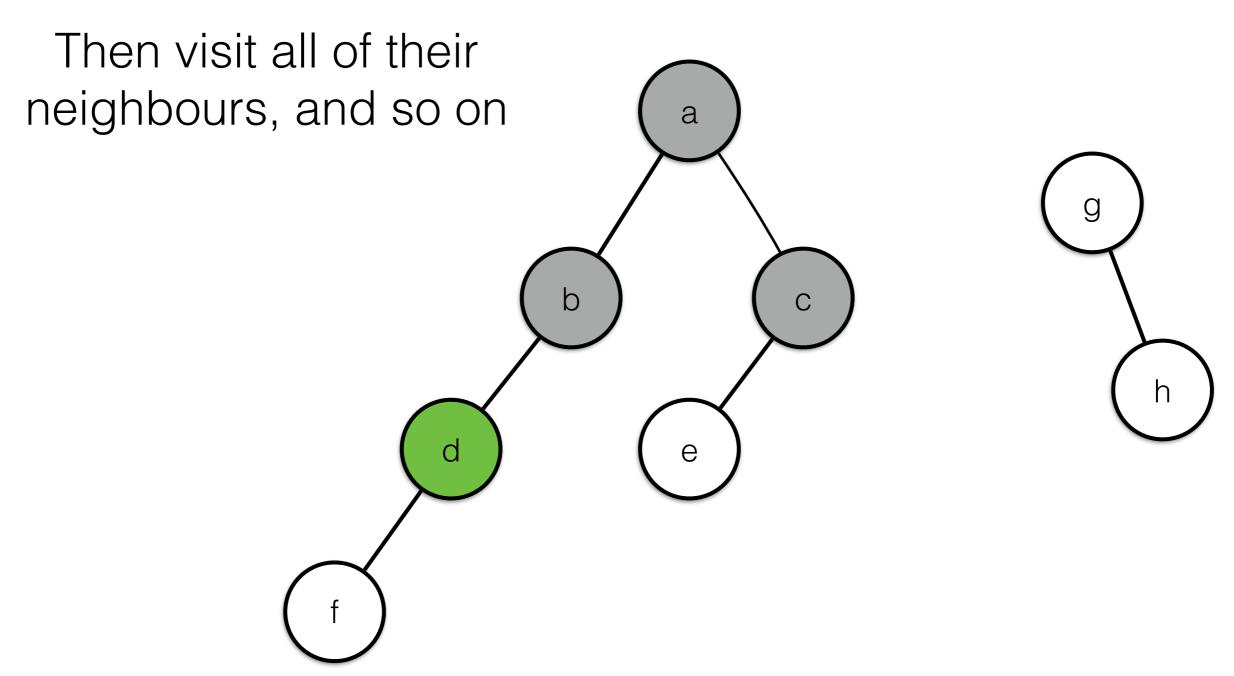


Nodes visited in this order: a b c



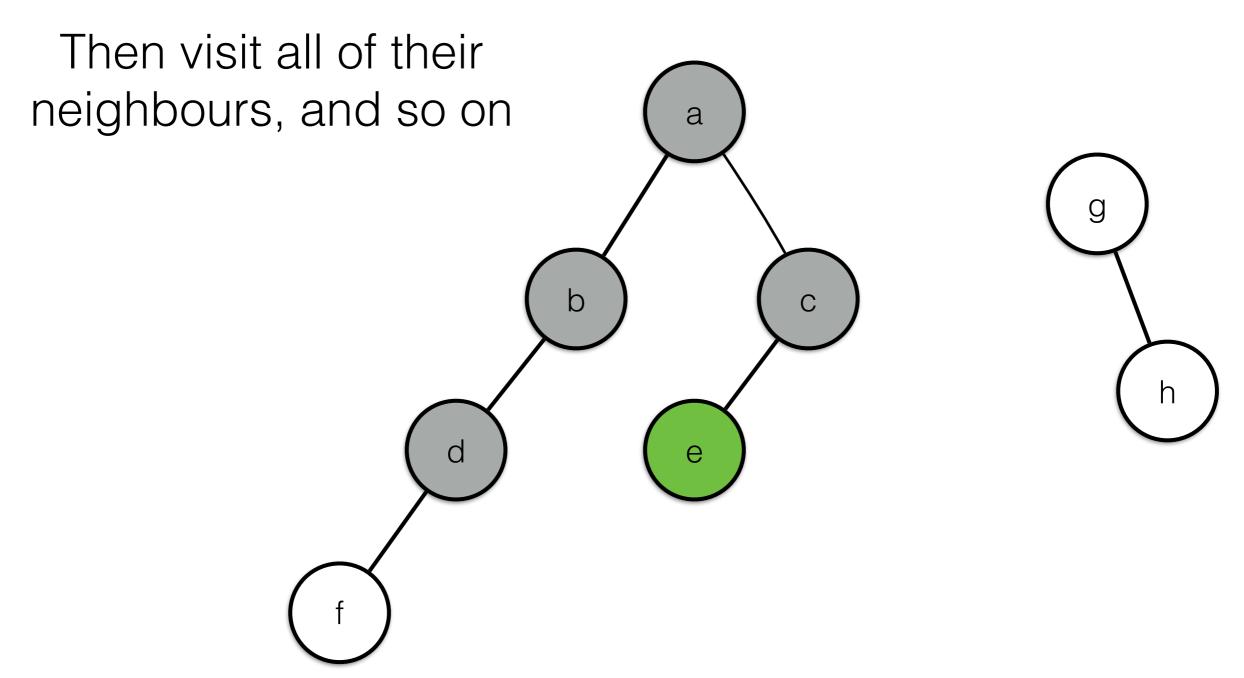


Nodes visited in this order: a b c d





Nodes visited in this order: a b c d e



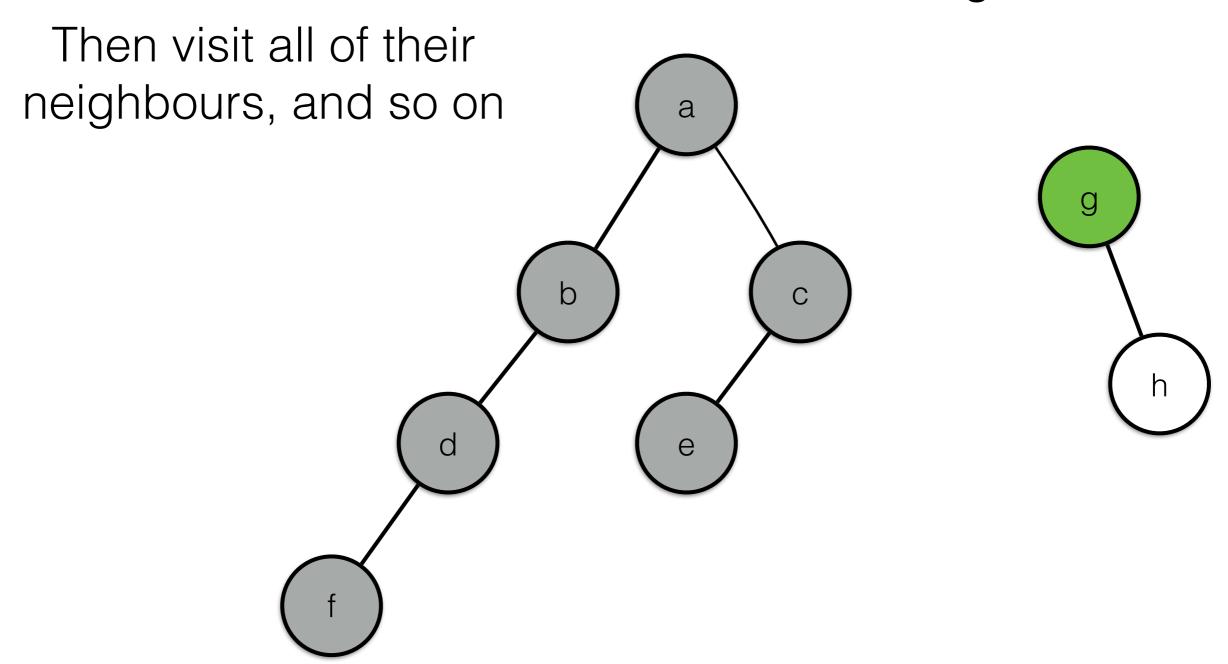


Nodes visited in this order: a b c d e f

Then visit all of their neighbours, and so on

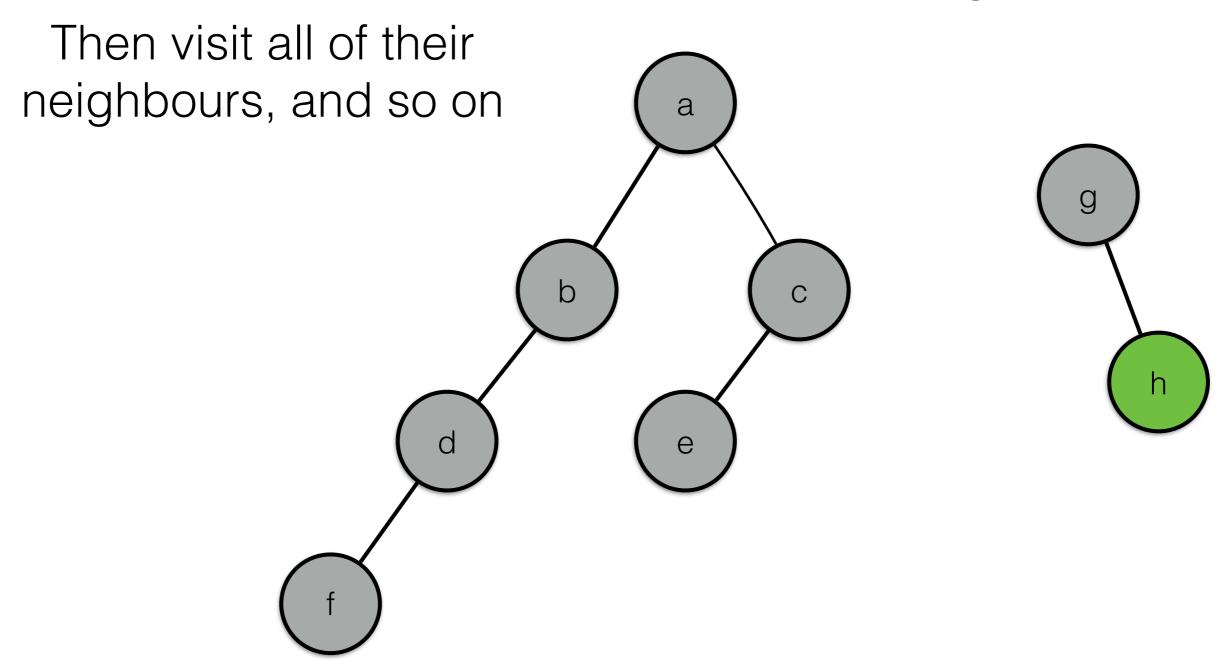


Nodes visited in this order: a b c d e f g

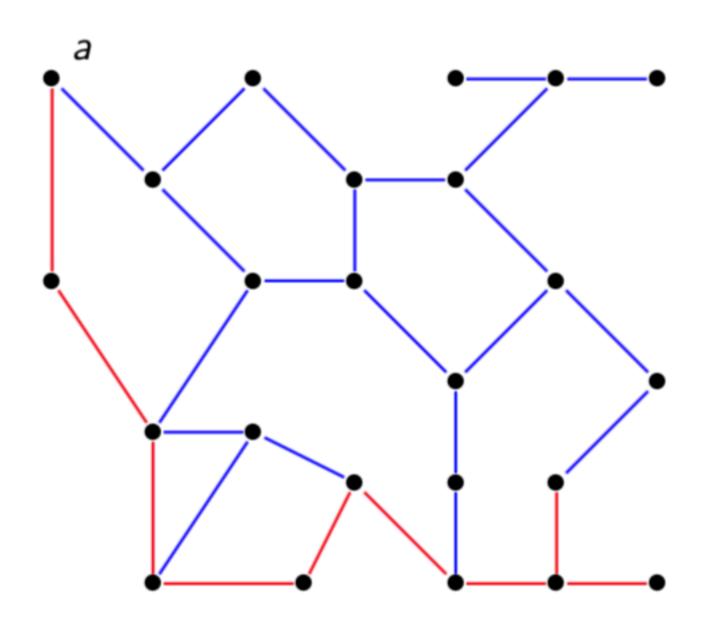




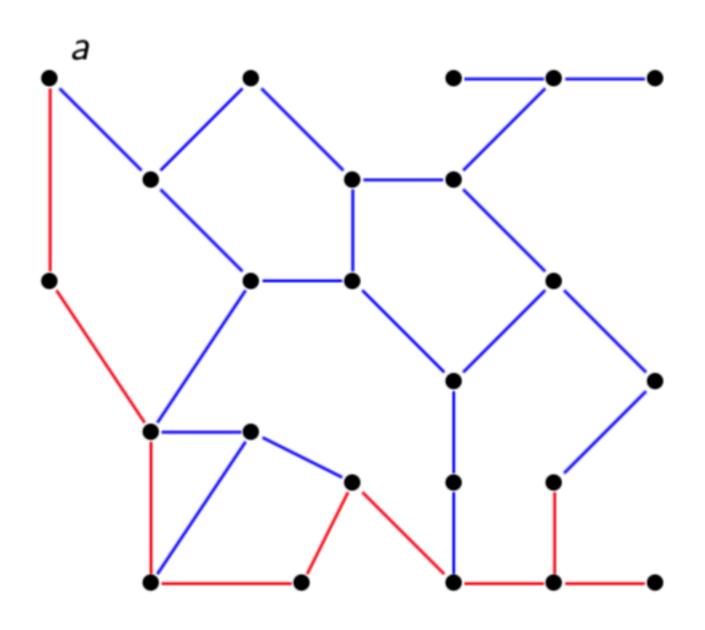
Nodes visited in this order: a b c d e f g h





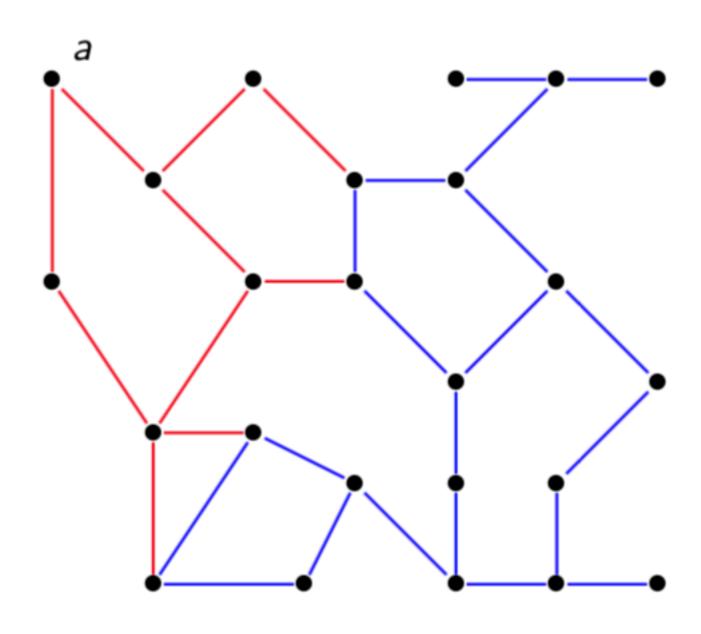




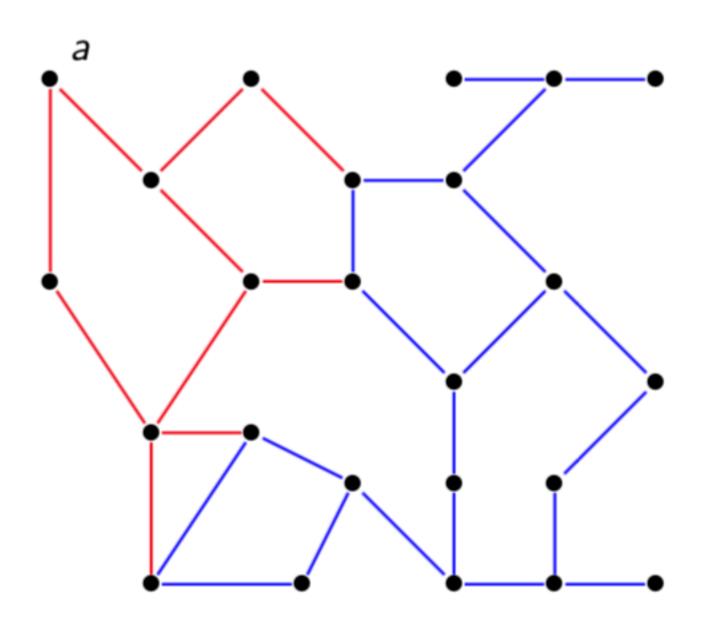


Typical Depth-First Search



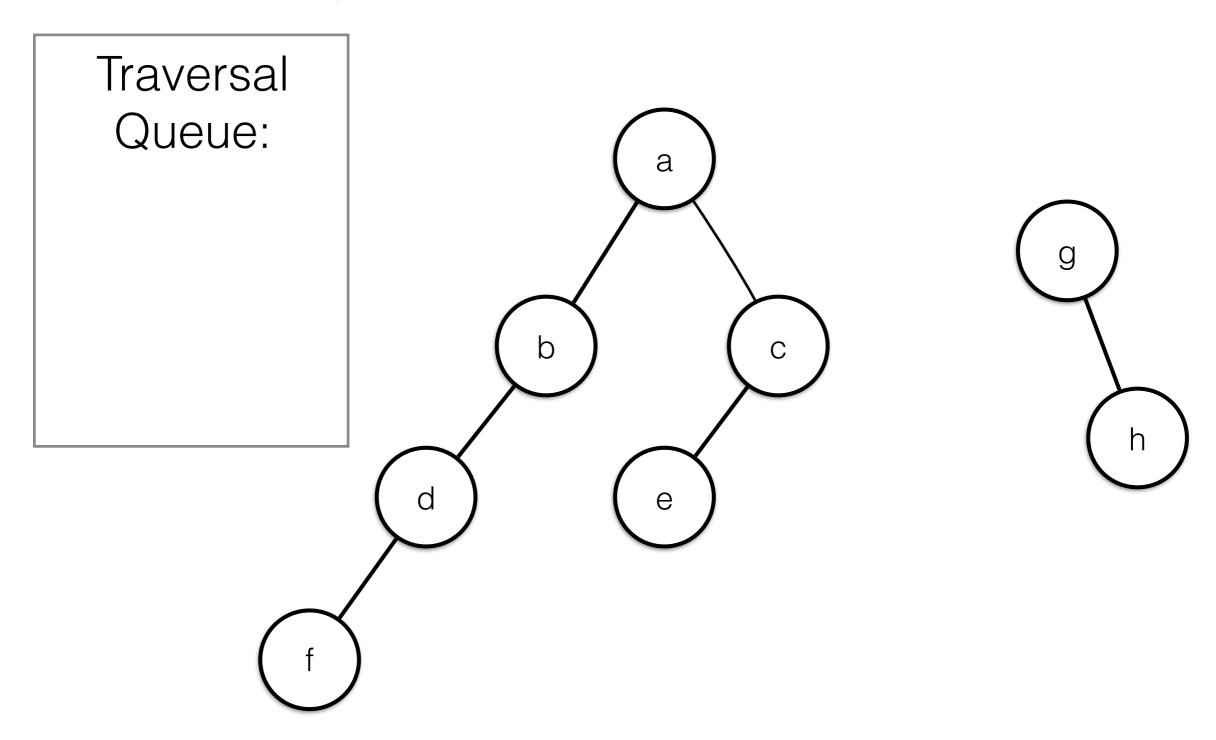






Typical Breadth-First Search

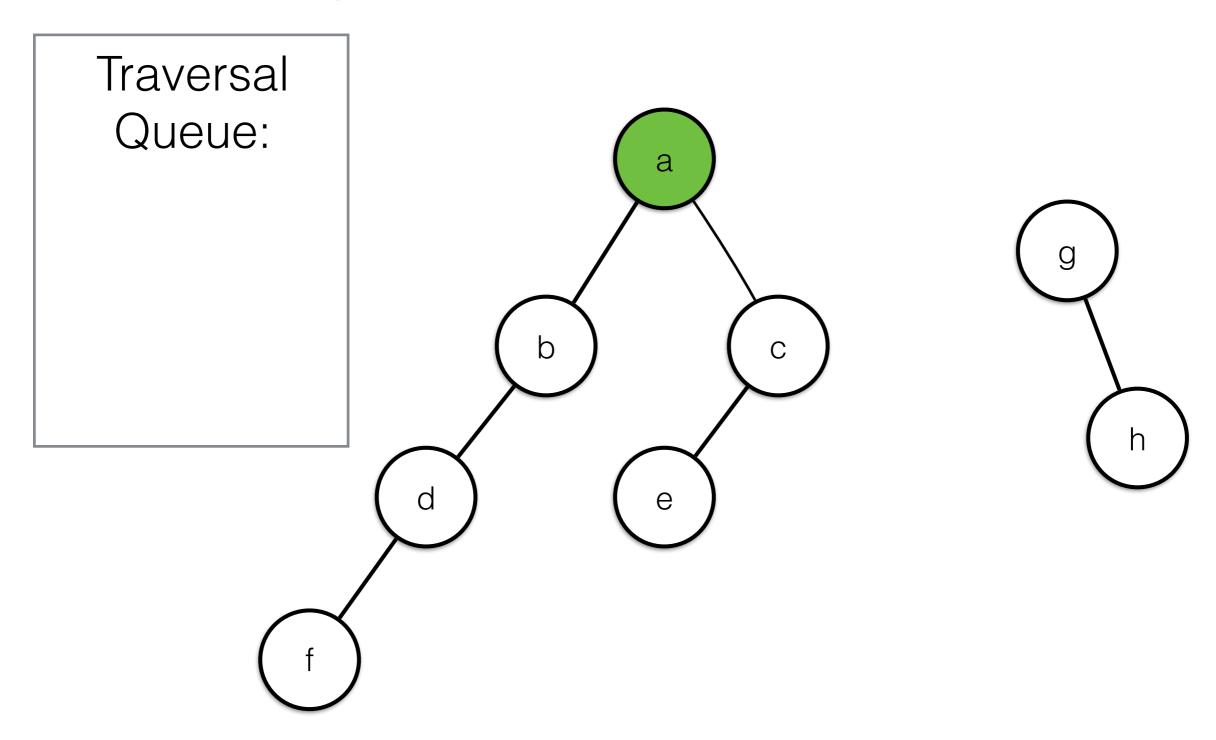




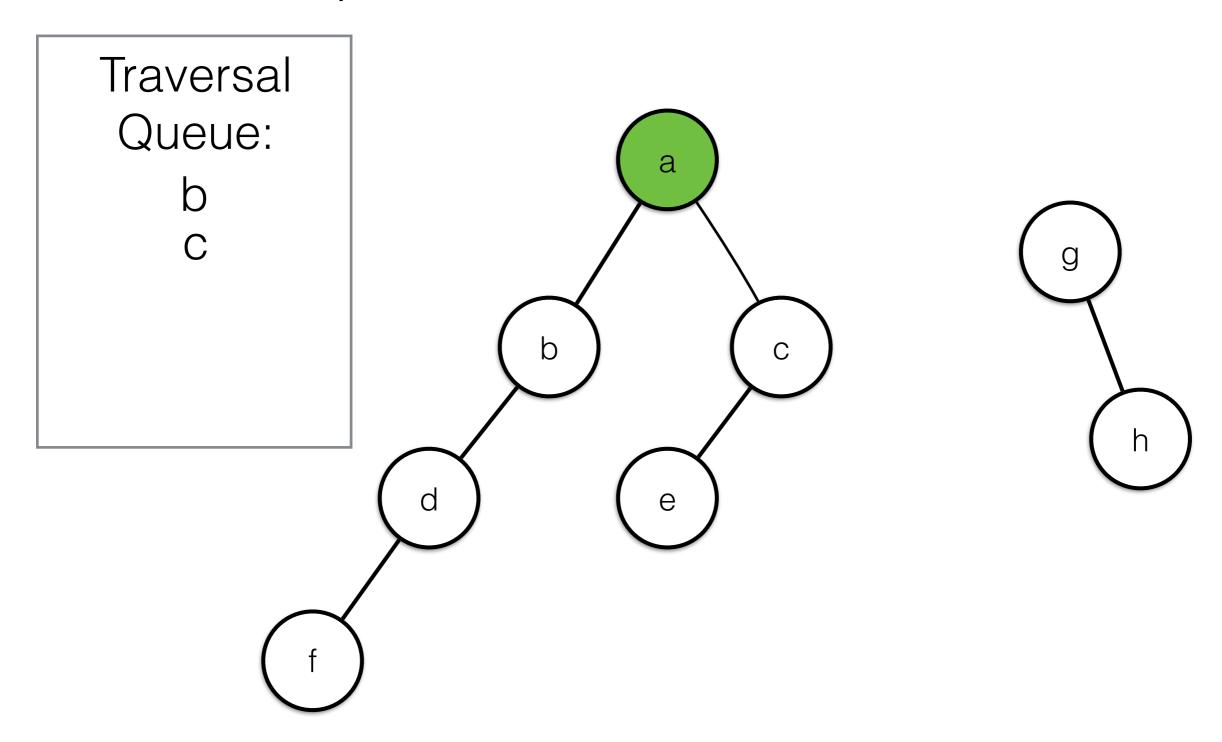


Traversal Queue: а a b

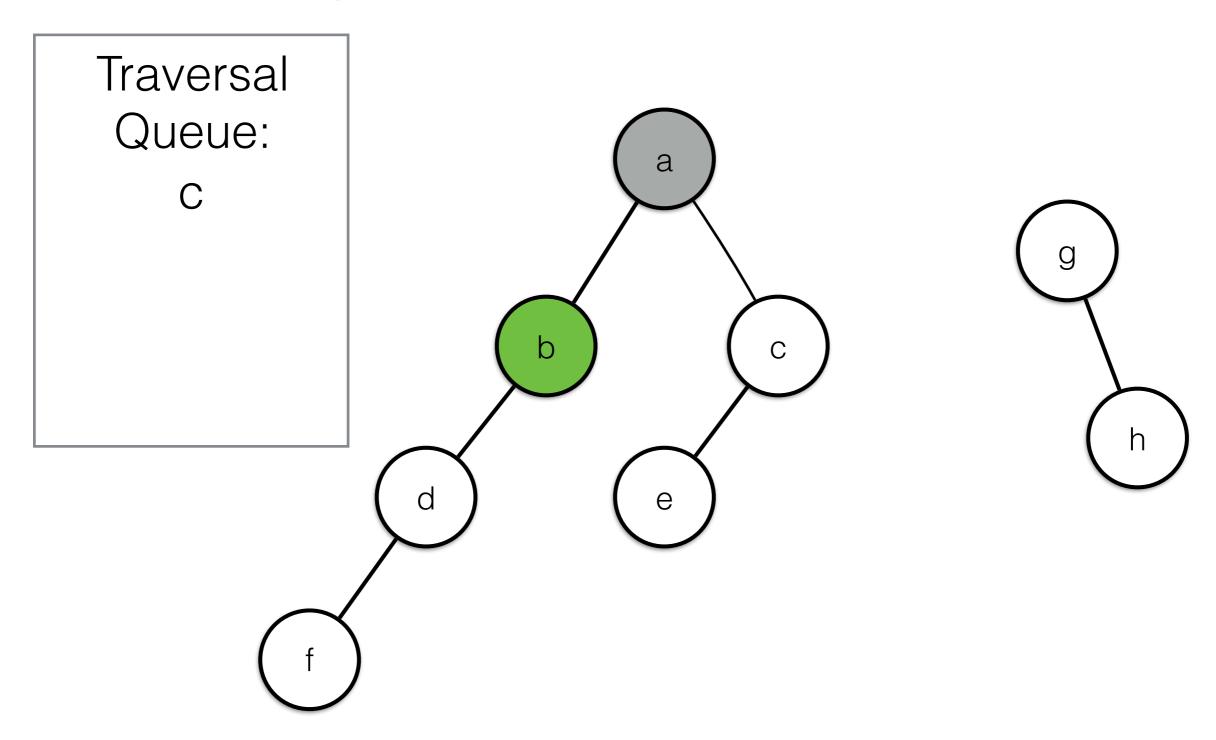




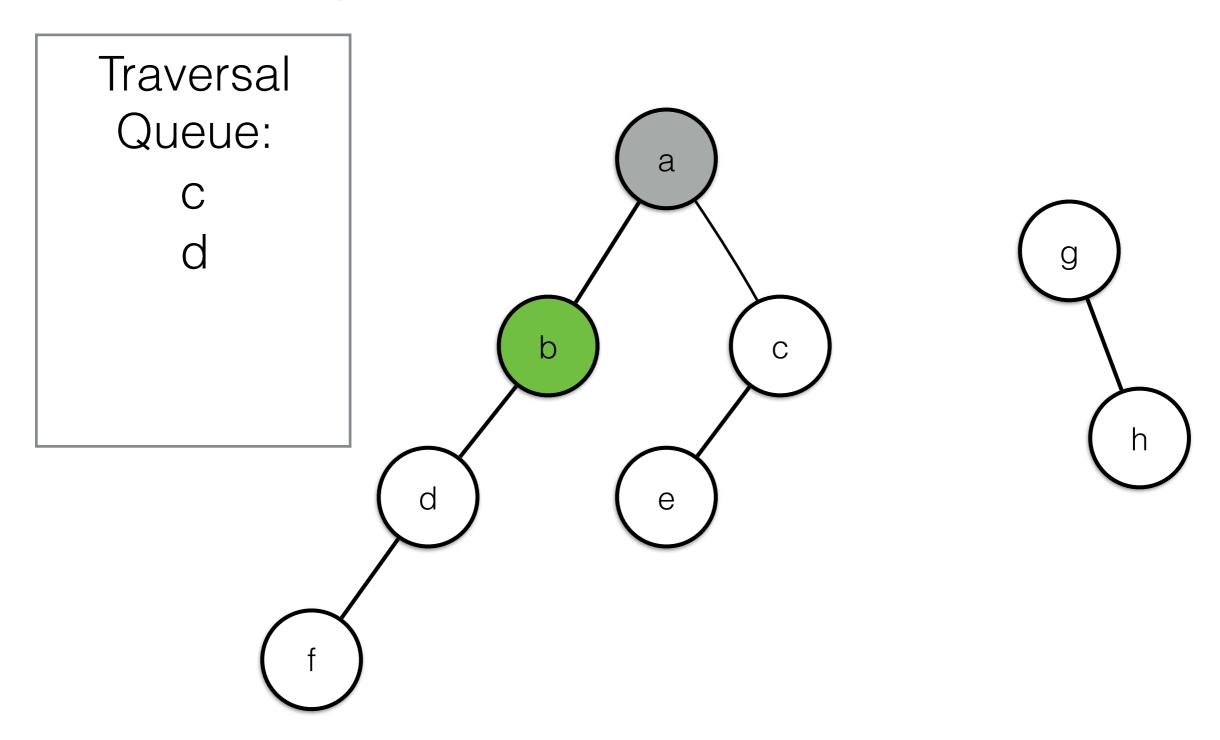




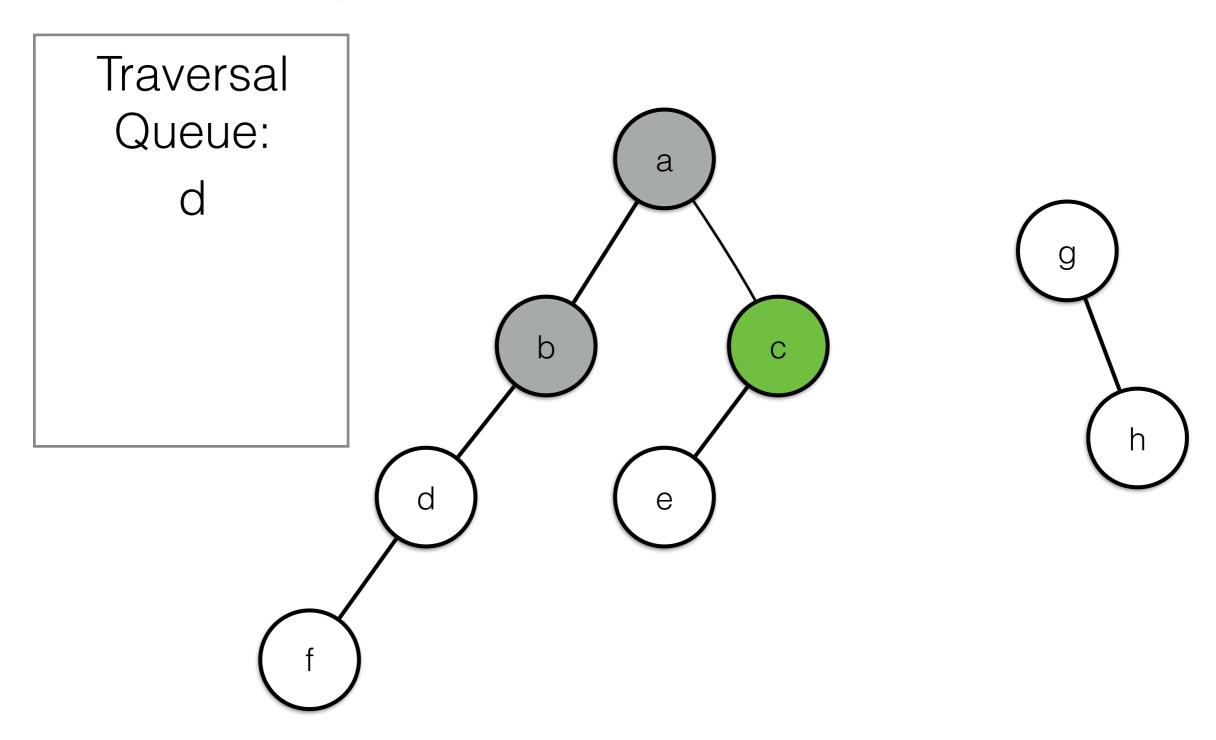




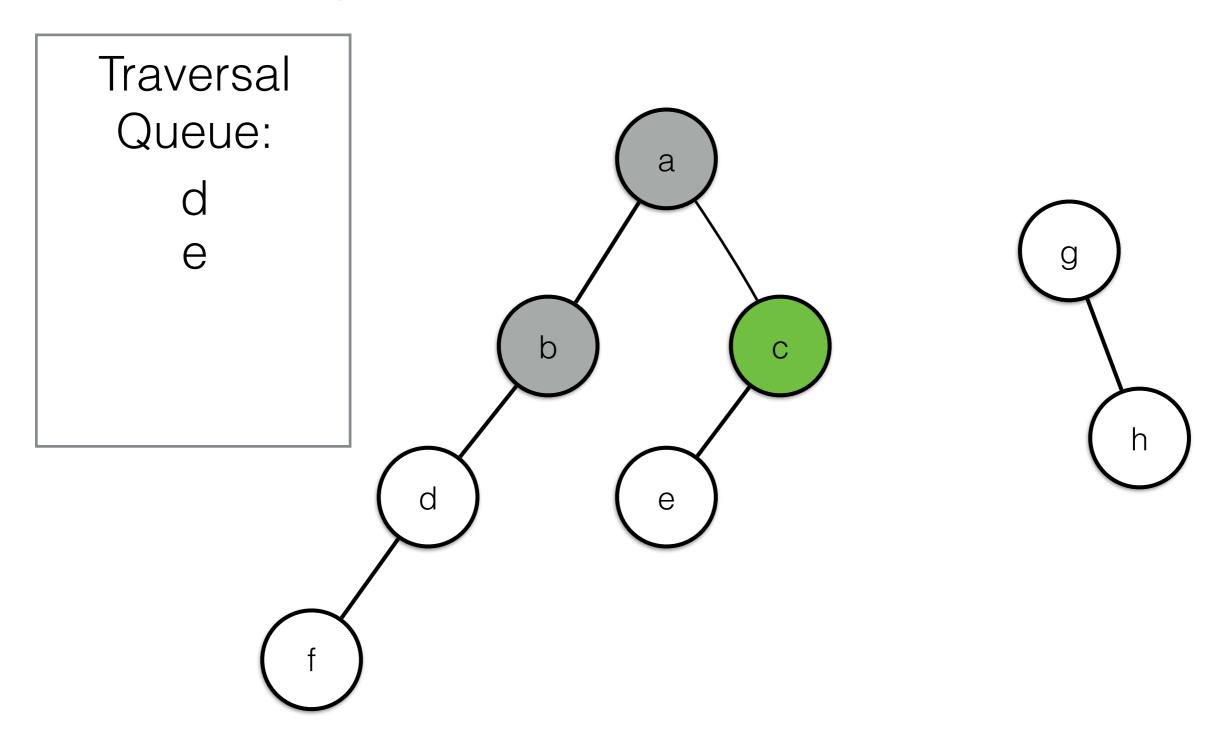




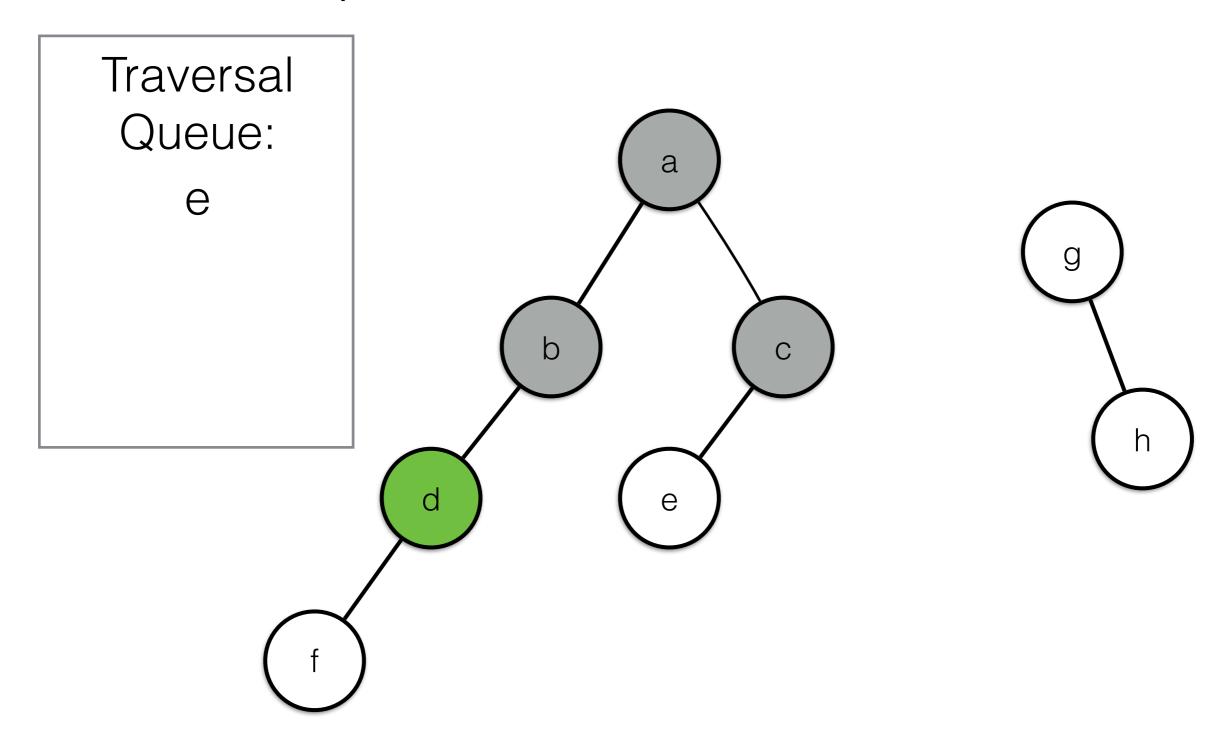






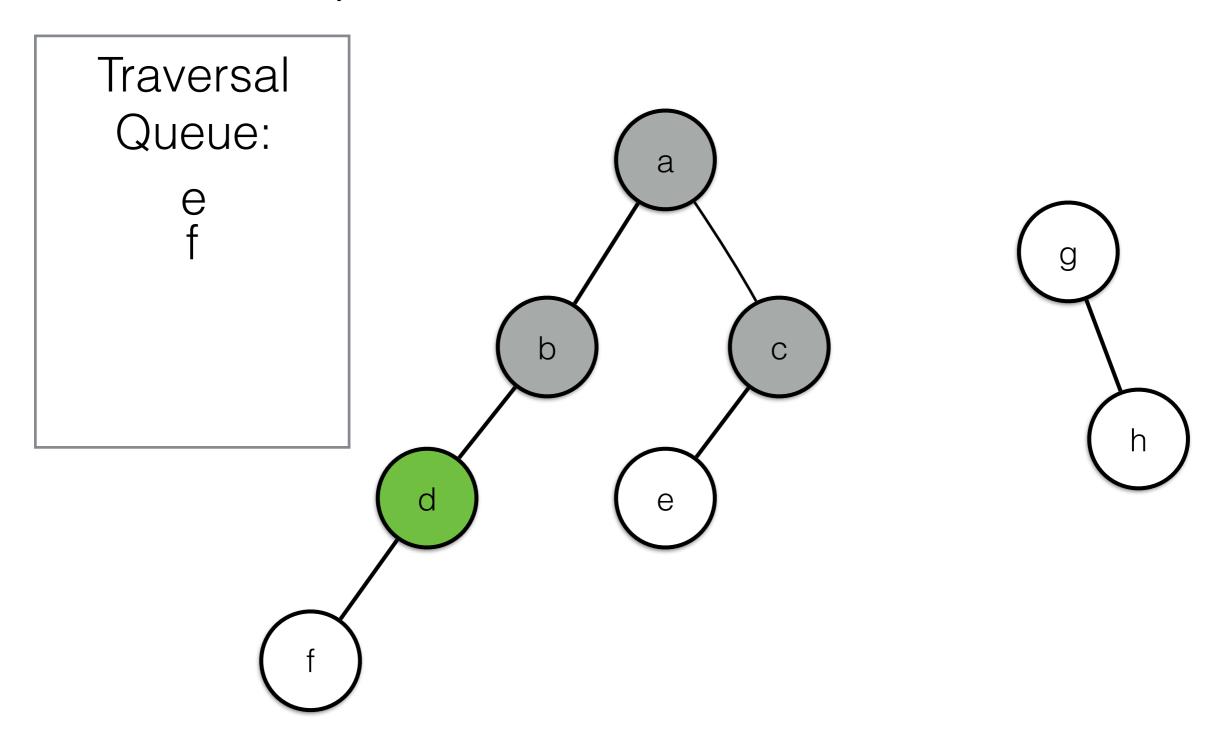




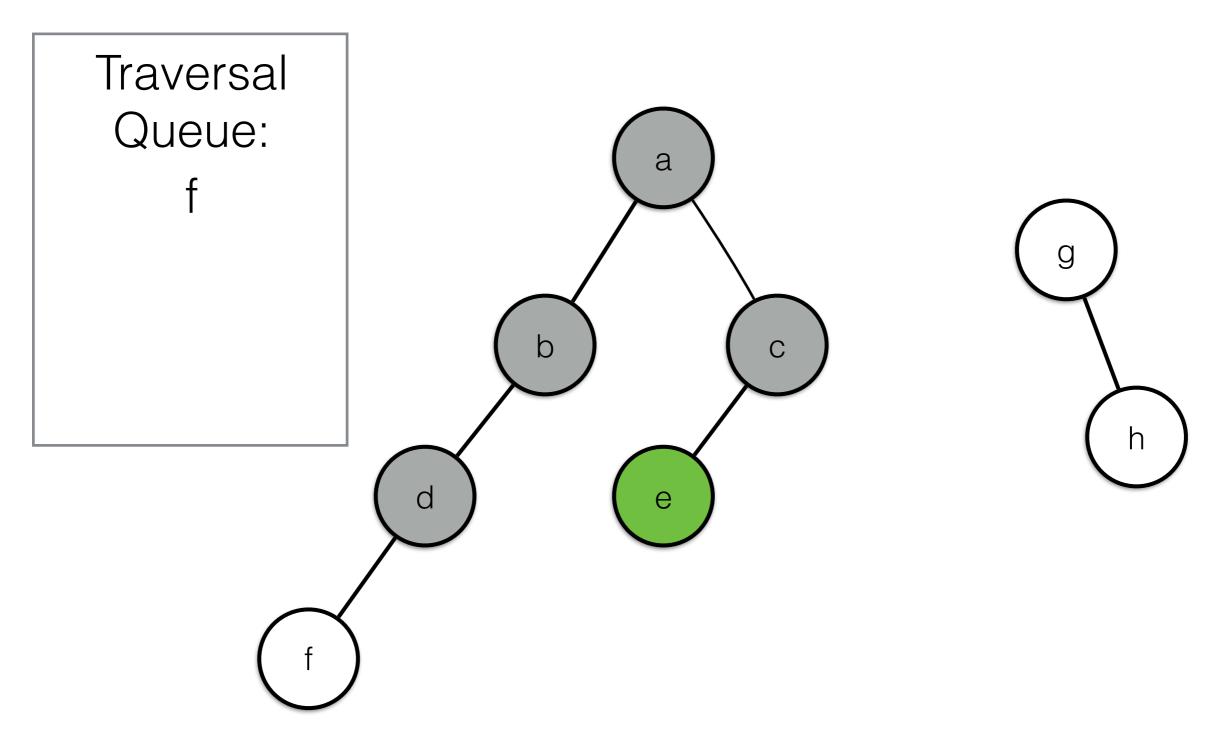


74

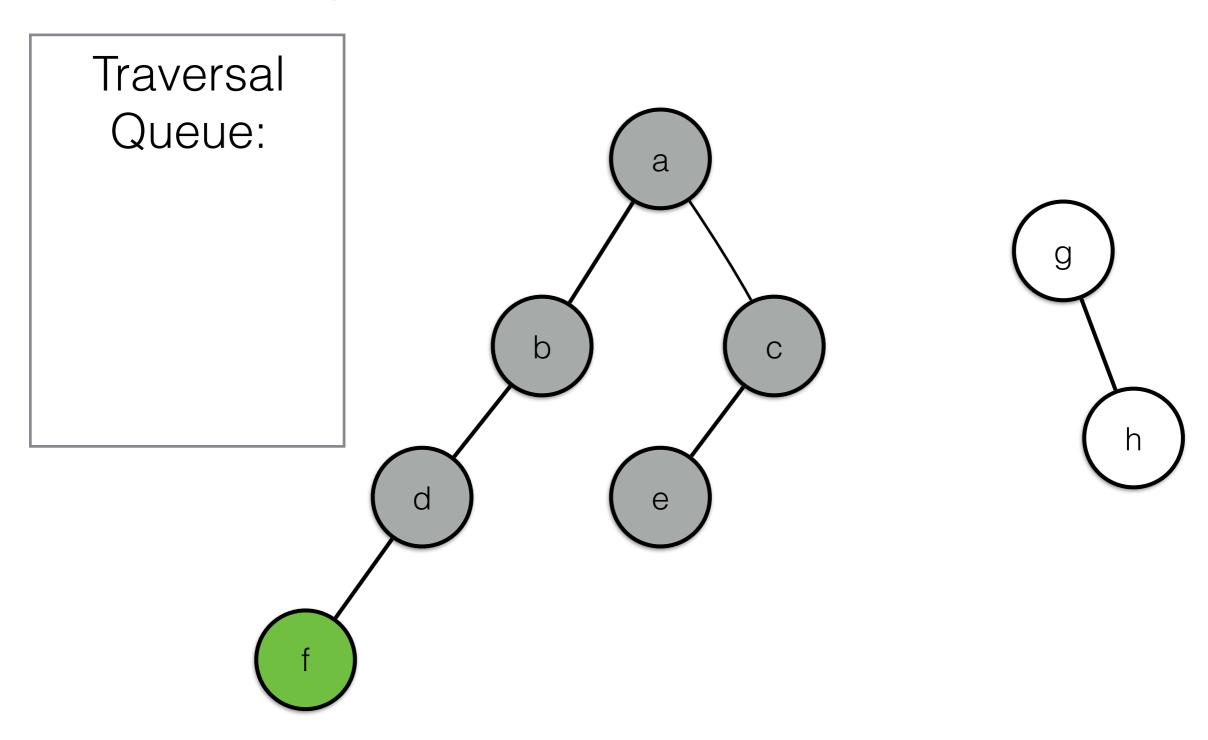




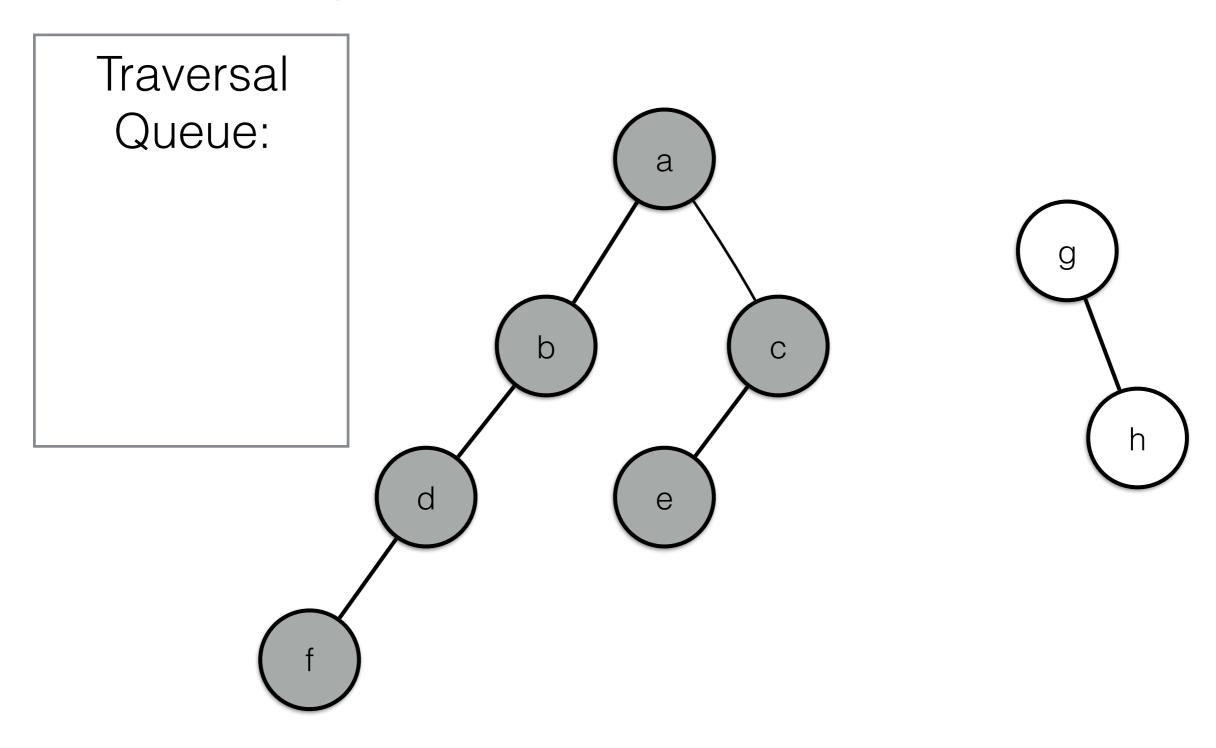




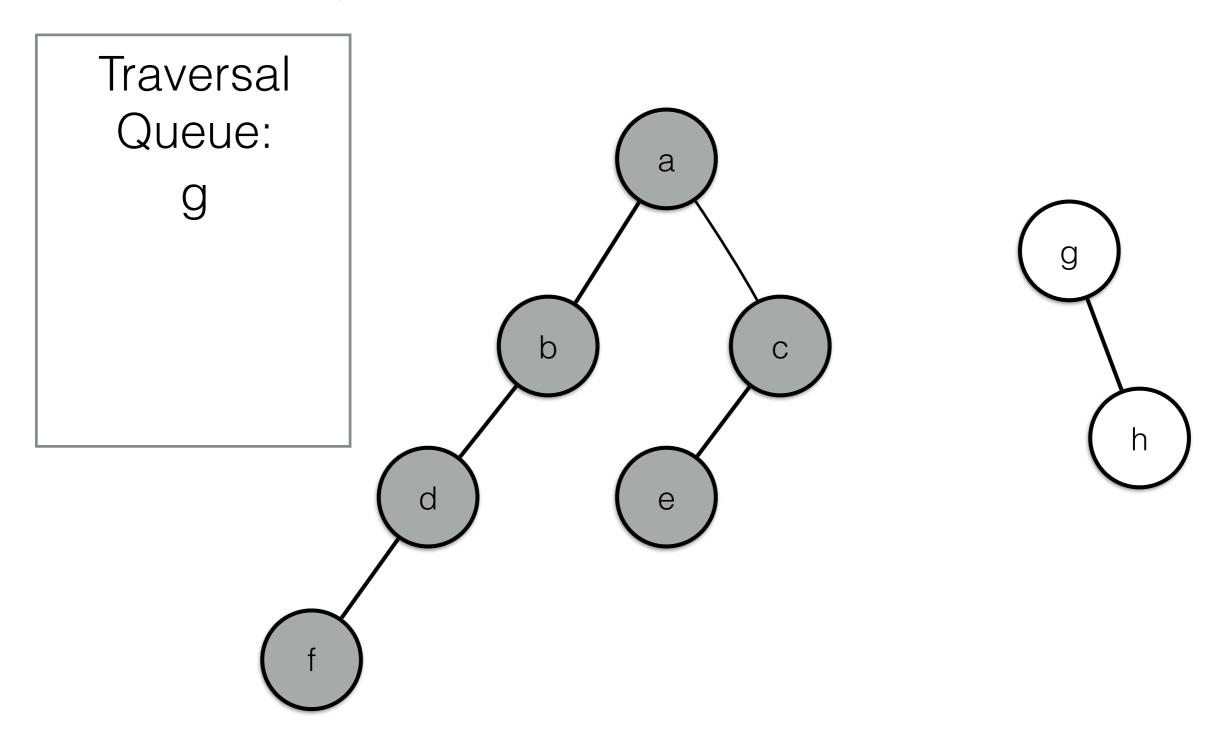




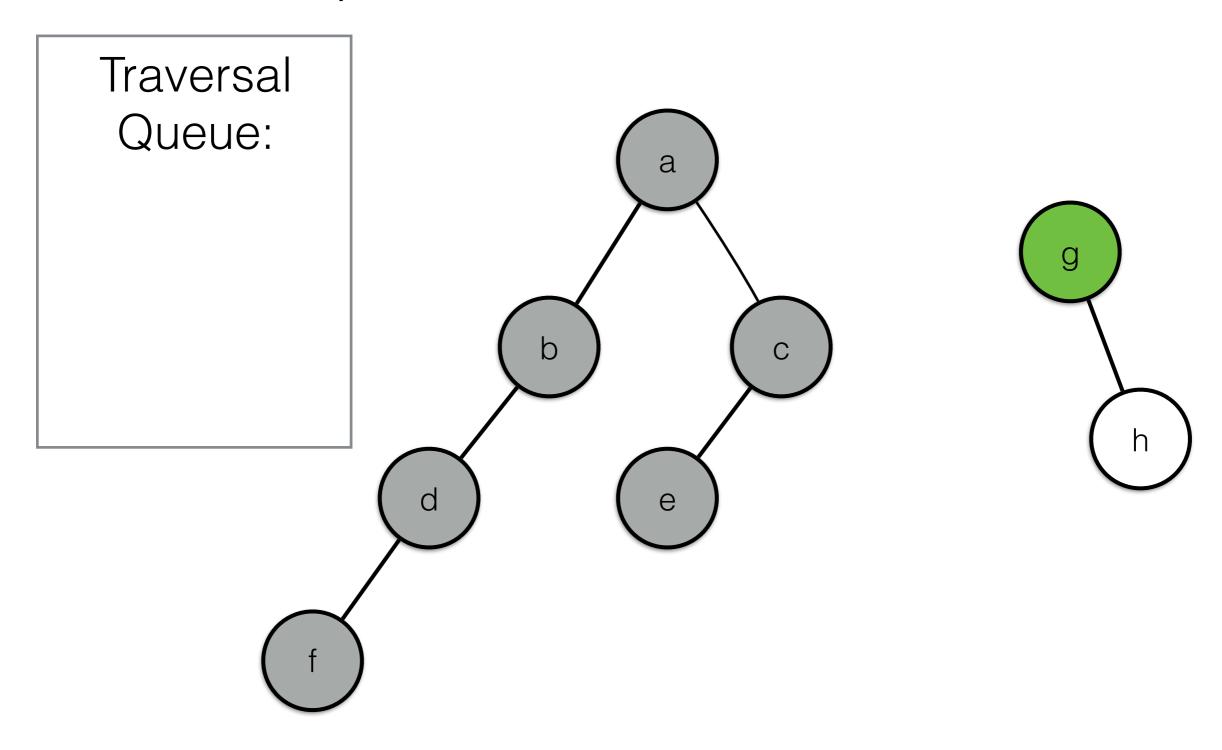




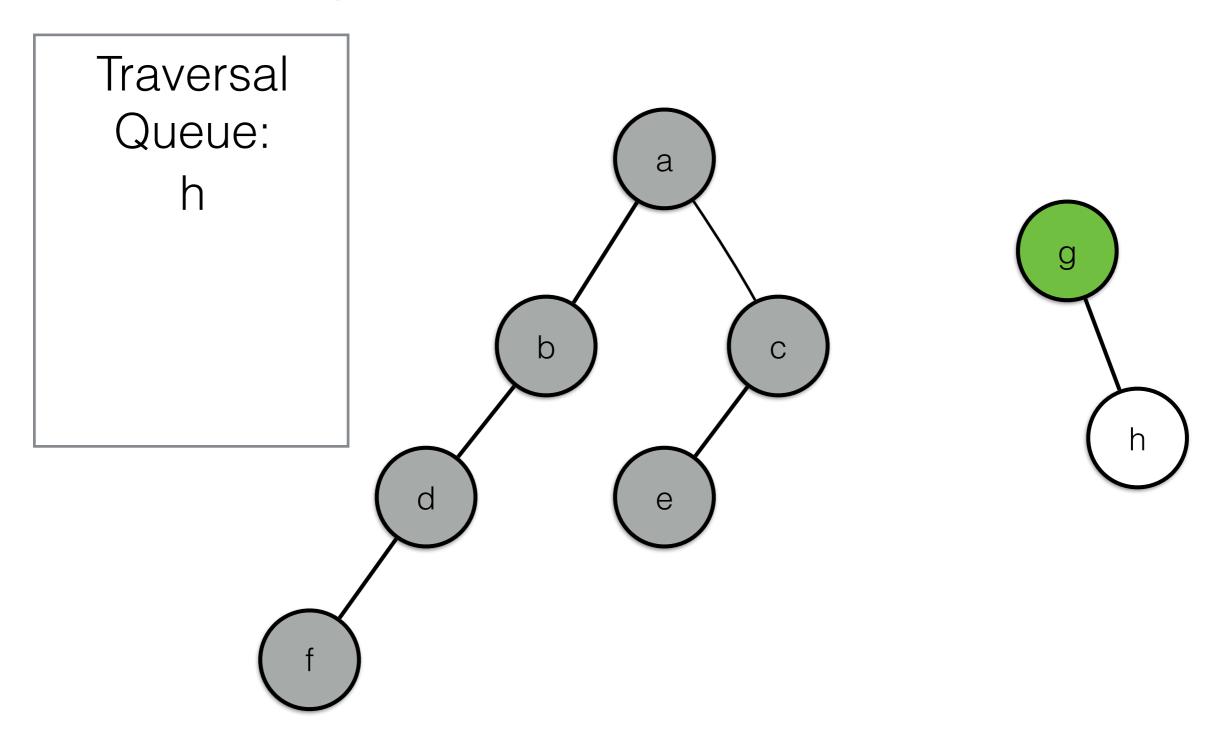




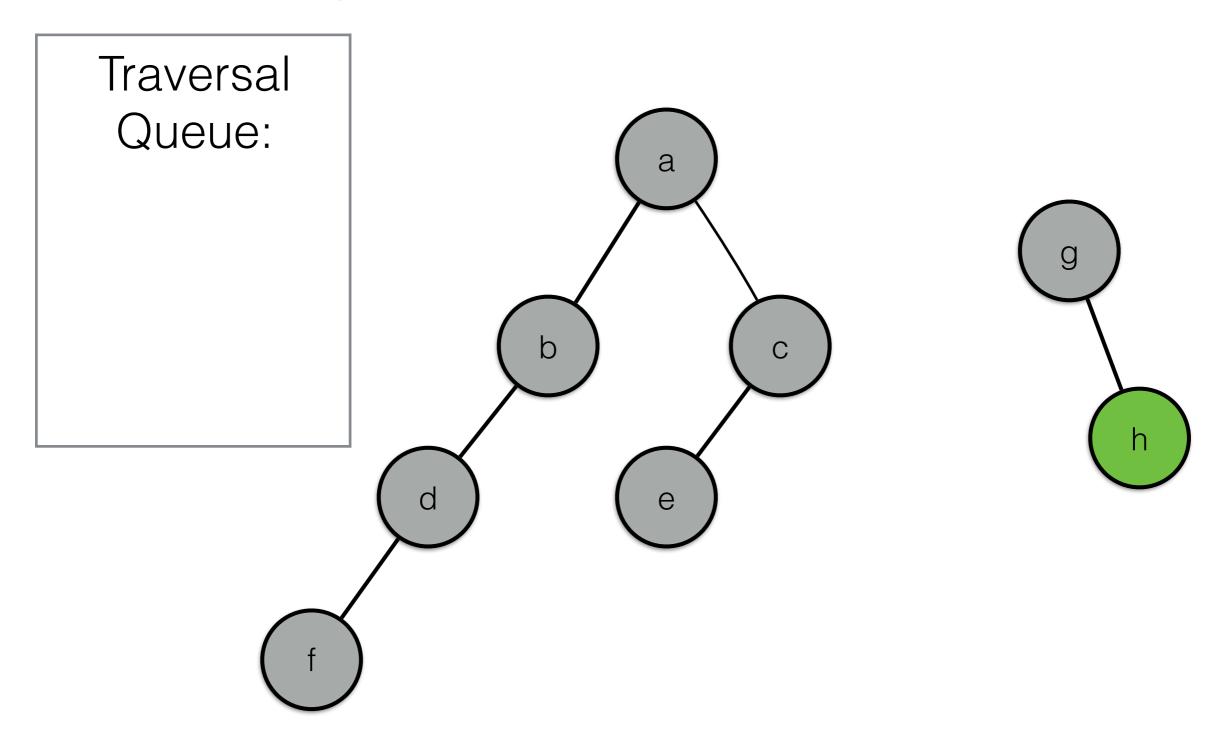












### Breadth-First Search Algorithm



```
function BFS(\langle V, E \rangle)
    mark each node in V with 0
    count \leftarrow 0, init(queue)
                                               create an empty queue
    for each v in V do
        if v is marked with 0 then
            count \leftarrow count + 1
            mark v with count
            inject(queue, v)

    □ queue containing just v

            while queue is non-empty do
                u \leftarrow eject(queue)

    b dequeues u

                for each edge (u, w) do
                                                   \triangleright w is u's neighbour
                    if w is marked with 0 then
                        count \leftarrow count + 1
                        mark w with count
                        inject(queue, w)
                                                            ▷ enqueues w
```

## BFS Algorithm Notes

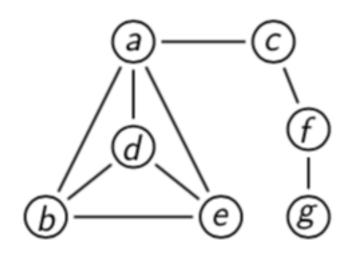


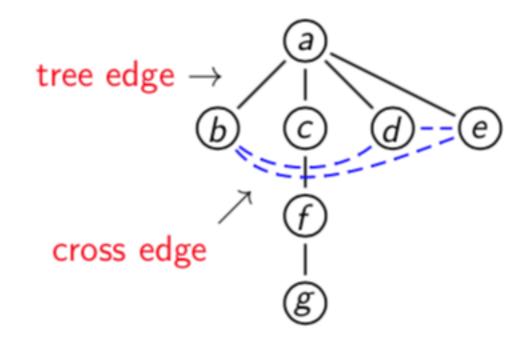
- BFS has the same complexity as DFS.
- Again, the same algorithm works for directed graphs as well.
- Certain problems are most easily solved by adapting BFS.
- For example, given a graph and two nodes, a and b in the graph, how would you find the length of the shortest path from a to b?

## Breadth-First Search Forest MELBOURNE



BFS **Tree** for this connected graph:





In general, we may get a **BFS Forest** 

## Topological Sorting

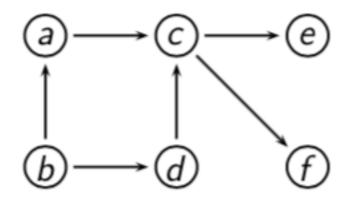


- We mentioned scheduling problems and their representation by directed graphs.
- Assume a directed edge from a to b means that task a must be completed before b can be started.
- The graph must be a dag; otherwise the problem cannot be solved.
- Assume the tasks are carried out by a single person, unable to multi-task.
- Then we should try to linearize the graph, that is, order the nodes as a sequence v<sub>1</sub>,v<sub>2</sub>,...,v<sub>n</sub> such that for each edge (v<sub>i</sub>,v<sub>j</sub>) ∈ E, we have v<sub>i</sub> comes before v<sub>j</sub> in the sequence (that is, v<sub>i</sub> is scheduled to happen before v<sub>j</sub>).

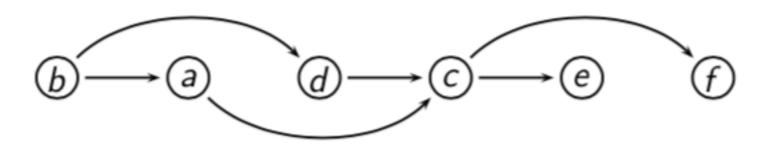
## Topological Sorting Example



There are 4 ways to linearise the following graph



Here is one:



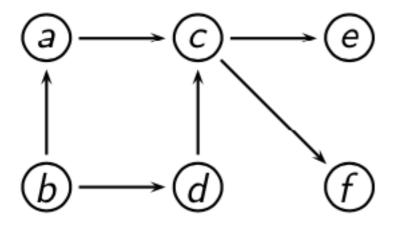
# Topological Sorting Algorithm 1



- We can solve the top-sort problem with depth-first search:
  - 1. Perform DFS and note the order in which nodes are popped off the stack.
  - 2. List the nodes in the reverse of that order.
- This works because of the stack discipline.
- If (u,v) is an edge then it is possible (given some way of deciding ties) to arrive at a DFS stack with u sitting below v.
- Taking the "reverse popping order" ensures that u is listed before v.



Using the DFS method and resolving ties by using alphabetical order, the graph gives rise to the traversal stack shown on the right (the popping order shown in red):



$$e_{3,1}$$
  $f_{4,2}$   $c_{2,3}$   $d_{6,5}$   $a_{1,4}$   $b_{5,6}$ 

Taking the nodes in reverse popping order yields b, d, a, c, f, e.

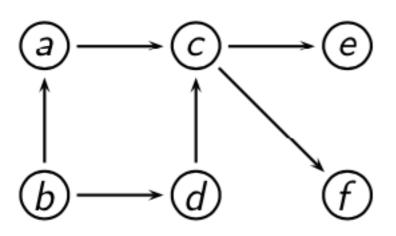
# Topological Sorting Algorithm 2



- An alternative method would be to repeatedly select a random **source** in the graph (that is, a node with no incoming edges), list it, and remove it from the graph (including removing its outgoing edges).
- This is a very natural approach, but it has the drawback that we repeatedly need to scan the graph for a source.
- However, it exemplifies the general principle of decrease-and-conquer.



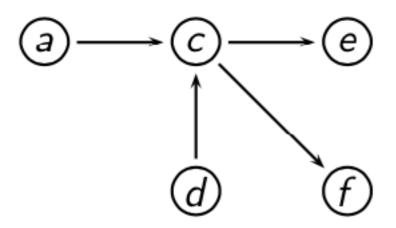
Using the source removal method (and resolving ties alphabetically):



Topological sorted order:



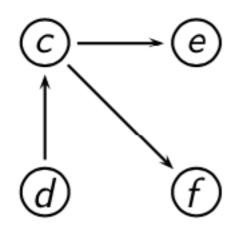
Using the source removal method (and resolving ties alphabetically):



Topological sorted order:



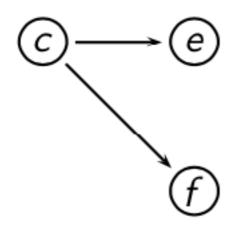
Using the source removal method (and resolving ties alphabetically):



Topological sorted order: b, a



Using the source removal method (and resolving ties alphabetically):



Topological sorted order: b, a, d



Using the source removal method (and resolving ties alphabetically):

(e)

(f)

Topological sorted order: b, a, d, c



Using the source removal method (and resolving ties alphabetically):



Topological sorted order: b, a, d, c, e



Using the source removal method (and resolving ties alphabetically):

Topological sorted order: b, a, d, c, e, f

#### Next time



 So next we turn our attention to the very useful "decrease and conquer" principle (Levitin Chapter 4).