CSCI 104
Graph Traversals and Other Algorithms
Mark Redekopp
BREADTH-FIRST SEARCH
Breadth-First Search

• Given a graph with vertices, V, and edges, E, and a starting vertex that we'll call u

• BFS starts at u (‘a’ in the diagram to the left) and fans-out along the edges to nearest neighbors, then to their neighbors and so on

• Goal: Find shortest paths (a.k.a. minimum number of hops or depth) from the start vertex to every other vertex
Breadth-First Search

- Given a graph with vertices, V, and edges, E, and a starting vertex, u
- BFS starts at u (‘a’ in the diagram to the left) and fans-out along the edges to nearest neighbors, then to their neighbors and so on
- Goal: Find shortest paths (a.k.a. minimum number of hops or depth) from the start vertex to every other vertex

Depth 0: a
Depth 1: c,e
Breadth-First Search

• Given a graph with vertices, V, and edges, E, and a starting vertex, u
• BFS starts at u (‘a’ in the diagram to the left) and fans-out along the edges to nearest neighbors, then to their neighbors and so on
• Goal: Find shortest paths (a.k.a. minimum number of hops or depth) from the start vertex to every other vertex

Depth 0: a
Depth 1: c,e
Depth 2: b,d,f,g
Breadth-First Search

- Given a graph with vertices, $V$, and edges, $E$, and a starting vertex, $u$
- BFS starts at $u$ (‘a’ in the diagram to the left) and fans-out along the edges to nearest neighbors, then to their neighbors and so on
- Goal: Find shortest paths (a.k.a. minimum number of hops or depth) from the start vertex to every other vertex

Depth 0: a
Depth 1: c,e
Depth 2: b,d,f,g
Depth 3: h
Developing the Algorithm

- Key idea: Must explore all nearer neighbors before exploring further-away neighbors
  - From ‘a’ we find ‘e’ and ‘c’
    - Computer can only do one thing at a time so we have to pick either e or c to explore from
    - Let’s say we pick e...we will find f
    - Now what vertex should we explore (i.e. visit neighbors) next?
    - C!! (if we don’t we won’t find shortest paths...e.g. d)
    - Must explore all vertices at depth i before any vertices at depth i+1

Depth 0: a
Depth 1: c,e
Depth 2: b,d,f,g
Depth 3: h
Developing the Algorithm

• Exploring all vertices at more shallow depths means is equivalent to exploring vertices in the order they are found
  – Keep a first-in / first-out queue (FIFO) of neighbors found
• Put newly found vertices in the back and pull out a vertex from the front to explore next
• We don’t want to put a vertex in the queue more than once...
  – ‘mark’ a vertex the first time we encounter it
  – only allow unmarked vertices to be put in the queue
• May also keep a ‘predecessor’ list that indicates how each vertex got discovered (i.e. which vertex caused this one to be found)
  – Allows us to find a shortest-path back to the start vertex
Breadth-First Search

Algorithm:

BFS(G,u)
1  for each vertex v
3  Q = new Queue
4  Q.enqueue(u), pred[u] = -2, d[u]=0
5  while Q is not empty
6     v = Q.front(); Q.dequeue()
7     foreach neighbor, w, of v:
8         if pred[w] == -1 // w not found
9             Q.enqueue(w)
10        pred[w] = v, d[w] = d[v] + 1
Breadth-First Search

Algorithm:

\[
\text{BFS}(G, u) \\
1 \quad \text{for each vertex } v \\
2 \quad \text{pred}[v] = -1, \text{d}[v] = \text{inf.} \\
3 \quad Q = \text{new Queue} \\
4 \quad Q.\text{enqueue}(u), \text{pred}[u] = -2, \text{d}[u] = 0 \\
5 \quad \text{while } Q \text{ is not empty} \\
6 \quad \quad v = Q.\text{front}(); Q.\text{dequeue}() \\
7 \quad \quad \text{foreach neighbor, } w, \text{ of } v: \\
8 \quad \quad \quad \text{if } \text{pred}[w] == -1 // w \text{ not found} \\
9 \quad \quad \quad \quad Q.\text{enqueue}(w) \\
10 \quad \quad \text{pred}[w] = v, \text{d}[w] = \text{d}[v] + 1
\]
Breadth-First Search

Algorithm:

```
BFS(G, u)
1 for each vertex v
3 Q = new Queue
4 Q.enqueue(u), pred[u] = -2, d[u]=0
5 while Q is not empty
6 v = Q.front(); Q.dequeue()
7 foreach neighbor, w, of v:
8 if pred[w] == -1 // w not found
9 Q.enqueue(w)
10 pred[w] = v, d[w] = d[v] + 1
```
Breadth-First Search

Algorithm:

BFS(G,u)
1  for each vertex v
3  Q = new Queue
4  Q.enqueue(u), pred[u] = -2, d[u]=0
5  while Q is not empty
6     v = Q.front(); Q.dequeue()
7     foreach neighbor, w, of v:
8           if pred[w] == -1 // w not found
9              Q.enqueue(w)
10           pred[w] = v,  d[w] = d[v] + 1
Breadth-First Search

Algorithm:

BFS(G,u)
1  for each vertex v
3  Q = new Queue
4  Q.enqueue(u), pred[u] = -2, d[u]=0
5  while Q is not empty
6     v = Q.front(); Q.dequeue()
7     foreach neighbor, w, of v:
8         if pred[w] == -1 // w not found
9             Q.enqueue(w)
10        pred[w] = v, d[w] = d[v] + 1
Breadth-First Search

Algorithm:

BFS(G,u)
1   for each vertex v
3   Q = new Queue
4   Q.enqueue(u), pred[u] = -2, d[u]=0
5 while Q is not empty
6   v = Q.front(); Q.dequeue()
7   foreach neighbor, w, of v:
8      if pred[w] == -1 // w not found
9         Q.enqueue(w)
10     pred[w] = v, d[w] = d[v] + 1
Breadth-First Search

Algorithm:

BFS(G,u)
1  for each vertex v
3  Q = new Queue
4  Q.enqueue(u), pred[u] = -2, d[u]=0
5  while Q is not empty
6     v = Q.front(); Q.dequeue()
7     foreach neighbor, w, of v:
8         if pred[w] == -1 // w not found
9             Q.enqueue(w)
10        pred[w] = v, d[w] = d[v] + 1
Breadth-First Search

Algorithm:

BFS(G,u)
1  for each vertex v
3  Q = new Queue
4  Q.enqueue(u), pred[u] = -2, d[u]=0
5  while Q is not empty
6     v = Q.front(); Q.dequeue()
7     foreach neighbor, w, of v:
8        if pred[w] == -1 // w not found
9           Q.enqueue(w)
10       pred[w] = v, d[w] = d[v] + 1
Breadth-First Search

Algorithm:

BFS(G,u)
1  for each vertex v
3  Q = new Queue
4  Q.enqueue(u), pred[u] = -2, d[u]=0
5  while Q is not empty
6     v = Q.front(); Q.dequeue()
7      foreach neighbor, w, of v:
8          if pred[w] == -1 // w not found
9              Q.enqueue(w)
10         pred[w] = v, d[w] = d[v] + 1
Breadth-First Search

• Shortest paths can be found by walking predecessor value from any node backward

• Example:
  – Shortest path from a to h
  – Start at h
  – Pred[h] = b (so walk back to b)
  – Pred[b] = c (so walk back to c)
  – Pred[c] = a (so walk back to a)
  – Pred[a] = -2 ... no predecessor, Done!!
Correctness

• Define
  – \( \text{dist}(s,v) = \) shortest distance
  – \( d[v] = \) BFS computed distance
  – \( p[v] = \) predecessor of \( v \)

• Loop invariant
  – What can we say about the nodes in the queue?

BFS(G,u)
1 for each vertex \( v \)
2 \( \text{pred}[v] = -1, \ d[v] = \text{inf.} \)
3 \( Q = \) new Queue
4 \( Q.\text{enqueue}(u), \ \text{pred}[u] = -2, \ d[u]=0 \)
5 while \( Q \) is not empty
6 \( v = Q.\text{front}(); \ Q.\text{dequeue}() \)
7 foreach neighbor, \( w, \) of \( v: \)
8 \( \text{if pred}[w] == -1 \ // w \) not found
9 \( Q.\text{enqueue}(w) \)
10 \( \text{pred}[w] = v, \ \ d[w] = d[v] + 1 \)
Correctness

• Define
  – \( \text{dist}(s,v) = \) shortest distance
  – \( d[v] = \) BFS computed distance
  – \( p[v] = \) predecessor of \( v \)

• Loop invariant
  – All vertices with \( \text{pred}[v] \neq -1 \) (i.e. already in the queue or popped from queue) have \( d[v] = \text{dist}(s,v) \)
  – The distance of the nodes in the queue are sorted
    • If \( Q = \{v_1, v_2, ..., v_r\} \) then \( d[v_1] \leq d[v_2] \leq ... \leq d[v_r] \)
  – The nodes in the queue are from 2 adjacent layers/levels
    • i.e. \( d[v_k] \leq d[v_1] + 1 \)
    • Suppose there is a node from a 3rd level \( (d[v_1] + 2) \), it must have been found by some, \( v_i \), where \( d[v_i] = d[v_1]+1 \)

BFS(\( G,u \))
1  for each vertex \( v \)
2    \( \text{pred}[v] = -1, \ d[v] = \text{inf.} \)
3  \( Q = \) new Queue
4  \( Q.\text{enqueue}(u), \ \text{pred}[u] = -2, \ d[u]=0 \)
5  while \( Q \) is not empty
6    \( v = Q.\text{front}(); Q.\text{dequeue()} \)
7    foreach neighbor, \( w, \) of \( v: \)
8      if \( \text{pred}[w] == -1 \) // \( w \) not found
9        \( Q.\text{enqueue}(w) \)
10     \( \text{pred}[w] = v, \ d[w] = d[v] + 1 \)
Breadth-First Search

• Analyze the run time of BFS for a graph with $n$ vertices and $m$ edges
  – Find $T(n)$

• How many times does loop on line 5 iterate?

• How many times loop on line 7 iterate?

BFS($G,u$)
1  for each vertex $v$
3  $Q$ = new Queue
4  $Q$.enqueue($u$), pred[$u$] = -2, $d[u] = 0$
5  while $Q$ is not empty
6   $v$ = $Q$.front(); $Q$.dequeue()
7  foreach neighbor, $w$, of $v$:
8     if pred[$w$] == -1 // $w$ not found
9        $Q$.enqueue($w$)
10    pred[$w$] = $v$, $d[w] = d[v] + 1$
### Breadth-First Search

- Analyze the run time of BFS for a graph with n vertices and m edges
  - Find $T(n)$
- How many times does loop on line 5 iterate?
  - $N$ times (one iteration per vertex)
- How many times loop on line 7 iterate?
  - For each vertex, $v$, the loop executes $\deg(v)$ times
    - $= \sum_{v \in V} \theta[1 + \deg(v)]$
    - $= \theta(\sum_v 1) + \theta(\sum_v \deg(v))$
    - $= \Theta(n) + \Theta(m)$
- Total = $\Theta(n+m)$

---

**BFS(G,u)**

1. for each vertex $v$
2. $\text{pred}[v] = -1$, $\text{d}[v] = \infty$
3. $Q = \text{new Queue}$
4. $Q.\text{enqueue}(u)$, $\text{pred}[u] = -2$, $\text{d}[u]=0$
5. while $Q$ is not empty
6. $v = Q.\text{front}(); Q.\text{dequeue}()$
7. foreach neighbor, $w$, of $v$:
8. if $\text{pred}[w] == -1$ // $w$ not found
9. $Q.\text{enqueue}(w)$
10. $\text{pred}[w] = v$, $\text{d}[w] = \text{d}[v] + 1$
DEPTH FIRST SEARCH
DFS Application: Topological Sort

• Breadth-first search doesn't solve all our problems.
• Given a graph of dependencies (tasks, prerequisites, etc.) topological sort creates a consistent ordering of tasks (vertices) where no dependencies are violated
• Many possible valid topological orderings exist
  – EE 109, EE 209, EE 354, EE 454, EE 457, CS104, PHYS 152, CS 201,…
  – CS 104, EE 109, CS 170, EE 209,…
Topological Sort

- Another example
  - Getting dressed
- More Examples:
  - Project management scheduling
  - Build order in a Makefile or other compile project
  - Cooking using a recipe
  - Instruction execution on an out-of-order pipelined CPU
  - Production of output values in a simulation of a combinational gate network

http://www.personal.kent.edu/~rmuhamma/Algorithms/MyAlgorithms/GraphAlgor/topoSort.htm
Topological Sort

- Does breadth-first search work?
  - No. What if we started at CS 170...
  - We'd go to CS 201L before CS 104
- All parent nodes need to be completed before any child node
- BFS only guarantees *some* parent has completed before child
- Turns out a Depth-First Search will be part of our solution
Depth First Search

- Explores ALL children before completing a parent
  - Note: BFS completes a parent before ANY children
- For DFS let us assign:
  - A start time when the node is first found
  - A finish time when a node is completed
- If we look at our nodes in reverse order of finish time (i.e. last one to finish back to first one to finish) we arrive at a...
  - Topological ordering!!!
DFS Algorithm

• Visit a node
  – Mark as visited (started)
  – For each visited neighbor, visit it and perform DFS on all of their children
  – Only then, mark as finished
• DFS is recursive!!
• If cycles in the graph, ensure we don’t get caught visiting neighbors endlessly
  – Color them as we go
  – White = unvisited,
  – Gray = visited but not finished
  – Black = finished

DFS-All (G)
1  for each vertex u
2    u.color = WHITE
3  finish_list = empty_list
4  for each vertex u do
5    if u.color == WHITE then
6      DFS-Visit (G, u, finish_list)
7  return finish_list

DFS-Visit (G, u)
1  u.color = GRAY
2  for each vertex v in Adj(u) do
3    if v.color == WHITE then
4      DFS-Visit (G, v)
5    u.color = BLACK
6  finish_list.append(u)
Depth First-Search

DFS-All (G)
1  for each vertex u
2   u.color = WHITE
3  finish_list = empty_list
4  for each vertex u do
5      if u.color == WHITE then
6         DFS-Visit (G, u, finish_list)
7  return finish_list

DFS-Visit (G, u)
1   u.color = GRAY
2   for each vertex v in Adj(u) do
3       if v.color = WHITE then
4          DFS-Visit (G, v)
5   u.color = BLACK
6   finish_list.append(u)
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6   DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4   DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
**Depth First-Search**

**DFS-All (G)**
1. for each vertex u
2. u.color = WHITE
3. finish_list = empty_list
4. for each vertex u do
5. if u.color == WHITE then
6. DFS-Visit (G, u, finish_list)
7. return finish_list

**DFS-Visit (G, u)**
1. u.color = GRAY
2. for each vertex v in Adj(u) do
3. if v.color = WHITE then
4. DFS-Visit (G, v)
5. u.color = BLACK
6. finish_list.append(u)
**Depth First-Search**

**DFS-All (G)**
1. for each vertex u
2. u.color = WHITE
3. finish_list = empty_list
4. for each vertex u do
5. if u.color == WHITE then
6. DFS-Visit (G, u, finish_list)
7. return finish_list

**DFS-Visit (G, u)**
1. u.color = GRAY
2. for each vertex v in Adj(u) do
3. if v.color = WHITE then
4. DFS-Visit (G, v)
5. u.color = BLACK
6. finish_list.append(u)
**Depth First-Search**

**DFS-All (G)**
1. for each vertex u
2. u.color = WHITE
3. finish_list = empty_list
4. for each vertex u do
5. if u.color == WHITE then
6. DFS-Visit (G, u, finish_list)
7. return finish_list

**DFS-Visit (G, u)**
1. u.color = GRAY
2. for each vertex v in Adj(u) do
3. if v.color = WHITE then
4. DFS-Visit (G, v)
5. u.color = BLACK
6. finish_list.append(u)

Finish_list: 

- h

**Graph Example**

- DFS-Visit(G,h):
- DFS-Visit(G,f):
- DFS-Visit(G,d):
- DFS-Visit(G,a):
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

**DFS-All (G)**
1. for each vertex u
2. u.color = WHITE
3. finish_list = empty_list
4. for each vertex u do
5. if u.color == WHITE then
6. DFS-Visit (G, u, finish_list)
7. return finish_list

**DFS-Visit (G, u)**
1. u.color = GRAY
2. for each vertex v in Adj(u) do
3. if v.color = WHITE then
4. DFS-Visit (G, v)
5. u.color = BLACK
6. finish_list.append(u)
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

DFS-All (G)
1  for each vertex u
2    u.color = WHITE
3  finish_list = empty_list
4  for each vertex u do
5    if u.color == WHITE then
6      DFS-Visit (G, u, finish_list)
7  return finish_list

DFS-Visit (G, u)
1    u.color = GRAY
2    for each vertex v in Adj(u) do
3      if v.color = WHITE then
4        DFS-Visit (G, v)
5    u.color = BLACK
6    finish_list.append(u)

Finish_list:
- h
- g
- f

DFS-Visit(G,f):
DFS-Visit(G,d):
DFS-Visit(G,a):
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

**DFS-All (G)**
1. for each vertex u
2. u.color = WHITE
3. finish_list = empty_list
4. for each vertex u do
5. if u.color == WHITE then
6. DFS-Visit (G, u, finish_list)
7. return finish_list

**DFS-Visit (G, u)**
1. u.color = GRAY
2. for each vertex v in Adj(u) do
3. if v.color = WHITE then
4. DFS-Visit (G, v)
5. u.color = BLACK
6. finish_list.append(u)

Finish_list:
- h,
- g,
- f,
- d

DFS-Visit(G,a):
- u
- v
- w
- x
- y
- z

DFS-Visit(G,b):
- a
- b
- c
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color == WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)

Finish_list:
- h
- g
- f
- d

DFS-Visit(G,e):
- f
- d

DFS-Visit(G,c):
- h
- g
- f
- d

DFS-Visit(G,a):
- h
- g
- f
- d

DFS-Visit(G,b):
- b

DFS-Visit(G,a):
- h
- g
- f
- d
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)

Finish_list:
- h
- g
- f
- d
- e

DFS-Visit(G,e):
- d
- e
- f
- g
- h

DFS-Visit(G,c):
- b
- d
- e
- f
- g
- h

DFS-Visit(G,a):
- a
- b
- d
- e
- f
- g
- h

DFS-Visit(G,c):
- b
- d
- e
- f
- g
- h

DFS-Visit(G,a):
- a
- b
- d
- e
- f
- g
- h
Depth First-Search

DFS-All (G)
1 for each vertex u
2 \( u.\text{color} \) = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if \( u.\text{color} \) == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 \( u.\text{color} \) = GRAY
2 for each vertex v in Adj(u) do
3 if \( v.\text{color} \) = WHITE then
4 DFS-Visit (G, v)
5 \( u.\text{color} \) = BLACK
6 finish_list.append(u)

Finish_list:
- h
- g
- f
- d
- e
- c

DFS-Visit(G,c):
- d
- e
- c

DFS-Visit(G,a):
- h
- g
- f
- d
- e
- c
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)

Finish_list:
- h, g, f, d, e, c, a

DFS-Visit(G,a):
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)

Finish_list:
- h, g, f, d, e, c, a

May iterate through many complete vertices before finding b to launch a new search from
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

DFS-All (G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)
7 return finish_list

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
With Cycles in the graph

ANOTHER EXAMPLE
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)

DFS-Visit(G,a):

[Diagram of a graph showing vertices a, b, c, d, e, f, g, h, and edges connecting them]
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

Toposort(G)
1  for each vertex u
2    u.color = WHITE
3  finish_list = empty_list
4  for each vertex u do
5    if u.color == WHITE then
6      DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1    u.color = GRAY
2    for each vertex v in Adj(u) do
3      if v.color = WHITE then
4        DFS-Visit (G, v)
5    u.color = BLACK
6    finish_list.append(u)
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6   DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4   DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
**Depth First-Search**

Toposort(G)
1. for each vertex u
2. u.color = WHITE
3. finish_list = empty_list
4. for each vertex u do
5. if u.color == WHITE then
6. DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1. u.color = GRAY
2. for each vertex v in Adj(u) do
3. if v.color = WHITE then
4. DFS-Visit (G, v)
5. u.color = BLACK
6. finish_list.append(u)
Depth First-Search

Toposort(G)
1 for each vertex u
2 \( u.\text{color} = \text{WHITE} \)
3 finish_list = empty_list
4 for each vertex u do
5 if \( u.\text{color} == \text{WHITE} \) then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 \( u.\text{color} = \text{GRAY} \)
2 for each vertex v in Adj(u) do
3 if v.\text{color} = \text{WHITE} then
4 DFS-Visit (G, v)
5 \( u.\text{color} = \text{BLACK} \)
6 finish_list.append(u)
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)

DFSQ:

d

DFS-Visit(G,d):
DFS-Visit(G,f):
DFS-Visit(G,g):
DFS-Visit(G,h):
DFS-Visit(G,b):
DFS-Visit(G,c):
DFS-Visit(G,a):
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)

DFSQ:

<table>
<thead>
<tr>
<th>DFS-Visit(G,e):</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFS-Visit(G,f):</td>
</tr>
<tr>
<td>DFS-Visit(G,g):</td>
</tr>
<tr>
<td>DFS-Visit(G,h):</td>
</tr>
<tr>
<td>DFS-Visit(G,b):</td>
</tr>
<tr>
<td>DFS-Visit(G,c):</td>
</tr>
<tr>
<td>DFS-Visit(G,a):</td>
</tr>
</tbody>
</table>
Depth First-Search

Toposort(G)
1  for each vertex u
2    u.color = WHITE
3  finish_list = empty_list
4  for each vertex u do
5     if u.color == WHITE then
6        DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1    u.color = GRAY
2    for each vertex v in Adj(u) do
3        if v.color = WHITE then
4           DFS-Visit (G, v)
5    u.color = BLACK
6    finish_list.append(u)

DFSQ:

<table>
<thead>
<tr>
<th>DFS-Visit(G,e)</th>
<th>DFS-Visit(G,f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFS-Visit(G,g)</td>
<td>DFS-Visit(G,h)</td>
</tr>
<tr>
<td>DFS-Visit(G,b)</td>
<td>DFS-Visit(G,c)</td>
</tr>
<tr>
<td>DFS-Visit(G,a)</td>
<td></td>
</tr>
</tbody>
</table>
Depth First-Search

Toposort(G)
1 for each vertex u
2 u.color = WHITE
3 finish_list = empty_list
4 for each vertex u do
5 if u.color == WHITE then
6 DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1 u.color = GRAY
2 for each vertex v in Adj(u) do
3 if v.color = WHITE then
4 DFS-Visit (G, v)
5 u.color = BLACK
6 finish_list.append(u)
Depth First-Search

Toposort(G)
1  for each vertex u
2    u.color = WHITE
3  finish_list = empty_list
4  for each vertex u do
5    if u.color == WHITE then
6      DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1  u.color = GRAY
2  for each vertex v in Adj(u) do
3    if v.color = WHITE then
4      DFS-Visit (G, v)
5  u.color = BLACK
6  finish_list.append(u)

DFS-Visit(G,a):

DFSQ:
Toposort(G)
1  for each vertex u
2     u.color = WHITE
3  finish_list = empty_list
4  for each vertex u do
5      if u.color == WHITE then
6         DFS-Visit (G, u, finish_list)

DFS-Visit (G, u)
1     u.color = GRAY
2     for each vertex v in Adj(u) do
3        if v.color = WHITE then
4            DFS-Visit (G, v)
5     u.color = BLACK
6     finish_list.append(u)
ITERATIVE VERSION
Depth First-Search

DFS (G,s)
1    for each vertex u
2        u.color = WHITE
3    st = new Stack
4    st.push_back(s)
5    while st not empty
6        u = st.back()
7        if u.color == WHITE then
8            u.color = GRAY
9            foreach vertex v in Adj(u) do
10                if v.color == WHITE
11                   st.push_back(v)
12            else if u.color != WHITE
13                u.color = BLACK
14        st.pop_back()
Depth First-Search

DFS (G, s)
1 for each vertex u
2 u.color = WHITE
3 st = new Stack
4 st.push_back(s)
5 while st not empty
6 u = st.back()
7 if u.color == WHITE then
8    u.color = GRAY
9    foreach vertex v in Adj(u) do
10       if v.color == WHITE
11          st.push_back(v)
12     else if u.color != WHITE
13        u.color = BLACK
14    st.pop_back()
Depth First-Search

DFS (G,s)
1    for each vertex u
2      u.color = WHITE
3    st = new Stack
4    st.push_back(s)
5    while st not empty
6      u = st.back()
7      if u.color == WHITE then
8          u.color = GRAY
9      foreach vertex v in Adj(u) do
10         if v.color == WHITE
11            st.push_back(v)
12      else if u.color != WHITE
13          u.color = BLACK
14      st.pop_back()
Depth First-Search

DFS (G,s)
1 for each vertex u
2 u.color = WHITE
3 st = new Stack
4 st.push_back(s)
5 while st not empty
6 u = st.back()
7 if u.color == WHITE then
8 u.color = GRAY
9 foreach vertex v in Adj(u) do
10 if v.color == WHITE
11 st.push_back(v)
12 else if u.color != WHITE
13 u.color = BLACK
14 st.pop_back()
Depth First-Search

DFS (G,s)
1    for each vertex u
2      u.color = WHITE
3    st = new Stack
4    st.push_back(s)
5    while st not empty
6      u = st.back()
7      if u.color == WHITE then
8        u.color = GRAY
9      foreach vertex v in Adj(u) do
10         if v.color == WHITE
11            st.push_back(v)
12      else if u.color != WHITE
13          u.color = BLACK
14      st.pop_back()
Depth First-Search

DFS (G,s)
1   for each vertex u
2     u.color = WHITE
3   st = new Stack
4   st.push_back(s)
5   while st not empty
6     u = st.back()
7     if u.color == WHITE then
8       u.color = GRAY
9     foreach vertex v in Adj(u) do
10        if v.color == WHITE
11           st.push_back(v)
12     else if u.color != WHITE
13       u.color = BLACK
14   st.pop_back()
Depth First-Search

**DFS (G,s)**

1. for each vertex u
2. \( u.color = \text{WHITE} \)
3. \( st = \text{new Stack} \)
4. \( st.push\_back(s) \)
5. while st not empty
6. \( u = st\_\text{back}() \)
7. if \( u.color == \text{WHITE} \) then
8. \( u.color = \text{GRAY} \)
9. foreach vertex v in Adj(u) do
10. if v.color == \text{WHITE} then
11. \( st.push\_\text{back}(v) \)
12. else if u.color != \text{WHITE} then
13. \( u.color = \text{BLACK} \)
14. \( st.pop\_\text{back}() \)

**Diagram:**

- **Stack (st):** a c e c f d g c h b c
- **Graph:**
  - a connected to b, c, e
  - b connected to a, c, d, h
  - c connected to a, d, f, g, h
  - d connected to b, c, f, e
  - e connected to c, g
  - f connected to c, d, e, h
  - g connected to c, e
  - h connected to b, c, d, f
Depth First-Search

DFS (G,s)
1. for each vertex u
2. u.color = WHITE
3. st = new Stack
4. st.push_back(s)
5. while st not empty
6. u = st.back()
7. if u.color == WHITE then
8. u.color = GRAY
9. foreach vertex v in Adj(u) do
10. if v.color == WHITE
11. st.push_back(v)
12. else if u.color != WHITE
13. u.color = BLACK
14. st.pop_back()
Depth First-Search

DFS (G,s)
1    for each vertex u
2       u.color = WHITE
3    st = new Stack
4    st.push_back(s)
5    while st not empty
6       u = st.back()
7       if u.color == WHITE then
8          u.color = GRAY
9       foreach vertex v in Adj(u) do
10          if v.color == WHITE
11              st.push_back(v)
12       else if u.color != WHITE
13          u.color = BLACK
14       st.pop_back()
Depth First-Search

DFS (G,s)
1   for each vertex u
2     u.color = WHITE
3   st = new Stack
4   st.push_back(s)
5   while st not empty
6     u = st.back()
7     if u.color == WHITE then
8       u.color = GRAY
9       foreach vertex v in Adj(u) do
10          if v.color == WHITE
11             st.push_back(v)
12     else if u.color != WHITE
13       u.color = BLACK
14     st.pop_back()
Depth First-Search

DFS (G,s)
1  for each vertex u
2    u.color = WHITE
3  st = new Stack
4  st.push_back(s)
5  while st not empty
6    u = st.back()
7    if u.color == WHITE then
8      u.color = GRAY
9      foreach vertex v in Adj(u) do
10         if v.color == WHITE
11            st.push_back(v)
12      else if u.color != WHITE
13         u.color = BLACK
14    st.pop_back()
Depth First-Search

DFS (G,s)
1 for each vertex u
2 u.color = WHITE
3 st = new Stack
4 st.push_back(s)
5 while st not empty
6 u = st.back()
7 if u.color == WHITE then
8 u.color = GRAY
9 foreach vertex v in Adj(u) do
10 if v.color == WHITE
11 st.push_back(v)
12 else if u.color != WHITE
13 u.color = BLACK
14 st.pop_back()
Depth First-Search

DFS (G,s)
1   for each vertex u
2     u.color = WHITE
3   st = new Stack
4   st.push_back(s)
5   while st not empty
6     u = st.back()
7     if u.color == WHITE then
8           u.color = GRAY
9           foreach vertex v in Adj(u) do
10              if v.color == WHITE
11                st.push_back(v)
12           else if u.color != WHITE
13                u.color = BLACK
14           st.pop_back()
BFS vs. DFS Algorithm

• BFS and DFS are more similar than you think
  – Do we use a FIFO/Queue (BFS) or LIFO/Stack (DFS) to store vertices as we find them

BFS-Visit (G, start_node)
1 for each vertex u
2 u.color = WHITE
3 u.pred = nil
4 bfsq = new Queue
5 bfsq.push_back(start_node)
6 while bfsq not empty
7 u = bfsq.pop_front()
8 if u.color == WHITE
9   u.color = GRAY
10 foreach vertex v in Adj(u) do
11   bfsq.push_back(v)

DFS-Visit (G, start_node)
1 for each vertex u
2 u.color = WHITE
3 u.pred = nil
4 st = new Stack
5 st.push_back(start_node)
6 while st not empty
7 u = st.pop_back()
8 if u.color == WHITE
9   u.color = GRAY
10 foreach vertex v in Adj(u) do
11   st.push_back(v)
PAGERANK ALGORITHM
PageRank

• Consider the graph at the right
  – These could be webpages with links shown in the corresponding direction
  – These could be neighboring cities

• PageRank generally tries to answer the question:
  – If we let a bunch of people randomly "walk" the graph, what is the probability that they end up at a certain location (page, city, etc.) in the "steady-state"

• We could solve this problem through Monte-Carlo simulation (similar to CS 103 Coin-flipping game assignment)
  – Simulate a large number of random walkers and record where each one ends to build up an answer of the probabilities for each vertex

• But there are more efficient ways of doing it
PageRank

• Let us write out the adjacency matrix for this graph
• Now let us make a weighted version by normalizing based on the out-degree of each node
  – Ex. If you're at node B we have a 50-50 chance of going to A or E
• From this you could write a system of linear equations (i.e. what are the chances you end up at vertex I at the next time step, given you are at some vertex J now
  – \( p_A = 0.5p_B \)
  – \( p_B = p_C \)
  – \( p_C = p_A + p_D + 0.5p_E \)
  – \( p_D = 0.5p_E \)
  – \( p_E = 0.5p_B \)
  – We also know: \( p_A + p_B + p_C + p_D + p_E = 1 \)
PageRank

- System of Linear Equations
  - \( p_A = 0.5 \cdot p_B \)
  - \( p_B = p_C \)
  - \( p_C = p_A + p_D + 0.5 \cdot p_E \)
  - \( p_D = 0.5 \cdot p_E \)
  - \( p_E = 0.5 \cdot p_B \)
  - We also know: \( p_A + p_B + p_C + p_D + p_E = 1 \)

- If you know something about linear algebra, you know we can write these equations in matrix form as a linear system
  - \( Ax = y \)

Weighted Adjacency Matrix

\[
\begin{array}{ccccc}
0 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0.5 \\
0 & 0 & 0 & 0 & 0.5 \\
0 & 0.5 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{ccccc}
0 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0.5 \\
0 & 0 & 0 & 0 & 0.5 \\
0 & 0.5 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{ccccc}
pA & pB & pC & pD & pE \\
pA & pB & pC & pD & pE \\
pA & pB & pC & pD & pE \\
pA & pB & pC & pD & pE \\
pA & pB & pC & pD & pE \\
\end{array}
\]

\[
\begin{array}{ccccc}
pA & pB & pC & pD & pE \\
pA & pB & pC & pD & pE \\
pA & pB & pC & pD & pE \\
pA & pB & pC & pD & pE \\
pA & pB & pC & pD & pE \\
\end{array}
\]

\[
\begin{array}{ccccc}
pA & = pA = 0.5PB \\
pB & = pB \\
pC & = pA + pD + 0.5 \cdot pE \\
pD & = 0.5 \cdot pE \\
pE & = 0.5 \cdot pB \\
\end{array}
\]
PageRank

- But remember we want the steady state solution
  - The solution where the probabilities don't change from one step to the next
- So we want a solution to: $A p = p$
- We can:
  - Use a linear system solver (Gaussian elimination)
  - Or we can just seed the problem with some probabilities and then just iterate until the solution settles down

\[
\begin{array}{cccccc}
0 & 0.5 & 0 & 0 & 0 & \text{pA} \\
0 & 0 & 1 & 0 & 0 & \text{pB} \\
1 & 0 & 0 & 1 & 0.5 & \text{pC} \\
0 & 0 & 0 & 0 & 0.5 & \text{pD} \\
0 & 0.5 & 0 & 0 & 0 & \text{pE} \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{pA} & \text{pA} \\
\text{pB} & \text{pB} \\
\text{pC} & \text{pC} \\
\text{pD} & \text{pD} \\
\text{pE} & \text{pE} \\
\end{array}
\]

\[
A = \begin{bmatrix}
0 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0.5 \\
0 & 0 & 0 & 0 & 0.5 \\
0 & 0.5 & 0 & 0 & 0 \\
\end{bmatrix}
\]

\[
p = \frac{1}{\text{degree}} \begin{bmatrix}
\text{pA} \\
\text{pB} \\
\text{pC} \\
\text{pD} \\
\text{pE} \\
\end{bmatrix}
\]

**Weighted Adjacency Matrix**

[Divide by $(a_{i,j})$/degree(j)]
Iterative PageRank

- But remember we want the steady state solution
  - The solution where the probabilities don't change from one step to the next

- So we want a solution to: $Ap = p$

- We can:
  - Use a linear system solver (Gaussian elimination)
  - Or we can just seed the problem with some probabilities and then just iterate until the solution settles down

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 0.5 0 0 0</td>
<td>.2</td>
<td>.1</td>
<td>?</td>
</tr>
<tr>
<td>0</td>
<td>0 0 1 0 0</td>
<td>.2</td>
<td>.2</td>
<td>0 0 1 0 0</td>
</tr>
<tr>
<td>1</td>
<td>0 0 0 1 0.5</td>
<td>.2</td>
<td>.5</td>
<td>1 0 0 1 0.5</td>
</tr>
<tr>
<td>0</td>
<td>0 0 0 0 0.5</td>
<td>.2</td>
<td>.1</td>
<td>0 0 0 0 0.5</td>
</tr>
<tr>
<td>0</td>
<td>0.5 0 0 0</td>
<td>.2</td>
<td>.1</td>
<td>0 0.5 0 0 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 1 Sol.</th>
<th>Step 2 Sol.</th>
<th>Actual PageRank Solution from solving linear system:</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0.5 0 0 0</td>
<td>.1</td>
<td>.1538</td>
</tr>
<tr>
<td>0 0 1 0 0</td>
<td>.2</td>
<td>.3077</td>
</tr>
<tr>
<td>1 0 0 1 0.5</td>
<td>.5</td>
<td>.3077</td>
</tr>
<tr>
<td>0 0 0 0 0.5</td>
<td>.1</td>
<td>.0769</td>
</tr>
<tr>
<td>0 0.5 0 0 0</td>
<td>.1</td>
<td>.1538</td>
</tr>
</tbody>
</table>
Additional Notes

• What if we change the graph and now D has no incoming links...what is its PageRank?
  – 0

• Most PR algorithms add a probability that someone just enters that URL (i.e. enters the graph at that node)
  – Usually define something called the damping factor, \( \alpha \) (often chosen around 0.85)
  – Probability of randomly starting or jumping somewhere = 1-\( \alpha \)

• So at each time step the next PR value for node \( i \) is given as:

\[
\text{Pr}(i) = \frac{1-\alpha}{N} + \alpha \sum_{j \in \text{Pred}(i)} \frac{\text{Pr}(j)}{\text{OutDeg}(j)}
\]

  – \( N \) is the total number of vertices
  – Usually run 30 or so update steps
  – Start each \( \text{Pr}(i) = 1/N \)
In a Web Search Setting

- Given some search keywords we could find the pages that have that matching keywords
- We often expand that set of pages by including all successors and predecessors of those pages
  - Include all pages that are within a radius of 1 of the pages that actually have the keyword
- Now consider that set of pages and the subgraph that it induces
- Run PageRank on that subgraph
Dijkstra's Algorithm

SINGLE-SOURCE SHORTEST PATH (SSSP)
SSSP

• Let us associate a 'weight' with each edge
  – Could be physical distance, cost of using the link, etc.

• Find the shortest path from a source node, 'a' to all other nodes

<table>
<thead>
<tr>
<th>List of Vertices</th>
<th>Adjacency Lists</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>(c,13),(e,4)</td>
</tr>
<tr>
<td>b</td>
<td>(c,5),(h,6)</td>
</tr>
<tr>
<td>c</td>
<td>(a,13),(b,5),(d,2),(e,8),(g,7)</td>
</tr>
<tr>
<td>d</td>
<td>(c,2),(f,1)</td>
</tr>
<tr>
<td>e</td>
<td>(a,4),(c,8),(f,3)</td>
</tr>
<tr>
<td>f</td>
<td>(d,1),(e,3),(g,4)</td>
</tr>
<tr>
<td>g</td>
<td>(c,7),(f,4),(h,14)</td>
</tr>
<tr>
<td>h</td>
<td>(b,6),(g,14)</td>
</tr>
</tbody>
</table>
### SSSP

- What is the shortest distance from 'a' to all other vertices?
- How would you go about computing those distances?

#### Adjacency Lists

<table>
<thead>
<tr>
<th>Vertices</th>
<th>List of Adjacencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>(c,13), (e,4)</td>
</tr>
<tr>
<td>b</td>
<td>(c,5), (h,6)</td>
</tr>
<tr>
<td>c</td>
<td>(a,13), (b,5), (d,2), (e,8), (g,7)</td>
</tr>
<tr>
<td>d</td>
<td>(c,2), (f,1)</td>
</tr>
<tr>
<td>e</td>
<td>(a,4), (c,8), (f,3)</td>
</tr>
<tr>
<td>f</td>
<td>(d,1), (e,3), (g,4)</td>
</tr>
<tr>
<td>g</td>
<td>(c,7), (f,4), (h,14)</td>
</tr>
<tr>
<td>h</td>
<td>(b,6), (g,14)</td>
</tr>
</tbody>
</table>

#### List of Vertices

<table>
<thead>
<tr>
<th>Vertices</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
</tr>
<tr>
<td>b</td>
</tr>
<tr>
<td>c</td>
</tr>
<tr>
<td>d</td>
</tr>
<tr>
<td>e</td>
</tr>
<tr>
<td>f</td>
</tr>
<tr>
<td>g</td>
</tr>
<tr>
<td>h</td>
</tr>
</tbody>
</table>

#### Distances

<table>
<thead>
<tr>
<th>Vert</th>
<th>Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
</tr>
<tr>
<td>g</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
</tr>
</tbody>
</table>
**Dijkstra's Algorithm**

- Dijkstra's algorithm is similar to a BFS but pulls out the smallest distance vertex (from the source) rather than pulling vertices out in FIFO order (as in BFS).
- Maintain a data structure that you can identify shortly.
  - We'll show it as a table of all vertices with their currently 'known' distance from the source.
    - Initially, a has dist=0.
    - All others = infinite distance.
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
6. if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
8. v = min(); PQ.remove_min()
9. for u in neighbors(v)
10. w = weight(v,u)
11. if(v.dist + w < u.dist)
12. u.pred = v
13. u.dist = v.dist + w;
14. PQ.decreaseKey(u, u.dist)
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
6. if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
8. v = min(); PQ.remove_min()
9. for u in neighbors(v)
10. w = weight(v,u)
11. if(v.dist + w < u.dist)
12. u.pred = v
13. u.dist = v.dist + w;
14. PQ.decreaseKey(u, u.dist)
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
6. if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
8. v = min(); PQ.remove_min()
9. for u in neighbors(v)
10. w = weight(v,u)
11. if(v.dist + w < u.dist)
12. u.pred = v
13. u.dist = v.dist + w;
14. PQ.decreaseKey(u, u.dist)
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
6. if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
8. v = min(); PQ.remove_min()
9. for u in neighbors(v)
10. w = weight(v,u)
11. if(v.dist + w < u.dist)
12. u.pred = v
13. u.dist = v.dist + w;
14. PQ.decreaseKey(u, u.dist)
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
6. if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
8. v = min(); PQ.remove_min()
9. for u in neighbors(v)
10. w = weight(v,u)
11. if(v.dist + w < u.dist)
12. u.pred = v
13. u.dist = v.dist + w;
14. PQ.decreaseKey(u, u.dist)

List of Vertices

<table>
<thead>
<tr>
<th>Vert</th>
<th>Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td>inf</td>
</tr>
<tr>
<td>c</td>
<td>12</td>
</tr>
<tr>
<td>d</td>
<td>8</td>
</tr>
<tr>
<td>e</td>
<td>4</td>
</tr>
<tr>
<td>f</td>
<td>7</td>
</tr>
<tr>
<td>g</td>
<td>11</td>
</tr>
<tr>
<td>h</td>
<td>inf</td>
</tr>
</tbody>
</table>
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
6.   if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
8.   v = min(); PQ.remove_min()
9.   for u in neighbors(v)
10.  w = weight(v,u)
11.   if(v.dist + w < u.dist)
12.     u.pred = v
13.     u.dist = v.dist + w;
14.     PQ.decreaseKey(u, u.dist)

List of Vertices

<table>
<thead>
<tr>
<th>Vert</th>
<th>Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td>inf</td>
</tr>
<tr>
<td>c</td>
<td>10</td>
</tr>
<tr>
<td>d</td>
<td>8</td>
</tr>
<tr>
<td>e</td>
<td>4</td>
</tr>
<tr>
<td>f</td>
<td>7</td>
</tr>
<tr>
<td>g</td>
<td>11</td>
</tr>
<tr>
<td>h</td>
<td>inf</td>
</tr>
</tbody>
</table>
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
   6. if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
   8. v = min(); PQ.remove_min()
   9. for u in neighbors(v)
      10. w = weight(v, u)
      11. if(v.dist + w < u.dist)
      12. u.pred = v
      13. u.dist = v.dist + w;
      14. PQ.decreaseKey(u, u.dist)
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
   6. if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
   8. v = min(); PQ.remove_min()
9. for u in neighbors(v)
   10. w = weight(v,u)
   11. if(v.dist + w < u.dist)
   12. u.pred = v
   13. u.dist = v.dist + w;
   14. PQ.decreaseKey(u, u.dist)
Dijkstra's Algorithm

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
   6. if v != s then v.dist = inf; PQ.insert(v)
7. while PQ is not empty
   8. v = min(); PQ.remove_min()
9. for u in neighbors(v)
   10. w = weight(v,u)
   11. if(v.dist + w < u.dist)
   12. u.pred = v
   13. u.dist = v.dist + w;
   14. PQ.decreaseKey(u, u.dist)
Another Example

• Try another example of Dijkstra's algorithm.

<table>
<thead>
<tr>
<th>Vert</th>
<th>Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
</tr>
</tbody>
</table>
Analysis

• What is the loop invariant? What can I say about the vertex I pull out from the PQ?
  – It is guaranteed that there is no shorter path to that vertex
  – UNLESS: negative edge weights

• Could use induction to prove
  – When I pull the first node out (it is the start node) its weight has to be 0 and that is definitely the shortest path to itself
  – I then "relax" (i.e. decrease) the distance to neighbors it connects to and the next node I pull out would be the neighbor with the shortest distance from the start
    • Could there be shorter path to that node?
  – No, because any other path would use some other edge from the start which is already greater than the edge to this node
Dijkstra's Run-time Analysis

- What is the run-time of Dijkstra's algorithm?
- How many times do you execute the while loop on 8?
- How many total times do you execute the for loop on 10?

1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
6. if v != s then v.dist = inf;
7. PQ.insert(v)
8. while PQ is not empty
9. v = min(); PQ.remove_min()
10. for u in neighbors(v)
11. w = weight(v,u)
12. if(v.dist + w < u.dist)
13. u.pred = v
14. u.dist = v.dist + w;
15. PQ.decreaseKey(u, u.dist)
Dijkstra's Run-time Analysis

- What is the run-time of Dijkstra's algorithm?
- How many times do you execute the while loop on 8?
  - $V$ total times because once you pull a node out each iteration
  - That node is guaranteed to have found its shortest distance to the start
  - What does each call to remove_min() cost...
    - $\ldots \log(V)$ [at most $V$ items in PQ]
- How many total times do you execute the for loop on 10?
  - $E$ total times: Visit each vertex's neighbors
    - Each iteration may call decreaseKey() which is $\log(V)$
- Total runtime = $V\log(V) + E\log(V) = (V+E)\log(V)$
  - This is usually dominated by $E\log(V)$

```
1. SSSP(G, s)
2. PQ = empty PQ
3. s.dist = 0; s.pred = NULL
4. PQ.insert(s)
5. For all v in vertices
6.   if v != s then v.dist = inf;
7.   PQ.insert(v)
8. while PQ is not empty
9.   v = min(); PQ.remove_min()
10. for u in neighbors(v)
11.   w = weight(v,u)
12.   if(v.dist + w < u.dist)
13.     u.pred = v
14.     u.dist = v.dist + w;
15.    PQ.decreaseKey(u, u.dist)
```
Tangent on Heaps/PQs

- Suppose min-heaps
  - Though everything we're about to say is true for max heaps but for increasing values
- We know insert/remove is log(n) for a heap
- What if we want to decrease a value already in the heap...
  - Example: Decrease 26 to 9
  - Could we find 26 easily?
    - No requires a linear search through the array/heap => O(n)
  - Once we find it could we adjust it easily?
    - Yes, just promote it until it is in the right location => O(log n)
- So currently decrease-key() would cost O(n) + O(log n) = O(n)
- Can we do better?
Tangent on Heaps/PQs

- Can we provide a decrease-key() that runs in $O(\log n)$ and not $O(n)$
  - Remember we'd have to first find then promote
- We need to know where items sit in the heap
  - Essentially we want to quickly know the location given the key (i.e. Map key => location)
  - Unfortunately storing the heap as an array does just the opposite (maps location => key)
- What if we maintained an alternative map that did provide the reverse indexing
  - Then I could find where the key sits and then promote it
- If I keep that map as a balanced BST can I achieve $O(\log n)$
  - No! each promotion swap requires update your location and your parents
  - $O(\log n)$ swaps each requiring lookup(s) in the location map $[O(\log n)]$ yielding $O(\log^2(n))$
Tangent on Heaps/PQs

• Am I out of luck then?
• No, try a hash map
  – O(1) lookup
• Now each swap/promotion up the heap only costs O(1) and thus I have:
  – Find => O(1)
    • Using the hashmap
  – Promote => O(log n)
    • Bubble up at most log(n) levels with each level incurring O(1) updates of locations in the hashmap
• Decrease-key() is an important operation in the next algorithm we'll look at
A* Search Algorithm

ALGORITHM HIGHLIGHT
Search Methods

• Many systems require searching for goal states
  – Path Planning
    • Roomba Vacuum
    • Mapquest/Google Maps
    • Games!!
  – Optimization Problems
    • Find the optimal solution to a problem with many constraints
Search Applied to 8-Tile Game

- 8-Tile Puzzle
  - 3x3 grid with one blank space
  - With a series of moves, get the tiles in sequential order
  - Goal state:

HW6 Goal State

Goal State for these slides
Search Methods

- **Brute-Force Search**: When you don’t know where the answer is, just search all possibilities until you find it.

- **Heuristic Search**: A heuristic is a “rule of thumb”. An example is in a chess game, to decide which move to make, count the values of the pieces left for your opponent. Use that value to “score” the possible moves you can make.
  - Heuristics are not perfect measures, they are quick computations to give an approximation (e.g. may not take into account “delayed gratification” or “setting up an opponent”)


Brute Force Search

- Brute Force Search Tree
  - Generate all possible moves
  - Explore each move despite its proximity to the goal node
Heuristics

- Heuristics are “scores” of how close a state is to the goal (usually, lower = better)
- These scores must be easy to compute (i.e. simpler than solving the problem)
- Heuristics can usually be developed by simplifying the constraints on a problem
- Heuristics for 8-tile puzzle
  - # of tiles out of place
    - Simplified problem: If we could just pick a tile up and put it in its correct place
  - Total x-, y- distance of each tile from its correct location (Manhattan distance)
    - Simplified problem if tiles could stack on top of each other / hop over each other
Heuristic Search

• Heuristic Search Tree
  – Use total x-/y-distance (Manhattan distance) heuristic
  – Explore the lowest scored states
Caution About Heuristics

• Heuristics are just estimates and thus could be wrong

• Sometimes pursuing lowest heuristic score leads to a less-than-optimal solution or even no solution

• Solution
  – Take # of moves from start (depth) into account
A-star Algorithm

- Use a new metric to decide which state to explore/expand

- Define
  - $h =$ heuristic score (same as always)
  - $g =$ number of moves from start it took to get to current state
  - $f =$ $g + h$

- As we explore states and their successors, assign each state its $f$-score and always explore the state with lowest $f$-score

- Heuristics should always underestimate the distance to the goal
  - If they do, $A^*$ guarantees optimal solutions
A-Star Algorithm

• Maintain 2 lists
  – Open list = Nodes to be explored (chosen from)
  – Closed list = Nodes already explored (already chosen)

• Pseudocode

```python
open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list
     (if tie in f-values, select one w/ larger g-value)
  2. Add s to closed list
  3a. if s = goal node then trace path back to start; STOP!
  3b. Generate successors/neighbors of s, compute their f values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
```
Path-Planning w/ A* Algorithm

- Find optimal path from S to G using A*
  - Use heuristic of Manhattan \((x-/y-)\) distance

```
open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. \(f\)-value state from open_list (if tie in \(f\)-values, select one w/ larger \(g\)-value)
   2. Add s to closed list
      3a. if s = goal node then
          trace path back to start; STOP!
      3b. else
          Generate successors/neighbors of s, compute their \(f\)-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller \(f\) value
```

**If implementing this for a programming assignment, please see the slide at the end about alternate closed-list implementation**
Path-Planning w/ A* Algorithm

- Find optimal path from S to G using A*
  - Use heuristic of Manhattan (x-/y-) distance

```
open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/ larger g-value)
  2. Add s to closed list
  3a. if s = goal node then
      trace path back to start; STOP!
  3b. else
      Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value

**If implementing this for a programming assignment, please see the slide at the end about alternate closed-list implementation**
Path-Planning w/ A* Algorithm

• Find optimal path from S to G using A*
  – Use heuristic of Manhattan (x-/y-) distance

```python
open_list.push(Start State)
while(open_list is not empty)
    1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/ larger g-value)
    2. Add s to closed list
    3a. if s = goal node then
        trace path back to start; STOP!
    3b. else
        Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
```
Path-Planning w/ A* Algorithm

- Find optimal path from S to G using A*
  - Use heuristic of Manhattan (x-/y-) distance

```python
open_list.push(Start State)  
while(open_list is not empty)  
    1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/larger g-value)  
    2. Add s to closed list  
    3a. if s = goal node then trace path back to start; STOP!  
    3b. else Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
```
Path-Planning w/ A* Algorithm

• Find optimal path from S to G using A*
  – Use heuristic of Manhattan (x-/y-) distance

```
open_list.push(Start State)
while(open_list is not empty)
    1. s ← remove min. f-value state from
       open_list (if tie in f-values, select one w/
       larger g-value)
    2. Add s to closed list
    3a. if s = goal node then
        trace path back to start; STOP!
    3b. else
        Generate successors/neighbors of s,
        compute their f-values, and add them to
        open_list if they are not in the closed_list
        (so we don’t re-explore), or if they are
        already in the open list, update them if
        they have a smaller f value
```
Path-Planning w/ A* Algorithm

• Find optimal path from S to G using A*
  – Use heuristic of Manhattan (x-/y-) distance

open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/larger g-value)
  2. Add s to closed list
  3a. if s = goal node then
      trace path back to start; STOP!
  3b. else
      Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
Path-Planning w/ A* Algorithm

- Find optimal path from S to G using A*
  - Use heuristic of Manhattan (x-/y-) distance

```
open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/larger g-value)
  2. Add s to closed list
  3a. if s = goal node then
      trace path back to start; STOP!
  3b. else
      Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
```
Path-Planning w/ A* Algorithm

- Find optimal path from S to G using A*
  - Use heuristic of Manhattan (x-/y-) distance

open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/ larger g-value)
  2. Add s to closed list
  3a. if s = goal node then trace path back to start; STOP!
  3b. else
     Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
Path-Planning w/ A* Algorithm

• Find optimal path from S to G using A*
  – Use heuristic of Manhattan (x-/y-) distance

open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/ larger g-value)
  2. Add s to closed list
  3a. if s = goal node then
      trace path back to start; STOP!
  3b. else
      Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
Path-Planning w/ A* Algorithm

- Find optimal path from S to G using A*
  – Use heuristic of Manhattan (x-/y-) distance

open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/larger g-value)
  2. Add s to closed list
  3a. if s = goal node then trace path back to start; STOP!
  3b. else
     Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
Path-Planning w/ A* Algorithm

- Find optimal path from $S$ to $G$ using A*
  – Use heuristic of Manhattan ($x$-/y-) distance

```
open_list.push(Start State)
while(open_list is not empty)
  1. $s \leftarrow$ remove min. f-value state from open_list (if tie in f-values, select one w/ larger g-value)
  2. Add $s$ to closed list
  3a. if $s$ = goal node then
      trace path back to start; STOP!
  3b. else
      Generate successors/neighbors of $s$, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
```
Path-Planning w/ A* Algorithm

• Find optimal path from S to G using A*
  – Use heuristic of Manhattan (x-/y-) distance

open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/ larger g-value)
  2. Add s to closed list
  3a. if s = goal node then
    trace path back to start; STOP!
  3b. else
    Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
Path-Planning w/ A* Algorithm

• Find optimal path from S to G using A*
  – Use heuristic of Manhattan (x-/y-) distance

open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list (if tie in f-values, select one w/larger g-value)
  2. Add s to closed list
  3a. if s = goal node then
      trace path back to start; STOP!
  3b. else
      Generate successors/neighbors of s, compute their f-values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value.
Path-Planning w/ A* Algorithm

- Find optimal path from S to G using A*
  – Use heuristic of Manhattan (x-/y-) distance

open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from
     open_list (if tie in f-values, select one w/ larger g-value)
  2. Add s to closed list
  3a. if s = goal node then
      trace path back to start; STOP!
  3b. else
      Generate successors/neighbors of s,
      compute their f-values, and add them to
      open_list if they are not in the closed_list
      (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
A* and BFS

- BFS explores all nodes at a shorter distance from the start (i.e. g value)
A* and BFS

- BFS explores all nodes at a shorter distance from the start (i.e. g value)
A* and BFS

- BFS is A* using just the g value to choose which item to select and expand
A* Analysis

- What data structure should we use for the open-list?
- What data structure should we use for the closed-list?
- What is the run time?
- Run time is similar to Dijkstra's algorithm...
  - We pull out each node/state once from the open-list so that incurs N*O(remove-cost)
  - We then visit each successor which is like O(E) and perform an insert or decrease operation which is like E*max(O(insert), O(decrease))
  - E = Number of potential successors and this depends on the problem and the possible solution space
  - For the tile puzzle game, how many potential boards are there?

```
open_list.push(Start State)
while(open_list is not empty)
    1. s ← remove min. f-value state from open_list
       (if tie in f-values, select one w/ larger g-value)
    2. Add s to closed list
    3a. if s = goal node then trace path back to start; STOP!
    3b. Generate successors/neighbors of s, compute their f
        values, and add them to open_list if they are
        not in the closed_list (so we don’t re-explore), or
        if they are already in the open list, update them if
        they have a smaller f value
```
Implementation Note

• When the distance to a node/state/successor (i.e. g value) is uniform, we can greedily add a state to the closed-list at the same time as we add it to the open-list.

```
open_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list
     (if tie in f-values, select one w/ larger g-value)
  2. Add s to closed list
  3a. if s = goal node then trace path back to start; STOP!
  3b. Generate successors/neighbors of s, compute their f values, and add them to open_list if they are not in the closed_list (so we don’t re-explore), or if they are already in the open list, update them if they have a smaller f value
```

• Non-uniform g-values

```
open_list.push(Start State)
Closed_list.push(Start State)
while(open_list is not empty)
  1. s ← remove min. f-value state from open_list
     (if tie in f-values, select one w/ larger g-value)
  3a. if s = goal node then trace path back to start; STOP!
  3b. Generate successors/neighbors of s, compute their f values, and add them to open_list and closed_list if they are not in the closed_list
```

• Uniform g-values

The first occurrence of a board has to be on the shortest path to the solution.
BETWEENNESS CENTRALITY
BC Algorithm Overview

• What's the most central vertex(es) in the graph below?
• How do we define "centrality"?
• Betweenness centrality defines "centrality" as the nodes that are between the most other pairs

Sample Graph

Graph 1

Graph 2
BC Algorithm Overview

- Betweenness centrality (BC) defines "centrality" as the nodes that are between (i.e. on the path between) the most other pairs of vertices.
- BC considers betweenness on only "shortest" paths!
- To compute centrality score for each vertex we need to find shortest paths between all pairs...
  - Use the Breadth-First Search (BFS) algorithm to do this.

Sample Graph

Original 1

Are these gray nodes 'between' a and e?

No, a-c-d-e is the shortest path?

Original w/ added path
BC Algorithm Overview

• Betweenness-Centrality determines "centrality" as the number of shortest paths from all-pairs upon which a vertex lies

• Consider the sample graph below
  – Each external vertex (a, b, e, f) lies is a member of only the shortest paths between itself and each other vertex
  – Vertices c and d lie on greater number of shortest paths and thus will be scored higher

Sample Graph

Image each vertex is a ball and each edge is a chain or string. What would this graph look like if you picked it up by vertex c? Vertex a?
BC Implementation

- Based on Brandes' formulation for unweighted graphs
  - Perform $|V|$ Breadth-first traversals
  - Traversals result in a subgraph consisting of shortest paths from root to all other vertices
  - Messages are then sent back up the subgraph from "leaf" vertices to the root summing the percentage of shortest-paths each vertex is a member of
  - Summing a vertex's score from each traversal yields overall BC result

Sample Graph with final BC scores

Traversals from selected roots and resulting partial BC scores (in this case, the number of descendants)
BC Implementation

- As you work down, track # of shortest paths running through a vertex and its predecessor(s)
- On your way up, sum the nodes beneath

Traversals from selected roots and resulting partial BC scores (in this case, the number of descendants)

# of shortest paths thru the vertex, [List of predecessor]

Score on the way back up (if multiple shortest paths, split the score appropriately)