

Brett Bernstein and Sean McIntyre Lecture 13: Graphs

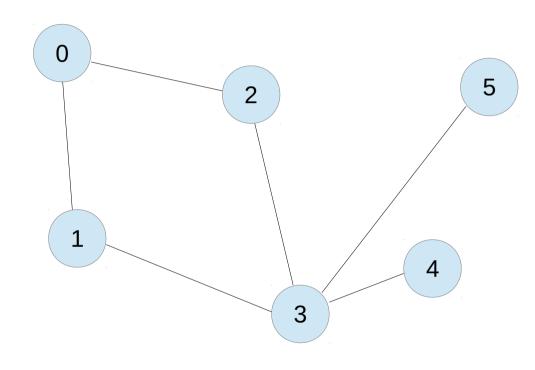
Graph Traversal Algorithms

- Many problems rely on traversing elements in a graph
 - e.g., UVa 469 Wetlands of Florida
 - You're given a 2D grid, each cell of which can be "water" or "land"
 - Cells adjacent on the major axes or diagonals are adjacent
 - For a given water (x, y) coordinate on the grid, determine the area of the connected water
 - These problems look hard if you're not familiar with graph traversals

Graph Traversal Algorithms

- Depth-first search
 - The first and most natural way to solve this problem is by visiting every node using recursion
 - As the name implies, visit the furthest nodes from the originating node
 - Perform backtracking





0: 1, 2

1: 1, 3

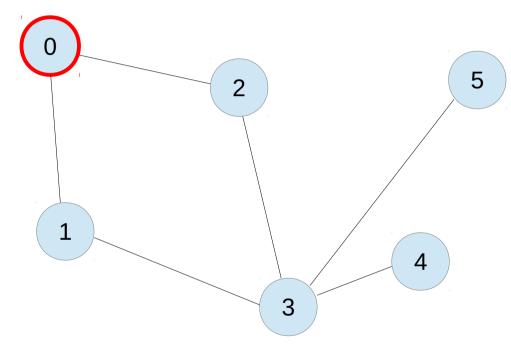
2: 0, 3

3: 1, 2, 4, 5

4: 3

5: 3





0: 1, 2

1: 1, 3

2: 0, 3

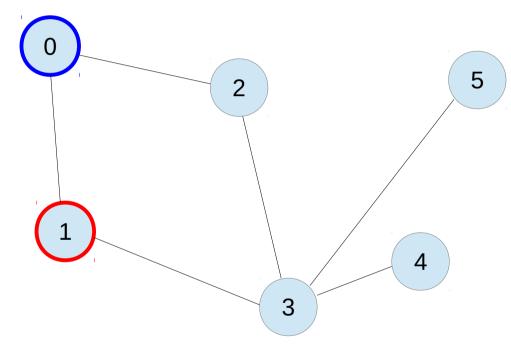
3: 1, 2, 4, 5

4: 3

5: 3

Stack dfs(0)





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

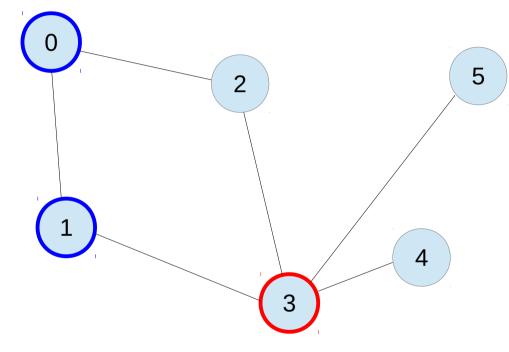
4: 3

5: 3

Stack

dfs(0) dfs(1)





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

4: 3

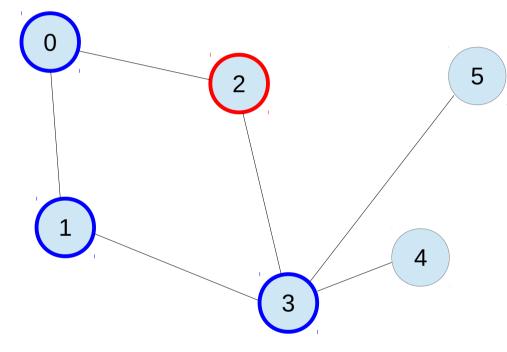
5: 3

Stack dfs(0)

dfs(1)

dfs(3)





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

4: 3

5: 3

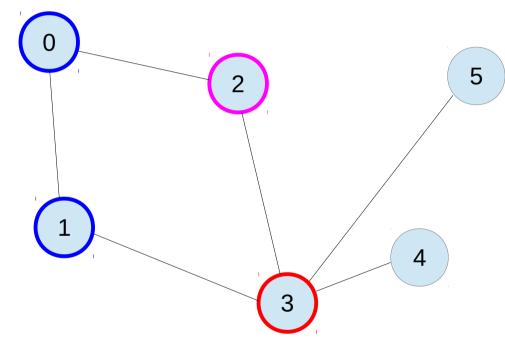
Stack

dfs(0) dfs(1)

dfs(3)

dfs(2)





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

4: 3

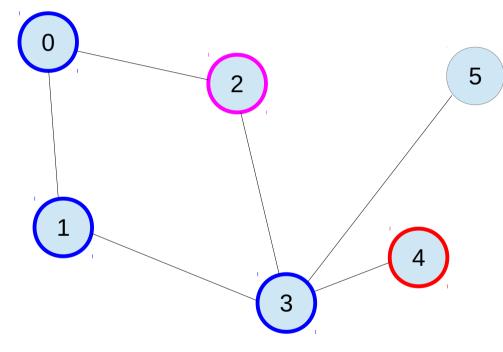
5: 3

Stack dfs(0)

dfs(1)

dfs(3)





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

4: 3

5: 3

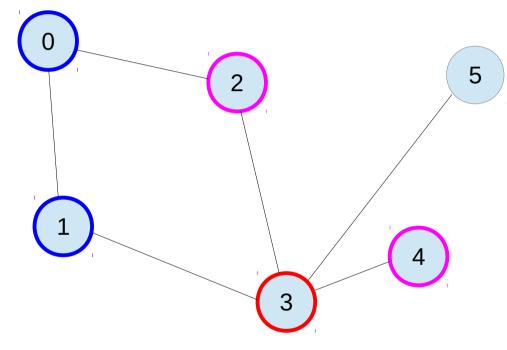
Stack dfs(0)

dfs(1)

dfs(3)

dfs(4)





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

4: 3

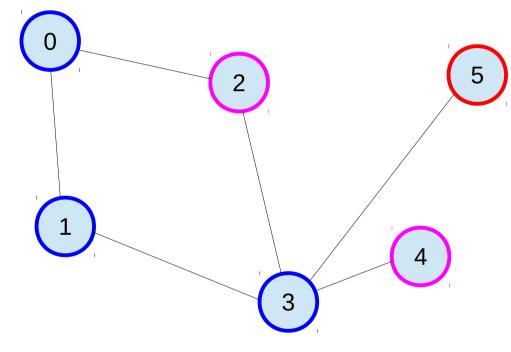
5: 3

Stack dfs(0)

dfs(1)

dfs(3)





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

4: 3

5: 3

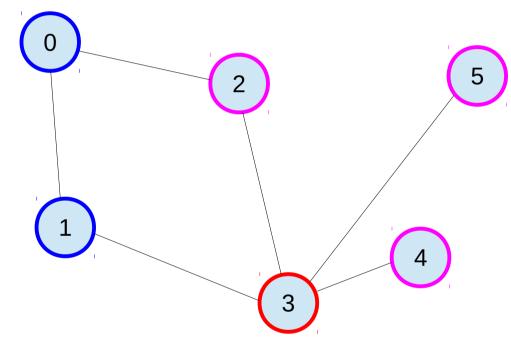
Stack dfs(0)

dfs(1)

dfs(3)

dfs(5)





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

4: 3

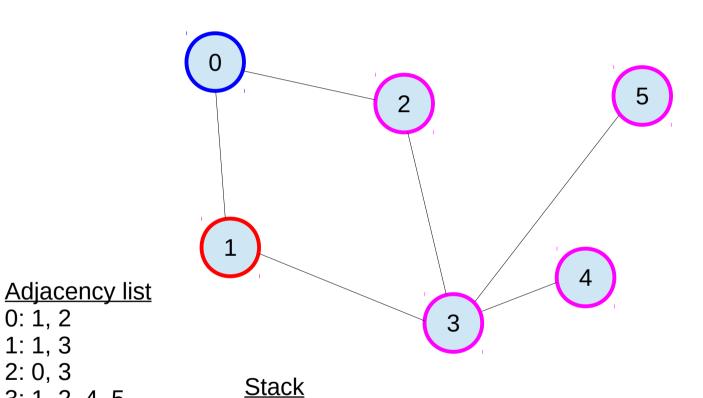
5: 3

Stack

dfs(0) dfs(1)

dfs(3)





dfs(0)

dfs(1)

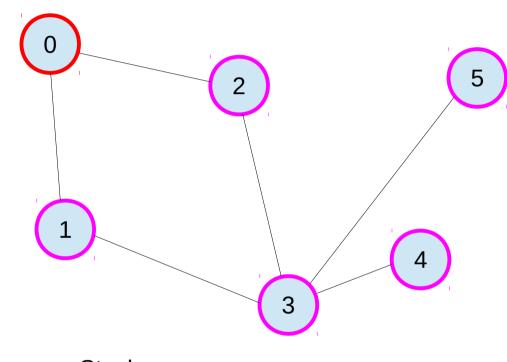
0: 1, 2 1: 1, 3 2: 0, 3

4: 3

5: 3

3: 1, 2, 4, 5





0: 1, 2

1: 1, 3

2: 0, 3

3: 1, 2, 4, 5

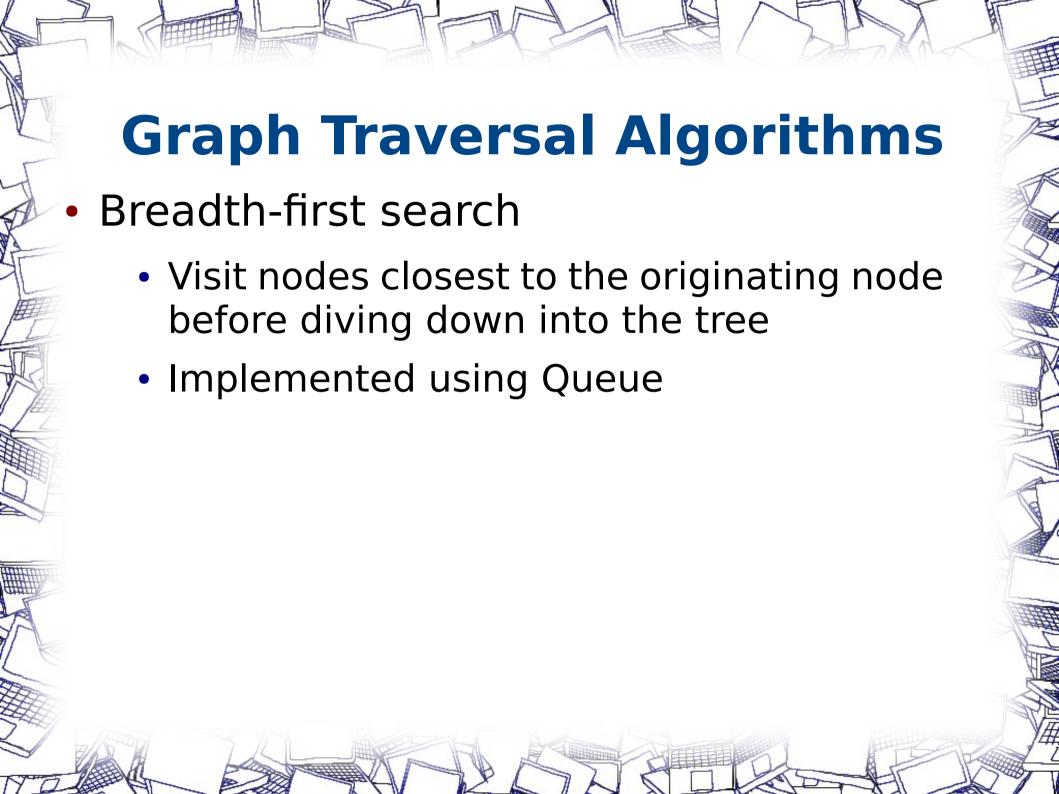
4: 3

5: 3

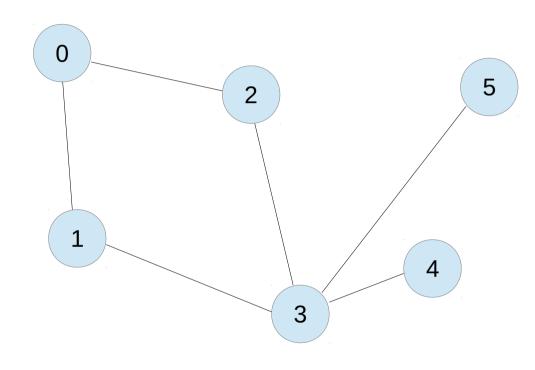
Stack dfs(0)

Graph Traversal Algorithms

```
ArrayList<ArrayList<Integer>> adjList; // prefilled with adjacents
int dfs(int node) { // returns # of nodes visited from node idx
   int res = 0;
   visited[node] = true; // mark this node as visited
   for (int i = 0; i < adjList.get(node).size(); i++) {</pre>
       int neighbor = adjList.get(node).get(i);
       if (!visited[neighbor]) {
           res += dfs(neighbor); // add number of dfs nodes visited
   return res+1; // the +1 refers to visiting the present node
public static void main(String[] args) {
   System.out.println(dfs(0));
```







0: 1, 2

1: 1, 3

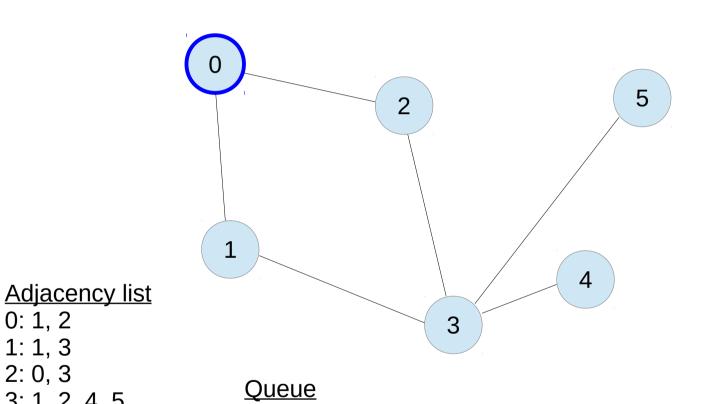
2: 0, 3

3: 1, 2, 4, 5

4: 3

5: 3





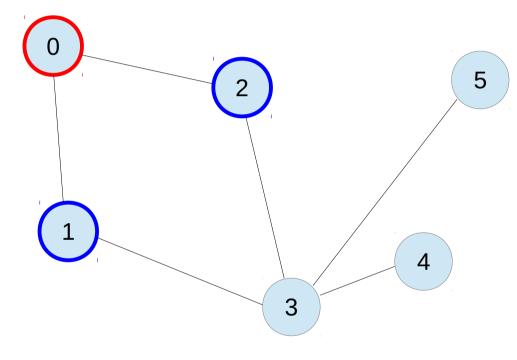
0: 1, 2

1: 1, 3 2: 0, 3

4: 3 5: 3

3: 1, 2, 4, 5





0: 1, 2

1: 1, 3

2: 0, 3

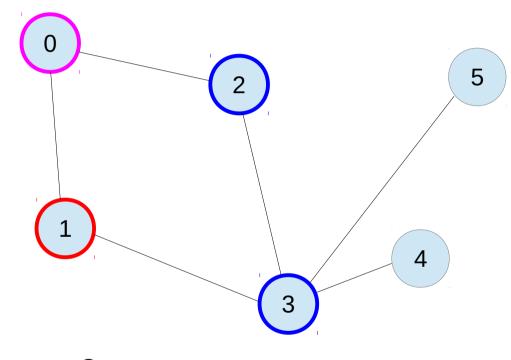
3: 1, 2, 4, 5

4: 3

5: 3

<u>Queue</u>





0: 1, 2

1: 1, 3

2: 0, 3

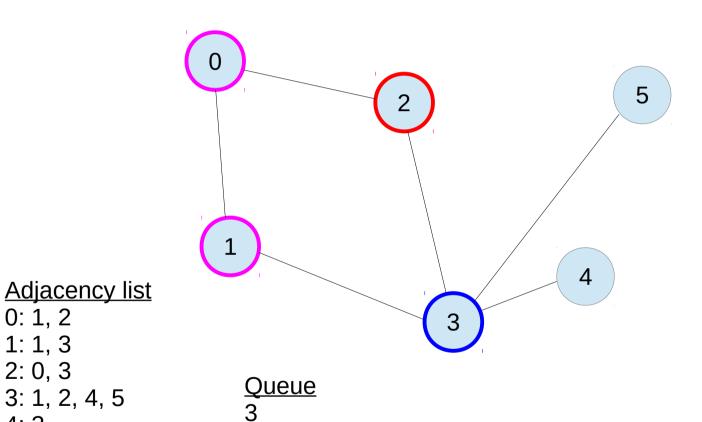
3: 1, 2, 4, 5

4: 3

5: 3

Queue 2





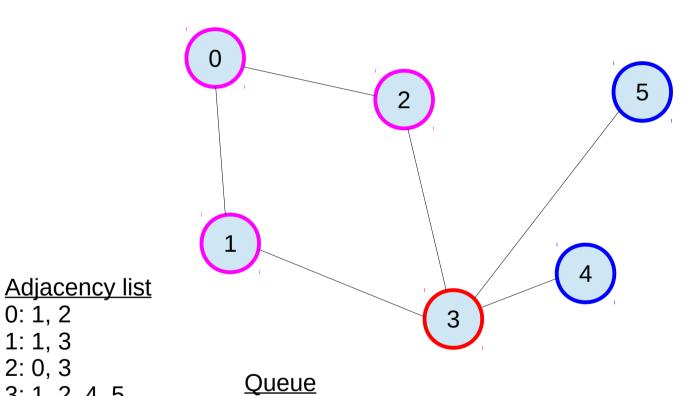
0: 1, 2

1: 1, 3 2: 0, 3

4: 3 5: 3

3: 1, 2, 4, 5





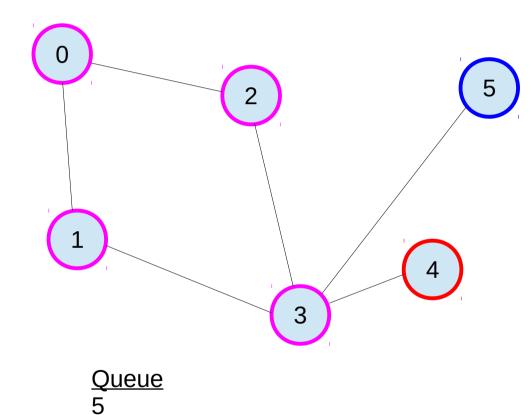
0: 1, 2

1: 1, 3 2: 0, 3

4: 3 5: 3

3: 1, 2, 4, 5





0: 1, 2

1: 1, 3

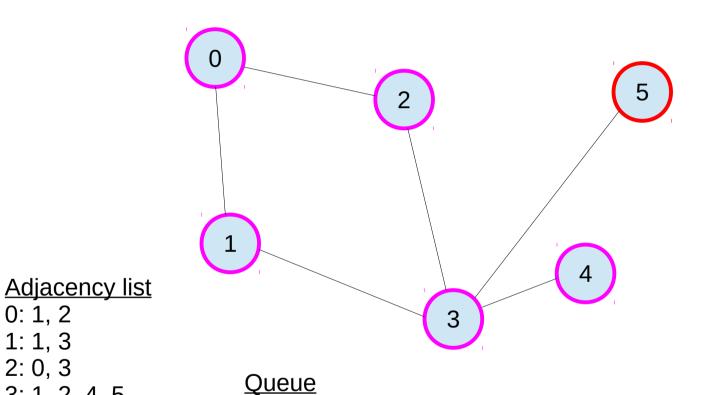
2: 0, 3

3: 1, 2, 4, 5

4: 3

5: 3





0: 1, 2

1: 1, 3 2: 0, 3

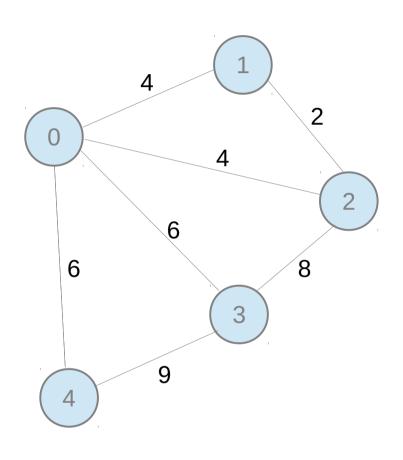
4: 3 5: 3

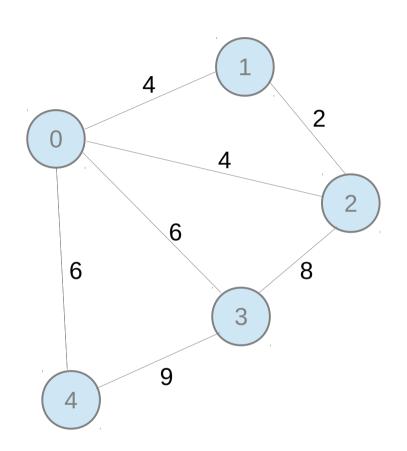
3: 1, 2, 4, 5

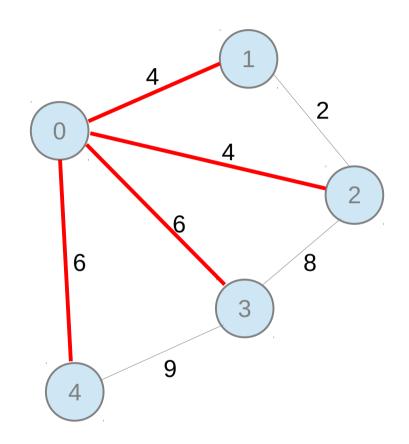
Graph Traversal Algorithms

```
int bfs(ArrayList<ArrayList<Integer>> adjList) {
   Queue<Integer> q = new LinkedList<Integer>();
    boolean visited[] = new boolean[N]; // keep track of visited nodes
   q.add(0); visited[0] = true; // add to traversal queue and mark
   int count = 1; // example: keep count of nodes traversed
   while (!q.isEmpty()) {
       int node = q.poll();
       for (int i = 0; i < adjList.get(node).size(); i++) {</pre>
            int neighbor = adjList.get(node).get(i);
            if (!visited[neighbor]) { // do not visit nodes twice
                q.add(neighbor); // add to traversal queue
               visited[neighbor] = true; // mark as visited
                count++; // visited a new node! Keep count
    return count;
```

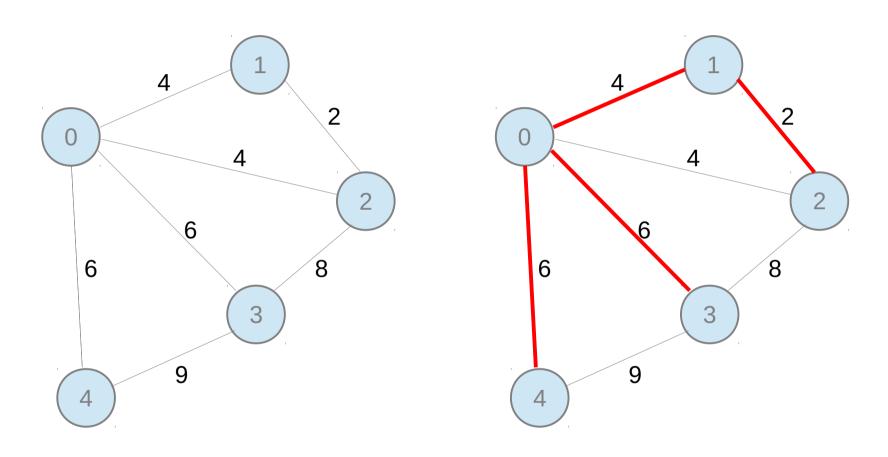
- Spanning tree
 - Given: a connected, undirected graph G = (V, E)
 - (*V* is the set of vertices, *E* is the set of edges)
 - A spanning tree is a set of edges that is a tree and "covers" all vertices V
 - There can be several trees
 - The spanning tree with the minimum cost (sum of edge weights) is called the Minimum Spanning Tree



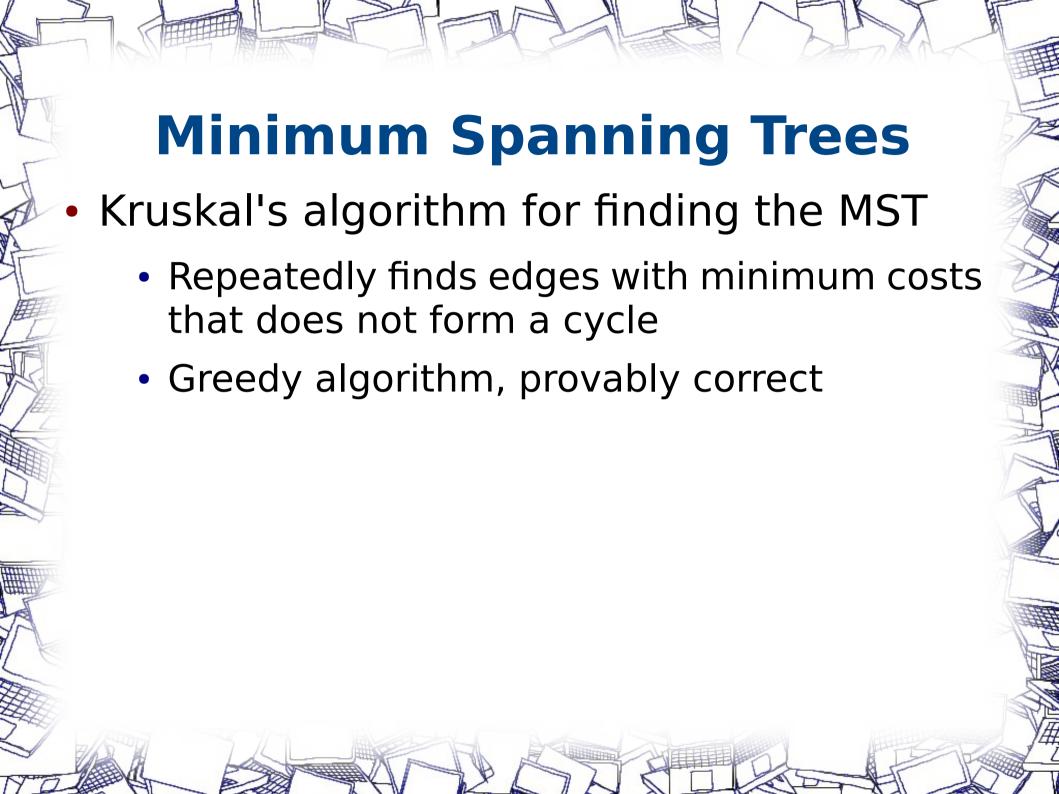


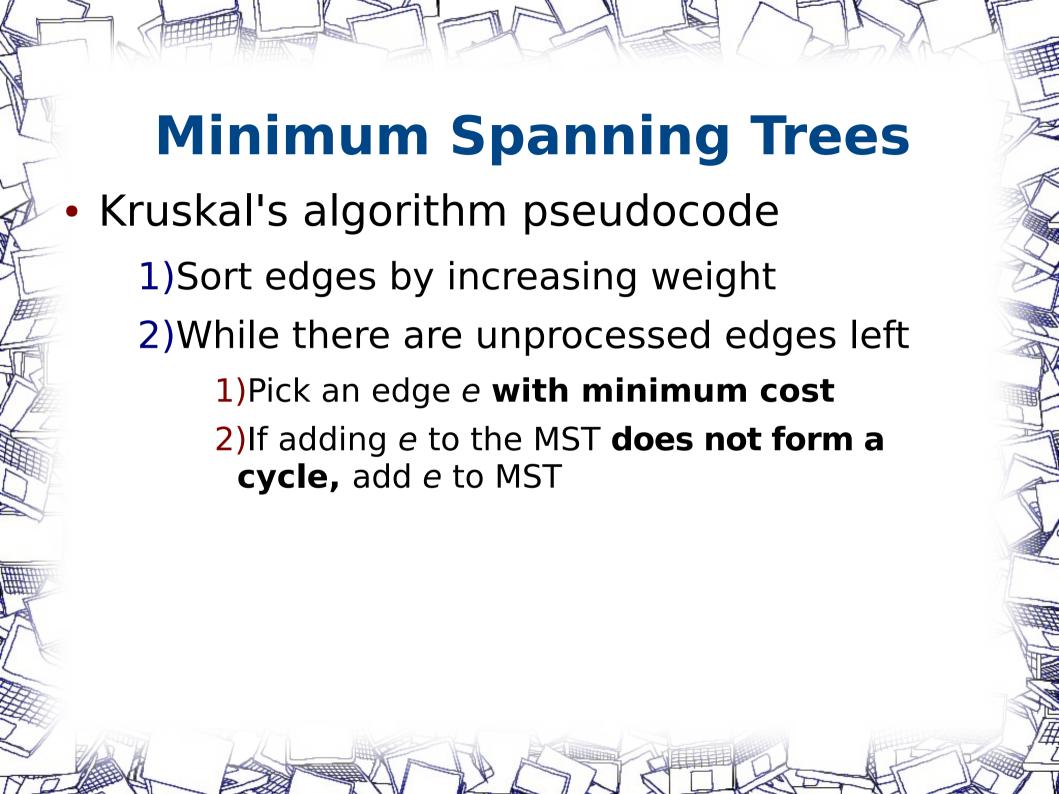


A spanning tree Cost: 4 + 4 + 6 + 6 = 20



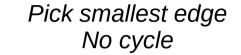
Minimum spanning tree Cost: 4 + 2 + 6 + 6 = 18

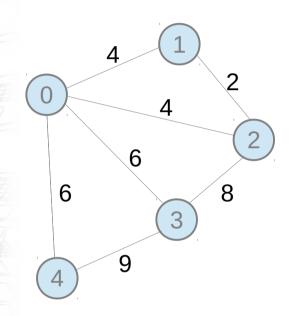


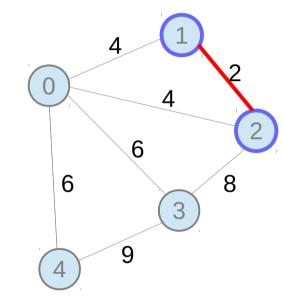


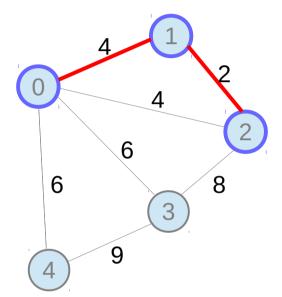
- Kruskal's algorithm pseudocode
 - How to store and sort edges?
 - Using an edge list and Collections.sort
 - How to test for cycles?
 - using disjoint sets and union-find
 - Runtime?
 - Sort: $O(|E| \log |E|)$; Processing: O(|E|)
 - Total: O(|E| log |E|) = O(|E| log |V|)







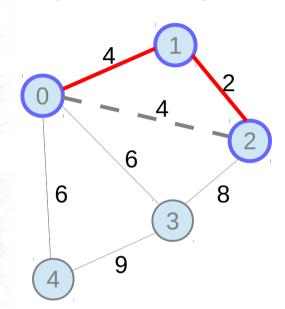




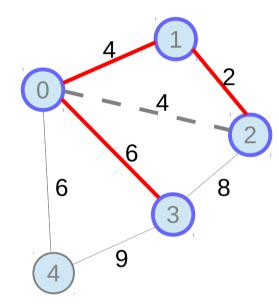
Weighted adjacency list by (index, weight)

- 0: (1, 4), (2, 4), (3, 6), (4, 6)
- 1: (0, 4), (2, 2)
- 2: (0, 4), (1, 2), (3, 8)
- 3: (0, 6), (2, 8), (4, 9)
- 4: (0, 6), (3, 9)

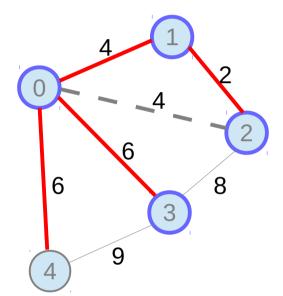
Pick smallest edge Cycle formed, ignore



Pick smallest edge No cycle



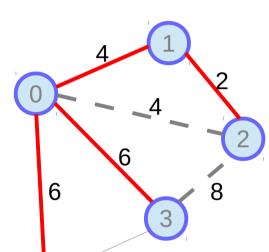
Pick smallest edge No cycle



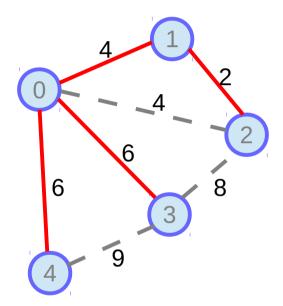
Weighted adjacency list by (index, weight)

- 0: (1, 4), (2, 4), (3, 6), (4, 6)
- 1: (0, 4), (2, 2)
- 2: (0, 4), (1, 2), (3, 8)
- 3: (0, 6), (2, 8), (4, 9)
- 4: (0, 6), (3, 9)

Algorithm not done! The edge list hasn't yet been exhausted Pick smallest edge Cycle formed, ignore

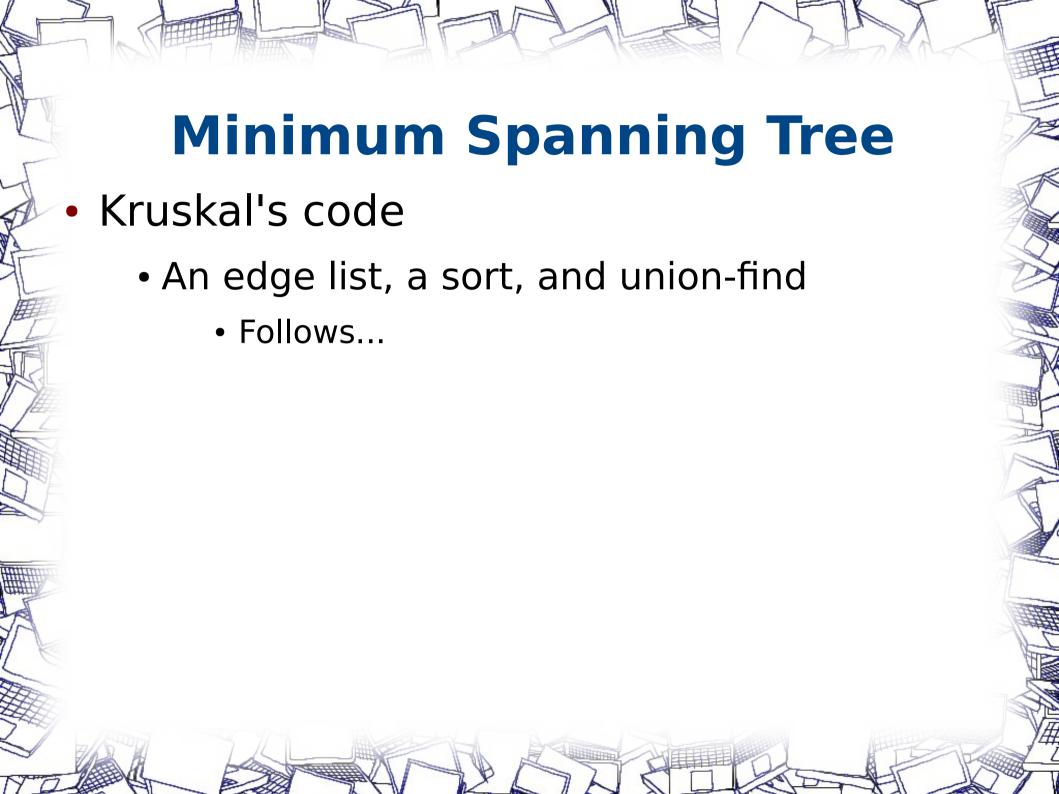


Pick smallest edge Cycle formed, ignore



Weighted adjacency list by (index, weight)

- 0: (1, 4), (2, 4), (3, 6), (4, 6)
- 1: (0, 4), (2, 2)
- 2: (0, 4), (1, 2), (3, 8)
- 3: (0, 6), (2, 8), (4, 9)
- 4: (0, 6), (3, 9)



Minimum Spanning Tree

```
class Edge implements Comparable < Edge > {
   int A, B, w;

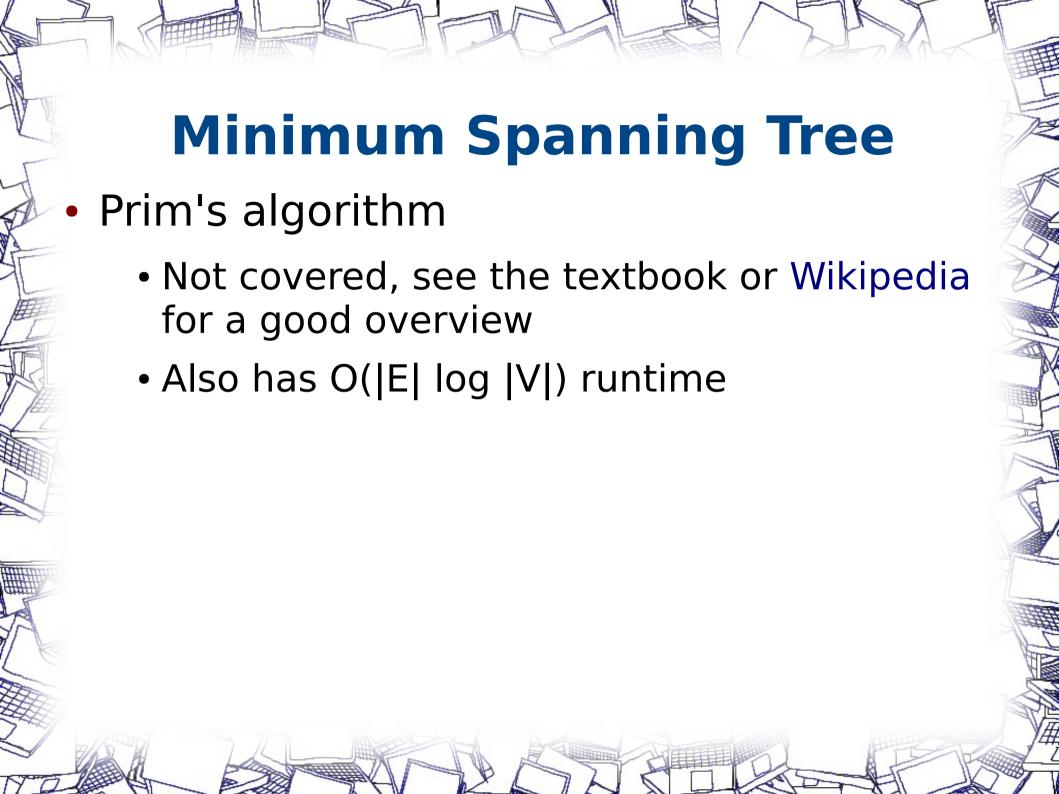
public Edge(int A, int B, int w) {
     this.A = Math.min(A, B);
     this.B = Math.max(A, B);
     this.w = w;
}

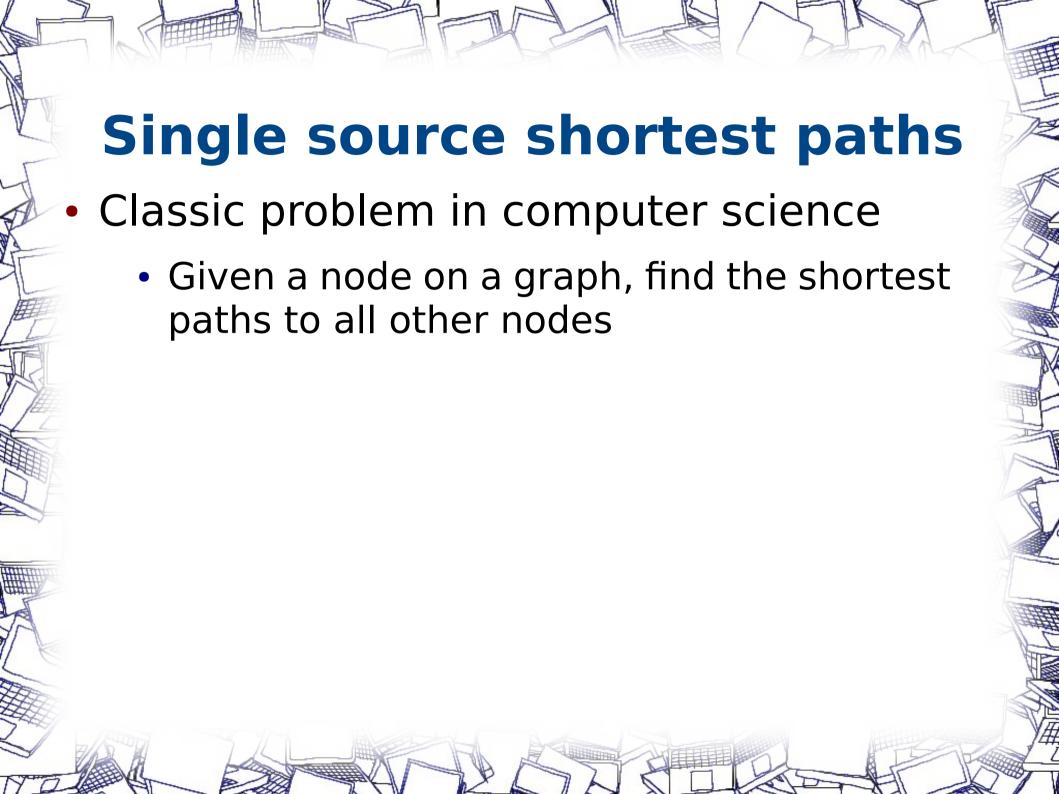
public int compareTo(Edge e) {
   if (w != e.w) {
     return w < e.w ? -1 : 1;
   } else {
     return 0;
   }
}</pre>
```

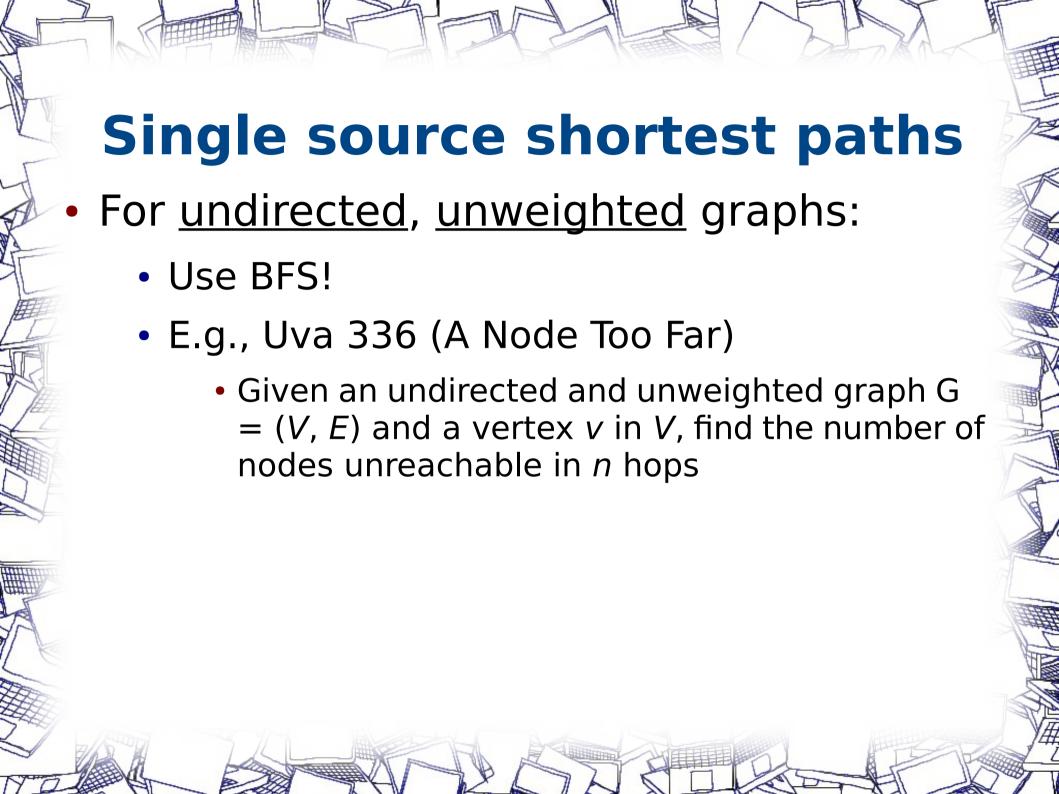
```
class UnionFind {
    int uf[];
    public UnionFind(int size) {
         uf = new int[size];
         for (int i = 0; i < size; i++) uf[i] = i;
    public boolean isSameSet(int A, int B) {
         return find(A) == find(B);
    public void union(int A, int B) {
         uf[find(A)] = find(B);
    public int find(int A) {
         int res = uf[A];
         while (uf[res] != res) res = uf[res];
         return uf[A] = res;
```

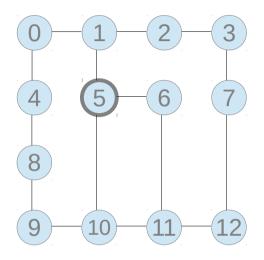
Minimum Spanning Tree

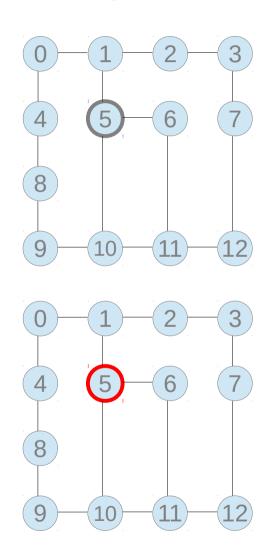
```
ArrayList<Edge> edgeList = parseEdgeList();
Collections.sort(edgeList);
int mstCost = 0;
UnionFind uf = new UnionFind(nVertices);
for (Edge e : edgeList) { // for each edge
   if (!uf.isSameSet(e.A, e.B)) { // if no cycle
      mstCost += e.w; // add it
      uf.union(e.A, e.B);
System.out.println(mstCost);
```





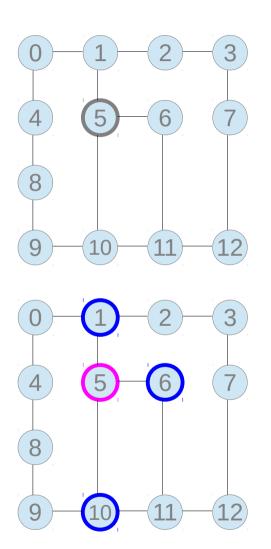




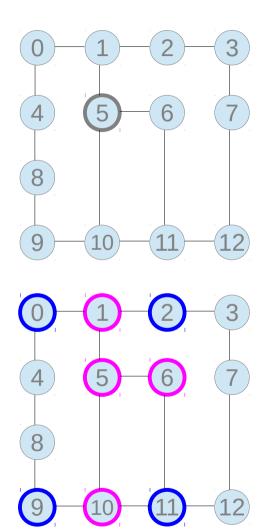


Queue 5

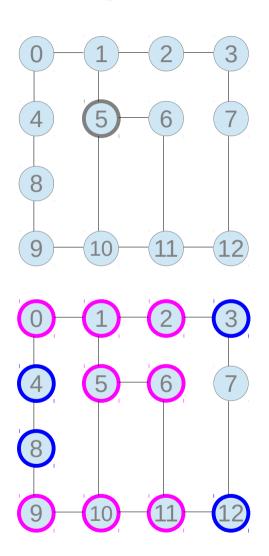
 $\frac{\text{Distances}}{D[5] = 0}$



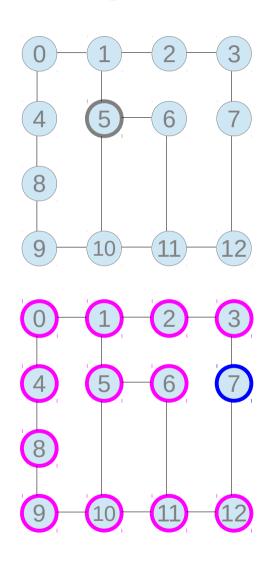
QueueDistances1D[5] = 06D[1] = D[5] + 1 = 110D[6] = D[5] + 1 = 1D[10] = D[5] + 1 = 1



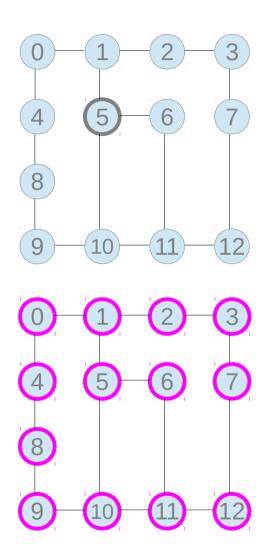
| <u>Queue</u> | <u>Distances</u> |
|--------------|-----------------------|
| 0 | D[5] = 0 |
| 2 | D[1] = 1 |
| 9 | D[6] = 1 |
| 11 | D[10] = 1 |
| | D[0] = D[1] + 1 = 2 |
| | D[2] = D[1] + 1 = 2 |
| | D[9] = D[10] + 1 = 2 |
| | D[11] = D[10] + 1 = 2 |



<u>Queue</u> Distances D[5] = 0D[1] = 18 D[6] = 112 D[10] = 1D[0] = 2D[2] = 2D[9] = 2D[11] = 2D[3] = D[2] + 1 = 3D[4] = D[0] + 1 = 3D[8] = D[9] + 1 = 3D[12] = D[11] + 1 = 3



Queue Distances7 D[5] = 0 D[1] = 1 D[6] = 1 D[10] = 1 D[0] = 2 D[2] = 2 D[9] = 2 D[11] = 2 D[3] = 3 D[4] = 3 D[8] = 3 D[12] = 3 D[7] = D[3] + 1 = 4



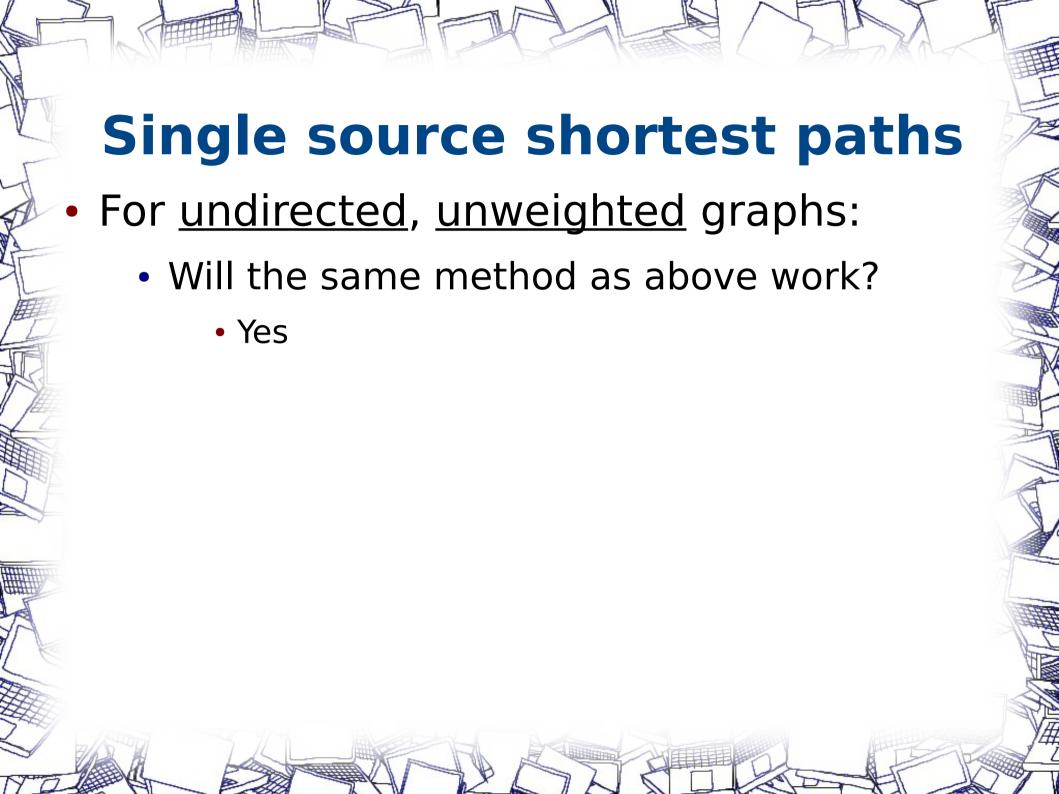
<u>Queue</u>

Distances
D[5] = 0
D[1] = 1
D[6] = 1
D[10] = 1
D[0] = 2
D[2] = 2
D[9] = 2
D[11] = 2
D[3] = 3
D[4] = 3
D[8] = 3

D[12] = 3

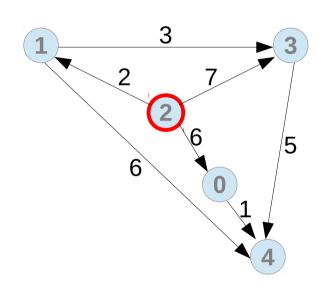
D[7] = 4

Answer: 1





- For <u>directed</u>, <u>weighted</u> graphs:
 - Without negative weights
 - Use Dijkstra's! O((|V| + |E|) log |V|)
 - This can be done using a priority queue
 - Works kind of like a greedy, modified BFS
 - Shown by example...



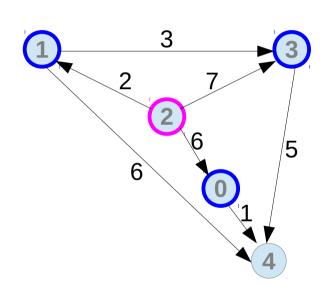
Question: Shortest paths from 2 to all other nodes?

Priority Queue (distance, node index)
{ (0, 2) }

| i | 0 | 1 | 2 | 3 | 4 |
|------|-----|-----|---|-----|-----|
| d[i] | INF | INF | 0 | INF | INF |

Distance table

Start from node 2



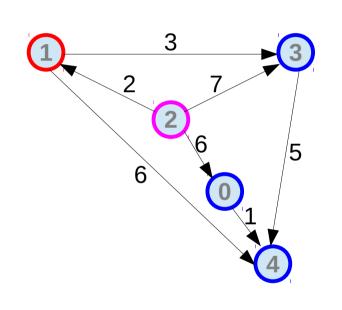
Question: Shortest paths from 2 to all other nodes?

Priority Queue (distance, node index) { (0, 2) } { (2, 1), (6, 0), (7, 3) }

| i | 0 | 1 | 2 | 3 | 4 |
|------|---|---|---|---|-----|
| d[i] | 6 | 2 | 0 | 7 | INF |

Distance table

Add all unvisited nodes from node 2 to the priority queue. The PQ sorts the distances so the "next closest" node floats to the top. Right now the closest node is 1, followed by 0, then 3.



Question: Shortest paths from 2 to all other nodes?

Priority Queue (distance, node index) { (0, 2) } { (2, 1), (6, 0), (7, 3) } { (5, 3), (6, 0), (7, 3), (8, 4) }

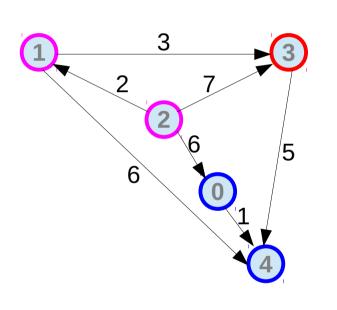
| i | 0 | 1 | 2 | 3 | 4 |
|------|---|---|---|---|---|
| d[i] | 6 | 2 | 0 | 5 | 8 |

Distance table

Poll from the PQ to get node 1.

Add all neighboring nodes to node 1 that haven't been polled yet.

BUT be sure to add all nodes that may already be in the queue with longer distances – there may be a shorter way to reach them



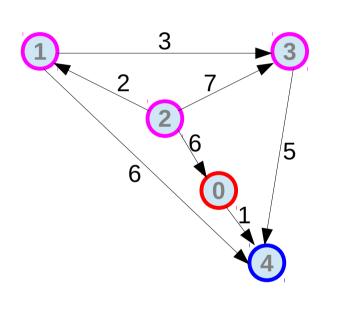
Question: Shortest paths from 2 to all other nodes?

Priority Queue (distance, node index) { (0, 2) } { (2, 1), (6, 0), (7, 3) } { (5, 3), (6, 0), (7, 3), (8, 4) } { (6, 0), (7, 3), (8, 4) }

| i | 0 | 1 | 2 | 3 | 4 |
|------|---|---|---|---|---|
| d[i] | 6 | 2 | 0 | 5 | 8 |

Distance table

Poll from the PQ to get node 3. Since we know there is a faster way to get node 4, don't bother adding node 4 to PQ



Question: Shortest paths from 2 to all other nodes?

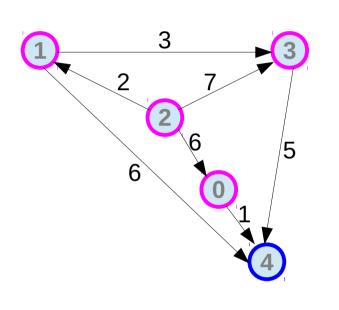
Priority Queue (distance, node index) { (0, 2) } { (2, 1), (6, 0), (7, 3) } { (5, 3), (6, 0), (7, 3), (8, 4) } { (6, 0), (7, 3), (8, 4) }

| i | 0 | 1 | 2 | 3 | 4 |
|------|---|---|---|---|---|
| d[i] | 6 | 2 | 0 | 5 | 7 |

 $\{ (7, 3), (7, 4), (8, 4) \}$

Distance table

Poll from the PQ to get node 0. Since we know there is a faster way to get node 4, don't bother adding node 4 to PQ



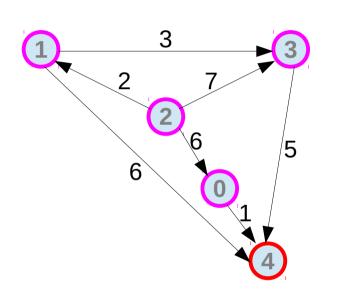
Question: Shortest paths from 2 to all other nodes?

Priority Queue (distance, node index) { (0, 2) } { (2, 1), (6, 0), (7, 3) } { (5, 3), (6, 0), (7, 3), (8, 4) } { (6, 0), (7, 3), (8, 4) } { (7, 3), (7, 4), (8, 4) } { (7, 4), (8, 4) }

| i | 0 | 1 | 2 | 3 | 4 |
|------|---|---|---|---|---|
| d[i] | 6 | 2 | 0 | 5 | 7 |

Distance table

Now the (7, 3) state is ignored because it's been determined that 7 is a longer path than another existing path to node 3



Question: Shortest paths from 2 to all other nodes?

Priority Queue (distance, node index) { (0, 2) } { (2, 1), (6, 0), (7, 3) }

 $\{\frac{(5, 3)}{(6, 0)}, (6, 0), (7, 3), (8, 4)\}$ $\{\frac{(6, 0)}{(7, 3)}, (7, 3), (8, 4)\}$ $\{\frac{(7, 3)}{(7, 4)}, (8, 4)\}$

{ (7, 4), (8, 4) } { (8, 4) }

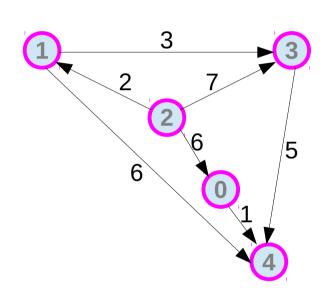
| i | 0 | 1 | 2 | 3 | 4 |
|------|---|---|---|---|---|
| d[i] | 6 | 2 | 0 | 5 | 7 |

Distance table

Nowhere to go, so nothing is added to the PQ

 $\{ (7, 4), (8, 4) \}$

{ (8, 4) }



Question: Shortest paths from 2 to all other nodes?

```
{ (0, 2) }
{ (2, 1), (6, 0), (7, 3) }
{ (5, 3), (6, 0), (7, 3), (8, 4) }
{ (6, 0), (7, 3), (8, 4) }
{ (7, 3), (7, 4), (8, 4) }
```

Priority Queue (distance, node index)

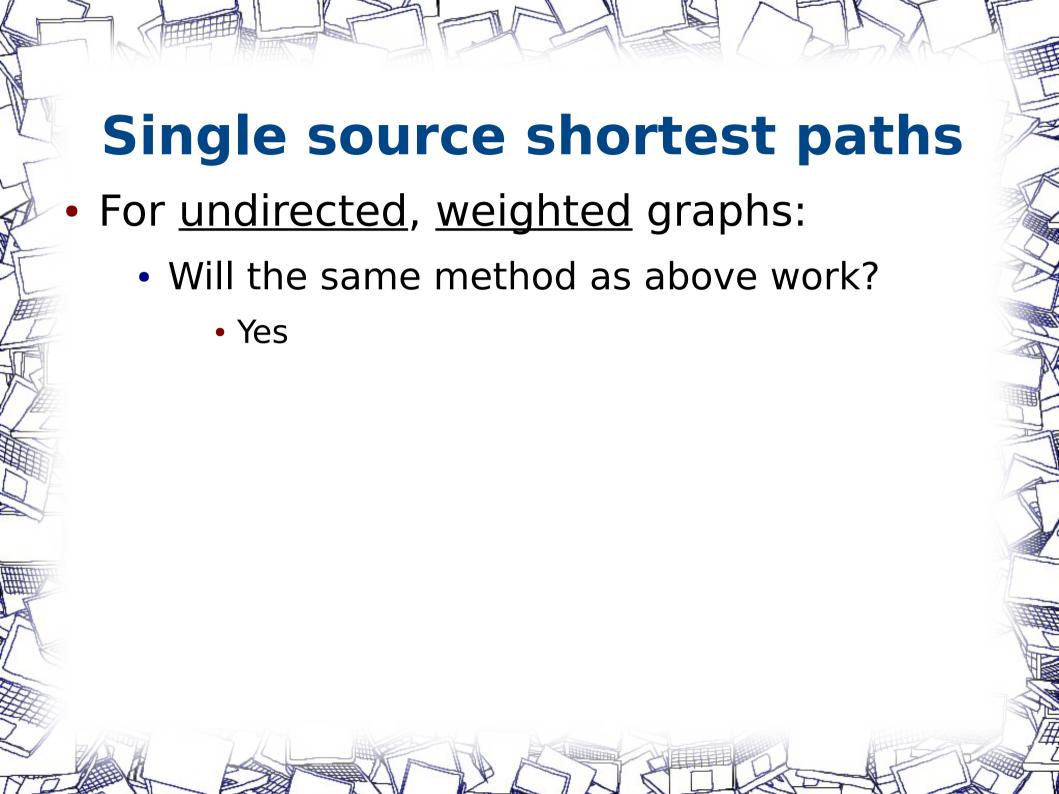
| i | 0 | 1 | 2 | 3 | 4 | |
|------|---|---|---|---|---|--|
| lilb | 6 | 2 | 0 | 5 | 7 | |

Distance table

State (8, 4) is ignored because 8 > 7

Dijkstra's in code

```
public static void main(String[] args) {
    LinkedHashMap<Integer, LinkedHashMap<Integer, Integer>> adj;
    // State is a pair (dist, index)
    PriorityQueue<State> pq = new PriorityQueue<State>();
    int dist[] = new int[V]; Arrays.fill(dist, 1 << 20); // INF</pre>
    pq.add(new State(2, 0)); // Initial state
    while (!pq.isEmpty()) {
        State s = pq.poll();
        if (s.dist == dist[s.index]) { // true if has not been updated
            LinkedHashMap<Integer, Integer> nbors = adj.get(s.index);
             for (Map.Entry<Integer, Integer> e : nbors.getEntrySet()) {
                 int nbor = e.getKey();
                 int nborDist = e.getValue();
                 if (nborDist + dist[s.index] < dist[nbor]) {</pre>
                     // have found a closer path
                     dist[nbor] = nborDist + dist[s.index];
                     pq.add(new State(nbor, nborDist + dist[s.index]));
```





- For <u>directed</u>, <u>weighted</u> graphs with negative weights:
 - Dijkstra's algorithm does not work and will get stuck in an infinite loop if there is a cycle, always finding a better path
 - Instead, use Bellman-Ford algorithm:
 - Repeat the "relaxing" part of Dijkstra's |V|-1 times, regardless of how close the nodes are

- Why Bellman-Ford works:
 - Performing relaxing on the graph |V|-1 times guarantees shortest paths are found
 - Proof omitted
 - If relaxing can happen after |V|-1 loops, then a negative cycle exists
 - And it runs in O(|V| * |E|) time with an adjacency list, much greater than Dijkstra's

All pairs shortest paths

- What happens if you want to find the shortest distance between all pairs of nodes?
 - On a weighted, connected graph, use Floyd Warshall algorithm
 - Implement in ~4 lines of code
 - O(V³) instead of N Dijkstra's algorithm, which would be O(V³ log V)

Floyd Warshall in code

```
// inside int main()
// precondition: m[i][j] contains the weight of edge (i, j)
// or INF if there is no such edge
// (m is an adjacency matrix)

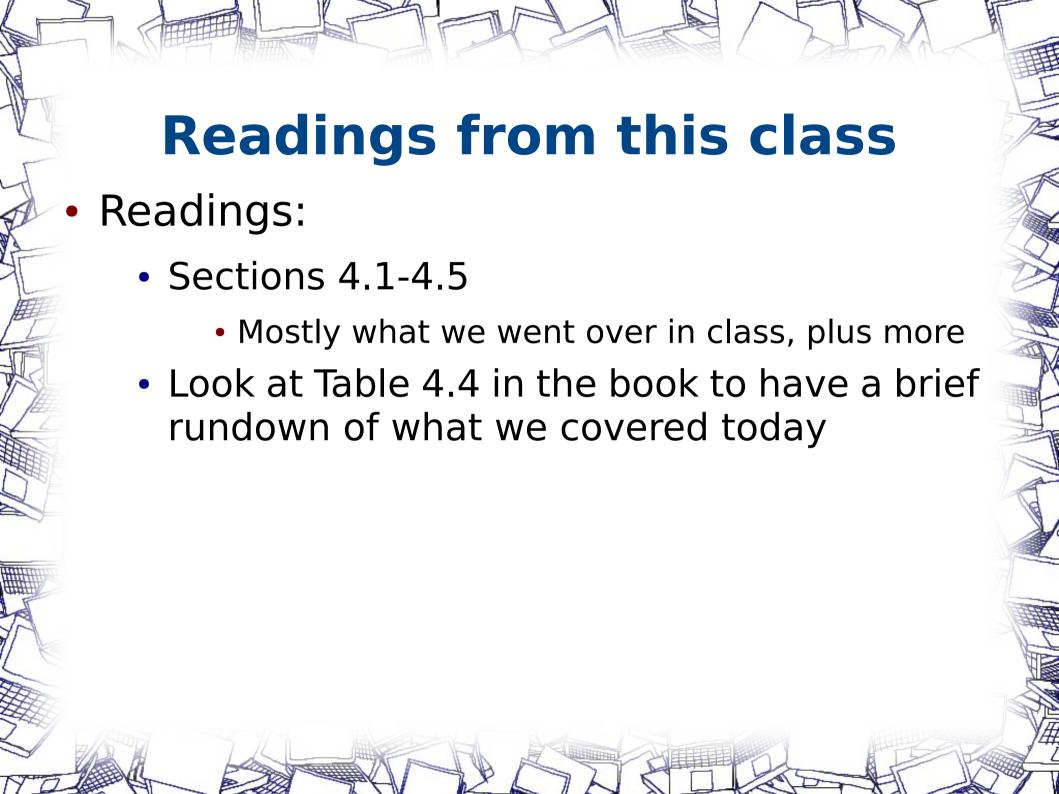
for (int k = 0; k < V; k++)
    for (int i = 0; i < V; i++)
        for (int j = 0; j < V; j++)
            m[i][j] = min(m[i][j], m[i][k] + m[k][j]);

// common error: remember that loop order is k->i->j
```

Graph algorithms

- Now you've seen the bread and butter of graph algorithms
 - There are many more problems associated with graphs
 - Find the width of a graph, find strongly connected components
 - There are special kinds of graphs and smarter algorithms for them
 - Trees, directed acyclic graphs (DAGs), bipartite graphs, eulerian graphs

Map problems tricks Demo in Eclipse

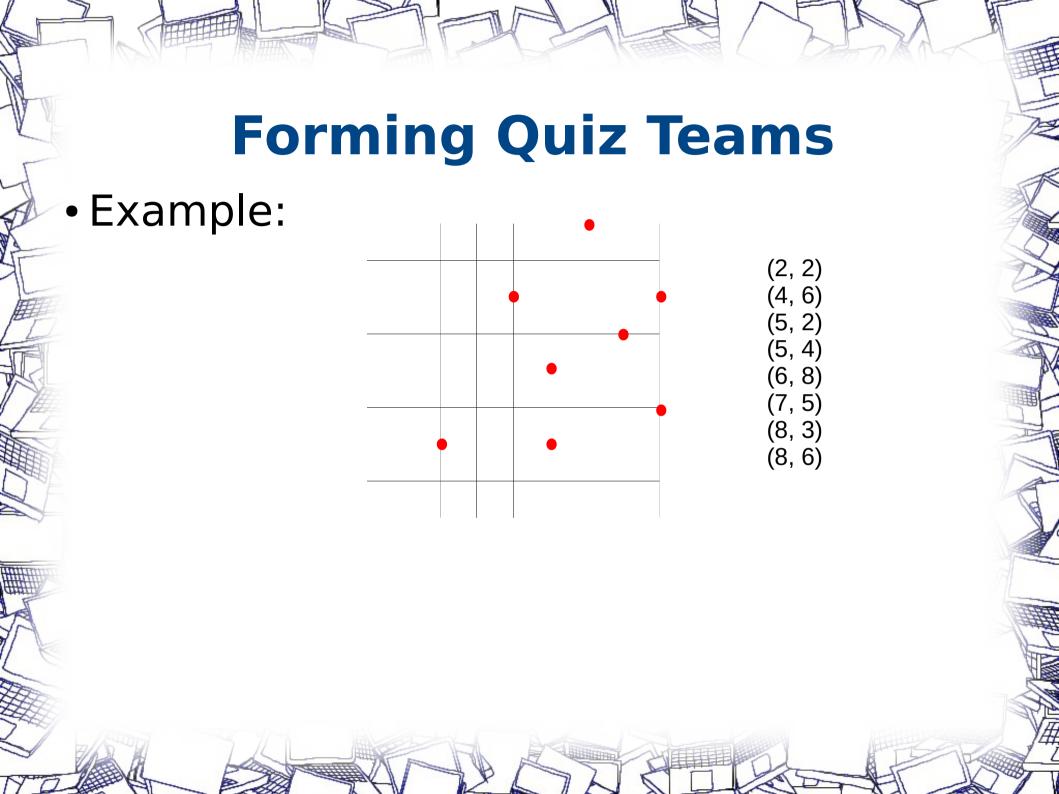


Forming Quiz Teams

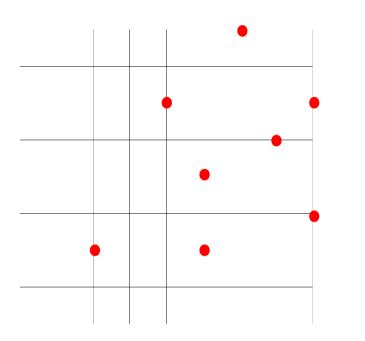
- Given 2*N points on a grid, make N pairs so that the sum of the distances of the paired points is minimized
 - 1 <= N <= 8, so 16 points

Forming Quiz Teams

- Recurrence:
 - dp[used grid points] = the minimum sum of distances between all remaining grid points
 - The answer is dp[all grid points]
 - The recursive step is to find the minimum sum by trying matching each pair of remaining grid points
 - There are a lot of overlapping states, so store the subresults



Example:



(2, 2)

(4, 6)

(5, 2)

(5, 4)

(6, 8)

(7, 5)

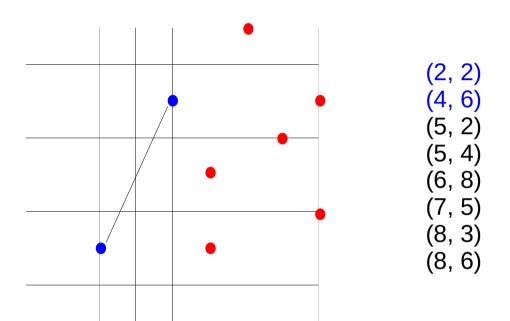
(8, 3)

(8, 6)

Recursion depth 0

What is the minimum sum for all the points?

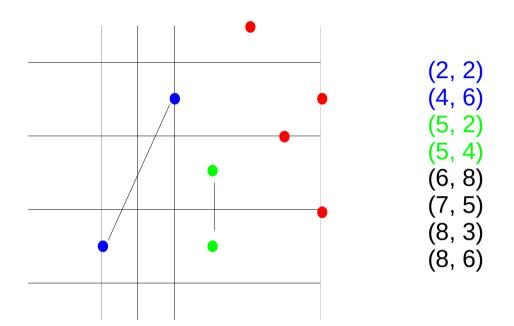
Example:



Recursion depth 1

What is the minimum sum for the 2nd, 3rd, 4th, 5th, 6th, and 7th points?

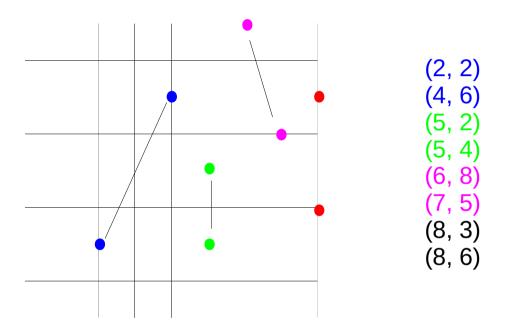
Example:



Recursion depth 2

What is the minimum sum for the 4th, 5th, 6th, and 7th points?

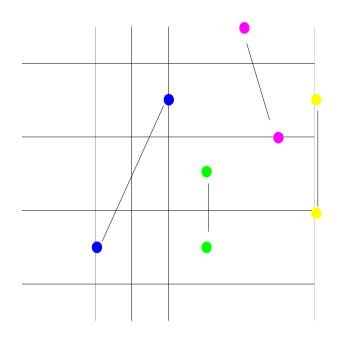
Example:



Recursion depth 3

What is the minimum sum for the 6th and 7th points?

Example:



(2, 2)

(4, 6)

(5, 2)

(5, 4)

(6.8)

(7, 5)

(8.3)

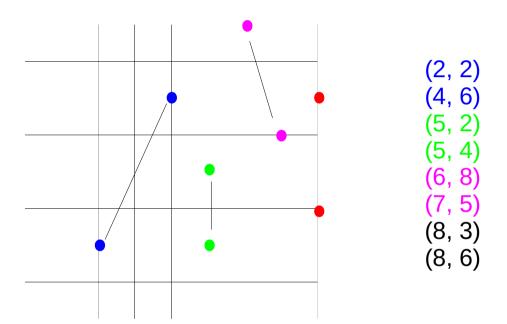
(8, 6)

Recursion depth 4

What is the minimum sum no points?

Answer: Base case, 0

Example:

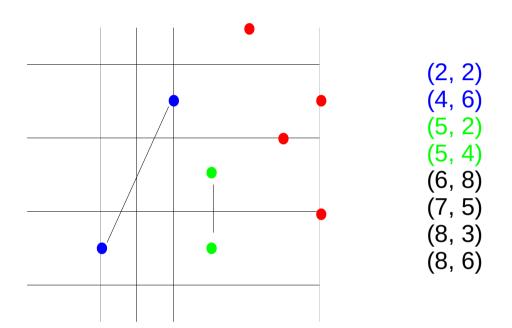


Recursion depth 3

What is the minimum sum for the 6th and 7th points?

Answer: 0 + dist(P[6], P[7]) = 3

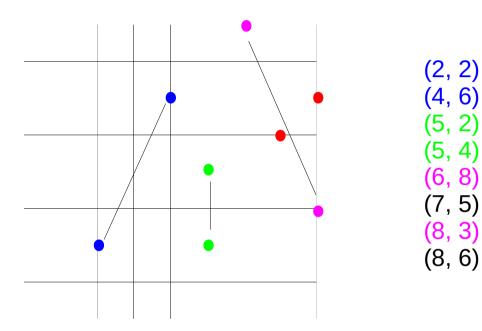
Example:



Recursion depth 2

What is the minimum sum for the 4th, 5th, 6th, and 7th points? So far: 3 + dist(P[4], P[5]) = 3 + 3.16 = 6.16

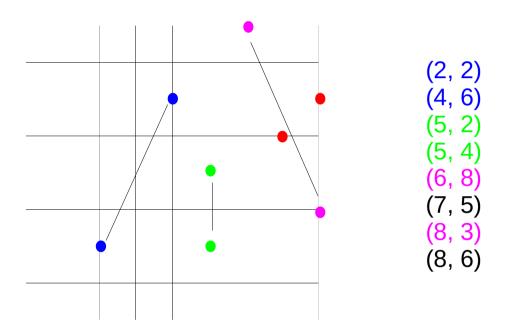
Example:



Recursion depth 3

What is the minimum sum for the 5th and 7th points?

Example:

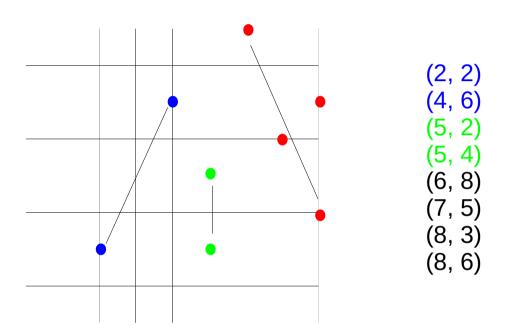


Recursion depth 3

What is the minimum sum for the 5th and 7th points?

Answer: 0 + dist(P[5], P[7]) = 1.41

Example:



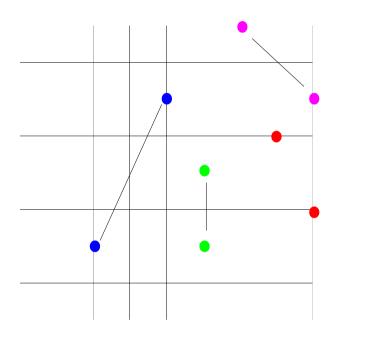
Recursion depth 2

What is the minimum sum for the 4th, 5th, 6th, and 7th points? From before: 6.16

Just computed: 1.41 + dist(P[4], P[6]) = 1.41 + 5.39 = 6.80

Still: 6.16

Example:



(2, 2) (4, 6)

(5, 2)

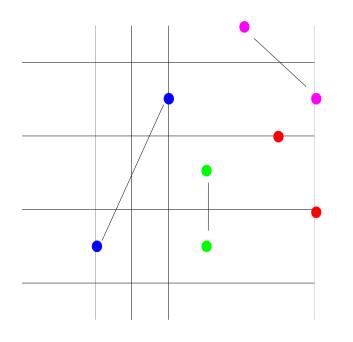
(5, 4) (6, 8)

(7, 5) (8, 3) (8, 6)

Recursion depth 3

What is the minimum sum for the 5th and 6th points?

Example:



(2, 2)

(4, 6)

(5, 2)

(5, 4)

(6, 8)

(7, 5)

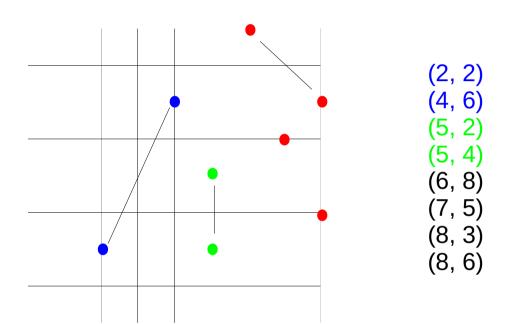
(8, 3)

(8, 6)

Recursion depth 3

What is the minimum sum for the 5th and 6th points? Answer: **2.24**

Example:



Recursion depth 2

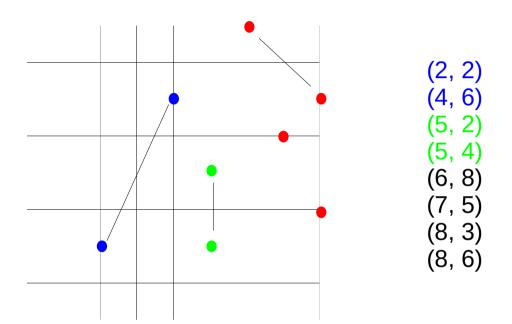
What is the minimum sum for the 4th, 5th, 6th, and 7th points?

From before: 6.16

Just computed: 2.24 + dist(P[4], P[7]) = 5.07

Now: 5.07

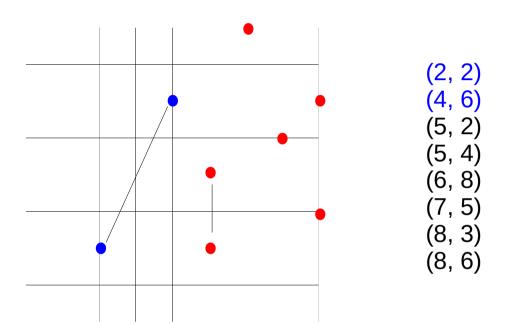
• Example:



Recursion depth 2

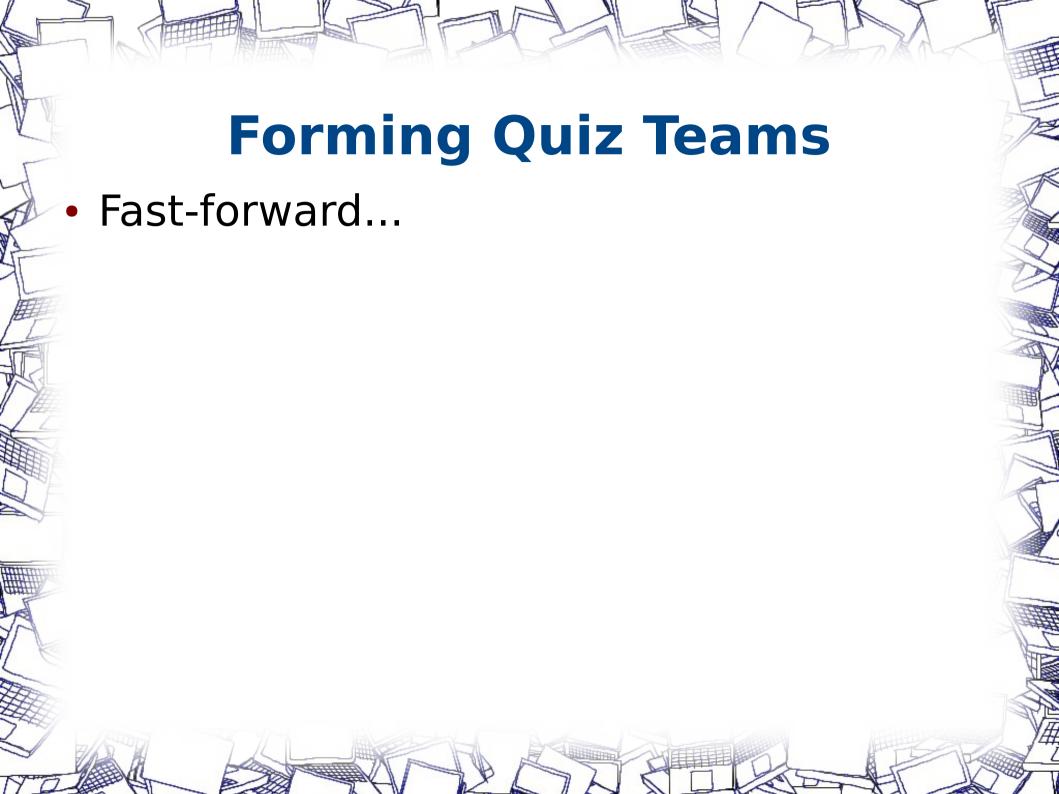
What is the minimum sum for the 4th, 5th, 6th, and 7th points? We have tried everything, so we can definitively answer 5.07.

Example:

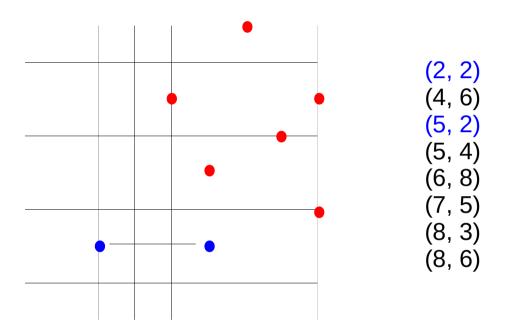


Recursion depth 1

What is the minimum sum for the 2nd, 3rd, 4th, 5th, 6th, and 7th points? So far: 5.07 + dist(P[2], P[3]) = 7.07



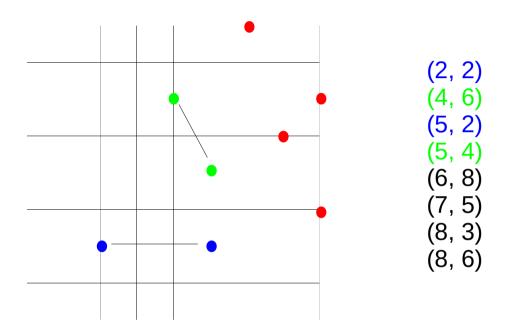
Example:



Recursion depth 1

What is the minimum sum for the 1st, 3rd, 4th, 5th, 6th, and 7th points?

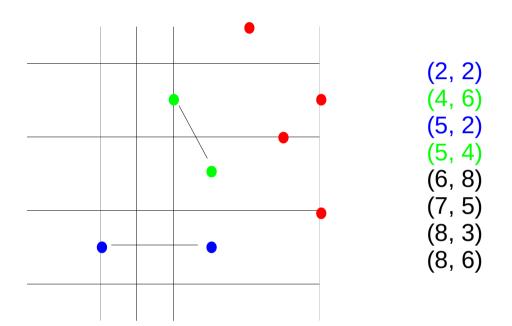
Example:



Recursion depth 2

What is the minimum sum for the 4th, 5th, 6th, and 7th points?

• Example:



What is the minimum sum for the 4th, 5th, 6th, and 7th points? **Answer: 5.07 (from the memoization table)**

- How to implement:
 - Use bitmasks!
 - Use a memoization table with 2^16 elements
 - Each entry in the table is considered a bitmask representing the set of all grid points chosen
- Another similar DP solution is the O(2ⁿ * n) solution to the Traveling Salesman Problem

```
int N:
int x[] = new int[16], y[] = new int[16]; // grid coordinates
double dp[] = new double[1 << 16]; // 2^16 entries
public double solve(int mask) {
   if (dp[mask] >= 0) return dp[mask]; // memoization step
   double res = INFINITY:
   for (int i = 0; i < 2*N; i++) {
      for (int j = i+1; j < 2*N; j++) { // filters out permutations
          if ((((1 << i) | (1 << j)) \& mask) == 0) { // unused set elmnts}
             double dist = sqrt(pow(x[i] - x[j], 2)
                            + pow(y[i] - y[j], 2));
             res = min(res, solve(mask | (1 << i) | (1 << j));
   return dp[mask] = res; // store the solution in memo table
dp[(1 \ll (N*2)) - 1] = 0.0; // base case: all points used = 0 min dist
   System.out.printf(solve(0)); // 0 = empty bit mask = all points rem
```