Lecture 9: Chapter 4 Part 1

The Greedy Approach CS3310

Optimization Problems

- An **optimization problem** can have more than one possible solution for an instance.
- Sorting is <u>not</u> an optimization problem.
 - o i.e. A list is either sorted or not, so there is only one solution.
- Finding a path from your home to school *is* an optimization problem, since many different routes exist.
 - However, only one of these routes is **optimal**.
 - Note: ties can exist (two or more different routes can be equally fast). In such a case, we arbitrarily pick one of the fastest routes.
- With *optimization problems* we are usually interested in maximizing or minimizing some value.

- A greedy algorithm calculates a solution to an optimization problem by making a sequence of choices, each of which seems best at the moment.
 - Each choice is called **locally optimal**.
- The goal is to find a solution that is **globally optimal** i.e. the best solution for the overall problem.
 - Some greedy algorithms accomplish this, others don't.
- Imagine you are picking a route to school. You must go through either points A-B-C-D or A-E-F-G to get here.
- A-B is much faster than A-E so you take that road; it is locally optimal.
- However, there is horrible traffic from C-D, so A-E-F-G would have been faster overall. This greedy algorithm didn't work because it didn't consider **global state**.

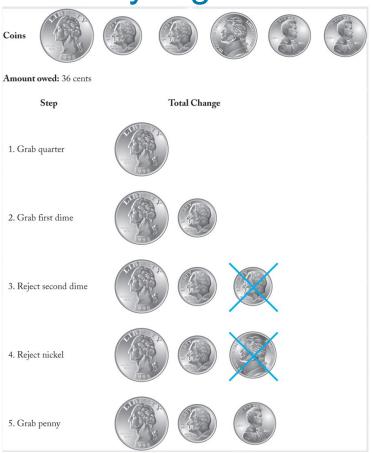
In the next few lectures, we'll discuss a few greedy algorithms that always find a globally optimal solution, unlike the previous example.

Greedy Algorithms repeat the following steps:

- 1. Selection Procedure: Decide what option is *locally* optimal.
- 2. Feasibility Check: Make sure the chosen value does not make us exceed our goal.
- **3. Solution Check**: See if the problem is solved. If not, return to step 1.

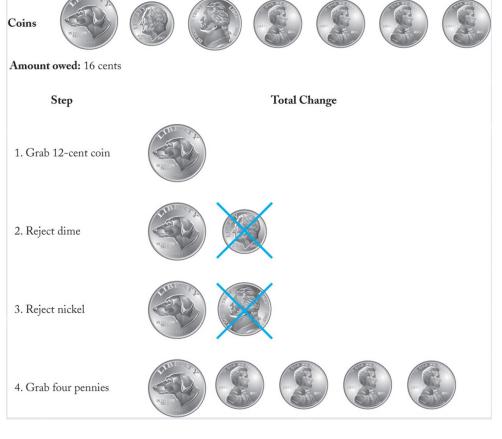
While using **U.S. coins**, a greedy algorithm <u>always</u> works for the problem of giving change.

• We want to calculate a given amount of change using as few coins as possible.



6

While the previous algorithm always generates a globally optimal solution with U.S. coins, it doesn't always do so if other values are added. Imagine we also have access to a 12 cent coin.

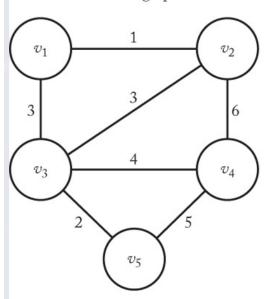


Graph Theory Refresher

- A **graph** is made up a **vertices** and **edges** which connect them.
- A graph is **undirected** if its edges have no direction
- A **path** in an undirected graph is a sequence of vertices such that there is an edge between each vertex and its successor
 - \circ i.e. $\{v_1, v_2, v_4\}$
- An undirected graph is **connected** if there is a path between every pair of vertices.
- A path from a vertex to itself, which contains at least three vertices and in which all intermediate vertices are distinct, is a simple **cycle**

$$\circ \quad \text{i.e. } \{ v_1, v_2, v_3, v_1 \}$$

(a) A connected, weighted, undirected graph *G*.



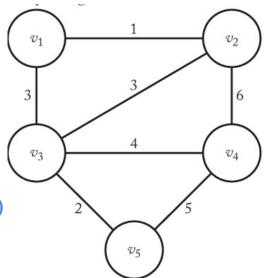
Graph Theory Refresher

An **undirected** graph G consists of a finite set V whose members are the vertices of G, together with a set E of pairs of vertices in V. These pairs are the **edges** of G. We denote G by G = (V, E).

• Members of E are denoted by (v_i, v_j)

This graph is made up of the following vertices and edges:

- $V = \{ v_1, v_2, v_3, v_4, v_5 \}$
- $E = \{ (v_1, v_2), (v_1, v_3), (v_2, v_3), (v_2, v_4), (v_3, v_4), (v_3, v_5), (v_4, v_5) \}$



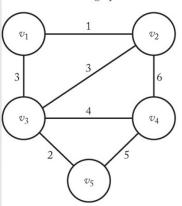
Minimum Spanning Trees

- A **free tree** is a connected, undirected graph such that no cycles exist.
- No vertex is designated the root in a free tree.
 - This is different from the more common **rooted tree**, such as a binary tree.

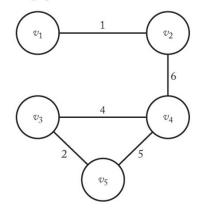
- Suppose we want to connect a group of certain cities with roads, but we also want to use as little road as possible.
 - If we create a **minimum spanning tree**, each city in the graph is connected to each other city, but not necessarily directly.

Minimum Spanning Trees

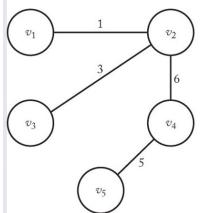
(a) A connected, weighted, undirected graph *G*.



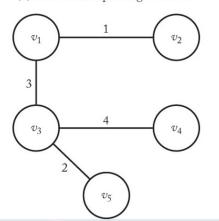
(b) If (v_4, v_5) were removed from this subgraph, the graph would remain connected.



(c) A spanning tree for G.



(d) A minimum spanning tree for G.



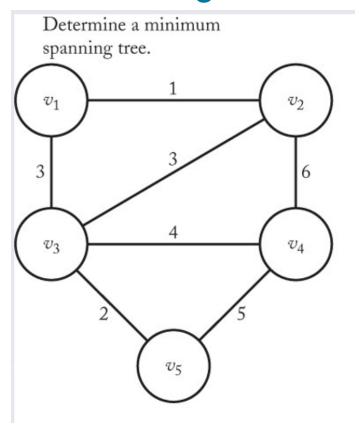
Minimum Spanning Trees

A very high-level greedy algorithm to find a minimum spanning tree.

Problem: Given a graph G = (V, E), determine a minimum spanning tree T = (Y, F).

- Initialize an empty subset F to contain the edges in the minimum spanning tree.
- Initialize a subset of vertices Y with one arbitrary vertex from V.
 - Once every vertex from V is in Y, the spanning tree is complete.
- A vertex **nearest** to Y is a vertex in V Y that is connected to a vertex in Y by an edge of minimum weight.
 - O In other words, select a vertex in V that is *not* in Y that is connected to a vertex in Y with an edge of the smallest possible weight.
 - This ensures that a cycle is not created.
 - Ties are broken arbitrarily.

```
F = \{\} // initialize set of edges to empty
                     // initialize set of vertices to contain first one
Y = \{ v_1 \}
while (the instance is not solved)
         // selection procedure and feasibility check:
          select a vertex in V - Y that is nearest to Y
         add the vertex to Y;
          add the edge to F;
          // solution check:
          if (Y == V)
                   Every vertex has been added to Y: the instance is solved.
```



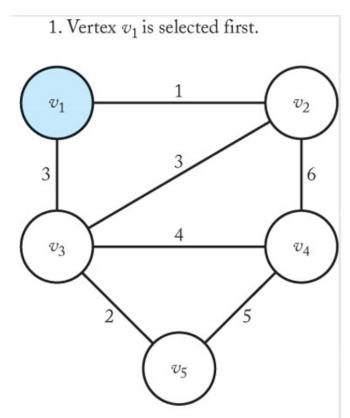
$$F = \{\}$$

$$Y = \{v_1\}$$

$$V = \{v_1, v_2, v_3, v_4, v_5\}$$

$$V - Y = \{v_2, v_3, v_4, v_5\}$$

Which vertex in V - Y is closest to a vertex in Y?



$$F = \{\}$$

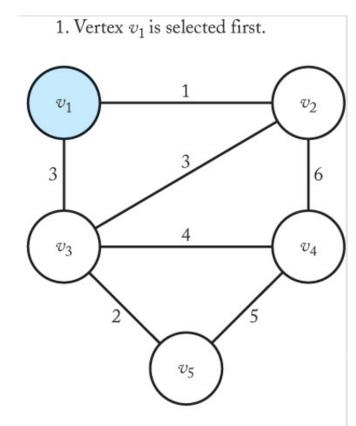
$$Y = \{v_1\}$$

$$V = \{v_1, v_2, v_3, v_4, v_5\}$$

$$V - Y = \{v_2, v_3, v_4, v_5\}$$

Which vertex in V - Y is closest to a vertex in Y?

- v_3 connects to v_1 with a weight of 3
- v_2 connects to v_1 with a weight of 1



$$F = \{(v_1, v_2)\}\$$

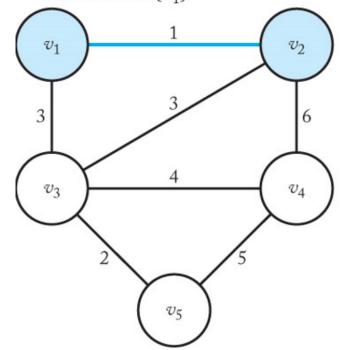
$$Y = \{v_1, v_2\}\$$

$$V = \{v_1, v_2, v_3, v_4, v_5\}\$$

$$V - Y = \{v_3, v_4, v_5\}\$$

Which vertex in V - Y is closest to a vertex in Y?

2. Vertex v_2 is selected because it is nearest to $\{v_1\}$.



$$F = \{(v_1, v_2)\}\$$

$$Y = \{v_1, v_2\}\$$

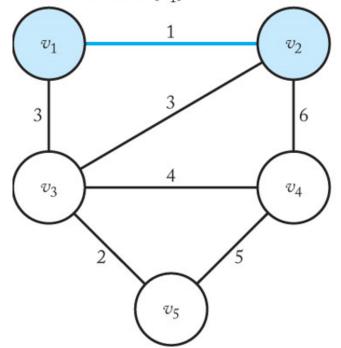
$$V = \{v_1, v_2, v_3, v_4, v_5\}\$$

$$V - Y = \{v_3, v_4, v_5\}\$$

Which vertex in V - Y is closest to a vertex in Y?

- v_4 connects to v_2 with a weight of 6.
- v_3 connects with a weight of 3 to both v_1 and v_2
 - \circ Tie is broken arbitrarily. We'll choose v_1

2. Vertex v_2 is selected because it is nearest to $\{v_1\}$.



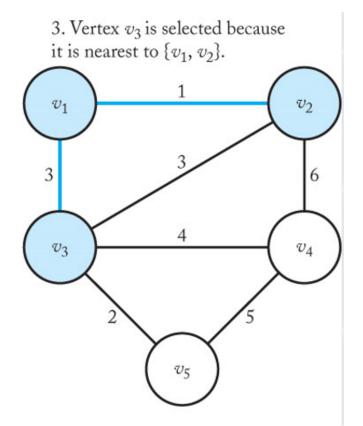
$$F = \{(v_1, v_2), (v_1, v_3)\}$$

$$Y = \{v_1, v_2, v_3\}$$

$$V = \{v_1, v_2, v_3, v_4, v_5\}$$

$$V - Y = \{v_4, v_5\}$$

Which vertex in V - Y is closest to a vertex in Y?



$$F = \{(v_1, v_2), (v_{1}, v_3)\}$$

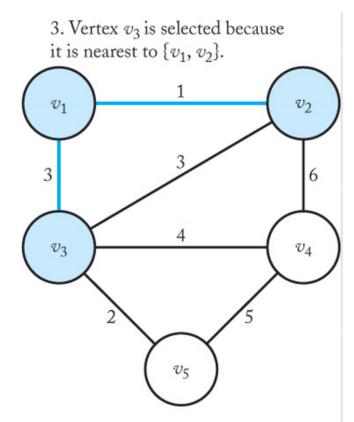
$$Y = \{v_1, v_2, v_3\}$$

$$V = \{v_1, v_2, v_3, v_4, v_5\}$$

$$V - Y = \{v_4, v_5\}$$

Which vertex in V - Y is closest to a vertex in Y?

- v_4 connects to v_2 with a weight of 6 and v_3 with a weight of 4.
- v_5 connects v_3 with a weight of 2.



$$F = \{(v_1, v_2), (v_1, v_3), (v_3, v_5)\}$$

$$Y = \{v_1, v_2, v_3, v_5\}$$

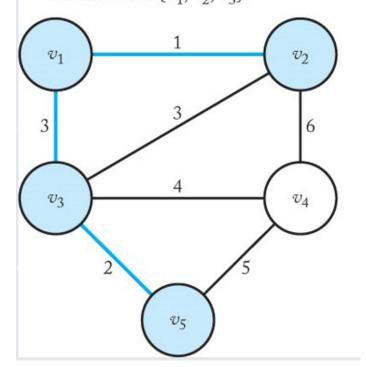
$$V = \{v_1, v_2, v_3, v_4, v_5\}$$

$$V - Y = \{v_4\}$$

Which vertex in V - Y is closest to a vertex in Y?

• v_4 is all that's left! The edge with the smallest weight between it and Y is the one to v_3

4. Vertex v_2 is selected because it is nearest to $\{v_1, v_2, v_3\}$.



$$F = \{(v_1, v_2), (v_1, v_3), (v_3, v_5), (v_3, v_4)\}$$

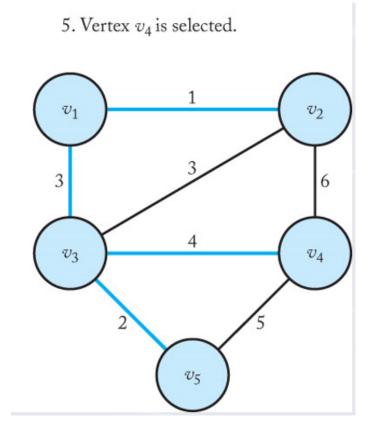
$$Y = \{v_1, v_2, v_3, v_4, v_5\}$$

$$V = \{v_1, v_2, v_3, v_4, v_5\}$$

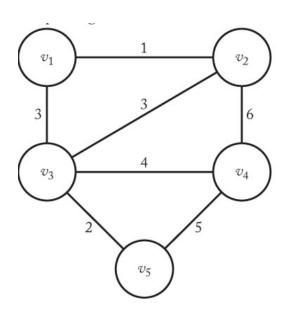
$$V - Y = \{\}$$

Once V - Y is empty, we have a spanning tree. Every vertex from V is in Y.

• T = (F, Y) is a minimum spanning tree with a total weight of 10.



While the previous steps are fairly straightforward for a human, we need a step-by-step procedure for a computer to follow.



We will represent the graph by an adjacency matrix

W[i][j] =

- 0 if i = j
- ∞ if there is no edge between i and j
- weight on edge if there is an edge between *i* and *j*

Along with the adjacency matrix, we keep two other arrays, nearest and distance.

```
for i = 2, ..., n:
```

- nearest[i] = vertex in Y nearest to v_i
- distance[i] = weight on edge between v_i and the vertex in nearest[i]

For example, if we have the following:

$$Y = \{1, 3\}$$

V - Y = \{2, 4, 5\}

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	0 1 3 ∞	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

- > nearest[4] = 3. This means 3 is the vertex in Y nearest to 4
- \rightarrow distance [4] = 4. This means the edge (3, 4) has a weight of 4

When we begin Prim's algorithm, $Y = \{1\}$

> Arbitrarily start with vertex 1 in Y (the spanning tree)

Determine:

- nearest[2]:
 - i.e. what vertex in Y is closest to 2?
- distance[2]:
 - i.e. what is the distance between nearest[2] and 2?

	1	2	3	4	5
1	0	1	3	∞	∞
2	0 1 3 ∞ ∞	0	3	6	∞
3	3	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

When we begin Prim's algorithm, $Y = \{1\}$

> Arbitrarily start with vertex 1 in Y (the spanning tree)

Determine:

- nearest[2]: 1, since it's the only vertex in Y
 - i.e. what vertex in Y is closest to 2?
- distance[2]:W[1][2] = 1
 - o i.e. what is the distance between nearest [2] and 2?

	1	2	3	4	5
1	0	1	3	∞	∞
2	0 1 3 ∞	0	3	6	∞
3	3	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

Prim Pseudocode

```
void prim (int n, const number W[][], set of edges& F)
         int vnear; // will contain the vertex in V - Y nearest to Y each iteration
         edge e;
         index nearest[2..n];
         number distance[2..n];
         F = \{ \}
         for (index i = 2; i <= n; i++)</pre>
                                           For all vertices other than 1,
                   nearest[i] = 1; initialize 1 to be its nearest vertex
                  distance[i] = W[1][i]; in Y. Initialize the distance from Y
                                           to the weight on the edge from to 1
                                           t o i
```

// continued on next slide

Prim Pseudocode

```
// continued from previous slide
repeat (n - 1 times)
         int min = \infty
         for (i = 2; i <= n; i++)
                  if (0 <= distance[i] < min)// if distance[i] is -1, i is already</pre>
in Y
                            min = distance[i];
                            vnear = i;
    e = edge connecting vnear and nearest[vnear]
    add e to F;
    distance[vnear] = -1;
    for (i = 2; i \le n; i++)// see if any vertices in V - Y are closer to vnear
         if (W[i][vnear] < distance[i])</pre>
                  distance[i] = W[i][vnear];
                  nearest[i] = vnear;
```

Prim's Algorithm Initialization

nearest:

distance:

min:

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	0 1 3 ∞	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

$$F = \{\}$$

Prim's Algorithm Initialization

```
nearest: \{-1, 1, 1, 1, 1\}
distance: \{-1, 1, 3, \infty, \infty\} i.e. (Y = \{v_1\})
```

min:

- Setting a distance value to -1 indicates that that vertex is in the partial spanning tree, Y.
- Every vertex in V Y is closest to v_1 in Y (because v_1 is the only vertex currently in Y!)
 - In the distance array, we set each index to contain the distance from v_1 to the corresponding vertex.

```
F = \{\}
```

	1	2	3	4	5
1	0	1	3	∞	∞
2	0 1 3 ∞	0	3	6	∞
3	3	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

```
nearest: \{-1, 1, 1, 1, 1\}
distance: \{-1, 1, 3, \infty, \infty\} (Y = \{v_1\})
```

min: distance[2] which is 1.

- The smallest value in distance is 1 in index 2.
- vnear indicates the vertex in V Y closest to Y. We set it to 2

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	3	1 0 3 6 ∞	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

$$F = \{\}$$

```
nearest: \{-1, 1, 1, 1, 1\}
distance: \{-1, 1, 3, \infty, \infty\} (Y = \{v_1\})
```

min: distance[2] which is 1.

- Add the edge (nearest[vnear], vnear) to F, which connects the following vertices:
 - nearest[vnear] = a vertex in Y
 - \circ vnear = a vertex in V Y
- set distance[vnear] to -1 to indicate vnear is now in Y

$$F = \{\}$$

	1	2	3	4	5
1	0	1	3	∞	∞
2	0 1 3 ∞	0	3	6	∞
3	3	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

```
nearest: \{-1, 1, 1, 1, 1\}
distance: \{-1, -1, 3, \infty, \infty\} (Y = \{v_1, v_2\})
```

min: distance[2] which is 1.

vnear: 2

- Add the edge (nearest[vnear], vnear) to F, which connects the following vertices:
 - o nearest[vnear] = a vertex in Y
 - \circ vnear = a vertex in V Y
- set distance[vnear] to -1 to indicate vnear is now in Y

$$F = \{ (1, 2) \}$$

Anything else we need to do?

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	1 0 3 6 ∞	3	6	∞
3	3	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

```
nearest: \{-1, 1, 1, 2, 1\}
distance: \{-1, -1, 3, 6, \infty\} (Y = \{v_1, v_2\})
```

min: distance[2] which is 1.

vnear: 2

- See if any vertices i in V Y are closer to vnear than nearest[i].
- i.e. are any vertices *i* not in the partial spanning tree closer to the vertex we just added than distance[*i*]?

4 is 6 away from 2. Update nearest[4] to 2 and distance[4] to 6

$$F = \{ (1, 2) \}$$

	1	2	3	4	5
1	0	1	3	∞	∞
2	0 1 3 ∞ ∞	0	3	6	∞
3	3	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

```
nearest: \{-1, 1, 1, 2, 1\}
distance: \{-1, -1, 3, 6, \infty\} (Y = \{v_1, v_2\})
```

min:

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	0 1 3 ∞	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

$$F = \{ (1, 2) \}$$

```
nearest: \{-1, 1, 1, 2, 1\}
distance: \{-1, -1, 3, 6, \infty\} (Y = \{v_1, v_2\})
```

min: distance[3] which is 3

- Add (nearest[vnear], vnear) to F
- Set distance[vnear] to -1

```
F = \{ (1, 2) \}
```

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	3	3	0	4	2
4	0 1 3 ∞ ∞	6	4	0	5
5	∞	∞	2	5	0

```
nearest: \{-1, 1, 1, 2, 1\}
distance: \{-1, -1, -1, 6, \infty\} (Y = \{v_1, v_2, v_3\})
```

min: distance[3] which is 3

vnear: 3

- Add (nearest[vnear], vnear) to F
- Set distance[vnear] to -1

Any remaining vertices i in V - Y closer to vnear than nearest[i]?

$$F = \{ (1, 2), (1, 3) \}$$

	1	2	3	4	5
1	0	1	3	∞	∞
2	0 1 3 ∞	0	3	6	∞
3	3	3	0	4	2
4	∞	6	4	0	5
5	∞	00	2	5	0

```
nearest: \{-1, 1, 1, 3, 3\}
distance: \{-1, -1, -1, 4, 2\} (Y = \{v_1, v_2, v_3\})
```

min: distance[3] which is 3

- 4 is 4 away from 3
- 5 is 2 away from 3

```
F = \{ (1, 2), (1, 3) \}
```

	1	2	3	4	5
1	0	1	3	∞	∞
2	0 1 3 ∞	0	3	6	∞
3	3	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

```
nearest: \{-1, 1, 1, 3, 3\}
distance: \{-1, -1, -1, 4, 2\} (Y = \{v_1, v_2, v_3\})
```

min:

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	3	3	0	4	2
4	0 1 3 ∞	6	4	0	5
5	∞	∞	2	5	0

$$F = \{ (1, 2), (1, 3) \}$$

```
nearest: \{-1, 1, 1, 3, 3\}
distance: \{-1, -1, -1, 4, 2\} (Y = \{v_1, v_2, v_3\})
```

min: distance[5] which is 2

- Add (nearest[vnear], vnear) to F
- Set distance[vnear] to -1

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	3	3	0	4	2
4	0 1 3 ∞	6	4	0	5
5	∞	∞	2	5	0

$$F = \{ (1, 2), (1, 3) \}$$

```
nearest: \{-1, 1, 1, 3, 3\}
distance: \{-1, -1, -1, 4, -1\} (Y = \{v_1, v_2, v_3, v_5\})
```

min: distance[5] which is 2

vnear: 5

- Add (nearest[vnear], vnear) to F
- Set distance[vnear] to -1

Any remaining vertices i in V - Y closer to *vnear* than nearest[i]?

$$F = \{ (1, 2), (1, 3), (3, 5) \}$$

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	3	3	0	4	2
4	0 1 3 ∞	6	4	0	5
5	∞	∞	2	5	0

nearest:
$$\{-1, 1, 1, 3, 3\}$$

distance: $\{-1, -1, -1, 4, -1\}$ $(Y = \{v_1, v_2, v_3, v_5\})$

min: distance[5] which is 2

vnear: 5

- Add (nearest[vnear], vnear) to F
- Set distance[vnear] to -1

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	0 1 3 ∞ ∞	3	0	4	2
4	∞	6	4	0	5
5	∞	∞	2	5	0

Any remaining vertices i in V - Y closer to *vnear* than nearest[i]? **No!**

$$F = \{ (1, 2), (1, 3), (3, 5) \}$$

Prim's Algorithm Fourth Step

```
nearest: { -1, 1, 1, 3, 3} distance: {-1, -1, -1, 4, -1}
```

$$(Y = \{v_1, v_2, v_3, v_5\})$$

min:

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	3	3	0	4	2
4	0 1 3 ∞	6	4	0	5
5	∞	00	2	5	0

$$F = \{ (1, 2), (1, 3), (3, 5) \}$$

Prim's Algorithm Fourth Step

```
nearest: \{-1, 1, 1, 3, 3\}

distance: \{-1, -1, -1, 4, -1\} (Y = \{v_1, v_2, v_3, v_5\})
```

min: distance[4] which is 4

- Add (nearest[vnear], vnear) to F.
- Set distance[vnear] to -1

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	3	3	0	4	2
4	0 1 3 ∞ ∞	6	4	0	5
5	∞	00	2	5	0

$$F = \{ (1, 2), (1, 3), (3, 5) \}$$

Prim's Algorithm Fourth Step

```
nearest: { -1, 1, 1, 3, 3} distance: {-1, -1, -1, -1, -1}
```

$$(Y = \{v_1, v_2, v_3, v_5\})$$

min: distance[4] which is 4

vnear: 4

- Add (nearest[vnear], vnear) to F.
- Set distance[vnear] to -1

	1	2	3	4	5
1	0	1	3	∞	∞
2	1	0	3	6	∞
3	3	3	0	4	2
4	0 1 3 ∞ ∞	6	4	0	5
5	∞	∞	2	5	0

47

Done!

$$F = \{ (1, 2), (1, 3), (3, 5), (3, 4) \}$$

Prim's Algorithm Analysis

```
repeat (n - 1 times)
          int min = \infty
          for (i = 2; i <= n; i++)
                     if (0 <= distance[i] < min) // if distance[i] is -1, i is already in</pre>
Y
                               min = distance[i];
                               vnear = i:
     e = edge connecting vnear and nearest[vnear]
     add e to F:
     distance[vnear] = -1;
     for (i = 2; i \le n; i++) // see if any vertices in V - Y are closer to vnear
          if (W[i][vnear] < distance[i])</pre>
                     distance[i] = W[i][vnear];
                     nearest[i] = vnear;
```

What is the basic operation?

Prim's Algorithm Analysis

```
repeat (n - 1 times)
          int min = \infty
          for (i = 2; i <= n; i++)
                     if (0 <= distance[i] < min) // if distance[i] is -1, i is already in</pre>
Y
                               min = distance[i];
                               vnear = i:
     e = edge connecting vnear and nearest[vnear]
     add e to F:
     distance[vnear] = -1;
     for (i = 2; i \le n; i++) // see if any vertices in V - Y are closer to vnear
          if (W[i][vnear] < distance[i])</pre>
                     distance[i] = W[i][vnear];
                     nearest[i] = vnear;
```

What is the basic operation? There are two: the **if** statements in each loop. How many times do they occur?

Prim's Algorithm Analysis

```
repeat (n - 1 times)
          int min = \infty
          for (i = 2; i <= n; i++)
                     if (0 <= distance[i] < min) // if distance[i] is -1, i is already in</pre>
Υ
                                min = distance[i];
                                vnear = i;
     e = edge connecting vnear and nearest[vnear]
     add e to F:
     distance[vnear] = -1;
     for (i = 2; i <= n; i++) // see if any vertices in V - Y are closer to vnear</pre>
          if (W[i][vnear] < distance[i])</pre>
                     distance[i] = W[i][vnear];
                     nearest[i] = vnear;
```

What is the basic operation? There are two: the if statements in each loop.

How many times do they occur? (n - 1) times each, plus (n - 1) for the outer loop:

$$2(n-1)(n-1) \in \Theta(n^2)$$

In-Class Exercise

1. Use Prim's Algorithm to find a minimum spanning tree for the following graph. Show the values in nearest, distance, and F for each step.

	1	2	3	4	5	6	
1	0	10	∞	30	45	∞	
2	10	0	50	∞	40	25	
3	∞	50	0	∞	35	15	
4	30	∞	∞	0	∞	20	
5	45	40			0		
6	∞	25	15	20	55	0	