Lecture 21: Disjoint Sets with Arrays

CSE 373: Data Structures and Algorithms

Implementation

Use Nodes?

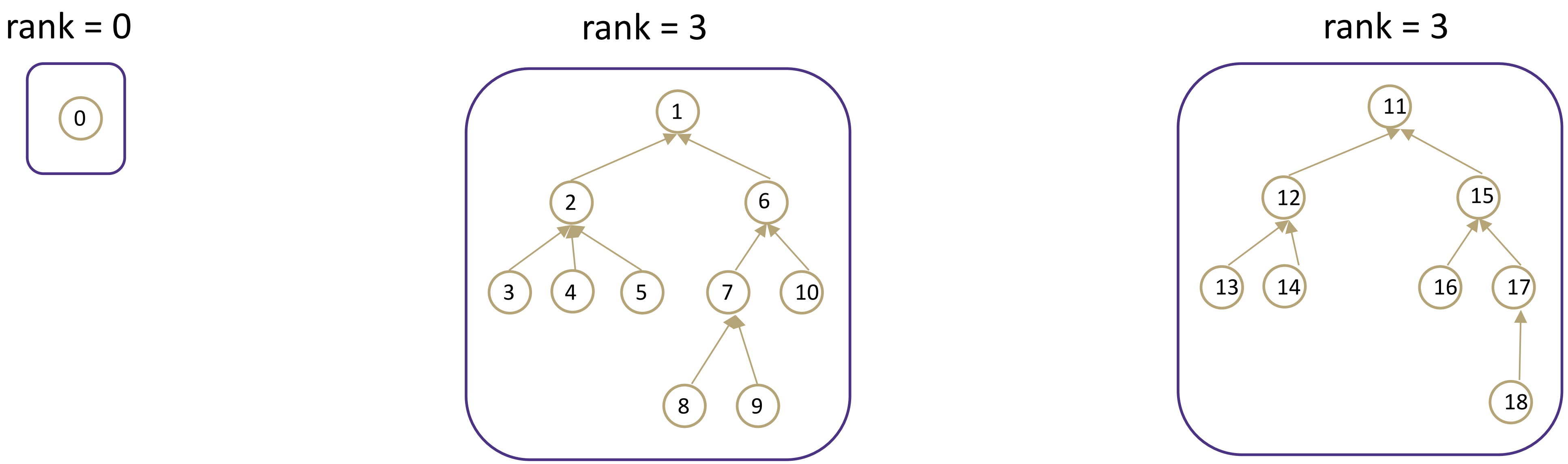
In modern Java (assuming 64-bit JDK) each object takes about 32 bytes

- int field takes 4 bytes
- Pointer takes 8 bytes
- Overhead ~ 16 bytes
- Adds up to 28, but we must partition in multiples of 8 => 32 bytes

Use arrays instead!

- Make index of the array be the vertex number
  - Either directly to store ints or representationally
  - We implement makeSet(x) so that **we** choose the representative
- Make element in the array the index of the parent

Array Implementation

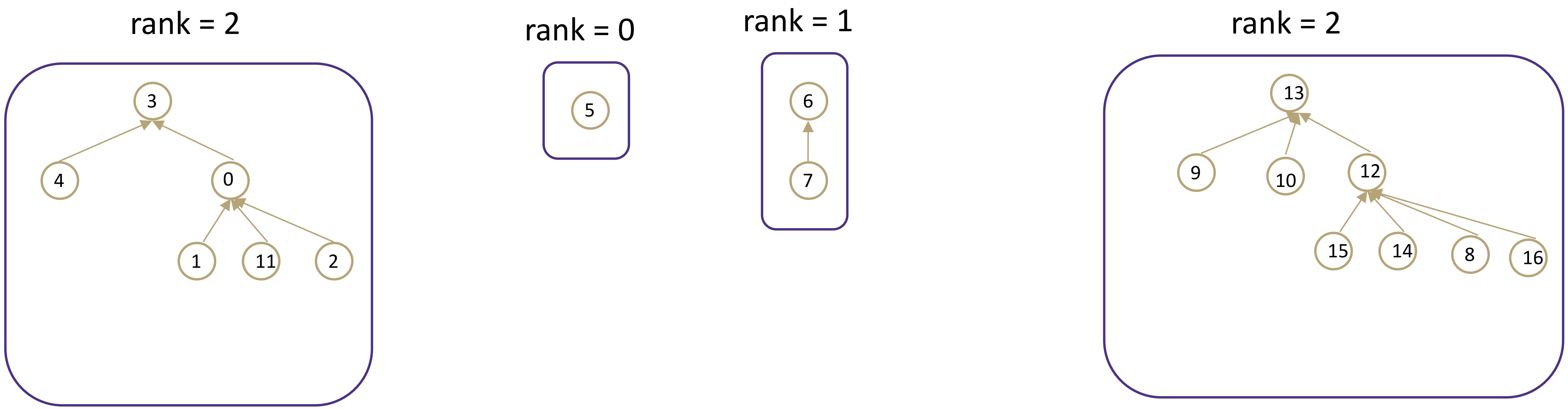


0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
-1	-4	1	2	2	2	1	6	7	7	6	-4	11	12	12	11	15	15	17

Store (rank * -1) - 1

Each “node” now only takes 4 bytes of memory instead of 32

Practice



0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
3	0	0	-3	3	-1	-2	6	12	13	13	0	13	-3	12	12	12

Array Method Implementation

makeSet(x)

add new value to array with a rank of -1

findSet(x)

Jump into array at index/value you're looking for, jump to parent based on element at that index, continue until you hit negative number

union(x, y)

findSet(x) and findSet(y) to decide who has larger rank, update element to represent new parent as appropriate

Graph Design

Graphs are about representing relationships…

Physical distances

Connections

Bloodlines

Probabilities

Sequences

States

Scenario #1

You are going to Disneyland for spring break! You've never been, so you want to make sure you hit ALL the rides.

Is there a graph algorithm that would help?

BFS or DFS

How would you draw the graph?

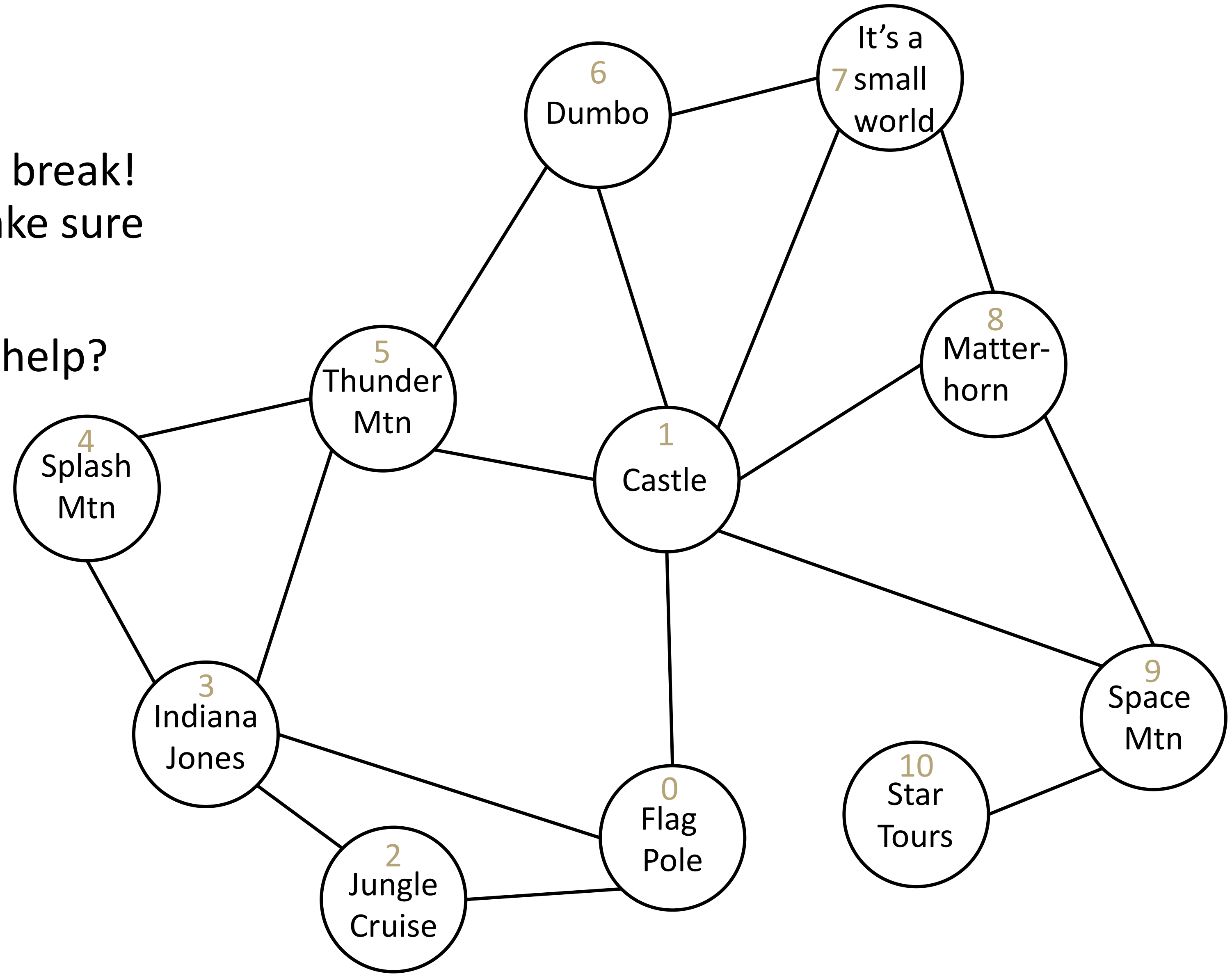
- What are the vertices?

Rides

- What are the edges?

Walkways

BFS = 0 1 2 3 5 6 7 8 9 4 10
DFS = 0 3 5 6 7 8 9 10 1 4 2



Scenario #1 continued

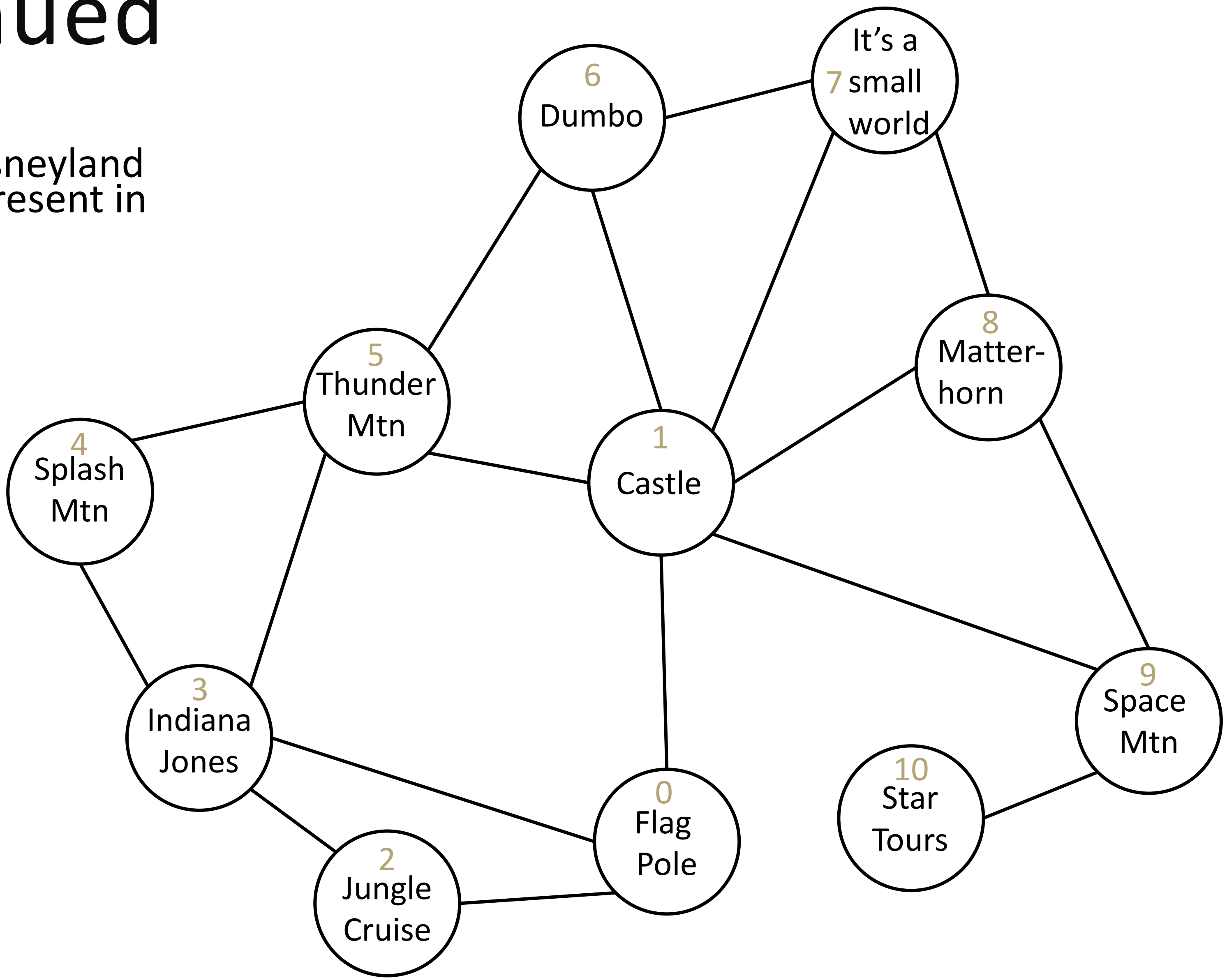
Now that you have your basic graph of Disneyland what might the following graph items represent in this context?

Weighted edges
- Walkway distances
- Walking times
- Foot traffic

Directed edges
- Entrances and exits
- Crowd control for fireworks
- Parade routes

Self Loops
- Looping a ride

Parallel Edges
- Foot traffic at different times of day
- Walkways and train routes



Scenario #2

You are a Disneyland employee and you need to rope off as many miles of walkways as you can for the fireworks while leaving guests access to all the rides.

Is there a graph algorithm that would help?

MST

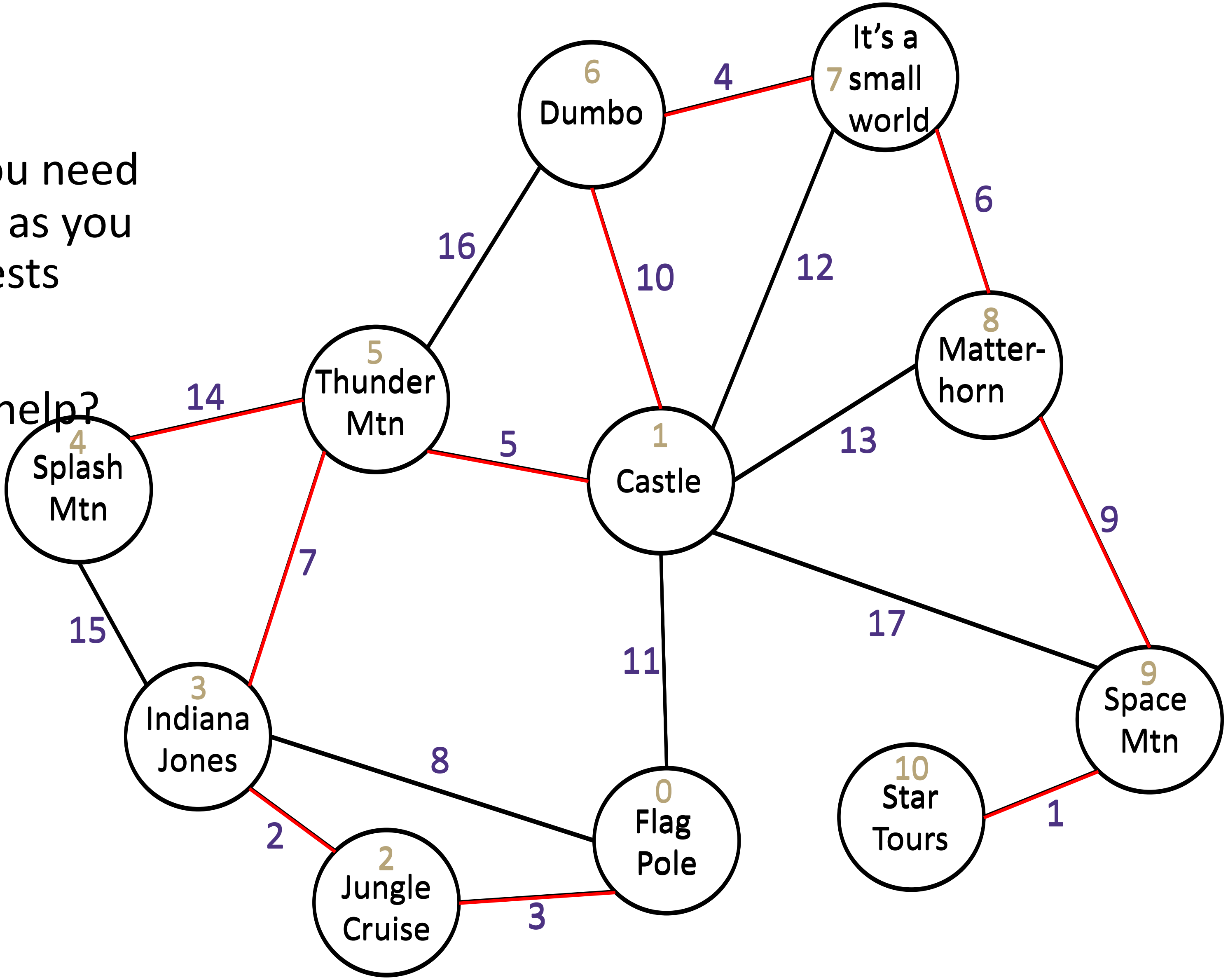
How would you draw the graph?

- What are the vertices?

Rides

- What are the edges?

Walkways with distances



Scenario #3

You arrive at Disneyland and you want to visit all the rides, but do the least amount of walking possible. If you start at the Flag Pole, plan the shortest walk to each of the attractions.

Is there a graph algorithm that would help?

Dijkstra's

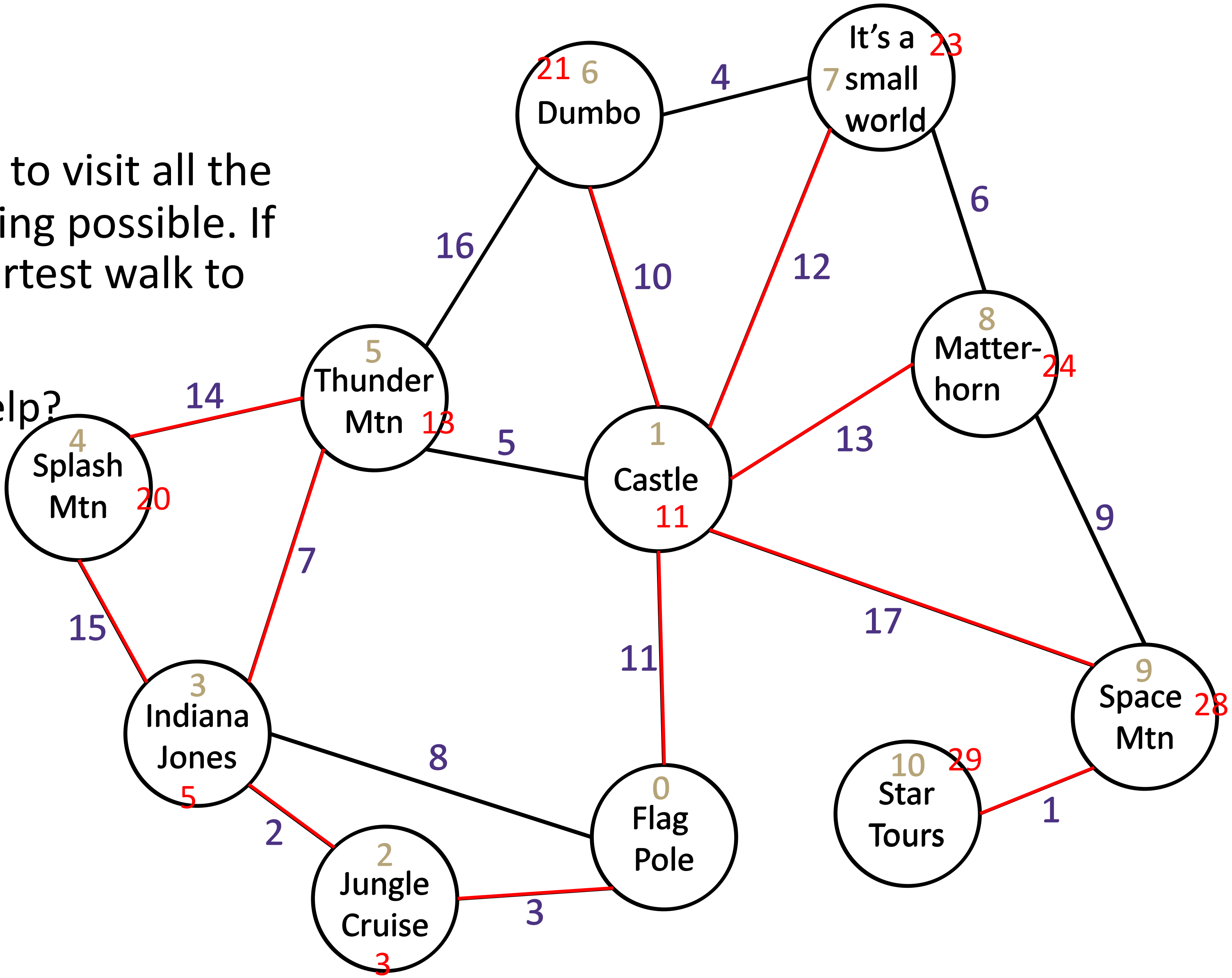
How would you draw the graph?

- What are the vertices?

Rides

- What are the edges?

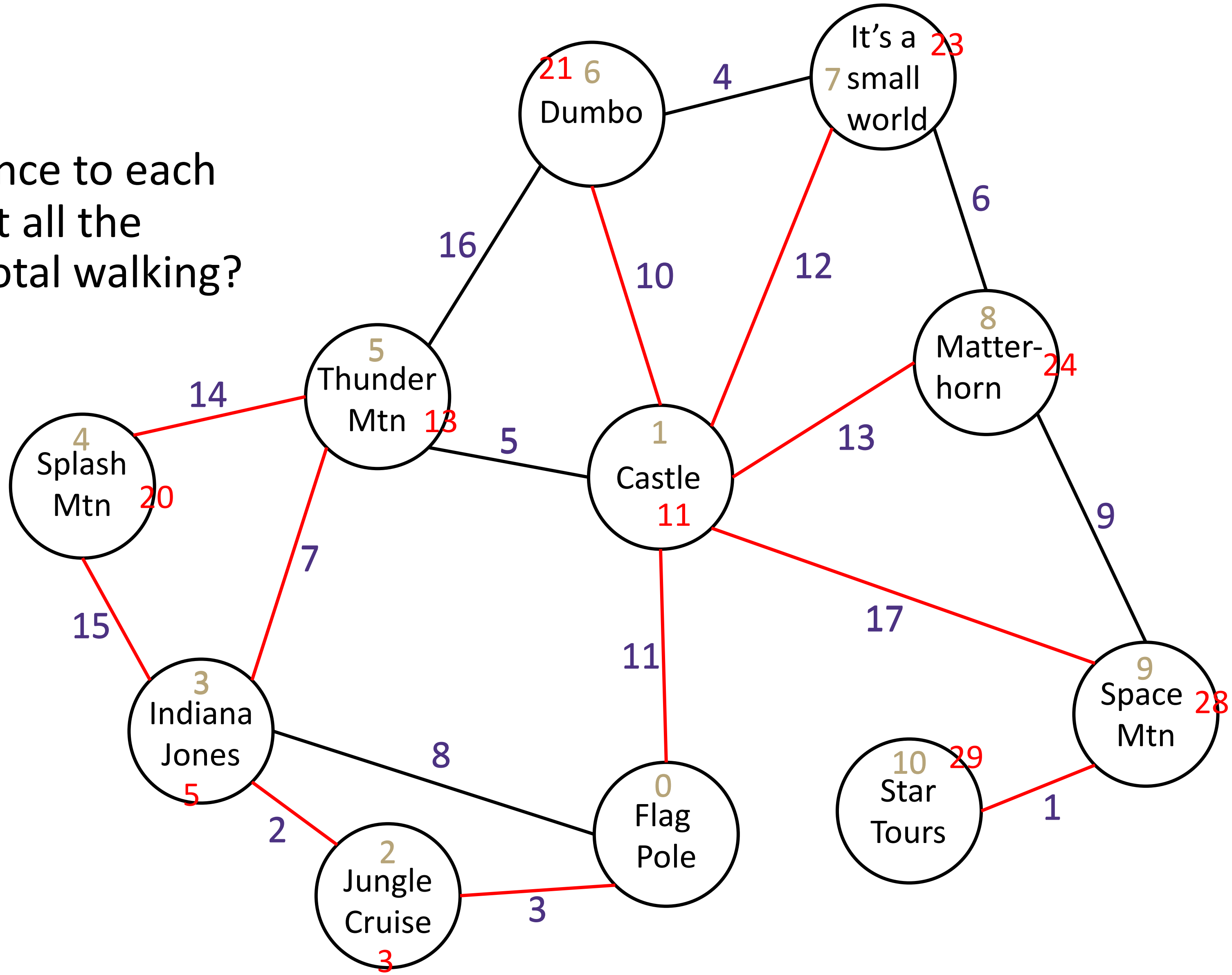
Walkways with distances



Scenario #2b

Now that you know the shortest distance to each attraction, can you make a plan to visit all the attractions with the least amount of total walking?

Nope! This is the travelling salesman problem which is much more complicated than Dijkstra's.
(NP Hard, more on this later)



Scenario #3

You have great taste so you are riding Space Mountain. Your friend makes poor choices so they are riding Splash Mountain. You decide to meet at the castle, how long before you can meet up?

Is there a graph algorithm that would help?

Dijkstra's

What information do our edges need to store?

Walking times

How do we apply the algorithm?

- Run Dijkstra's from Splash Mountain.
- Run Dijkstra's from Space Mountain.
- Take the larger of the two times.