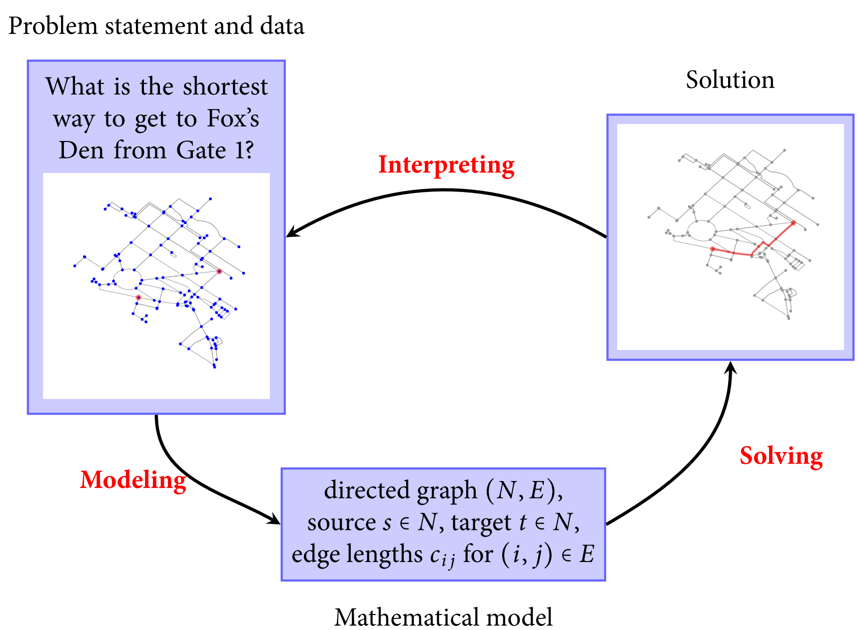
Lesson 1. Introduction and the Shortest Path Problem

1 Goals for this course

- A course in **operations research**: the discipline of applying advanced mathematical methods to help make better decisions
- Formulate **mathematical models** for real-world **decision-making** problems:
  - The shortest path problem
  - Dynamic programming – deterministic and stochastic
- Use **computational tools** to solve these models with medium-to-large scale data
  - Python and its many data science packages (e.g. `pandas`, `networkx`)
  - Focus on
    - setting up models with the help of design patterns
    - analyzing and interpreting solutions
- Analyze and interpret solutions to these models



- Detailed topic list and schedule on the syllabus

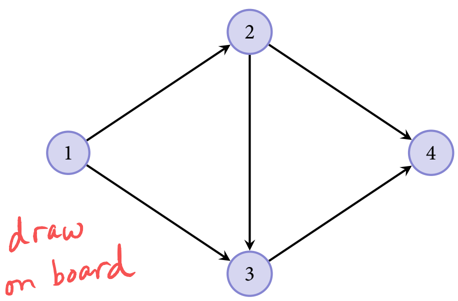
2 This lesson...

- What is the shortest way to get from Point A to Point B?

3 Graphs and networks

- **Graphs** model how various entities are connected
- A **directed graph** (also known as a **digraph**) (N, E) consists of
  - set of **nodes** N (also known as **vertices**)
  - set of **edges** E (also known as **arcs**)
    - edges are directed from one node to another
    - edge from node i to node j is denoted by (i, j)

Example 1.



draw on board

$N = \{1, 2, 3, 4\}$

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

4 Graphs are everywhere

- Physical networks – e.g. road networks
- Abstract networks – e.g. organizational charts
- Others?

5 Paths

- A **path** is a sequence of edges connecting two specified nodes in a graph:
  - Each edge must have exactly one node in common with its predecessor in the sequence
  - Edges must be passed in the forward direction
  - <u>No node may be visited more than once</u>

Example 2. Give some examples of paths from node 1 to node 4 in the network in Example 1.

Paths from node 1 to node 4:

- $(1,2), (2,4)$
- $(1,2), (2,3), (3,4)$
- $(1,3), (3,4)$

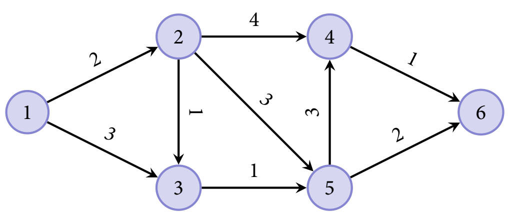
6 The shortest path problem

The shortest path problem

sometimes: "target node"

- Data:
  - Digraph (N, E)
  - **Source node** $s \in N$ and **sink node** $t \in N$ $(s \neq t)$
  - Each edge (i, j) in E has a **length** c_{ij}
- The **length of a path** is the sum of the lengths of the edges in the path
- Problem: What is the shortest path from s to t?

Example 3. Consider the digraph below. The labels next to each edge represent that edge's length. What is the shortest path from node 1 to node 6?



Shortest path:
(1, 3), (3, 5), (5, 6)
also: (1, 2), (2, 3), (3, 5), (5, 6)

Shortest path length: 6

- Natural applications of the shortest path problem:
  - Transportation (road networks, air networks)
  - Telecommunications (computer networks)
- Our focus: not-so-obvious applications of the shortest path problem
- In order to formulate a problem as a shortest path problem, we must specify:
  - (i) a digraph (nodes and edges)
  - (ii) a source and target node
  - (iii) the length of each edge
  - (iv) <u>how any path from the source to the target translates into a solution to the problem</u>

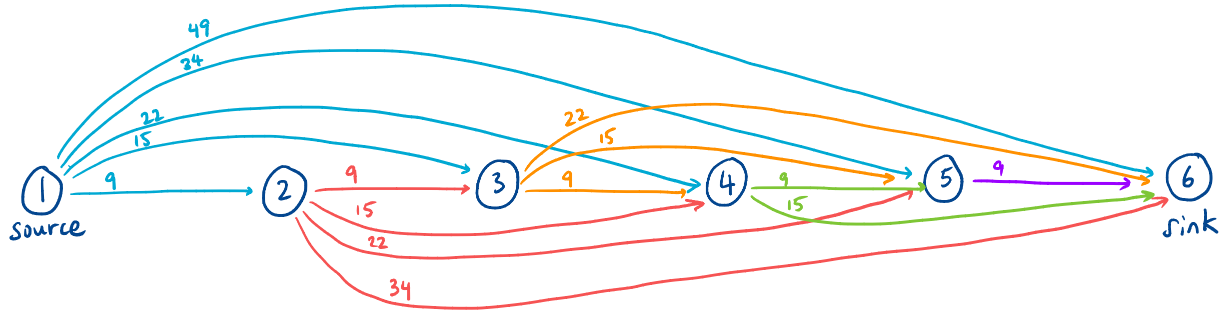
Example 4. You have just purchased a new car for \$22,000. The cost of maintaining a car during a year depends on its age at the beginning of the year:

Age of car (years)	0	1	2	3	4
Annual maintenance cost (\$)	2,000	3,000	4,000	8,000	12,000

To avoid the high maintenance costs associated with an older car, you may trade in your car and purchase a new car. The price you receive on a trade-in depends on the age of the car at the time of the trade-in:

Age of car (years)	1	2	3	4	5
Trade-in price (\$)	15,000	12,000	9,000	5,000	2,000

For now, assume that at any time, it costs \$22,000 to purchase a new car. Your goal is to minimize the net cost (purchasing costs + maintenance costs − money received in trade-ins) incurred over the next five years. Formulate your problem as a shortest path problem.



Node i corresponds to the beginning of year i.

Edge (i,j) corresponds to buying a new car at the beginning of year i and trading it in at the beginning of year j.

Each path from node 1 to node 6 corresponds to a purchasing and maintenance plan over the next 5 years. Each node in a path corresponds to when to buy a new machine. The length of each path is the total cost incurred over 5 years. We find the plan with the minimum total cost by finding the shortest path from node 1 to node 6.

$C_{12} = 22 + 2 - 15 = 9$

$C_{13} = 22 + 2 + 3 - 12 = 15$

$C_{14} = 22 + 2 + 3 + 4 - 9 = 22$

$C_{15} = 22 + 2 + 3 + 4 + 8 - 5 = 34$

$C_{16} = 22 + 2 + 3 + 4 + 8 + 12 - 2 = 49$

$C_{34} = 22 + 2 - 15 = 9$

$C_{35} = 22 + 2 + 3 - 12 = 15$

$C_{36} = 22 + 2 + 3 + 4 - 9 = 22$

$C_{23} = 22 + 2 - 15 = 9$

$C_{24} = 22 + 2 + 3 - 12 = 15$

$C_{25} = 22 + 2 + 3 + 4 - 9 = 22$

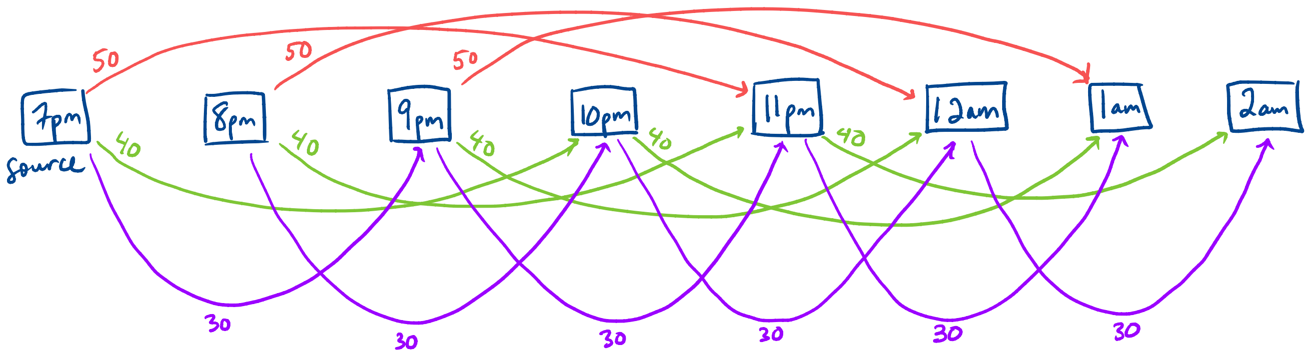
$C_{26} = 22 + 2 + 3 + 4 + 8 - 5 = 34$

$C_{45} = 22 + 2 - 15 = 9$

$C_{46} = 22 + 2 + 3 - 12 = 15$

$C_{56} = 22 + 2 - 15 = 9$

Example 5. The Simplexville College campus shuttle bus begins running at 7:00pm and continues until 2:00am. Several drivers will be used, but only one should be on duty at any time. If a shift starts at or before 9:00pm, a regular driver can be obtained for a 4-hour shift at a cost of \$50. Otherwise, part-time drivers need to be used. Several part-time drivers can work 3-hours shifts at \$40, and the rest are limited to 2-hour shifts at \$30. The college's goal is to schedule drivers in a way that minimizes the total cost of staffing the shuttle bus. Formulate this problem as a shortest path problem.

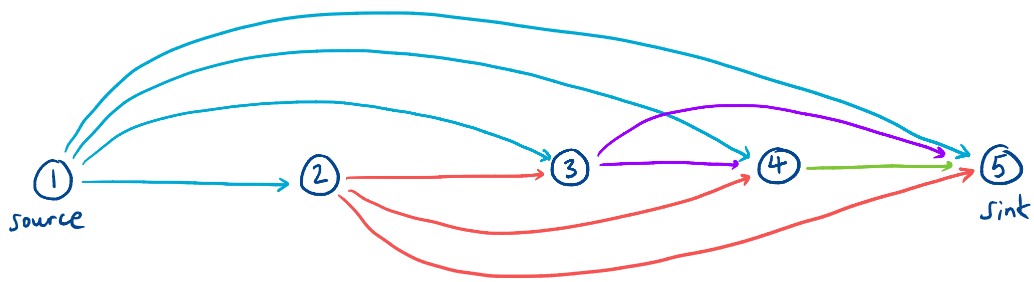


Node i corresponds to starting a new shift at time i

Edge (i,j) corresponds to a shift starting at time i and ending at time j.

Each path from 7pm to 2am corresponds to a shift schedule that covers all hours 7pm - 2am. Each edge in the path corresponds to the shifts that should be used. The length of the path is the total cost of the shifts used. We find a minimum cost shift schedule by finding a shortest path from 7pm to 2am.

Example 6. The Dijkstra Brewing Company is planning production of its new limited run beer, Primal Pilsner. The company must supply 30 batches in the next quarter, then 25, 10, and 35 in successive quarters. Each quarter in which the company produces the beer requires a factory setup cost of \$100,000. Each batch of beer costs \$3,000 to produce. Batches can be held in inventory, but due to refrigeration requirements, the cost is a high \$5,000 per batch per quarter. The company wants to find a production plan that minimizes its total cost. Formuate this problem as a shortest path problem.



Node i represents quarter i

Edge (i,j) represents producing beer in quarter i to meet demand in quarters $i, i+1, \ldots, j-1$.

Each path from node 1 to node 5 represents a production plan over the next 4 quarters. Each edge in the path tells us when to produce beer and for which quarters. The length of the path is the total cost of the plan. We find a minimum cost plan by finding a shortest path from 1 to 5.

$$C_{12} = 100 + 3(30) = 190$$

$$C_{13} = 100 + 3(30) + (3+5)(25) = 390$$

$$C_{14} = 100 + 3(30) + (3+5)(25) + (3+2(5))(10) = 520$$

$$C_{15} = 100 + 3(30) + (3+5)(25) + (3+2(5))(10) + (3+3(5))(35) = 1150$$

$$C_{23} = 100 + 3(25) = 175$$

$$C_{24} = 100 + 3(25) + (3+5)(10) = 255$$

$$C_{25} = 100 + 3(25) + (3+5)(10) + (3+2(5))(35) = 710$$

$$C_{34} = 100 + 3(10) = 130$$

$$C_{35} = 100 + 3(10) + (3+5)(35) = 255$$

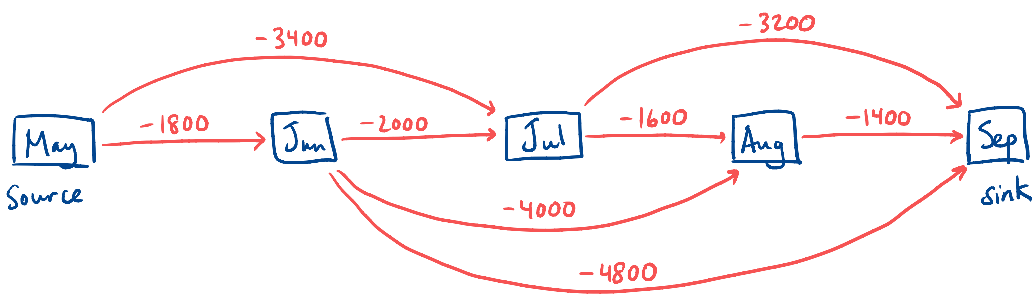
$$C_{45} = 100 + 3(35) = 205$$

Example 7. Beverly owns a vacation home in Cape Fulkerson that she wishes to rent for the summer season (May 1 to September 1). She has solicited bids from eight potential renters:

Renter	Rental start date	Rental end date	Amount of bid ($)
1	May 1	June 1	1800
2	May 1	July 1	3400
3	June 1	July 1	2000
4	June 1	August 1	4000
5	June 1	September 1	4800
6	July 1	August 1	1600
7	July 1	September 1	3200
8	August 1	September 1	1400

A rental starts at 15:00 on the start date, and ends at 12:00 on the end date. As a result, one rental can end and another rental can start on the same day. However, only one renter can occupy the vacation home at any time.

Beverly wants to identify a selection of bids that would maximize her total revenue. Formulate Beverly's problem as a shortest path problem.



Node i represents month i.

Edge (i, j) represents a rental that starts in month i and ends in month j.

Each path from May to Sep represents a set of rentals that covers the entire summer season (with only one rental at a time).

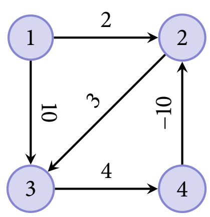
The length of the path is the negative of the total bids for that set of rentals.

By finding the shortest path from May to Sep, we find the rentals with the maximum total bids.

7 Longest paths and negative cycles

- We saw in the previous example that formulating a shortest path problem with negative edge lengths often makes sense, especially when a problem is naturally formulated as a **longest path problem**
- This can sometimes be problematic!

Example 8. Find the shortest path from node 1 to node 4 in the following digraph:



Shortest path : (1,2), (2,3), (3,4)

Shortest path length : 2 + 3 + 4 = 9

- Remember that a path can visit each node at most once
- A **cycle** in a digraph is a path from a source node s to a target node t plus an arc (t, s)
- A **negative cycle** has negative total length
  - For example: $(2, 3)$, $(3, 4)$, $(4, 2)$ in the digraph above

This cycle has length = −3

- Negative cycles make things complicated: if we traverse a negative cycle, we can reduce the cost of getting point A to point B infinitely
- Shortest path problems with negative cycles harder to solve
  - Standard shortest path algorithms fail when the digraph has a negative cycle
- Having a negative cycle in your shortest path problem might indicate (i) your problem will be hard to solve, or (ii) there is a mistake in your formulation!