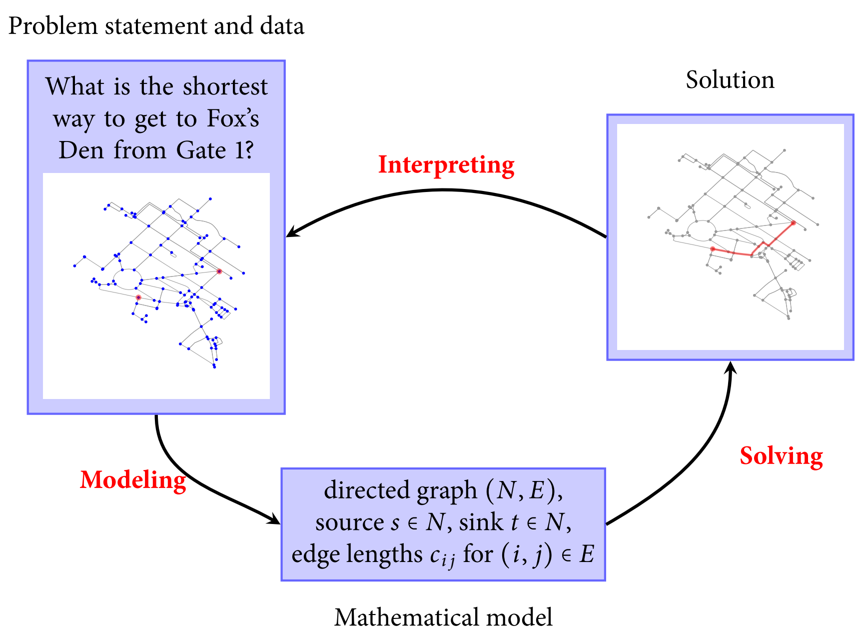
Lesson 1. Introduction and the Shortest Path Problem

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

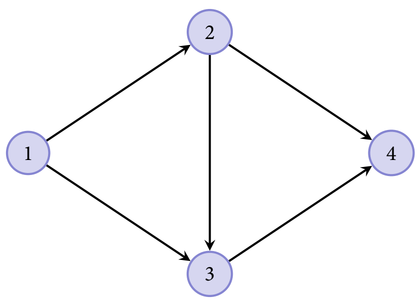
2 This lesson...

- 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.



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

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

- Graphs are everywhere!
  - Physical networks – e.g. road networks
  - Abstract networks – e.g. organizational charts
  - Others?

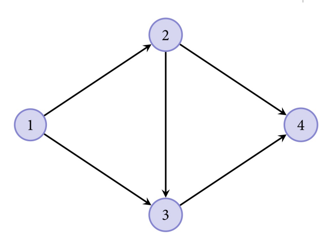
4 Paths

- 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
  - No node may be visited more than once

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:

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

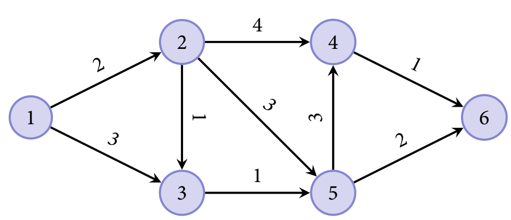


5 The shortest path problem

The shortest path problem

- Data:
  - Digraph (N, E)
  - **Source node** $s \in N$ and **sink node** $t \in N$ $(s \neq t)$
    (The sink node is often called the **target node**)
  - 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:
(1,2), (2,3), (3,5), (5,6)
also: (1,3), (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:
  - a digraph (nodes and edges)
  - a source and sink node
  - the length of each edge
  - how to translate (i) the length of a shortest path and (ii) the nodes/edges in a shortest path into a solution to the problem

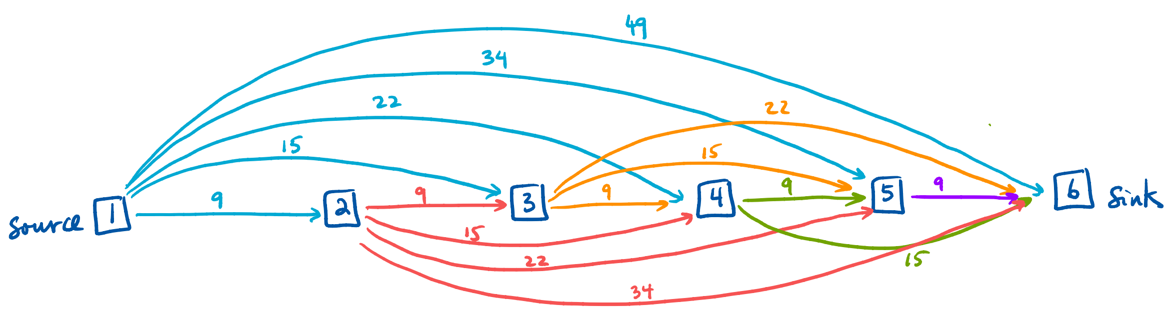
Example 4. You have just purchased a new car for \$22,000. The cost of maintaining a car during a year depends on its age at the beginning of the year:

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 $\leftrightarrow$ beginning of year i

Edge (i,j) $\leftrightarrow$ buying a new car at the beginning of year i, and trading it in at the beginning of year j

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

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

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

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

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

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

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

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

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

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

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

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

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

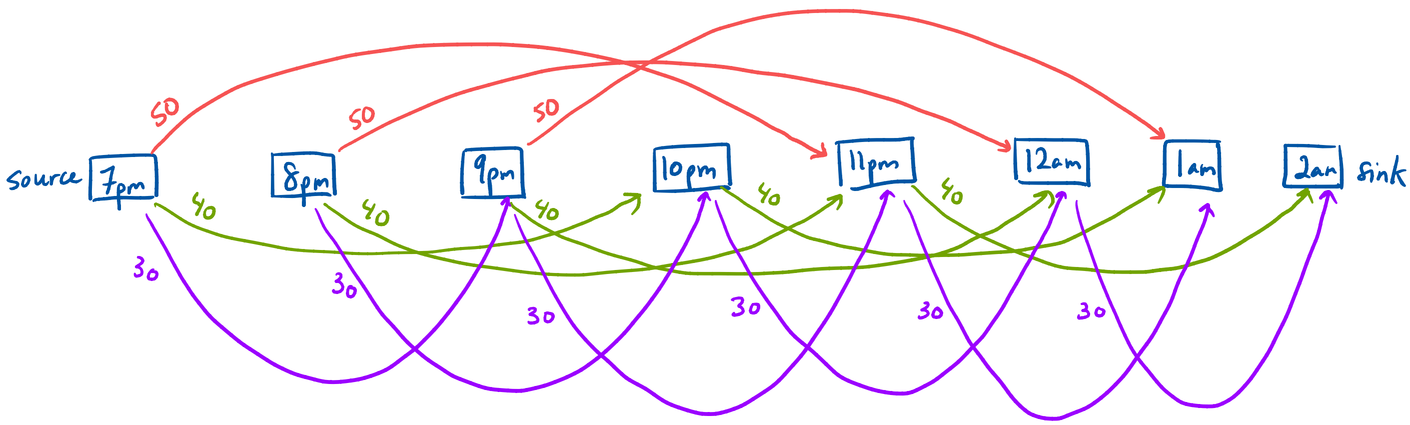
$c_{46} = 22 + 2 + 3 - 12 = 15$

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

Shortest path length $\leftrightarrow$ minimum total net cost incurred over the 5 year period.

Nodes in SP $\leftrightarrow$ when to buy a new car. For example, if $(1,2), (2,4), (4,6)$ is a SP, then we should buy a new car in years $1, 2, 4$.

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 $\leftrightarrow$ starting a shift at time i

Edge (i,j) $\leftrightarrow$ shift that starts at time i and ends at time j.

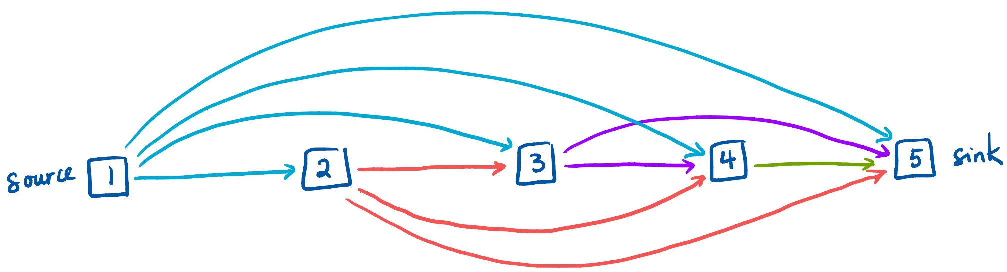
SP length $\leftrightarrow$ minimum total cost of staffing the bus.

Edges in a SP $\leftrightarrow$ which shifts should be used

For example, suppose (7pm, 11pm), (11pm, 2am) is a SP.

Then, the college should use 2 shifts: 4-hr shift starting at 7pm
3-hr shift starting at 11pm

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 $i \leftrightarrow$ beginning of quarter i

Edge $(i,j) \leftrightarrow$ producing beer in quarter i to meet demand in quarters $i, \dots, j-1$

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

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

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

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

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

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

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

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

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

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

SP length $\leftrightarrow$ minimum total production cost

Edges in SP $\leftrightarrow$ when to produce beer and for which quarters

For example, suppose $(1,2), (2,4), (4,5)$ is a SP.

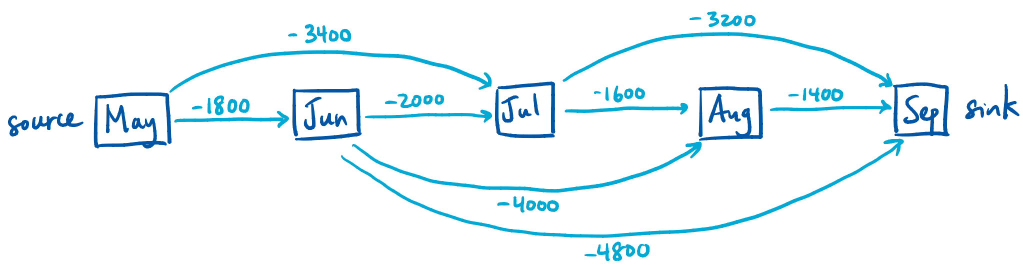
Then, the company should produce: in Q1 for Q1
in Q2 for Q2 + Q3
in Q4 for Q4

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 ↔ month i

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

SP length ↔ the negative of the maximum total revenue from rentals

Edges in SP ↔ which bids to take

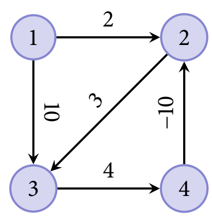
For example, suppose (May, Jul), (Jul, Aug), (Aug, Sep) is a SP.

Then Beverly should accept bids from renters 2, 6, 8

6 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:



SP: (1,2), (2,3), (3,4)

SP 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 sink node t plus an arc (t, s)
- A **negative cycle** has negative total length

  this cycle has length $3 + 4 - 10 = -3$

  - For example: $(2,3), (3,4), (4,2)$ in the digraph above
- Negative cycles make things complicated: if we traverse a negative cycle, we can reduce the cost of getting from 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!

A Problems

Problem 1 (Primal Praline Company)**.** The Primal Praline Company needs to have a working candy making machine during each of the next six years. Currently, it has a new machine. At the beginning of each year, the company may keep the machine or sell it and buy a new one. A new machine costs $5000, and cannot be kept for more than three years. The revenues earned by a machine, the cost of maintaining it, and the salvage value that can be obtained by selling it at the end of a year depend on the age of the machine:

	Age of machine at the beginning of the year		
	0 years	1 years	2 years
Revenues	4,500	3,000	1,500
Operating costs	500	700	1,100
Salvage value at the end of year	3,000	1,800	500

The company's problem is to maximize the net profit it earns over the next six years. Formulate this problem as a shortest path problem.

Problem 2 (Shapley Sneakers)**.** The Shapley Sneaker Company is opening a new factory in Simplexville. One of its major purchases will be a leather cutting machine. The cost of maintaining a machine depends on its age as follows:

Age at beginning of year (years)	0	1	2	3	4
Maintenance cost for next year ($)	38,000	50,000	97,000	182,000	304,000

The cost of purchasing a machine at the beginning of each year is:

Year	1	2	3	4	5
Purchase cost ($)	170,000	190,000	210,000	250,000	300,000

There is no trade-in value when a machine is replaced. The company's goal is to minimize the total cost (purchase plus maintenance) of having a machine for five years. Formulate this problem as a shortest path problem.