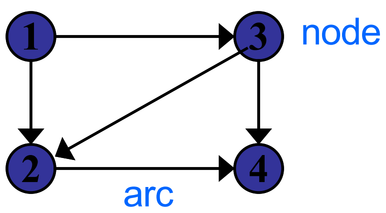
Introduction to Optimization
Lecture 5
Network Optimization

Outline

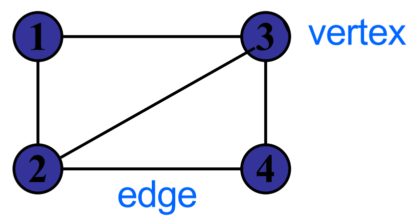
- Networks and Graphs
  - Terminology and Applications
- The General Min-Cost Network Flow (MCNF) formulations
  - Transportation, Assignment, Transhipment, Max Flow, and Shortest Path
- Currency Arbitrage
  - Flow problem

Network terminology

- Network terminology is not standardized
  - Directed and undirected networks
  - The terms ‘network’ and ‘graph’ are often used interchangeably



Directed



Undirected

$G = (V,E)$ – Network

$V = \{1, 2, 3, 4\}$ – Node/vertex set

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

Some application areas

Applications	**Nodes**	**Arcs**	**Flow**
Communication systems	Phone exchanges, Computers, Transmission facilities, Satellites	Cables, fiber optic links, microwave relay links	Voice messages, data, video
Hydraulic systems	Pumping stations reservoirs, lakes	Pipelines	Water, gas, oil, hydraulic fluids
Integrated computer circuits	Gates, registers, processors	Wires	Electrical current
Mechanical systems	Joints	Rods, beams, springs	Heat, energy
Transportation systems	Intersections, airports, rail stations	Highways, airline routes, tracks	Passengers, freight, vehicles, operators

A Transportation Problem

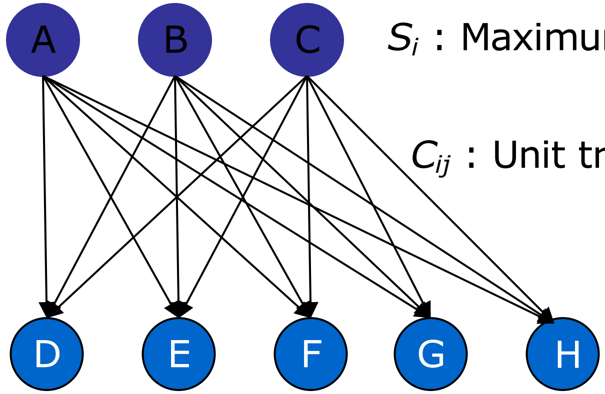
Costs	D	E	F	G	H	*Supply*
A	16	7	17	14	19	70
B	9	11	16	10	5	50
C	10	18	6	13	8	30
Demand	30	24	12	42	6	

$S_A=70$

i=A i=B i=C

Supply index '*i*'

Supply



S_i : Maximum supply available at '*i*'

C_{ij} : Unit transportation from *i* to *j*

D_j : Demand at '*j*'

$D_H=6$

Demand

j=D j=E j=F j=G j=H

Demand index '*j*'

The Transportation Problem in English

- What is the problem statement?
  – "Minimize transportation costs ensuring supply capacities and demand requirements are not violated"

- What are the decision variables?

- What are the constraints?

The complete LP formulation

- Decision variables
  - X_{ij} : The amount supplied from supply location '*i*' to demand location '*j*' where $i = A, B, or\ C$ and $j = D, E, F, G, or\ H$
- The objective function $\left(\min \sum_{i,j} C_{ij} X_{ij}\right)$
  - $16\ X_{AD} + 7\ X_{AE} + 17\ X_{AF} + 14\ X_{AG} + 19\ X_{AH} + 9\ X_{BD} + 11\ X_{BE} + 16\ X_{BF} + 10\ X_{BG} + 5\ X_{BH} + 10\ X_{CD} + 18\ X_{CE} + 6\ X_{CF} + 13\ X_{CG} + 8\ X_{CH}$

The complete LP formulation

- Supply capacity constraints

$$\sum_j X_{ij} \le S_i \quad \text{for all } i = A, B, C$$

  - $X_{AD}+X_{AE}+X_{AF}+X_{AG}+X_{AH} <= 70$
  - $X_{BD}+X_{BE}+X_{BF}+X_{BG}+X_{BH} <= 50$
  - $X_{CD}+X_{CE}+X_{CF}+X_{CG}+X_{CH} <= 30$

- Demand constraints

$$\sum_i X_{ij} \ge D_j \quad \text{for all } j = D, \ldots, H$$

  - $X_{AD}+X_{BD}+X_{CD} >= 30$
  - $X_{AE}+X_{BE}+X_{CE} >= 24$
  - $X_{AF}+X_{BF}+X_{CF} >= 12$
  - $X_{AG}+X_{BG}+X_{CG} >= 42$
  - $X_{AH}+X_{BH}+X_{CH} >= 6$

- Non-negativity constraints

$$X_{ij} \ge 0 \quad \text{for all } i = A, B, C \text{ and } j = D, \ldots, H$$

Optimal Solution

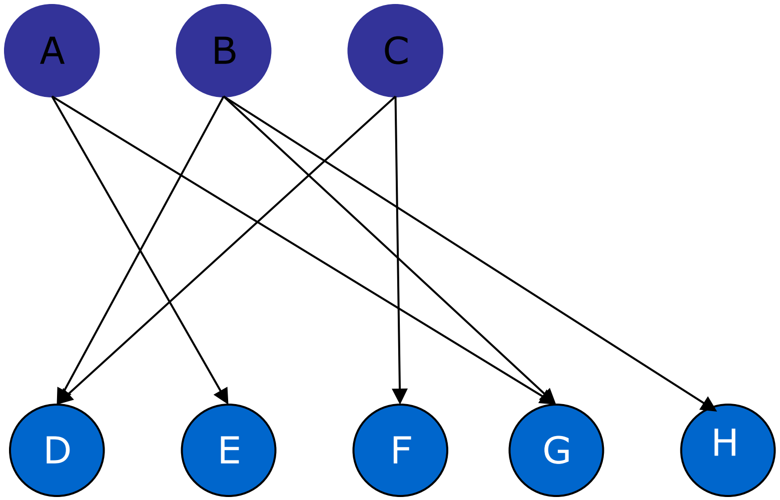
	D	E	F	G	H	*Sent*	*Supply*	Z = Total (
A	**0**	**24**	**0**	**10**	**0**	**34**	70	**1018**
B	**12**	**0**	**0**	**32**	**6**	**50**	50	
C	**18**	**0**	**12**	**0**	**0**	**30**	30	
Received	**30**	**24**	**12**	**42**	**6**			
Demand	30	24	12	42	6			

- What does the solution look like?
  - Do we need to use all available routes? Why?
- Which constraints are tight?
  - Demand constraint? Why?
  - Supply constraints? Why?

Understanding the optimal solution

	A	B	C
Supply	70	50	30

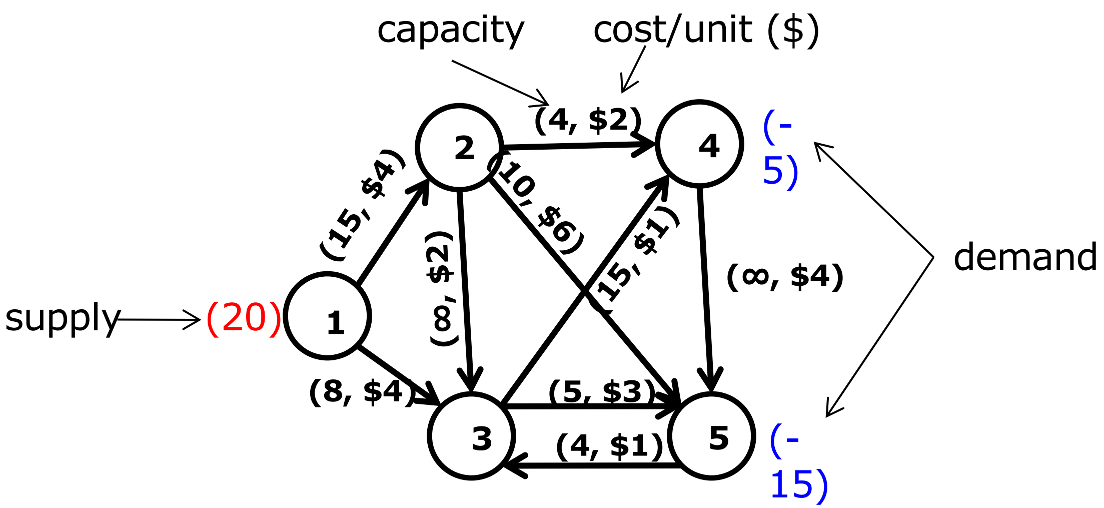
ONLY 7 arcs required.



	D	E	F	G	H
Demand	30	24	12	42	6

Generalized Min-Cost Flow / Capacitated Transshipment

- Consider a general network with arc capacities, costs, supplies and demands. Our goal is to satisfy the demands from the supply node. We use the –ve sign for demand as a convention.

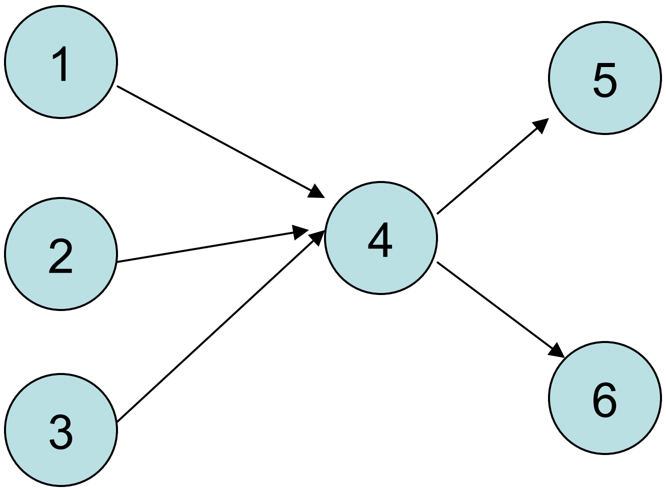


Flow conservation

- Three types of nodes:
  - Transhipment node
    - Flow conservation. Net flow is zero. $b_i = 0$
  - Sink node
    - Positive flow out of a sink node $b_i > 0$
  - Source node
    - Positive flow into a source node $b_i < 0$

Flow conservation

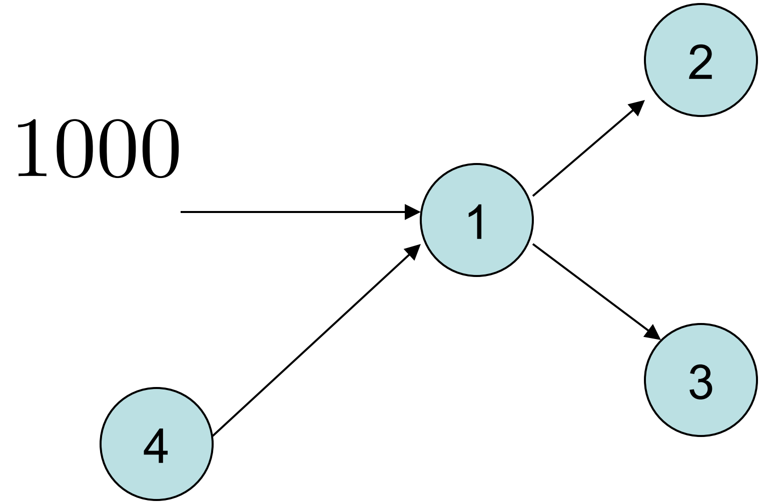
- Transhipment node



$$x_{14} + x_{24} + x_{34} - x_{45} - x_{46} = 0 : \text{Node } 4$$

Flow conservation

- Source nodes

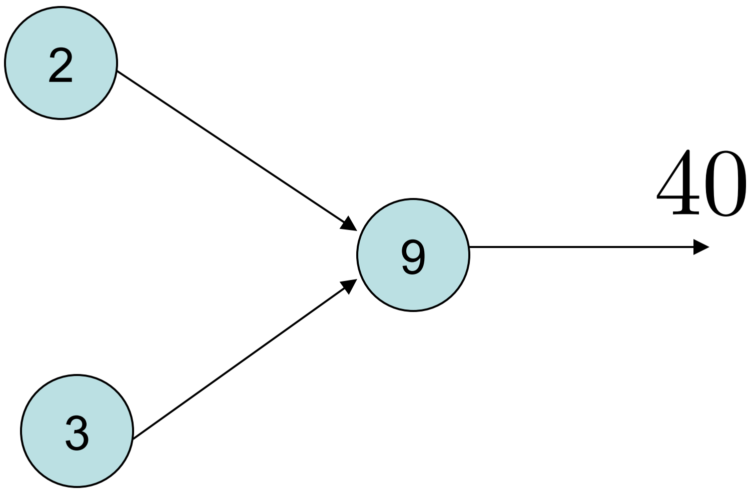


$$1000 + x_{41} - x_{12} - x_{13} = 0 : \text{Node } 1$$

$$x_{41} - x_{12} - x_{13} = -1000 : \text{Node } 1$$

Flow conservation

- Sink nodes



$$x_{29} + x_{39} - 40 = 0 : \text{Node } 9$$

$$x_{29} + x_{39} = 40 : \text{Node } 9$$

Min-Cost Flow Decision Variables

- What are the decision variables?
  - As with the transportation problem, the decision variables here are the amount we ship from node '*i*' to node '*j*' and is denoted by X_{ij}
  - X_{12} , X_{13} , X_{23} , X_{24} , X_{25} , X_{34} , X_{35} , X_{45} , X_{53}
- What are the constraints?
  - *Flow balance* constraints (*Flow conservation* at each node)
  - *Upper bound on flow constraints*
  - *Non-negativity constraints*

The MCNF LP Formulation

$$
\begin{aligned}
\min \quad & 4X_{12} + 4X_{13} + 2X_{23} + 2X_{24} + 6X_{25} + X_{34} + 3X_{35} + 2X_{45} + X_{53} \\
s.t. \quad & X_{12} + X_{13} = 20 \\
& X_{23} + X_{24} + X_{25} - X_{12} = 0 \\
& X_{34} + X_{35} - X_{13} - X_{23} - X_{53} = 0 \\
& X_{45} - X_{24} - X_{34} = -5 \\
& X_{53} - X_{25} - X_{35} - X_{45} = -15 \\
& X_{12} \le 15,\ X_{13} \le 8,\ X_{24} \le 4,\ X_{25} \le 10 \\
& X_{34} \le 15,\ X_{35} \le 5,\ X_{53} \le 4 \\
& X_{12},\ X_{13},\ X_{24},\ X_{25},\ X_{34},\ X_{35},\ X_{53} \ge 0
\end{aligned}
$$

Intermediary (transit) nodes: rhs = 0

Flow balance (One constraint for every node)

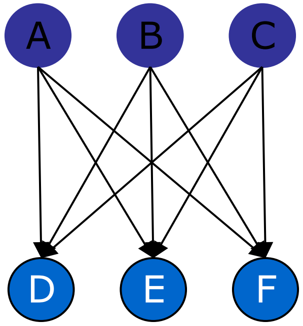
Capacity bound (One constraint for every arc)

A Special Transportation Problem

- Suppose all the supply and demands are exactly 1 unit and the number of supply nodes is exactly equal to the number of demand nodes

$S_i = 1$

Supply



Demand

$D_j = 1$

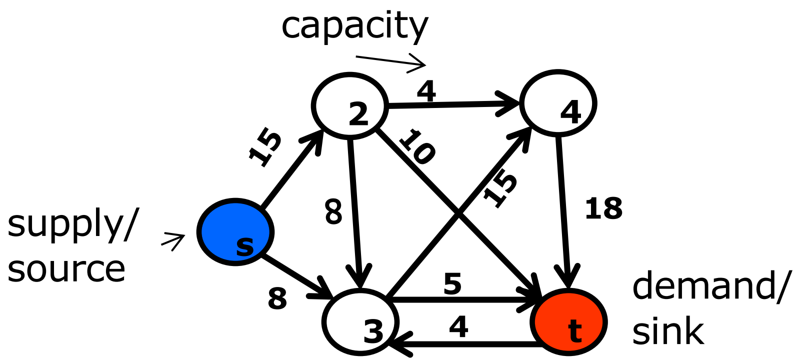
C_{ij} : Unit cost corresponding to arc (i,j)

Supply at each supply-node is 1 unit

Demand at each demand-node is 1 unit

Special Cases of the MCNF Problems: Max Flow

- Suppose a network has only one source node (blue node ‘s’) and one sink node (red node ‘t’). All other nodes are intermediary nodes (transit nodes). Furthermore, no costs are associated with the flows on each arc. We wish to send maximum possible material from ‘s’ to ‘t’ subject to the capacity constraints on each arc. This is a *Max-Flow problem.*



Max-Flow

- In this case the maximum flow achievable from '*s*' to '*t*' is also a variable (we wish to maximize this flow)
  - Let us denote this flow as '*f*'
  - This will be the total outflow from '*s*' and will be the total in-flow into '*t*'
- Thus, the MCNF problem now transforms to:

Max-Flow Formulation

$$
\begin{aligned}
\max \quad & f \\
s.t. \quad & X_{s2} + X_{s3} = f \\
& \boxed{\begin{aligned} & X_{23} + X_{24} + X_{2t} - X_{12} = 0 \\ & X_{34} + X_{35} - X_{13} - X_{23} - X_{t3} = 0 \\ & X_{4t} - X_{24} - X_{34} = 0 \end{aligned}} \\
& X_{t3} - X_{2t} - X_{3t} - X_{4t} = -f \\
& X_{s2} \leq 15,\ X_{s3} \leq 8,\ X_{24} \leq 4,\ X_{2t} \leq 10 \\
& X_{34} \leq 15,\ X_{3t} \leq 5,\ X_{t3} \leq 4 \\
& X_{s2},\ X_{s3},\ X_{24},\ X_{23},\ X_{2t},\ X_{34},\ X_{3t},\ X_{4t},\ X_{t3} \geq 0
\end{aligned}
$$

Flow out of 's'
(flow balance at 's')

Flow balance at intermediate nodes 2, 3, and 4

Flow into 't'
(flow balance at 't')

Capacity constraints

NOTICE: The only difference from the earlier formulations is in the RHS of the constraints and the objective function.

Integral solutions in network flows models

- **Theorem:** Suppose the RHS values of the network LPs are integers, then there exists an optimal solution whose values are all integers.
  - Not necessary to enforce integrality constraints as it comes for free!!
  - **Implication:** Network flows are special integer programs (IP) that are easy to solve.

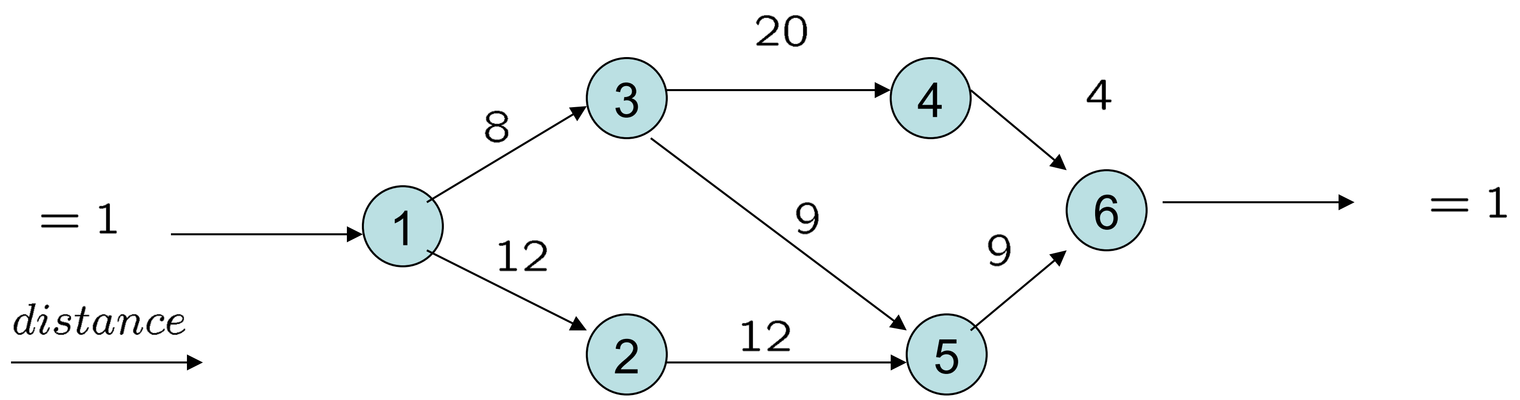
Example: Passenger Routing

- United Airlines has seven daily flights from BOS to SFO, every two hours, starting at 7am.
- Capacities are 100, 100, 100, 150, 150,150 and ∞
- Passengers sufferring from over booking are diverted to later flights.
- Delayed passengers get \$200 plus \$20 for every hour of delay.
- Suppose that today the first six flights have 110, 160, 103, 149, 175, and 140 confirmed reservations.

Determine the most economical passenger routing strategy.

Shortest Path Problem

- Shortest path problem can be formulated as a network flow problem
  - Source node: Start of City. 1 unit of in flow
  - Sink node: End of City. 1 unit of out flow
  - Cost of Arc: distance of arc linking the cities

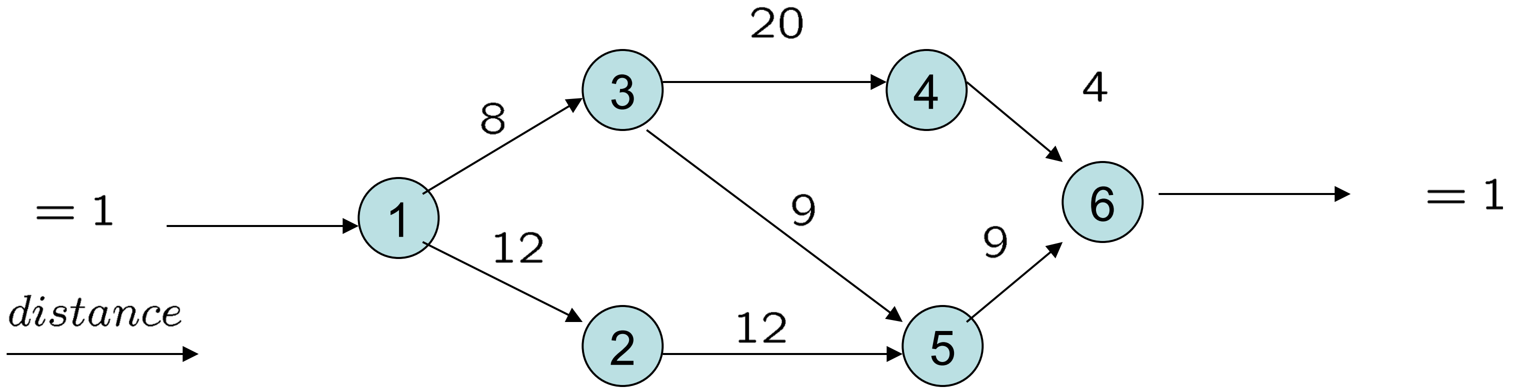


NUS National University of Singapore | NUS BUSINESS SCHOOL

Shortest Path Problem

- Network flow formation
  - Costs must be positive

$$
\begin{aligned}
\min \quad & \sum_{(i,j)\in A} c_{ij} x_{ij} \\
\text{s.t.} \quad & \sum_{j:(j,i)\in A} x_{ji} - \sum_{j:(i,j)\in A} x_{ij} = \begin{cases} -1 & \text{if } i = s \\ 1 & \text{if } i = t \\ 0 & \text{otherwise} \end{cases} & \forall i \in N \\
& x_{ij} \geq 0 & \forall (i,j) \in A
\end{aligned}
$$



20
8
4
3
4
1
6
= 1
= 1
12
9
9
distance
2
12
5

Longest Path Problem?

- Can this be used to find the longest path?
  - Generally no. If the network contains a cycle, the objective will be infinite!

$$
\begin{aligned}
\max \quad & \sum_{(i,j)\in A} c_{ij} x_{ij} \\
\text{s.t.} \quad & \sum_{j:(j,i)\in A} x_{ji} - \sum_{j:(i,j)\in A} x_{ij} = \begin{cases} -1 & \text{if } i = s \\ 1 & \text{if } i = t \\ 0 & \text{otherwise} \end{cases} && \forall i \in N \\
& x_{ij} \geq 0 && \forall (i,j) \in A
\end{aligned}
$$

Project Management

- A project consists of a set of jobs and a set of precedence relations
- Given a set A of jobs pairs (i, j) indicating that job j cannot start before job i is completed.
- c_i duration of job i
- Find the least possible duration of the project.

Project Management

- Introduce two artificial jobs s and t, of zero duration, that signify the begining and completion of the project. Add (s,i) and (i,t) to A
- p_i time at job i begins
- $(i,j) \in A : p_j \geq p_i + c_i$
- Project duration: $p_t - p_s$

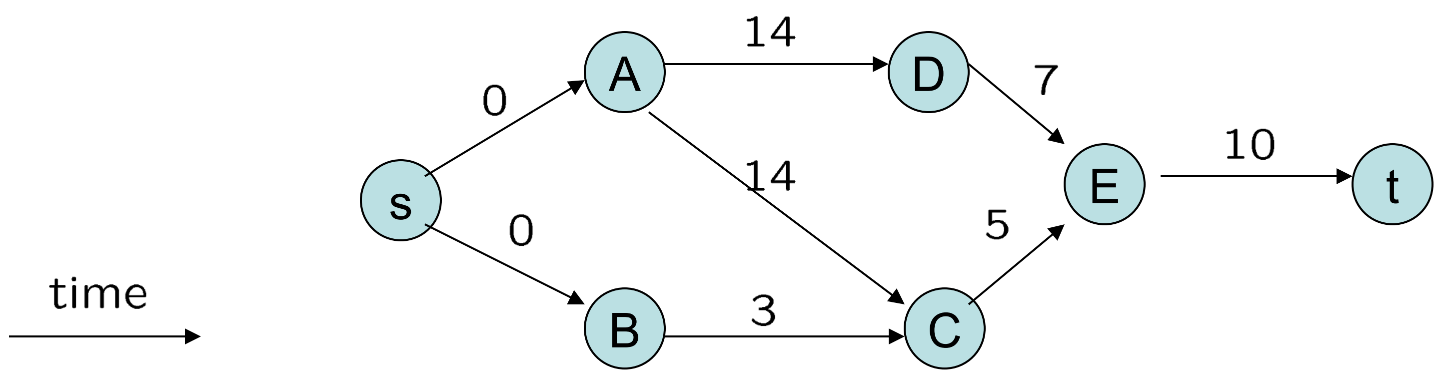
Project Management

- Completion time of project

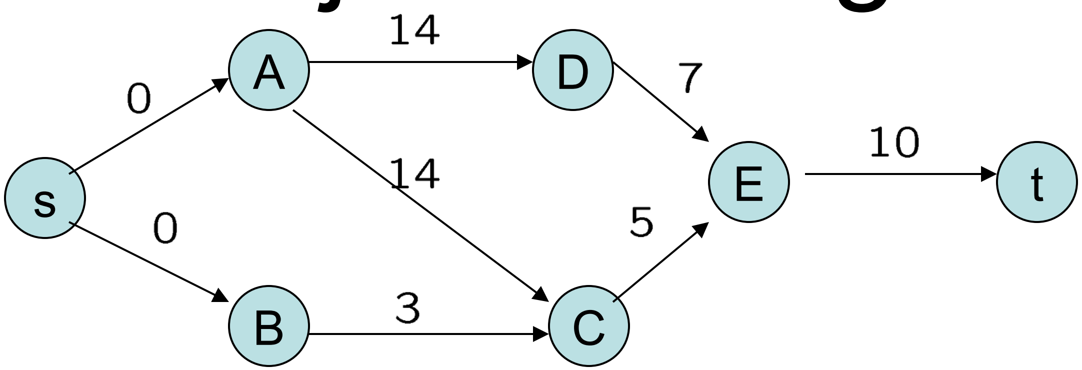
$$\begin{aligned} \min \quad & p_t - p_s \\ \text{s.t.} \quad & p_j - p_i \geq c_i \quad \forall (i,j) \in A \end{aligned}$$

Example: Project Management

Activity	Immediate Predecessor	Time (c_i)
s		0
A	s	14
B	s	3
C	A,B	5
D	A	7
E	C,D	10
t	E	0



Example: Project Management



$$\begin{array}{ll} \min & p_t - p_s \\ \text{s.t.} & p_j - p_i \geq c_i \quad \forall (i,j) \in A \end{array}$$

Project Management

- Completion time of project
  - Primal formulation:

$$\begin{array}{ll} \min & p_t - p_s \\ \text{s.t.} & p_j - p_i \geq c_i \quad \forall (i,j) \in A \end{array}$$

  - Dual formulation

$$\begin{array}{lll} \max & \displaystyle\sum_{(i,j)\in A} c_i x_{ij} & \\ \text{s.t.} & \displaystyle\sum_{j:(j,i)\in A} x_{ji} - \sum_{j:(i,j)\in A} x_{ij} = \begin{cases} -1 & \text{if } i = s \\ 1 & \text{if } i = t \\ 0 & \text{otherwise} \end{cases} & \forall i \in N \\ & x_{ij} \geq 0 & \forall (i,j) \in A \end{array}$$

Project Management

- The dual problem corresponds to a longest path problem
  - It works here because the network is acyclic!! Why?

$$
\begin{aligned}
\max \quad & \sum_{(i,j)\in A} c_i x_{ij} \\
\text{s.t.} \quad & \sum_{j:(j,i)\in A} x_{ji} - \sum_{j:(i,j)\in A} x_{ij} = \begin{cases} -1 & \text{if } i = s \\ 1 & \text{if } i = t \\ 0 & \text{otherwise} \end{cases} && \forall i \in N \\
& x_{ij} \geq 0 && \forall (i,j) \in A
\end{aligned}
$$

Project Management

- The longest path is important because it reveals the set of activities that are critical to project.
  – Any delay to an activity in the critical path may increase the completion time of the project.