15.053/8 **April 23, 2013**

The Minimum Cost Flow Problem

Quotes of the day

"A process cannot be understood by stopping it. Understanding must move with the flow of the process, must join it and flow with it."

-- Frank Herbert

"No question is so difficult to answer as that to which the answer is obvious."

-- George Bernard Shaw

Overview of lecture

- More examples of networks

- Examples of flows
  - movement of goods from one location to another.

- Why flows is such an important example of linear and/or integer programs
  - integrality property

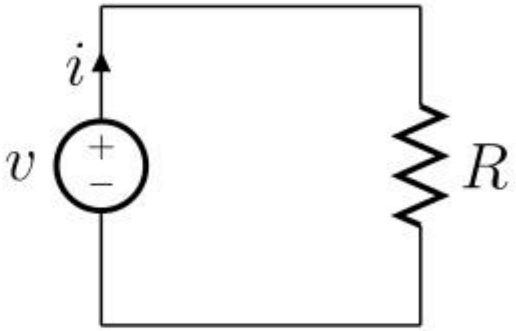
- Coverage of lecture is for broader knowledge than is covered on the quiz on networks.

Networks are everywhere

- Physical networks
- Time space networks
- Connections between concepts
- Social networks
- Network flows: model movements in networks

Road Network

Photo courtesy of Derrick Coetzee on Flickr.

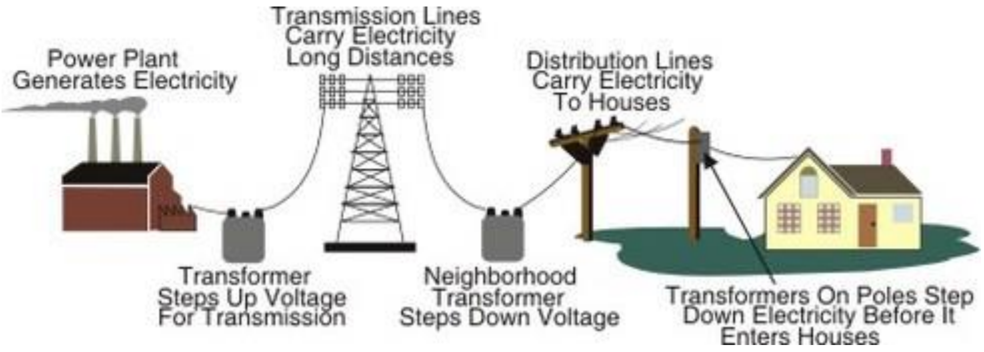


Electrical Network

Public domain image (Wikimedia Commons)

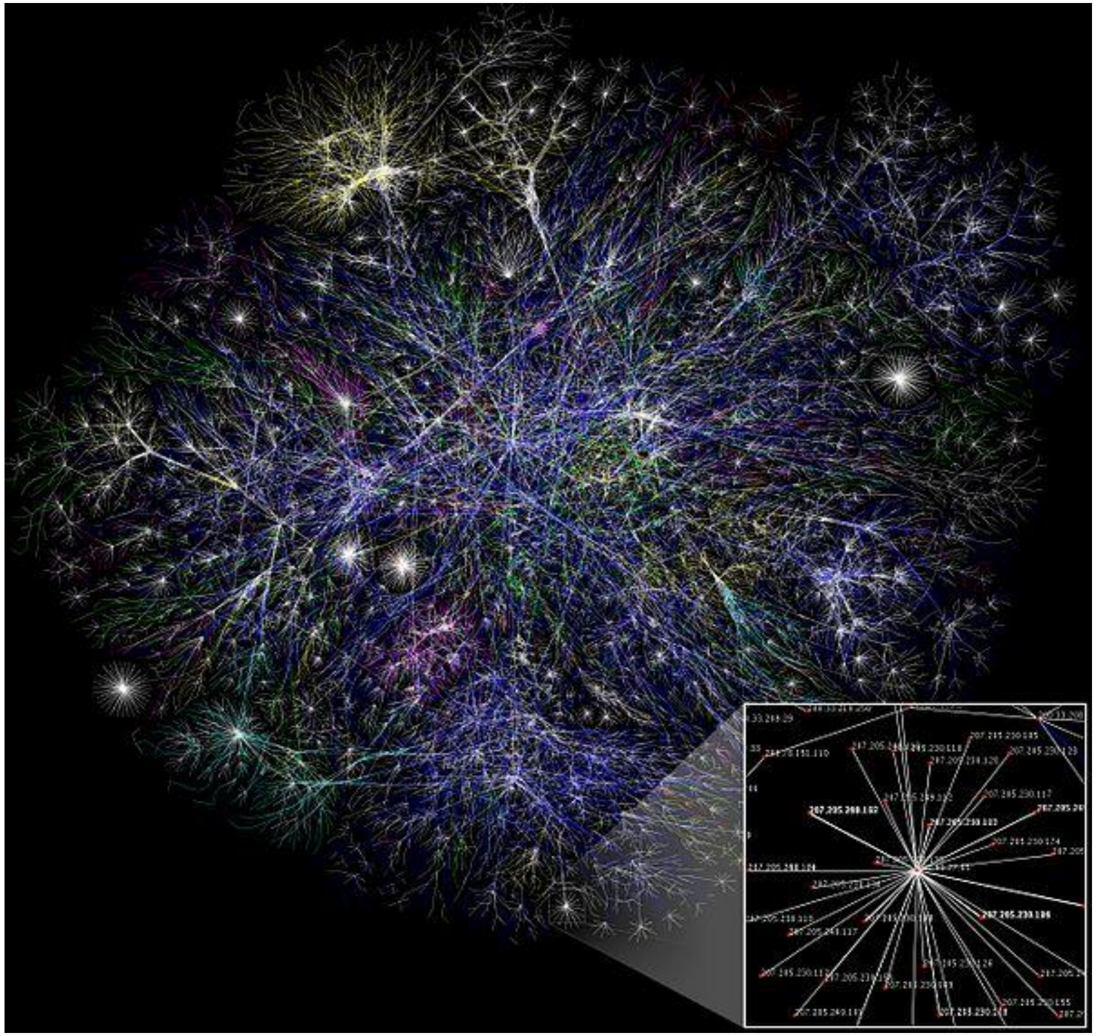
Image removed due to copyright restrictions.

Boston MBTA Train Map



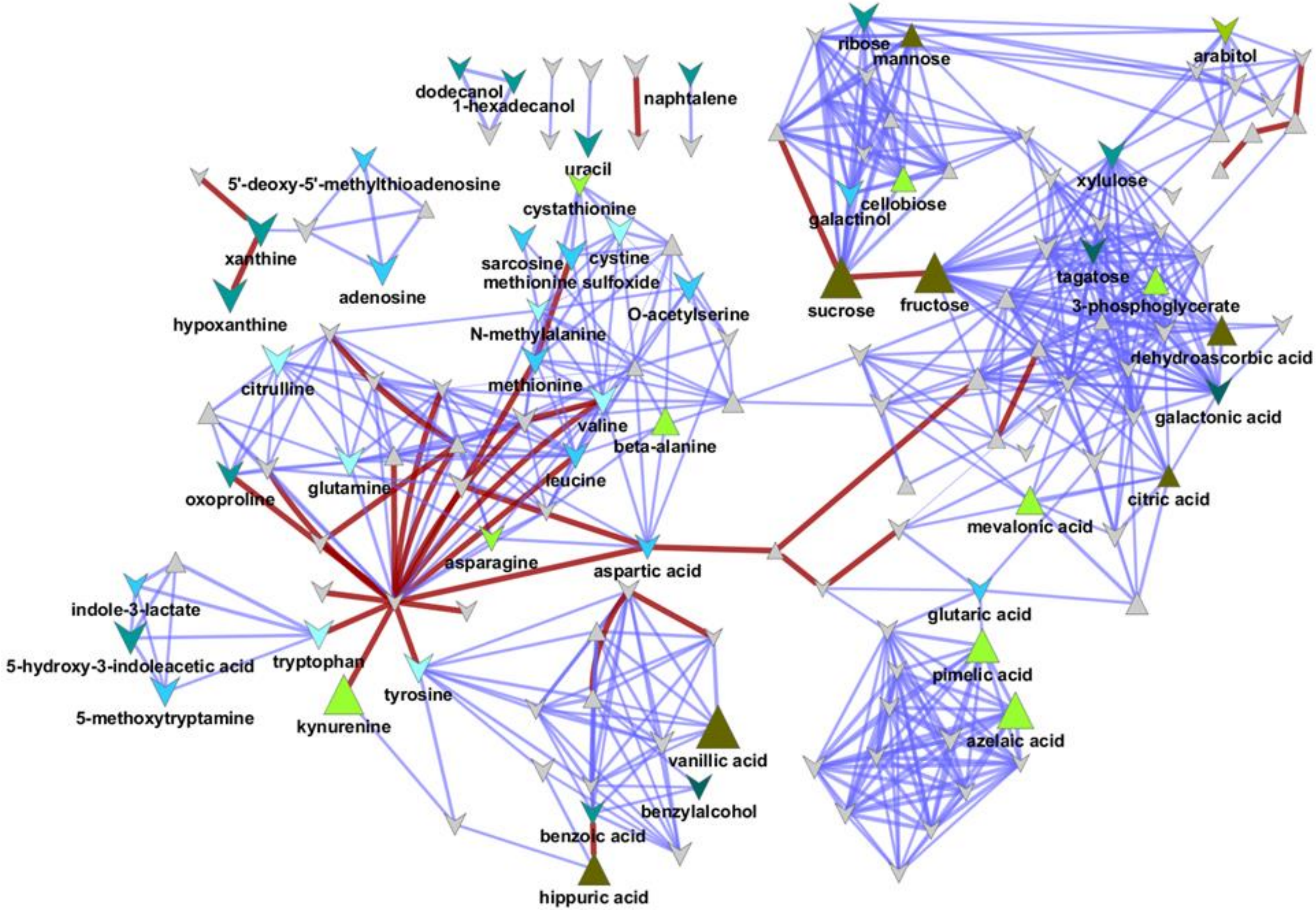
Power Grid

Public domain image (EIA.gov)



Internet

from wikipedia

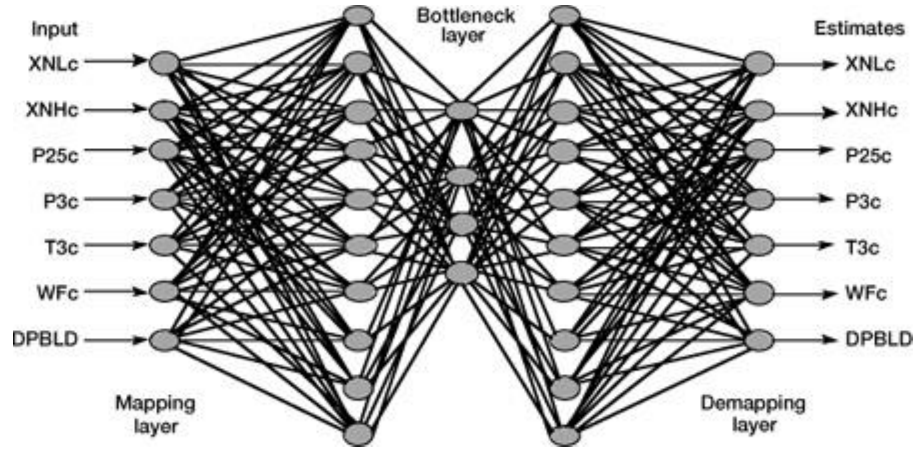


Biological Network

imdevsoftware.wordpress.com

Biological neural network

Public domain (NIH)



Computer neural network

Public domain (NASA)

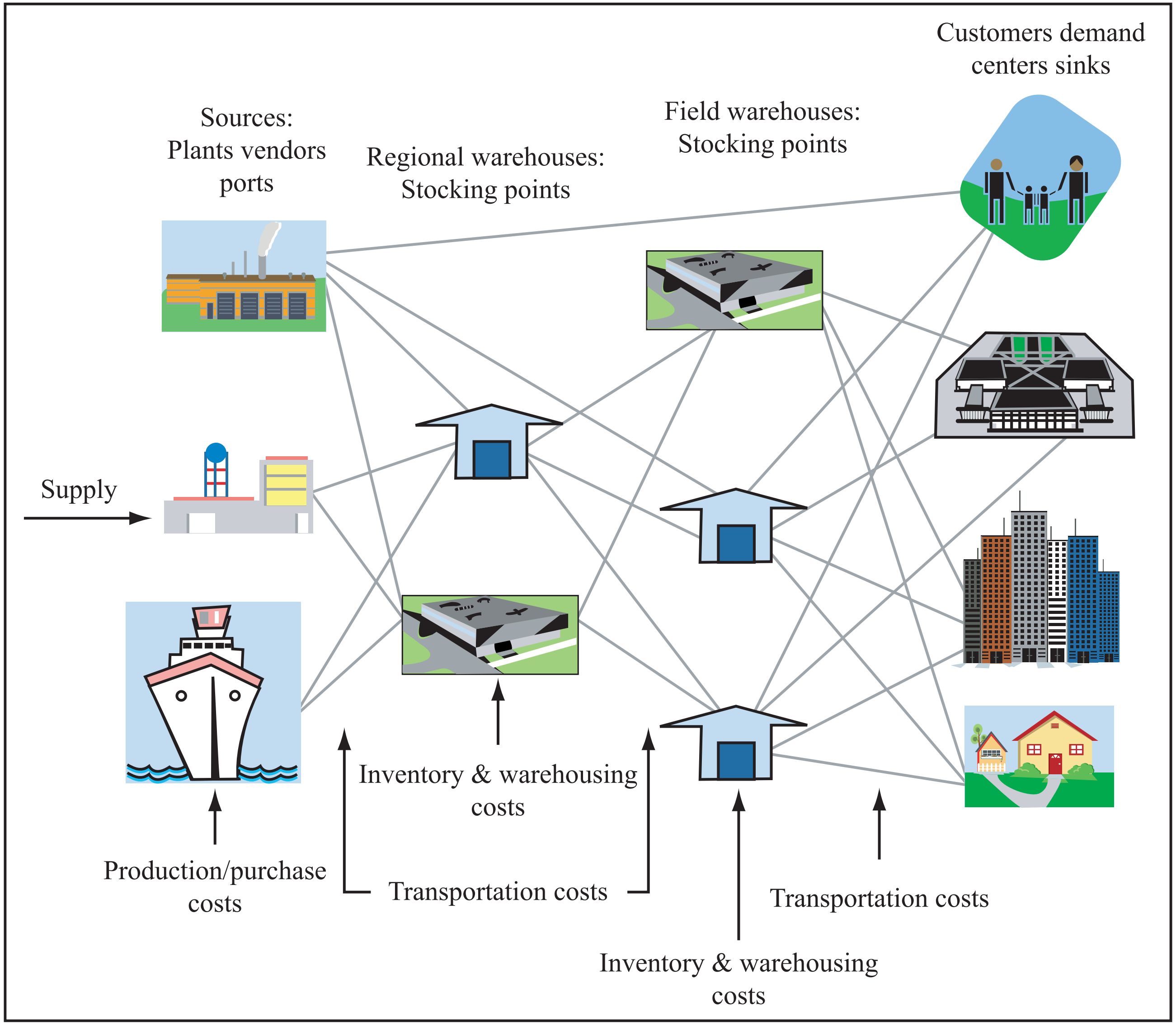
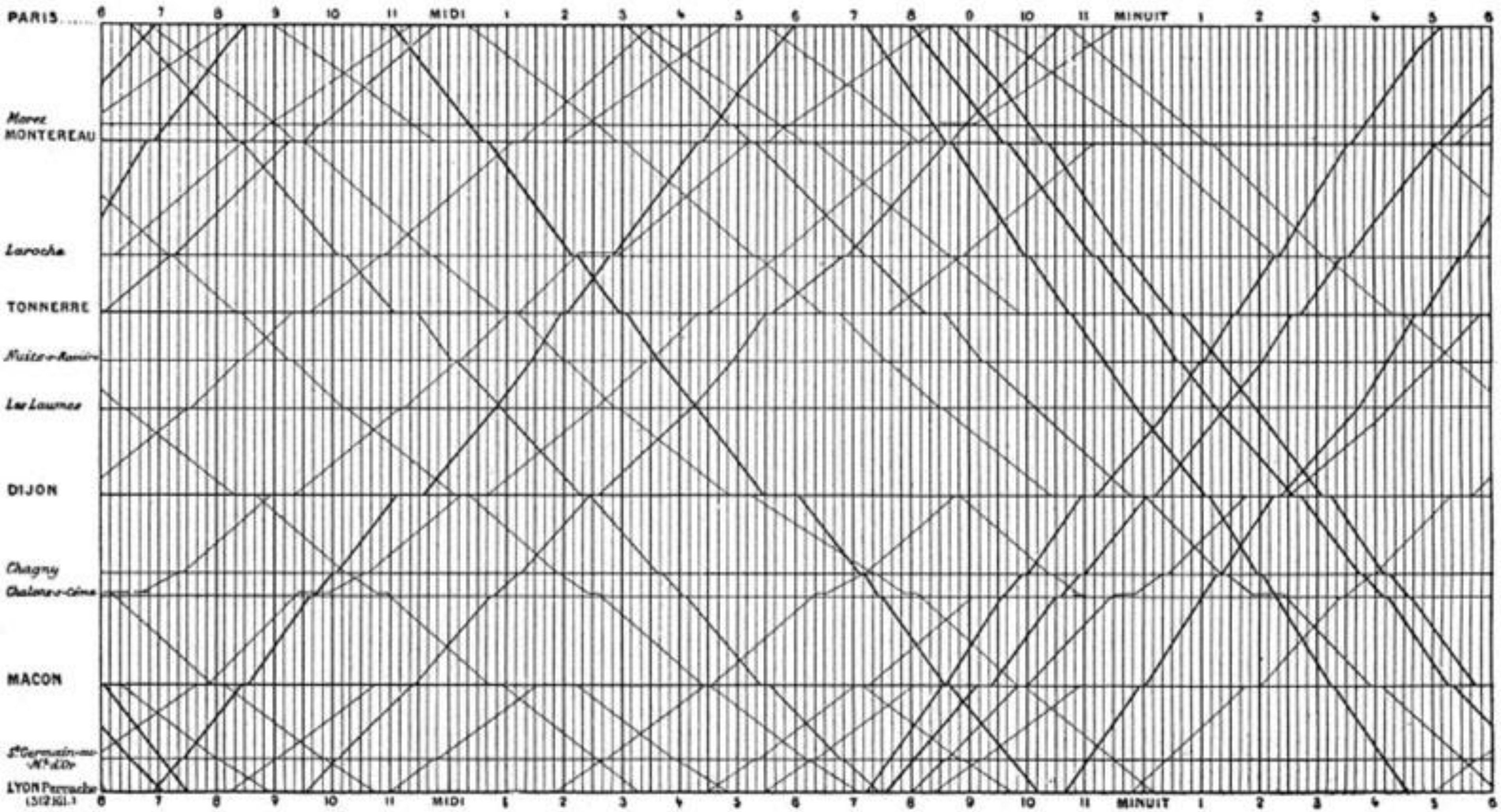


Image by MIT OpenCourseWare.

Supply chain



Train schedule

Public domain image: Paris-Lyon, 1885

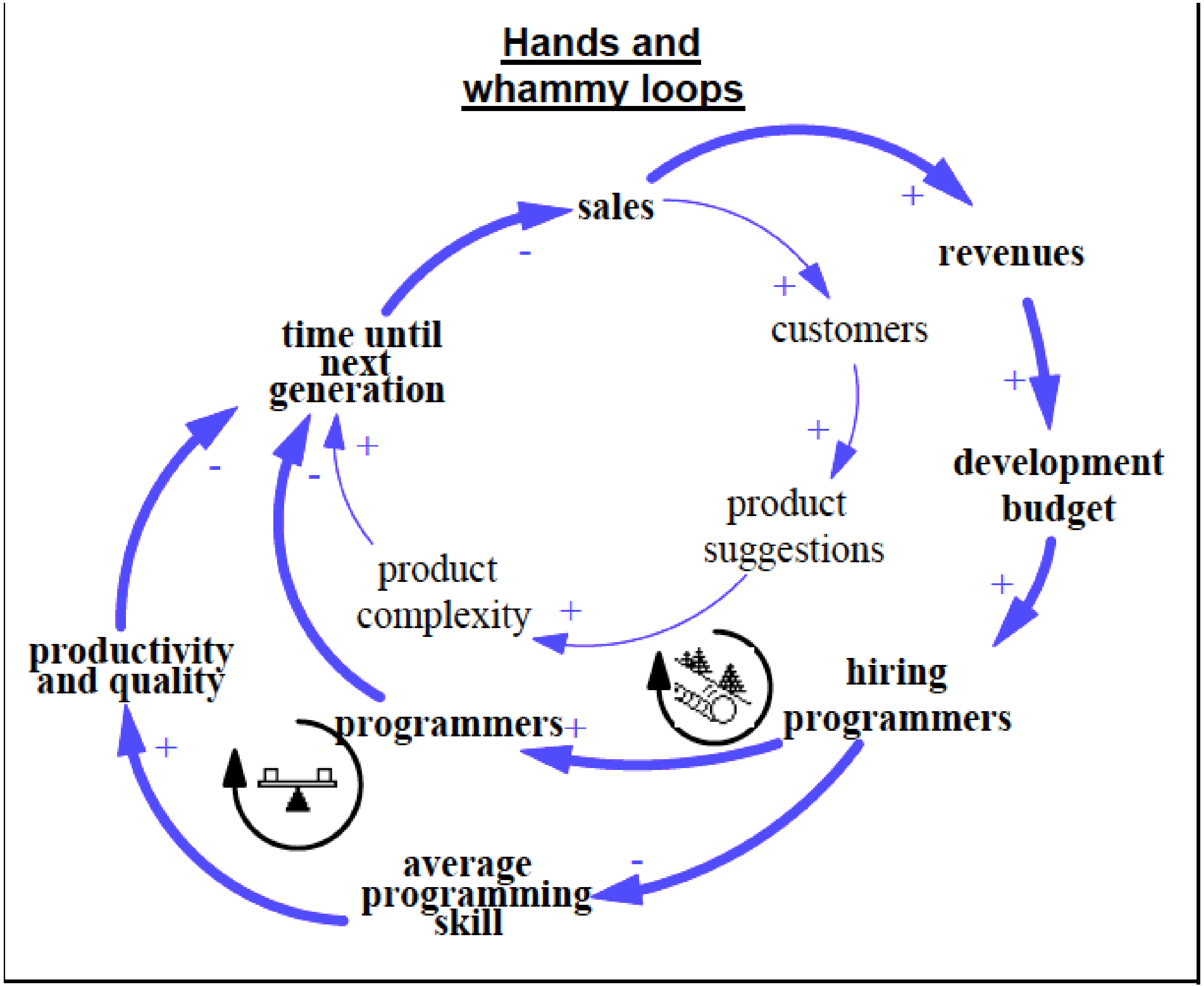
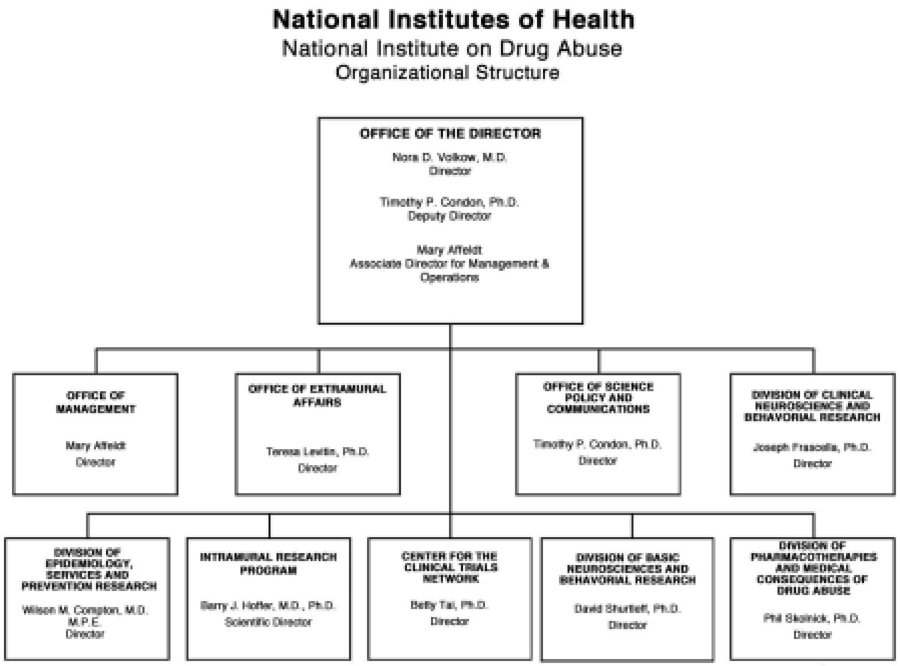
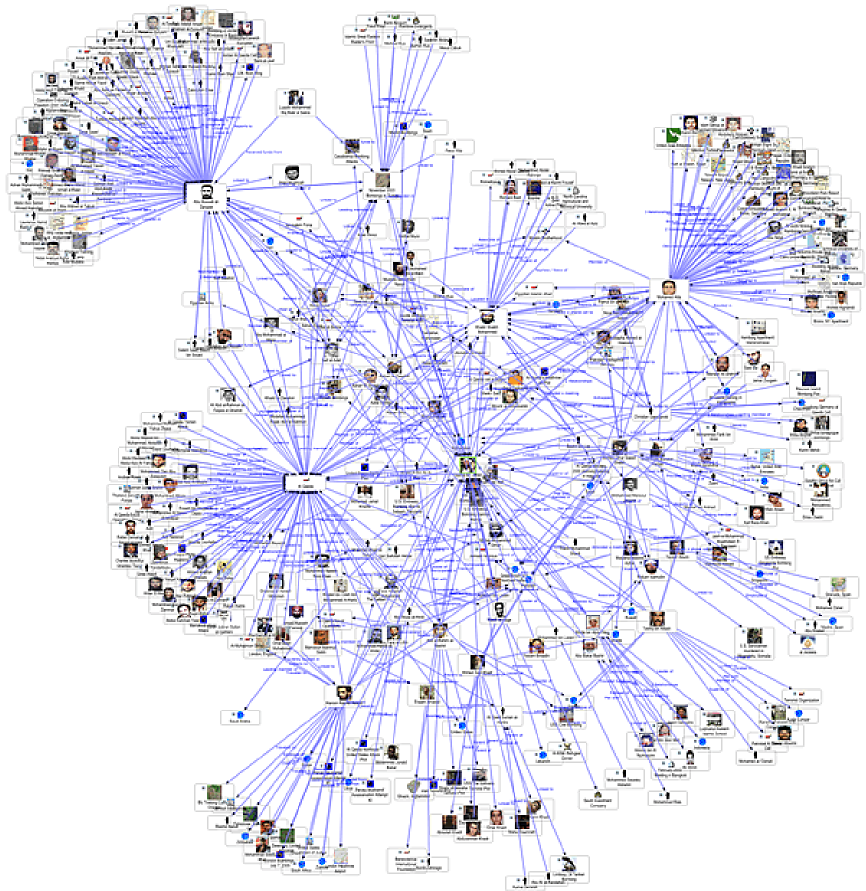


diagram: systems dynamics





Organizational Chart

Public domain (NIH)

Social Networks

Flows in networks

- Shipping from warehouses to retailers
- The min cost flow problem
- A remarkable theorem

The transportation problem

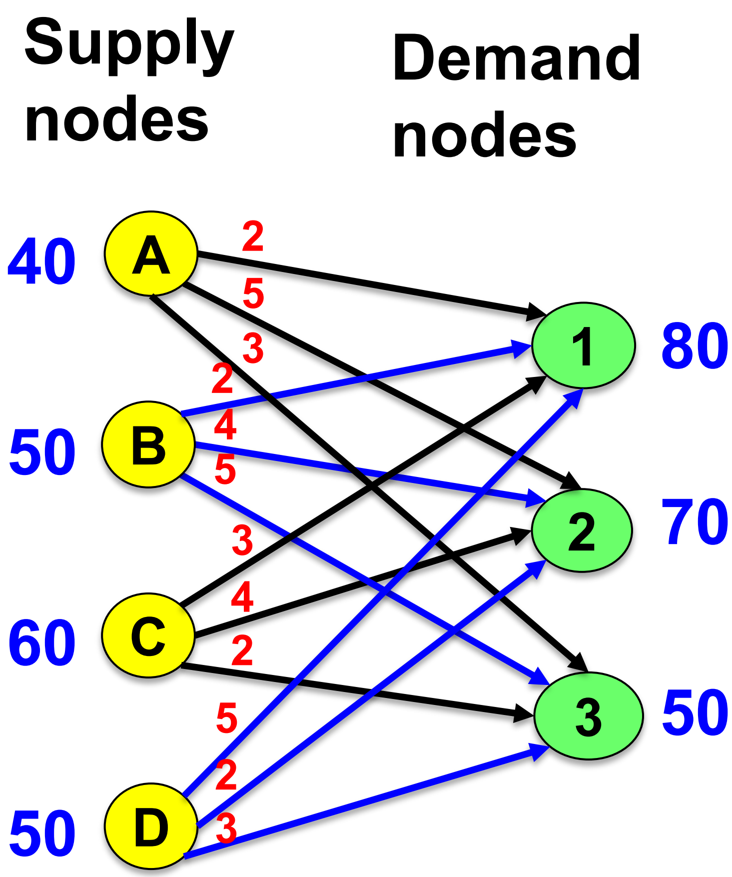
Warehouse	Supply
A	40
B	50
C	60
D	50

From/To	R1	R2	R3
A	2	5	3
B	2	4	5
C	3	4	2
D	5	2	3

Matrix of linear arc costs

Region	Demand
R1	80
R2	70
R3	50

The Transportation Problem



The ***transportation problem*** is a min cost flow problem with the following two properties:

- **All arcs are directed from "supply nodes" to "demand nodes."**
- **Arcs have costs, but there is no upper bound on arc flows.**

An LP formulation

Let x_{ij} = amount shipped from i to j
assigned to task j.

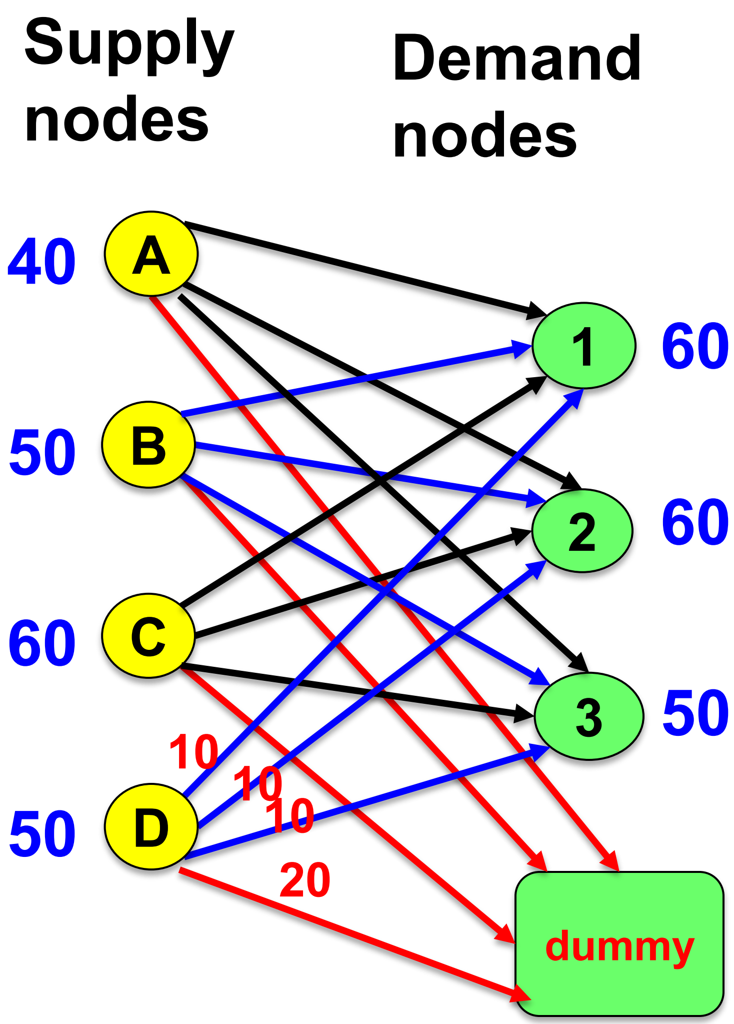
	x_{A1}	x_{A2}	x_{A3}	x_{B1}	x_{B2}	x_{B3}	x_{C1}	x_{C2}	x_{C3}	x_{D1}	x_{D2}	x_{D3}	RHS
A	1	1	1	0	0	0	0	0	0	0	0	0	40
B	0	0	0	1	1	1	0	0	0	0	0	0	50
C	0	0	0	0	0	0	1	1	1	0	0	0	60
D	0	0	0	0	0	0	0	0	0	1	1	1	50
1	1	0	0	1	0	0	1	0	0	1	0	0	80
2	0	1	0	0	1	0	0	1	0	0	1	0	70
3	0	0	1	0	0	1	0	0	1	0	0	1	50

An equivalent formulation

Sometimes it is better to treat supply as positive and demand as negative.

	x_{A1}	x_{A2}	x_{A3}	x_{B1}	x_{B2}	x_{B3}	x_{C1}	x_{C2}	x_{C3}	x_{D1}	x_{D2}	x_{D3}	RHS
A	1	1	1	0	0	0	0	0	0	0	0	0	40
B	0	0	0	1	1	1	0	0	0	0	0	0	50
C	0	0	0	0	0	0	1	1	1	0	0	0	60
D	0	0	0	0	0	0	0	0	0	1	1	1	50
1	-1	0	0	-1	0	0	-1	0	0	-1	0	0	-80
2	0	-1	0	0	-1	0	0	-1	0	0	-1	0	-70
3	0	0	-1	0	0	-1	0	0	-1	0	0	-1	-50

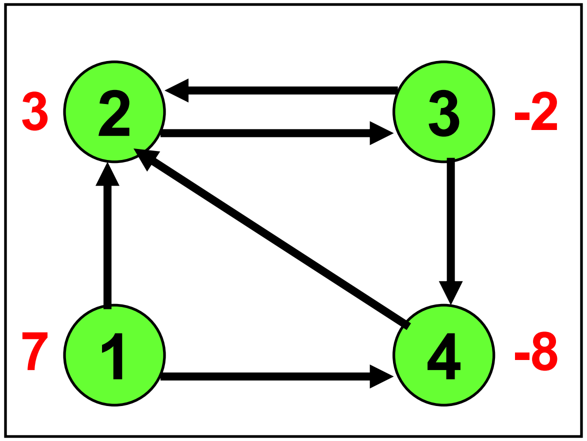
When supply exceeds demand



If supply exceeds demand:

- **add a dummy node**
- **Flow into the dummy node is the slack variable.**

Supply/demand constraints



- x_{ij} = flow in (i, j)
- node supply/demands b_i

Flow out of node i
\- Flow into node i
= b_i

Example: Node 4

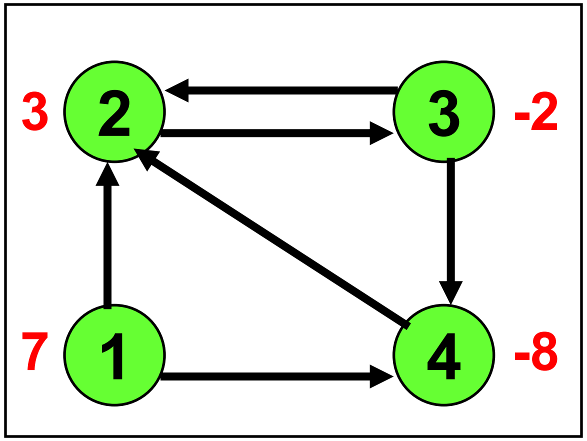
$x_{42} - x_{14} - x_{34} = -8$

We always assume that $\Sigma\, b_i = 0$.

That is, the available supply equals the required demand.

This is wlog. More on that latter.

Formulating a min cost flow problem



- x_{ij} = flow in (i, j)
- node supply/demands b_i
- arc costs c_{ij}
- arc capacities u_{ij}

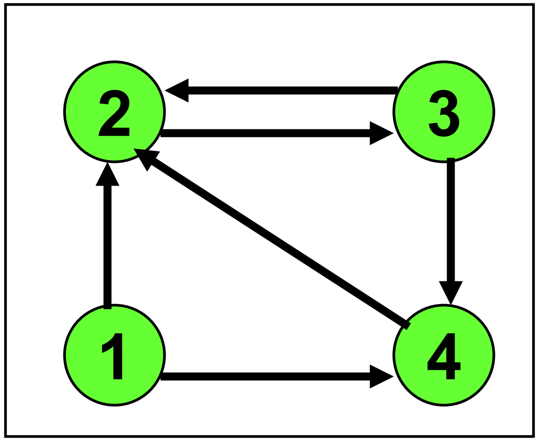
Minimize $$\sum_{(i,j)\in A} c_{ij} x_{ij}$$

x_{12}	x_{14}	x_{23}	x_{32}	x_{34}	x_{42}		RHS
						=	
						=	
						=	
						=	

$0 \leq x_{ij} \leq u_{ij}$

for all arcs (i, j) ∈ A

An LP formulation of the min cost flow problem



- x_{ij} = flow in (i, j)
- arc costs c_{ij}
- arc capacities u_{ij}
- node supply/demands b_i

Minimize $$\sum_{(i,j)\in A} c_{ij} x_{ij}$$

subject to $$\sum_{k:(j,k)\in A} x_{jk} - \sum_{j:(i,j)\in A} x_{ij} = b_j \qquad \text{for } j \in N$$

$$0 \le x_{ij} \le u_{ij} \qquad \text{for all arcs } (i, j) \in A$$

The amount shipped out of a node minus the amount shipped in to the node is the supply.

$$\sum_{i=1}^{n} b_i = 0$$

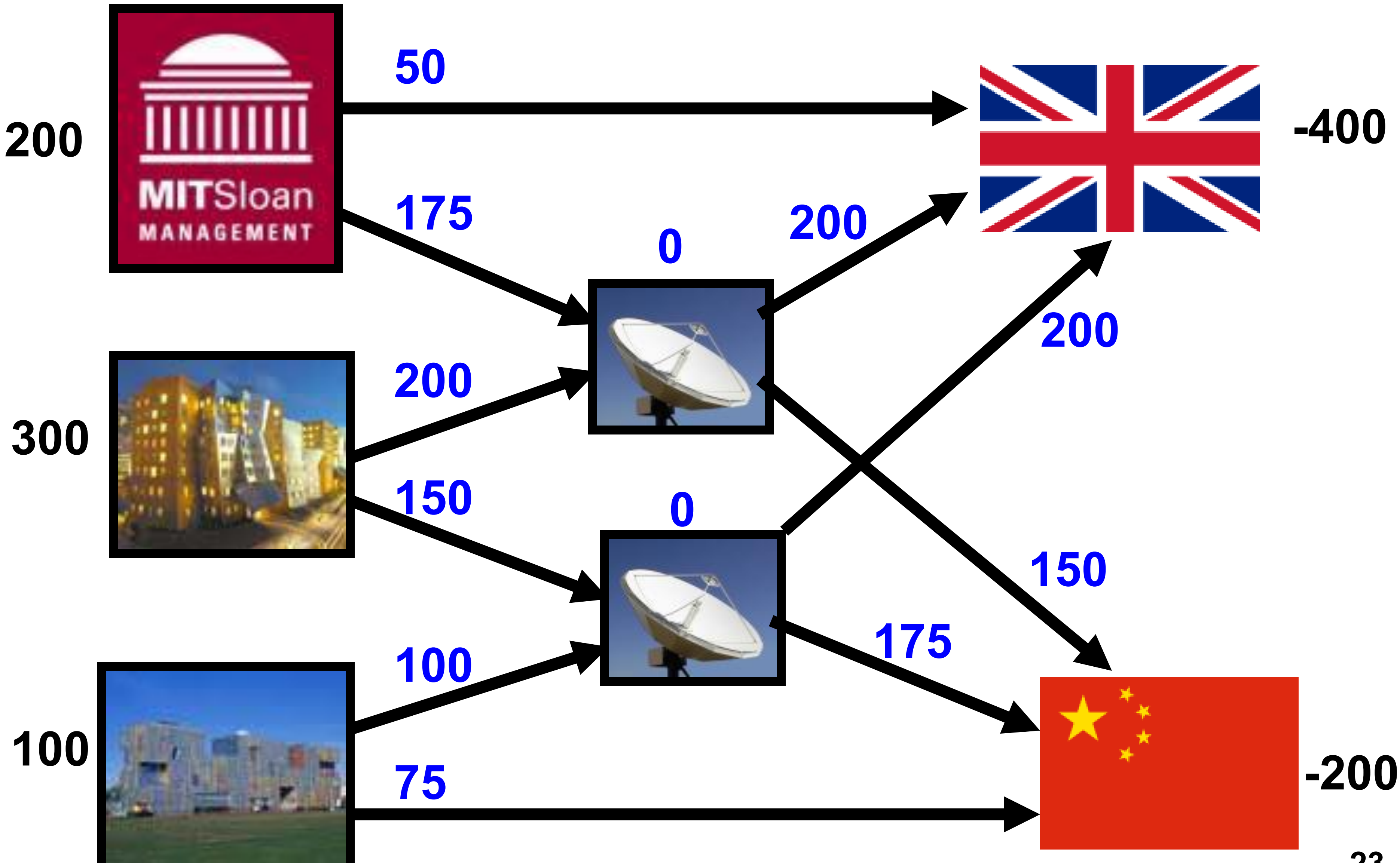
A communication problem

The year is 2013, and there is incredible demand for videos of MIT class lectures . MIT has set up three sites to handle the incredible load.

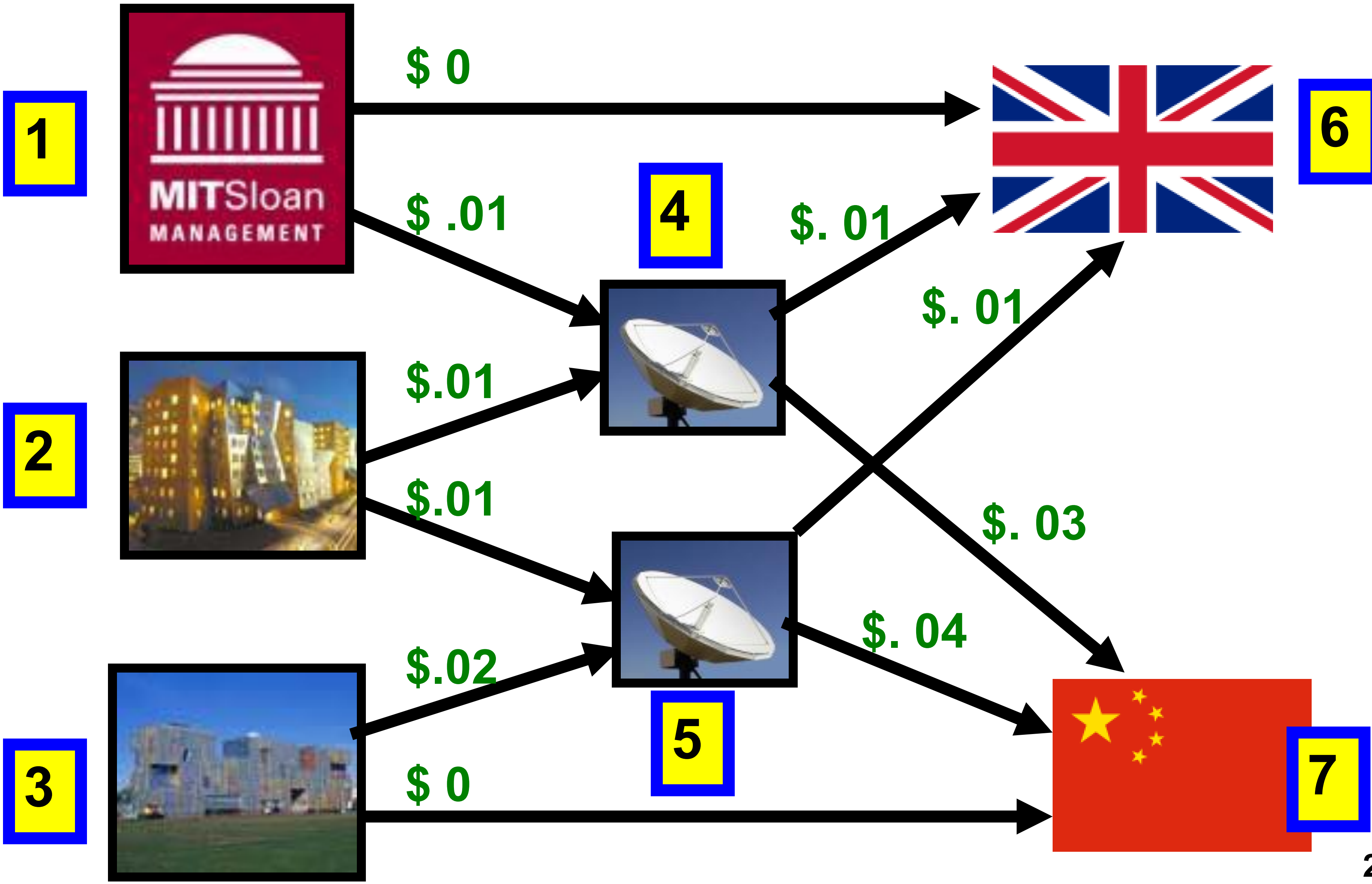
The major demand for the lectures are in London and China. There are some direct links from each of the three MIT sites, and the lectures can be sent through two intermediate satellite dishes as well.

Each node has a supply (or a demand) indicating how much should be shipped from (or to) the node. Each link (arc) has a unit cost of shipping flow, and a capacity on how much can be sent per second. What is the cheapest way of handling the required load?

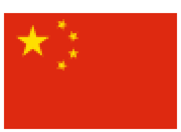
Supplies, Demands, and Capacities



Costs per megabyte. And node labels

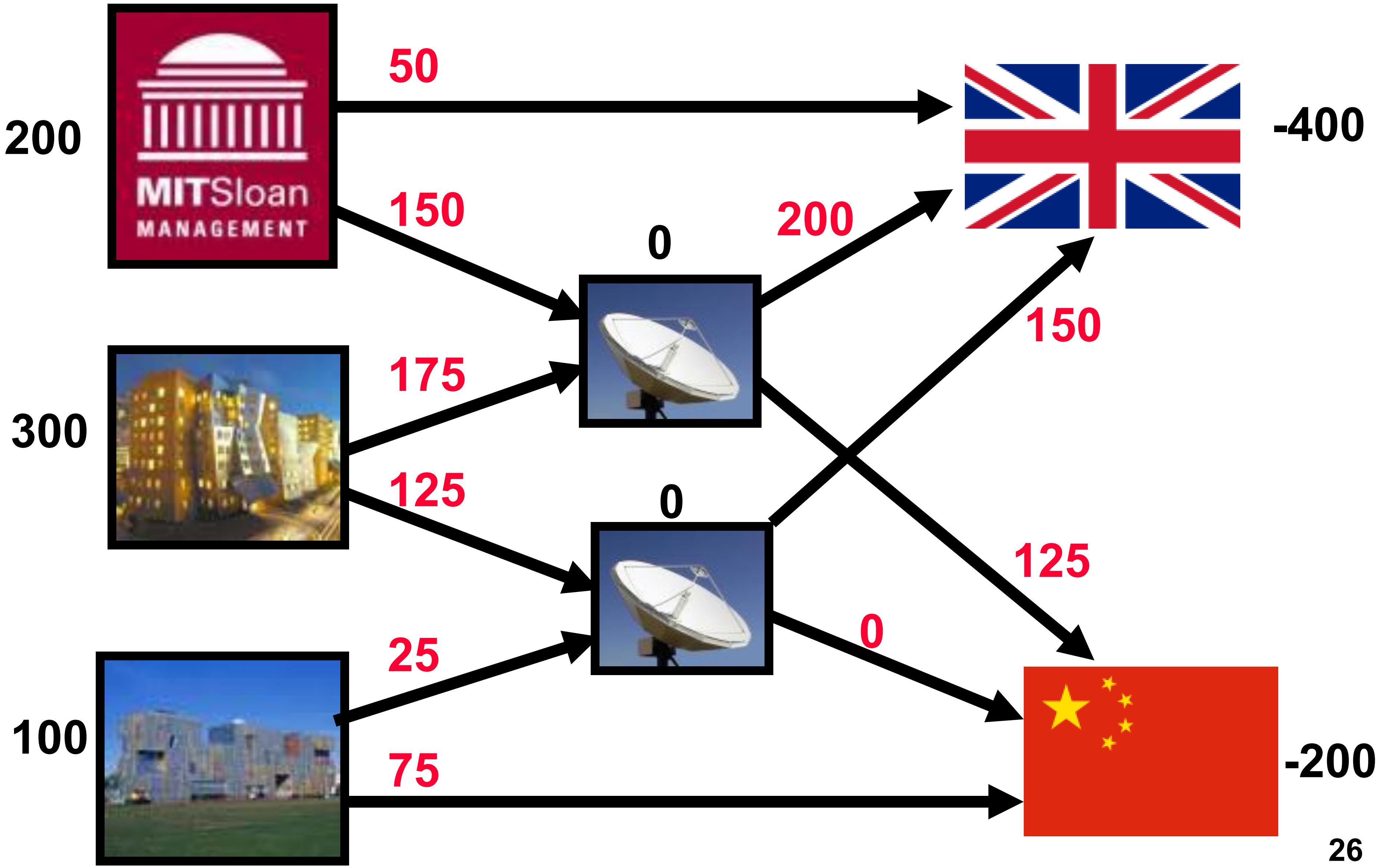


The Supply/Demand Constraints of the LP

	1-6	1-4	2-4	2-5	3-5	3-7	4-6	4-7	5-6	5-7		
	1	1									=	200
			1	1							=	300
					1	1					=	100
		-1	-1				1	1			=	0
				-1	-1				1	1	=	0
	-1						-1		-1		=	-400
						-1		-1		-1	=	-200

There is redundant constraint for the min cost flow problem.

The Optimal Flows



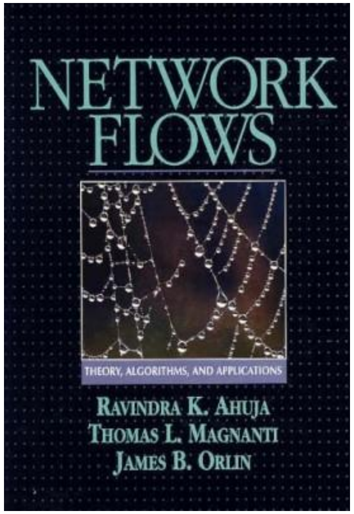
Which of the following is false?

1. If we consider all of the supply/demand constraints of a min cost flow problem, then each column has a coefficient that is 1, a coefficient that is -1, and all other coefficient are 0.
2. There is always a feasible solution for a min cost flow problem.
3. The supplies/demands sum to 0 for a min cost flow problem that is feasible.
4. At least one of the constraints of the min cost flow problem is redundant.

Mental Break

Why study the min cost flow problem

- Flows are everywhere
  - communication systems
  - manufacturing systems
  - transportation systems
  - energy systems
  - water systems
- Unifying Problem
  - shortest path problem
  - max flow problem
  - transportation problem
  - assignment problem
- Integrality Property
- Can be solved efficiently.
- Professor Orlin co-authored a textbook on network flows.



A Remarkable Theorem. (Integrality Theorem)

If the supplies, demands, and capacities of a minimum cost flow problem are all integral, then every basic feasible solution is integer valued. Therefore, the simplex method will provide an integer optimal solution.

Note: Most linear programs can have fractional solutions.

$x + y = 1$, $x - y = 0$. Unique solution (.5, .5)

Reason: The coefficients in the LHS of the constraints in the tableau remain as 0, 1 or -1.

Microsoft ®
Excel

Which of the following is false about the integrality theorem for min cost flows?

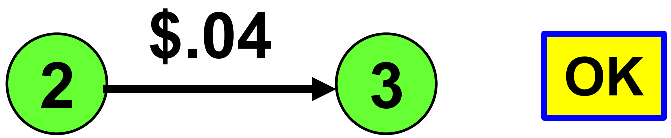
1. It is remarkable.
2. It is in contrast to the fact that most linear programs are not guaranteed to have integer valued bfs's.
3. It is remarkable.
4. It can be very useful in solving integer programs.
5. It was first proved to be true by Professor Orlin.

More on the integrality theorem

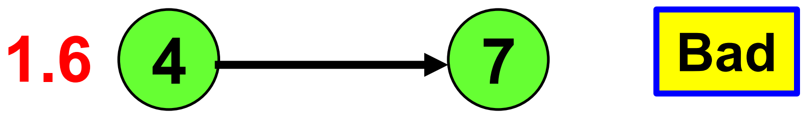
Valid for LPs in which the coefficients in each column (ignoring the objective coefficients and the RHS) have at most one 1 and at most one -1, with all other elements being 0.

0	1	0	-1	0	-1
1	0	0	0	-1	0
0	-1	1	1	0	0
0	0	-1	0	1	0

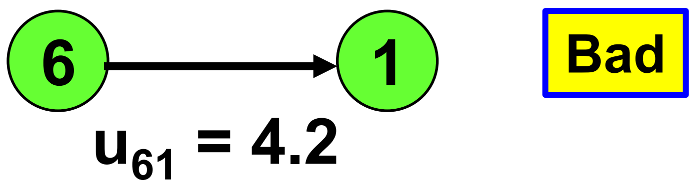
Does not depend on the costs.



Does depend RHS being integral.



Does depend on upper and lower bounds on variables being integer.



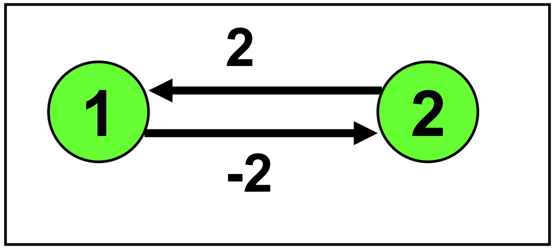
It is a theorem about basic feasible flows.

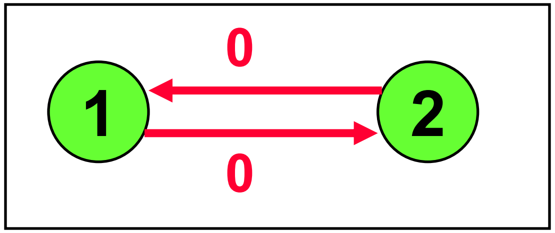
Non-basic flows can be fractional.

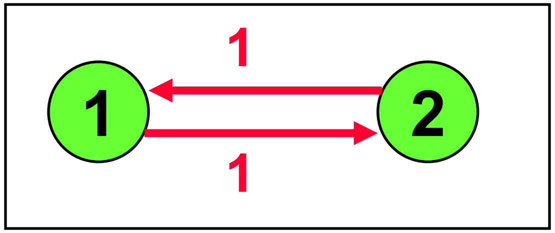
A network with arc costs.

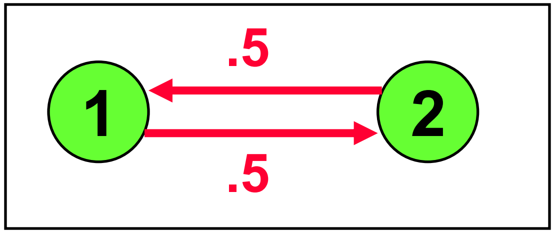
Suppose u_{ij} = 1 for all (i, j)

Suppose b(i) = 0 for all i.









Optimal Flow 1
cost = 0

Optimal Flow 2
cost = 0

Optimal Flow 3
cost = 0

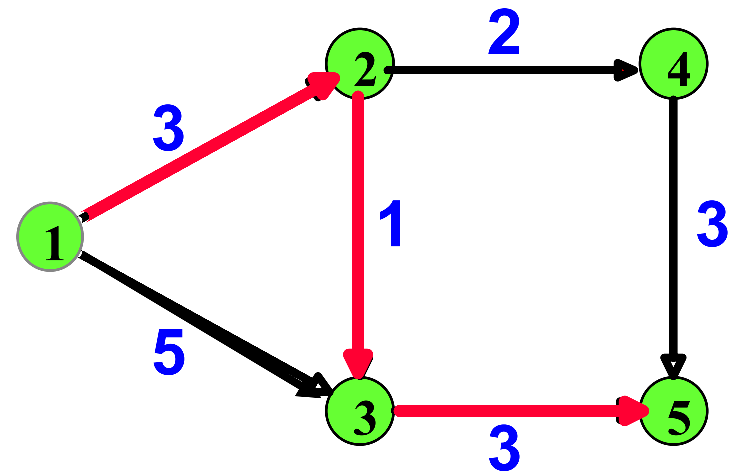
Which is <u>not</u> needed to guarantee that each bfs for a minimum cost flow problem has integer solutions?

1. Supplies/Demands are all integer valued
2. Capacities are integer valued
3. Costs are integer valued
4. All the above are needed.

Special cases of the min cost flow problem

- Shortest path problem
- Maximum flow problem
- Assignment problem

The shortest path problem with nonnegative arc lengths

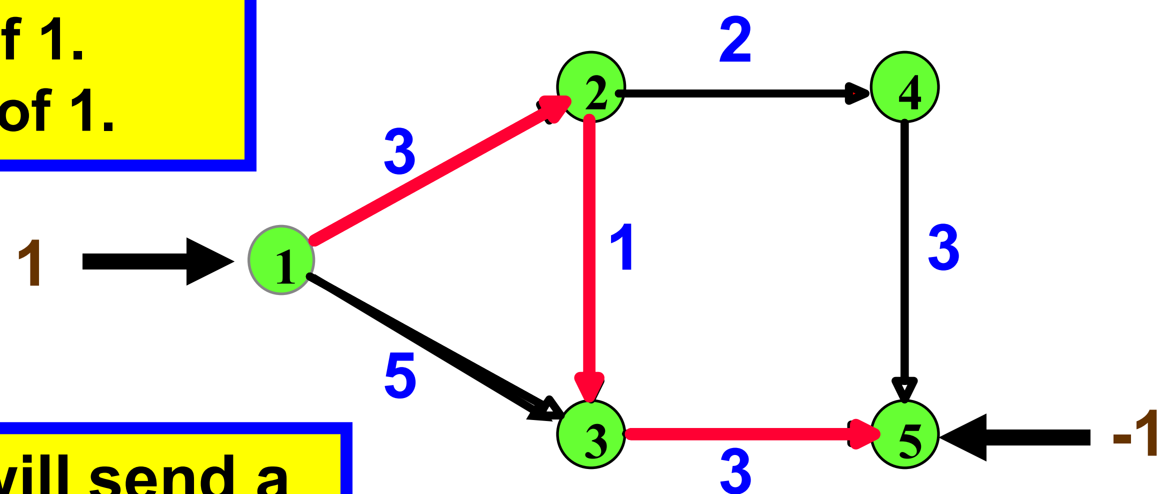


Find the shortest path from node 1 to node 5.

Translation to flow problem:

Node 1 has a supply of 1.

Node 5 has a demand of 1.

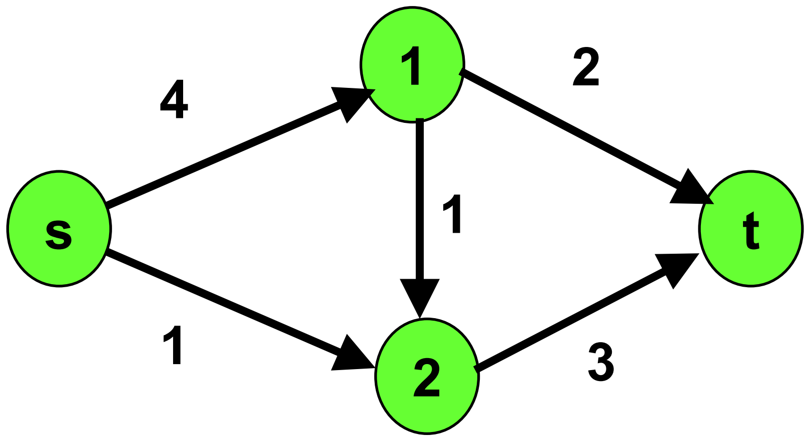


The optimal solution will send a flow of 1 unit along the shortest path from node 1 to node 5.

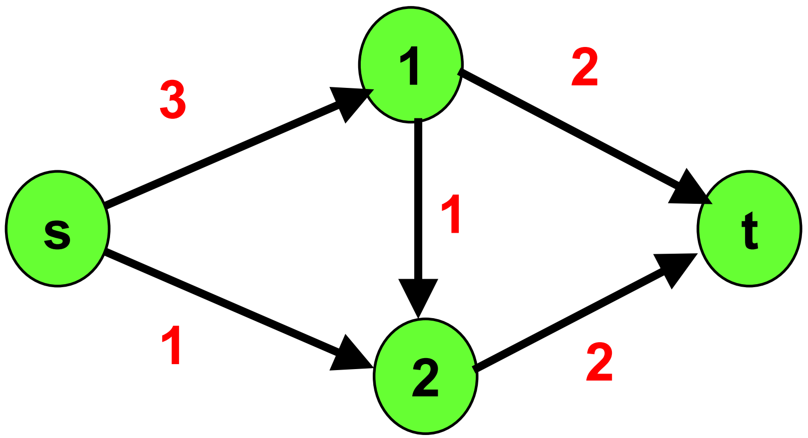
The Maximum Flow Problem

Directed Graph G = (N, A).

- **Source s**
- **Sink t**
- **Capacities u_{ij} on arc (i,j)**
- **Maximize the flow out of s, subject to**
- **Flow out of i = Flow into i, for i ≠ s or t.**

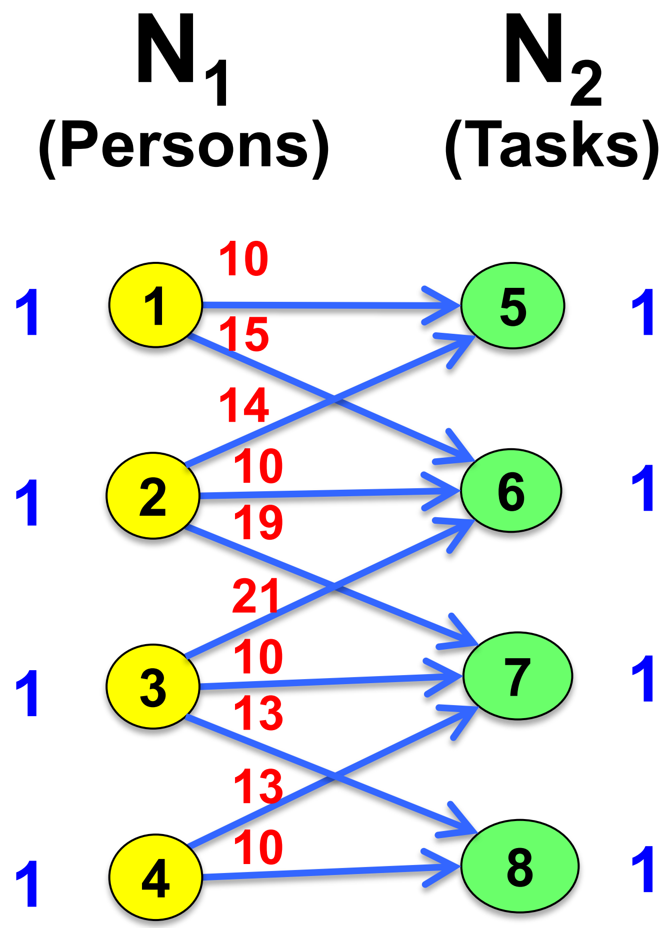


A Network with arc capacities



The maximum flow

The Assignment Problem



The ***assignment problem*** is the special case of transportation problem in which all supplies and demands are 1.

Usually, but not always, $|N_1| = |N_2|$.

Usually, **max utility** instead of **min cost**.

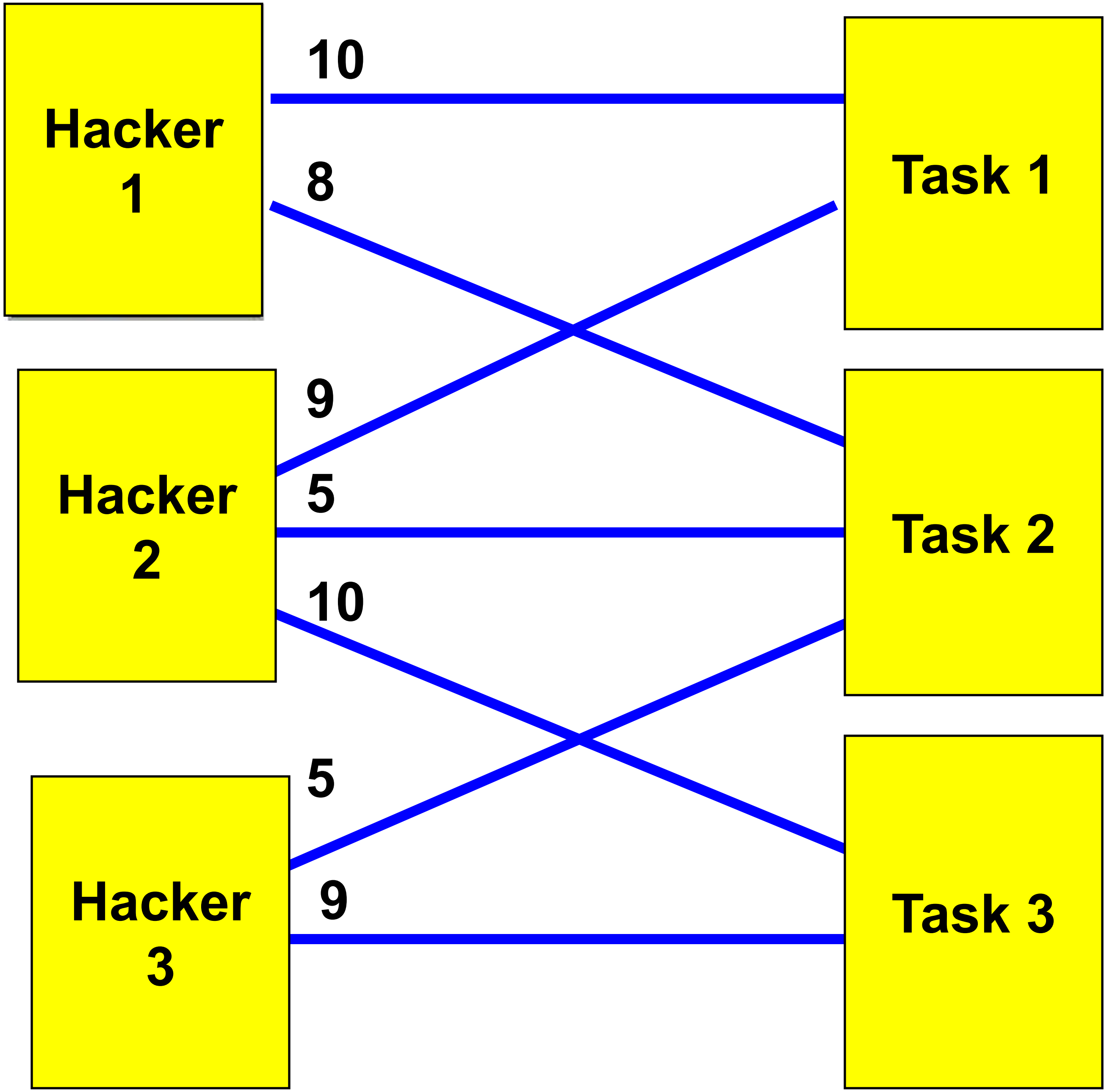
An assignment problem

- Three MIT hackers have decided to make the great dome look like R2D2, in honor of the hack from 5/17/99.

- Tasks.
  - Putting the sheets on the great dome
  - Ladder holder
  - Lookout

  - Objective: find the optimal allocation of persons to tasks.

  - What is the optimal assignment of hackers to tasks.



The arc numbers are utilities.

The goal is to find an assignment with maximum total utility.

An LP formulation

- Let x_{ij} = proportion of time that hacker i is assigned to task j.

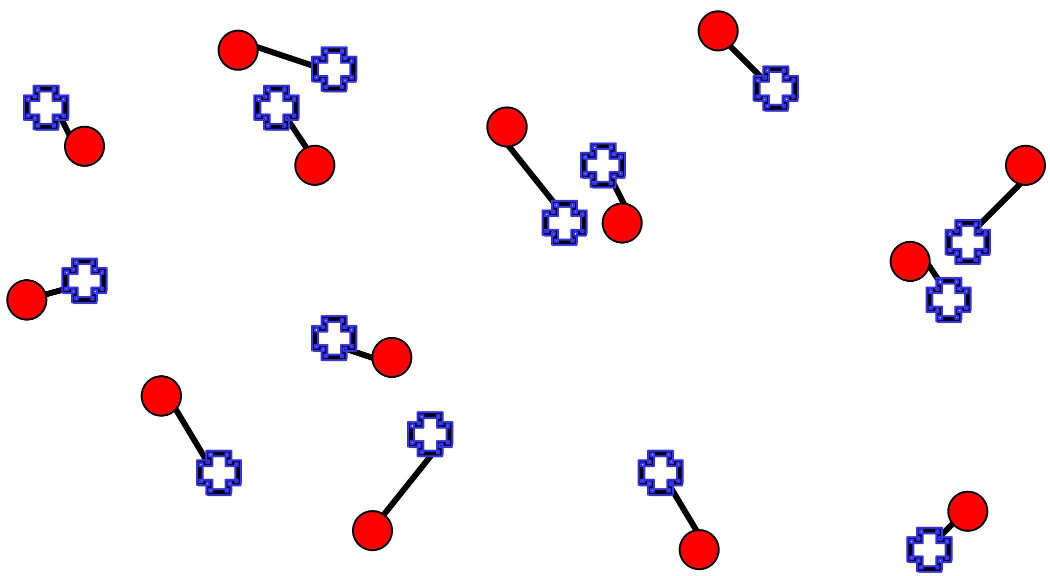
x_{11}	x_{12}	x_{21}	x_{22}	x_{23}	x_{32}	x_{33}	RHS	
								Hacker 1
								Hacker 2
								Hacker 3
								Task 1
								Task 2
								Task 3

An Application of the Assignment Problem

Suppose that there are moving targets in space.

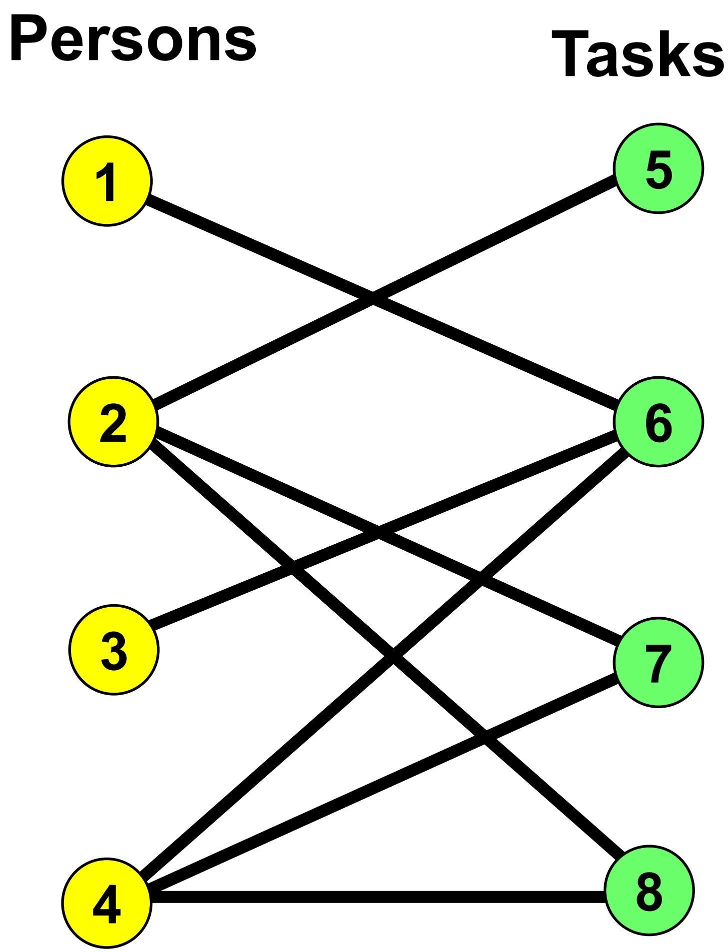
You can identify each target as a pixel on a radar screen.

Given two successive pictures, identify how the targets have moved.



This may be the most efficient way of tracking items.

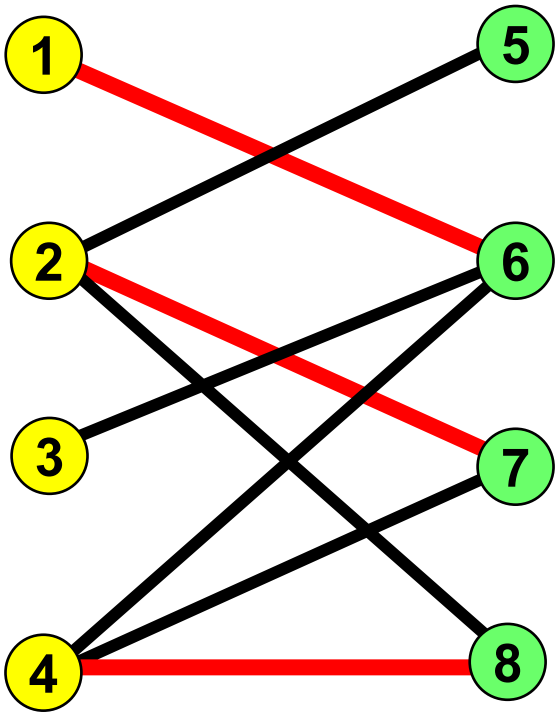
The matching problem: what is the maximum number of persons who can be matched to tasks?



	5	6	7	8
1	0	1	0	0
2	1	0	1	1
3	0	1	0	0
4	0	1	1	1

An adjacency matrix

The matching problem



	5	6	7	8
1	0	1	0	0
2	1	0	1	1
3	0	1	0	0
4	0	1	1	1

An adjacency matrix

A matching of cardinality three corresponds to three 1's of the adjacency matrix, no two of which are in the same row or column.

Independent 1's and line covers

	5	6	7	8
1	0	1	0	0
2	1	0	1	1
3	0	1	0	0
4	0	1	1	1

An adjacency matrix

Max-Matching Min-Cover

The minimum number of lines to cover all of the 1's of a matrix is equal to the max number of 1's no two of which are on a line.

Matrix rounding

				row sums
	.3	0	.4	.7
	.5	.7	0	1.2
	.2	.6	.2	1.0
col sums	1.0	1.3	.6	

Round coefficients of the matrix up or down so that the row sums and columns sum are also rounded.

	x_{11}	x_{12}	x_{13}	
	x_{21}	x_{22}	x_{23}	
	x_{31}	x_{32}	x_{33}	
col sums				

$|x_{11} - .3| < 1$

$|x_{21} - .5| < 1$

$|x_{31} - .2| < 1$

$|x_{11} + x_{21} + x_{31} - 1.0| < 1$

Application to matrix rounding

				row sums
	$0 \le x_{11} \le 1$	$x_{12} = 0$	$0 \le x_{13} \le 1$	**0 or 1**
	$0 \le x_{21} \le 1$	$0 \le x_{22} \le 1$	$x_{23} = 0$	**1 or 2**
	$0 \le x_{31} \le 1$	$0 \le x_{32} \le 1$	$0 \le x_{33} \le 1$	**1**
col sums	**1**	**1 or 2**	**0 or 1**	

Round coefficients of the matrix up or down so that the row sums and columns sum are also rounded.

$$x_{11} + x_{21} + x_{31} = 1$$

An LP formulation

Let x_{ij} = value in row i and column j.

Let r_1, r_2, c_2 and c_3 be slack variables.

x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}	x_{31}	x_{32}	x_{33}	r_1	r_2	c_2	c_3	RHS
1	1	1	0	0	0	0	0	0	1	0	0	0	1
0	0	0	1	1	1	0	0	0	0	1	0	0	2
0	0	0	0	0	0	1	1	1	0	0	0	0	1

1	0	0	1	0	0	1	0	0	0	0	0	0	1
0	1	0	0	1	0	0	1	0	0	0	1	0	2
0	0	1	0	0	1	0	0	1	0	0	0	1	1

Baseball Elimination Problem

	Games Won	Games Left
Bos	82	8
NY	77	8
Balt	80	8
Tor	79	8
Tamp	74	9

	Bos	NY	Balt	Tor	Tamp
Bos	--	1	4	1	2
NY	1	--	0	3	4
Balt	4	0	--	1	0
Tor	1	3	1	0	3
Tamp	2	4	0	3	0

Has Tampa already been eliminated from winning in this hypothetical season finale?

http://riot.ieor.berkeley.edu/~baseball/

Is there a way for Tampa to be tied for the lead (or winning) at the end of the season?

Assume that Tampa wins all of their games.

- **If they can’t lead the division after winning all of their games, they certainly can’t lead if they lose one or more games.**

Question: is it possible to assign wins and losses to all remaining games so that Tampa ends up in first place?

The results if Tampa wins all its games.

	Games Won	Games Left
Bos	82	6
NY	77	4
Balt	80	6
Tor	79	5
Tamp	83	0

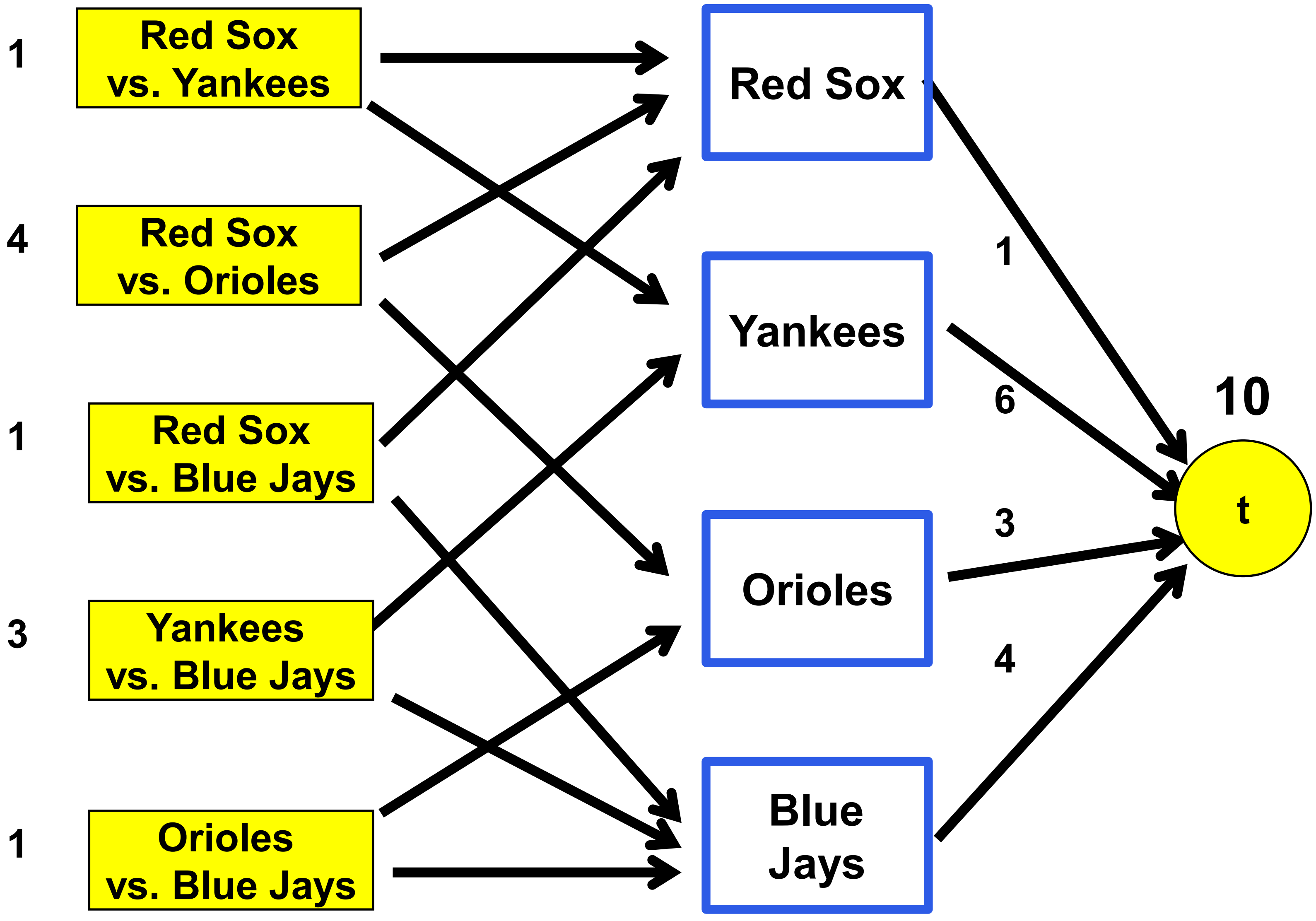
	Bos	NY	Balt	Tor	Tamp
Bos	--	1	4	1	0
NY	1	--	0	3	0
Balt	4	0	--	1	0
Tor	1	3	1	0	0
Tamp	0	0	0	0	0

Now check to see if the remaining games can be played so that no team wins more than 83 games.

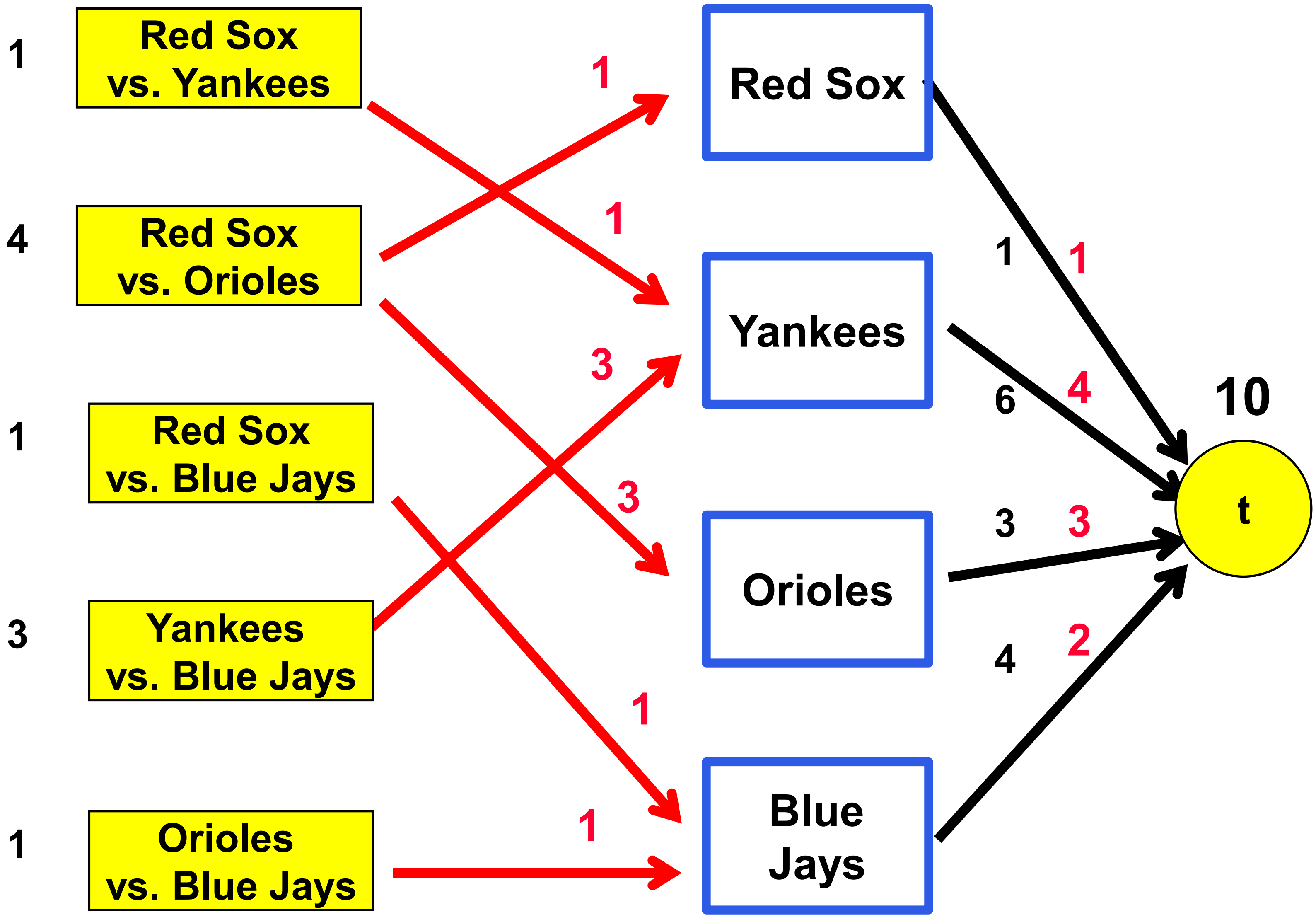
Constraints

- Upper bound on the number of games won by each team (except Tampa).
- Each game is won by one of the two teams playing the game.
- The flow on an arc is the number of games won.

Flow on (i,j) is interpreted as games won.



The optimum flow



Some Information on the Min Cost Flow Problem

- Reference text: Network Flows: Theory, Algorithms, and Applications by Ahuja, Magnanti, and Orlin [1993]
- 15.082J/6.885J: Network Optimization
- Polynomial time simplex algorithm (Orlin [1997])
- Basic feasible solutions of a minimum cost flow problem are integer valued (assuming that the data is integer valued)
- Very efficient solution techniques in practice