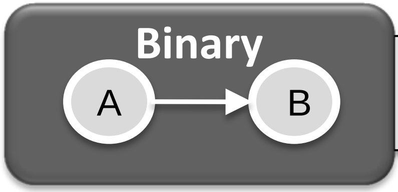
Working with network data

COMM 645: Communication Networks

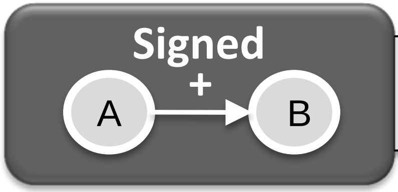
Katherine Ognyanova (Katya)

Fall semester, 2012

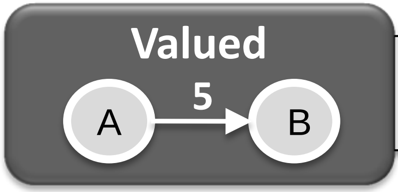
Link types



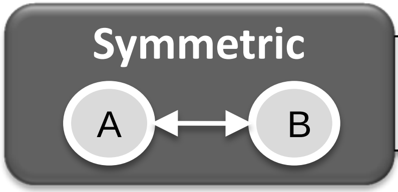
- **Link or no link** (1 or 0)



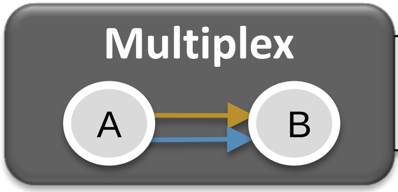
- **Positive or negative** (+, - or 0)



- **Weighted links** (each link is assigned a value)



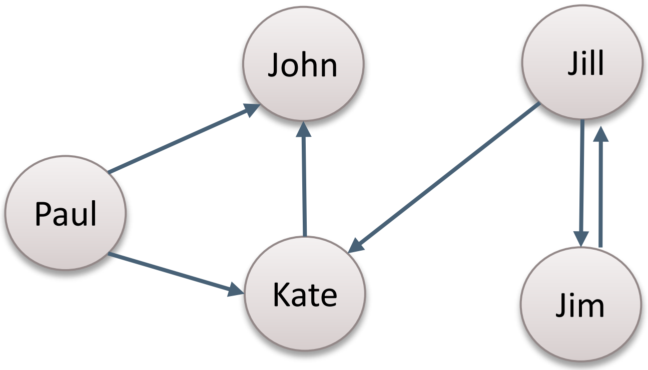
- **Directed vs. undirected** (undirected = symmetric)



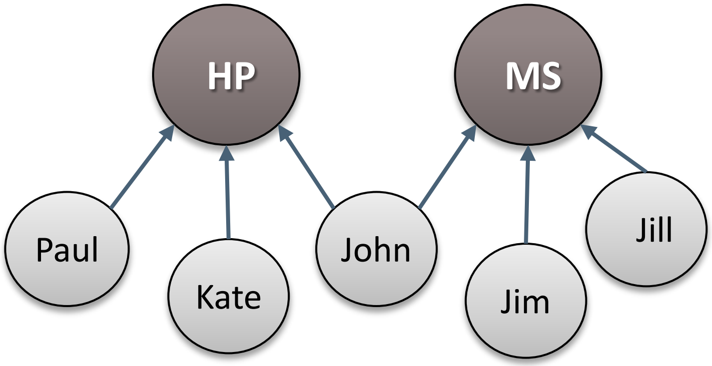
- **Multiplex** (more than one link type)

One and two-mode networks

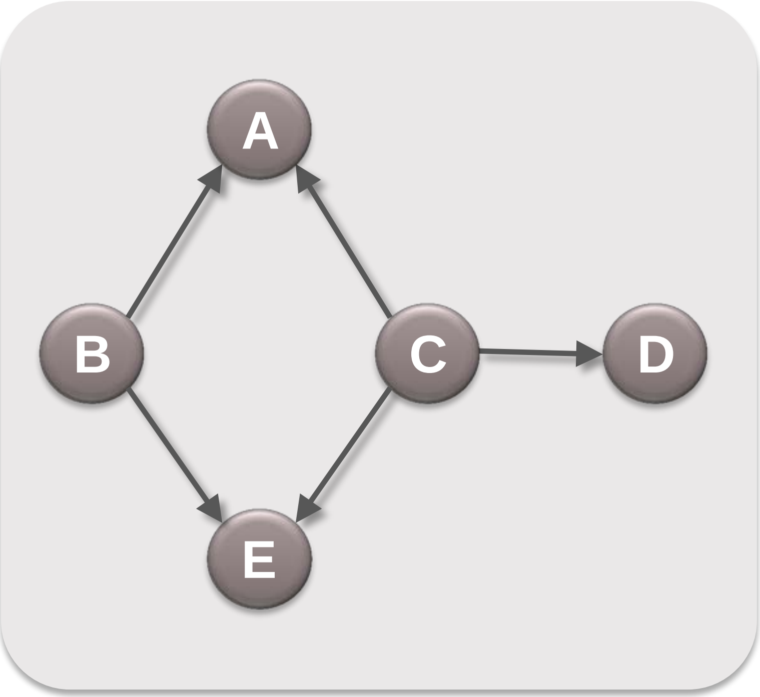
Adjacency
(e.g. friendship nets)



Affiliation
(e.g. employer-employee nets)

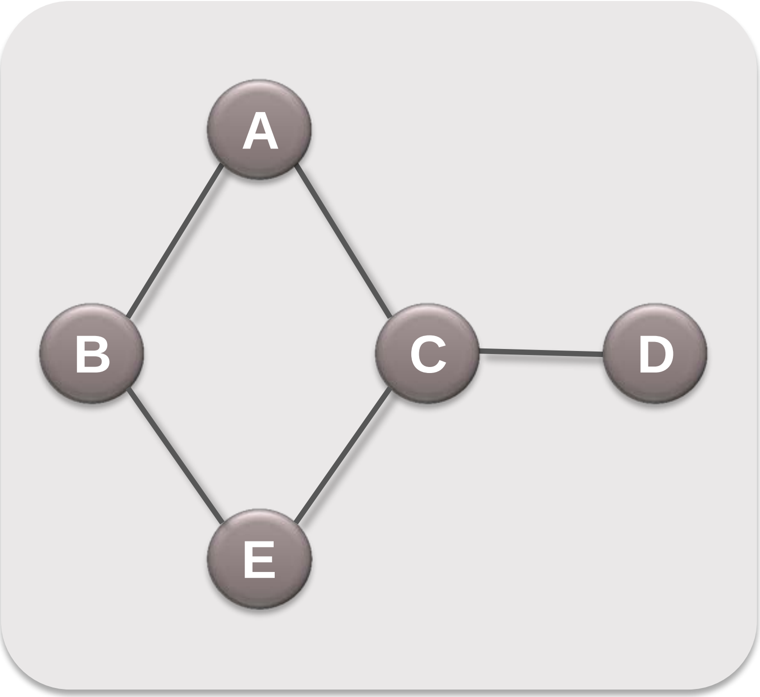


Matrix Representation: Directed Networks



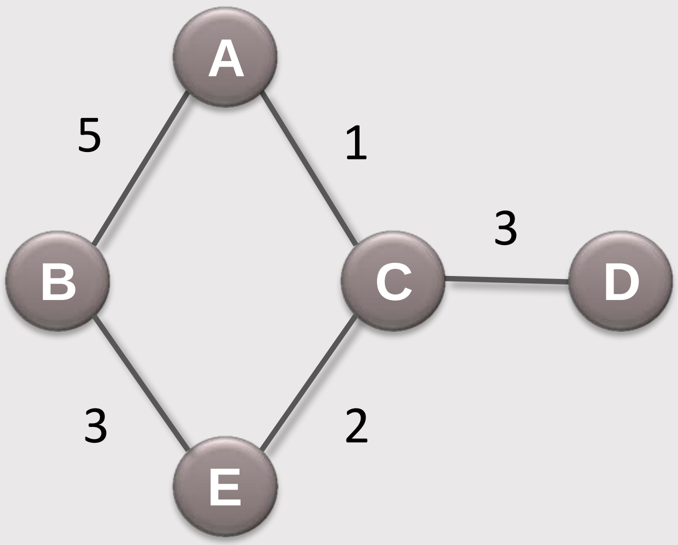
	A	B	C	D	E
A	0	0	0	0	0
B	1	0	0	0	1
C	1	0	0	1	1
D	0	0	0	0	0
E	0	0	0	0	0

Matrix Representation: Symmetric Networks



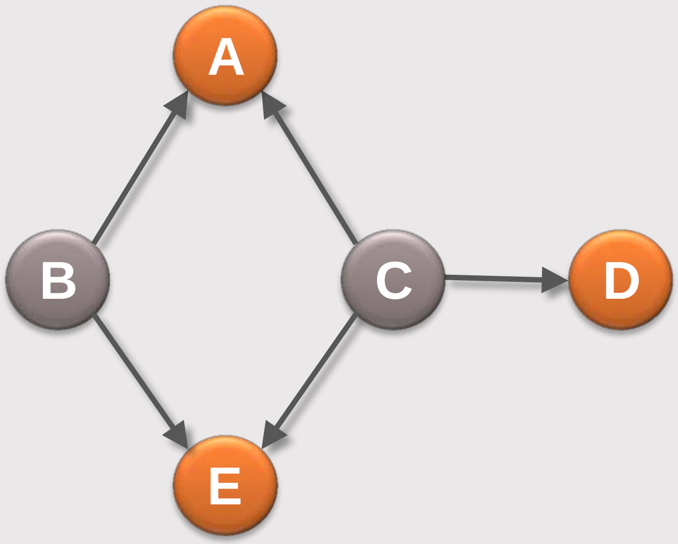
	A	B	C	D	E
A	0	1	1	0	0
B	1	0	0	0	1
C	1	0	0	1	1
D	0	0	1	0	0
E	0	1	1	0	0

Matrix Representation: Valued Networks



	A	B	C	D	E
A	0	5	1	0	0
B	5	0	0	0	3
C	1	0	0	3	2
D	0	0	3	0	0
E	0	3	2	0	0

Matrix Representation: Affiliation Data



	A	E	D
B	1	1	0
C	1	1	1

A quick detour to simple matrix algebra

Addition and Subtraction

$$A=\begin{bmatrix} 3 & -4 \\ 5 & 2 \end{bmatrix} \qquad B=\begin{bmatrix} 1 & 7 \\ 0 & 3 \end{bmatrix} \qquad C=\begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

$$A+B=\begin{bmatrix} 4 & 3 \\ 5 & 5 \end{bmatrix}$$

$A+C=$!

$$A + B = B + A$$

$$(A + B) + C = A + (B + C)$$

A quick detour to simple matrix algebra (Cont.)

Matrix multiplication

$$A=\begin{bmatrix} 1 & -4 \\ 0 & 3 \end{bmatrix} \quad B=\begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad C=\begin{bmatrix} 1 & 7 \end{bmatrix}$$

$$A*B=\begin{bmatrix} 1*3-4*2 \\ 0*3+3*2 \end{bmatrix}=\begin{bmatrix} -5 \\ 6 \end{bmatrix} \qquad C*A=\begin{bmatrix} 1 & 17 \end{bmatrix}$$

A*C= ! B*A= ! A*B ≠ B*A

If dim(A)=m x n, dim(B)=n x p then dim (A*B)=m x p

A quick detour to simple matrix algebra (Cont.)

Scalar multiplication

$$5* \begin{bmatrix} 1 & -4 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5 & -20 \\ 0 & 15 \end{bmatrix}$$

Element-wise matrix multiplication

$$\begin{bmatrix} 1 & -4 \\ 0 & 3 \end{bmatrix} \circ \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ 0 & 0 \end{bmatrix}$$

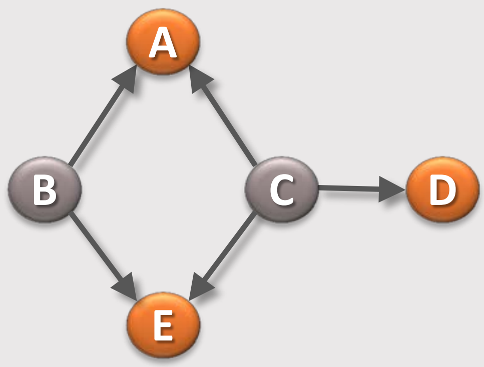
A quick detour to simple matrix algebra (Cont.)

Transposing matrices

$$A = \begin{bmatrix} 1 & -4 \\ 0 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 7 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 1 & 0 \\ -4 & 3 \end{bmatrix} \qquad B^T = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

Back to affiliation data



M

	A	E	D
B	1	1	0
C	1	1	1

M^T

	B	C
A	1	1
E	1	1
D	0	1

$M*M^T =$

	B	C
B	2	2
C	2	3

$M^T *M=$

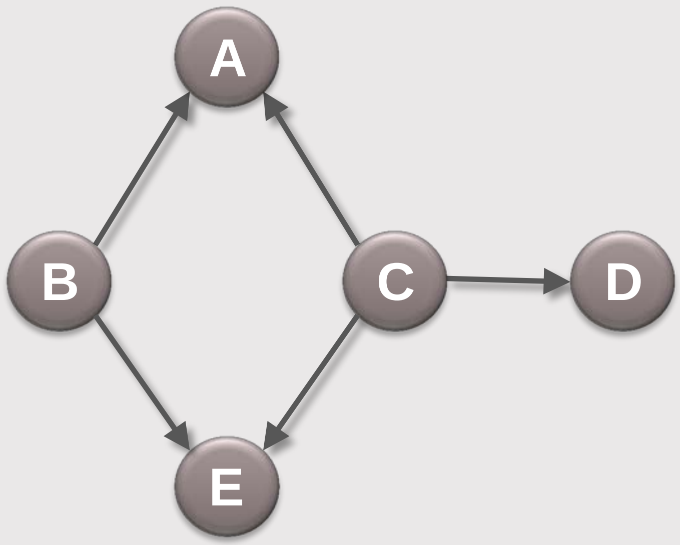
	A	E	D
A	2	2	1
E	2	2	1
D	1	1	1

Issues with matrix data

- Your dataset will likely contain network data in a non-matrix format .

- Large, sparse networks take way too much space if kept in a matrix format.

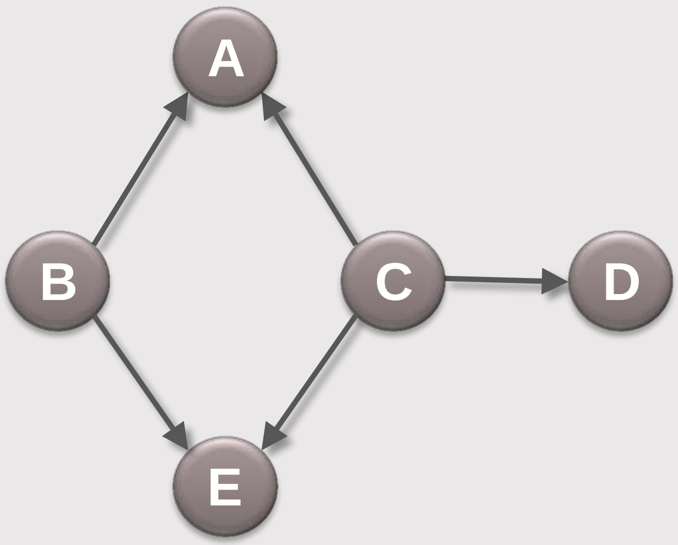
Edgelist Data Format



Source	Destination	Weight
B	A	1
B	E	1
C	A	1
C	E	1
C	D	1

Note: Weights are optional.

Nodelist Data Format



Source Destinations

B A E

C A D E

Some strategies for network data collection

Ego Networks	• Can use standard sampling techniques (e.g. random sample) • Each respondent describes their own relationships (name generators).
Complete Networks	• Boundary specification? • Each respondent reports their own relationships within the network. • Could use a roster that people use to identify contacts.
Cognitive Social Structures	• Ask not only for a person's own relationships, but also for perceived relationships between other people in your population.
Snowball Sampling	• Individuals included in the sample identify contacts (friends, sexual partners, etc.) who are added to the study at the next step. • Often used in preventive medicine.
Secondary Data	• Digital traces, social media, hyperlink networks and many more.