Representation and Data (mathematical models)

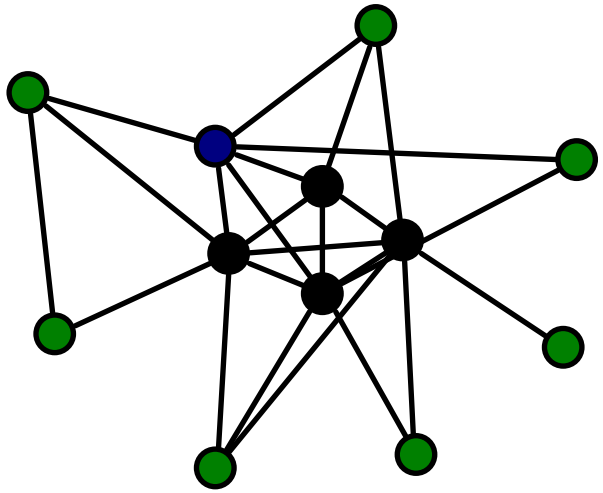
Steve Borgatti

Three Representations

- Network/relational data typically represented in one of three ways
  - Graphs
    - Graphs vs digraphs
  - Matrices
  - Relations on sets

(proper) Graphs

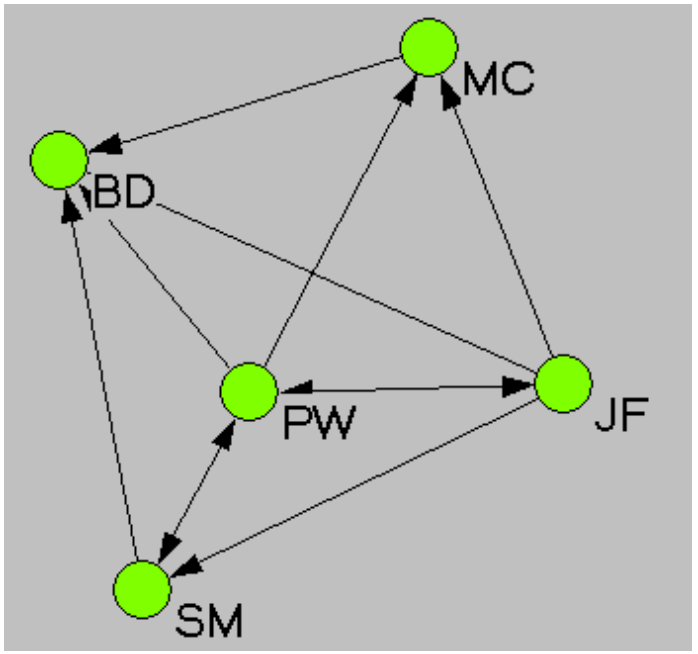
- A graph G(V,E) consists of …
  - Set of nodes|vertices V representing actors
  - Set of lines|edges E representing ties
    - An edge is an unordered pair of nodes (u,v)
    - Nodes u and v adjacent if (u,v) $\in$ E
    - So E is subset of set of all pairs of nodes



- Drawn without arrow heads
  - Sometimes with dual arrow heads
- Represent logically symmetric social relations
  - In communication with; attending same meeting as

Digraphs

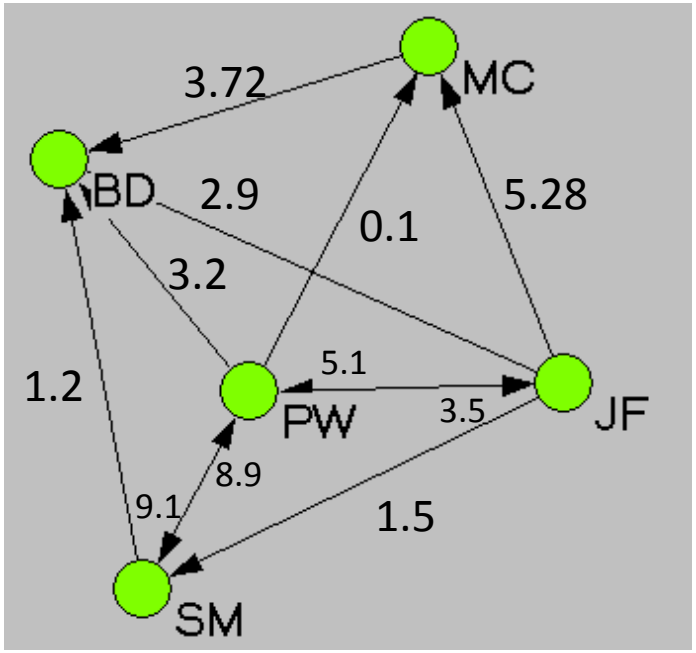
- Digraph G(V,E) consists of …
  - Set of nodes V
  - Set of directed arcs E
    - An arc is an ordered pair of nodes (u,v)
    - (u,v) $\in$ E indicates u sends arc to v
    - (u,v) $\in$ E does not imply that (v,u) $\in$ E



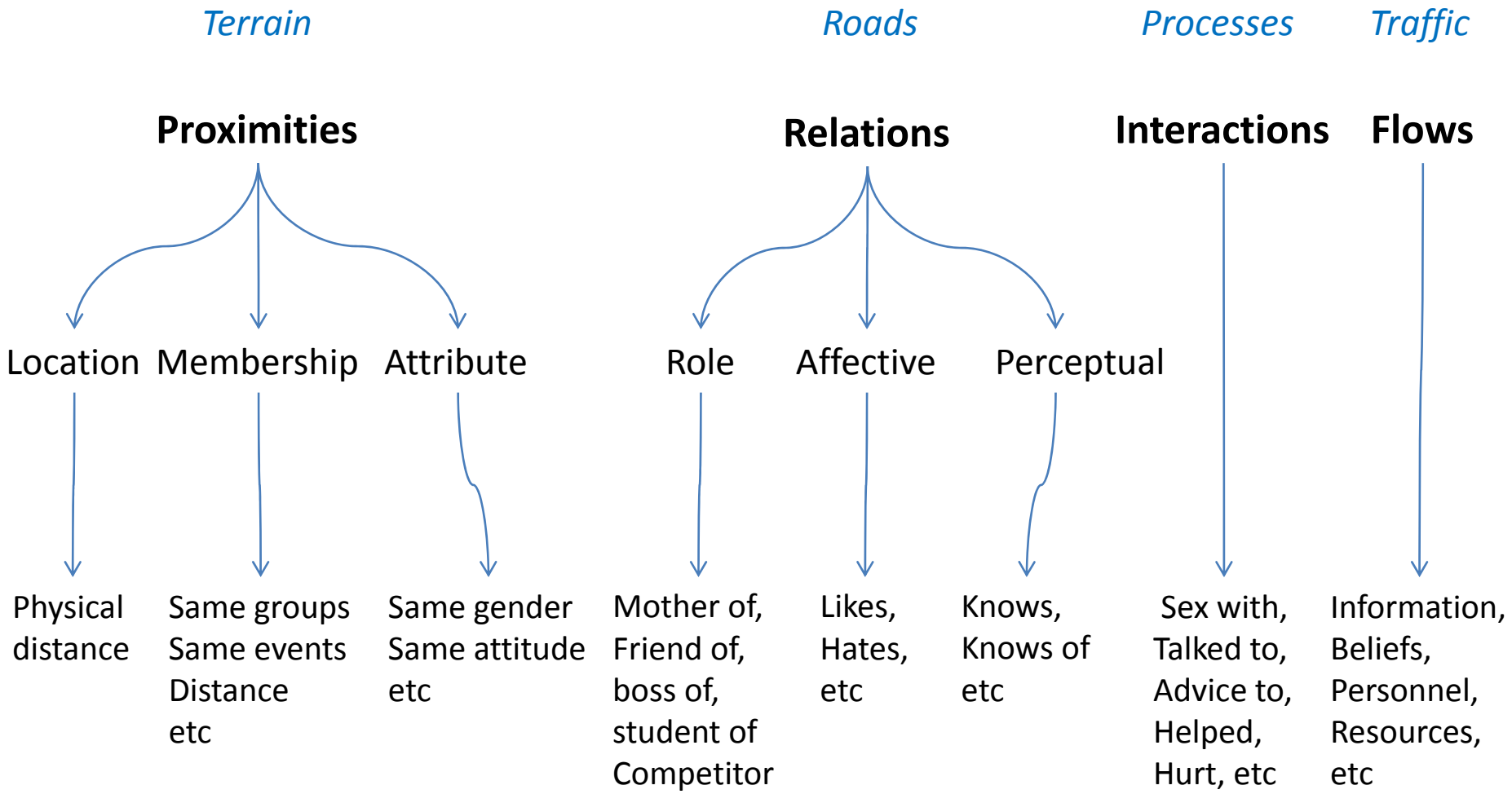
- Ties drawn with arrow heads, which can be in both directions
- Represent logically non-symmetric or anti-symmetric social relations
  - Lends money to

Valued Digraphs (vigraphs)

- A valued digraph G(V,E,W) consists of …
  - Set of nodes V
  - Set of directed arcs E
    - An arc is an ordered pair of nodes (u,v)
    - (u,v) ∈ E indicates u sends arc to v
    - (u,v) ∈ E does not imply that (v,u) ∈ E
  - Mapping W of arcs to real values
- Values can represent such things as
  - Strength of relationship
  - Information capacity of tie
  - Rates of flow or traffic across tie
  - Distances between nodes
  - Probabilities of passing on information
  - Frequency of interaction



Dyadic Relationships among Nodes



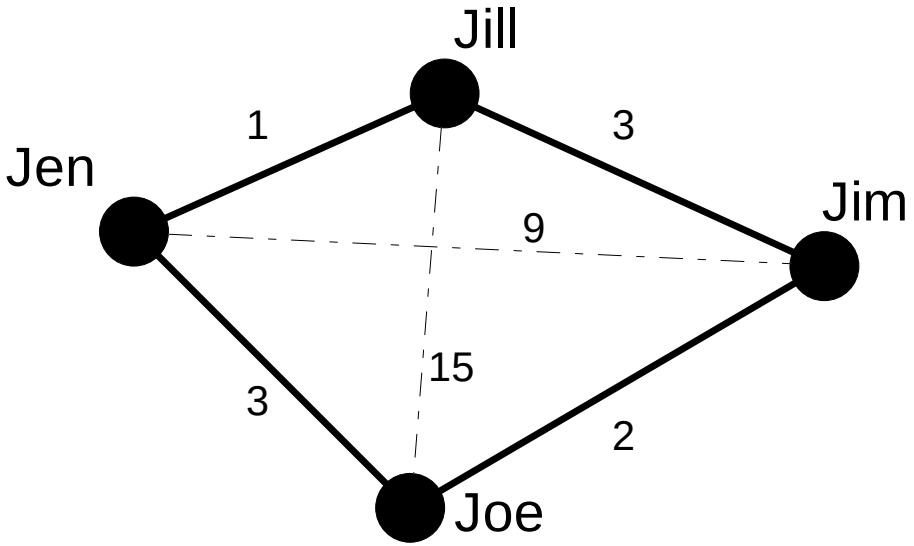
Adjacency Matrices

Friendship

	Jim	Jill	Jen	Joe
Jim	-	1	0	1
Jill	1	-	1	0
Jen	0	1	-	1
Joe	1	0	1	-

Proximity

	Jim	Jill	Jen	Joe
Jim	-	3	9	2
Jill	3	-	1	15
Jen	9	1	-	3
Joe	2	15	3	-



Matrices

- Matrix A
- Cell a_{ij}
  - First subscript is row, second is column

	Age	Gender	Income
Mary	32	1	90,000
Bill	50	2	45,000
John	12	2	0
Larry	20	2	8,000

A

$a_{12} = 1$
$a_{43} = 8K$

Ways and Modes

- Ways are the dimensions of a matrix.
- Modes are the sets of entities indexed by the ways of a matrix

	Event 1	Event 2	Event 3	Event 4
EVELYN	1	1	1	1
LAURA	1	1	1	0
THERESA	0	1	1	1
BRENDA	1	0	1	1
CHARLO	0	0	1	1
FRANCES	0	0	1	0
ELEANOR	0	0	0	0
PEARL	0	0	0	0
RUTH	0	0	0	0
VERNE	0	0	0	0
MYRNA	0	0	0	0

2-way, 2-mode

	Mary	Bill	John	Larry
Mary	0	1	0	1
Bill	1	0	0	1
John	0	1	0	0
Larry	1	0	1	0

2-way, 1-mode

Mainstream Logical Data Structure

- 2-mode rectangular matrices
  - Rows (cases) are entities, e.g., persons
  - Columns (variables) are attributes of the cases
- Analysis consists of correlating columns
  - Typically identify one column as the thing to be explained
  - We explain one attribute as a function of the others

Variables (attributes)

Cases (entities)

	Age	Sex	Education	Income
1001				
1002				
1003				
1004				
1005				
…				

Network Logical Data Structures

Adjacency matrices

Friendship

	Jim	Jill	Jen	Joe
Jim	-	1	0	1
Jill	1	-	1	0
Jen	0	1	-	1
Joe	1	0	1	-

Proximity

	Jim	Jill	Jen	Joe
Jim	-	3	9	2
Jill	3	-	1	15
Jen	9	1	-	3
Joe	2	15	3	-

Incidence matrix

	Friendship	Proximity
Jim - Jill	1	3
Jim - Jen	0	9
Jim - Joe	1	2
Jill - Jen	1	1
Jill - Joe	0	15
Jen - Joe	1	3

- Multiple relations for same set of actors
- Each relation is a (dyadic) variable
  - But can also be aggregated to node/group level
- Cases are pairs of actors
- Some hypotheses can be phrased in terms of correlations between relations