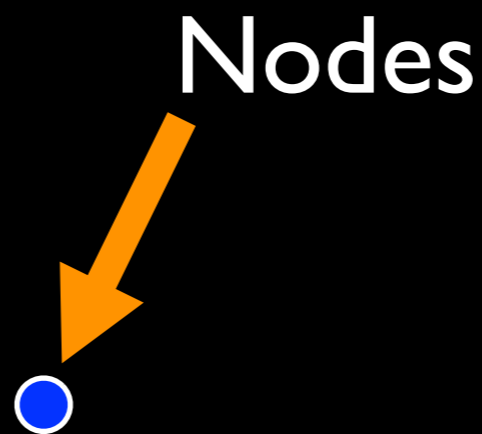
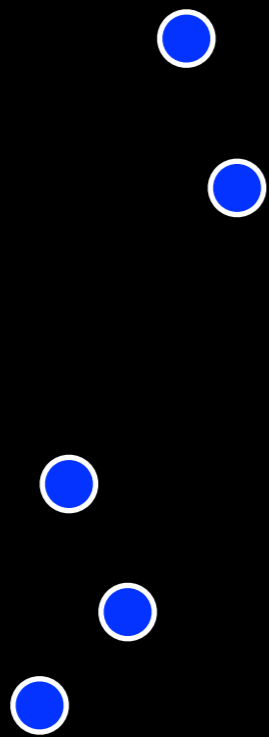


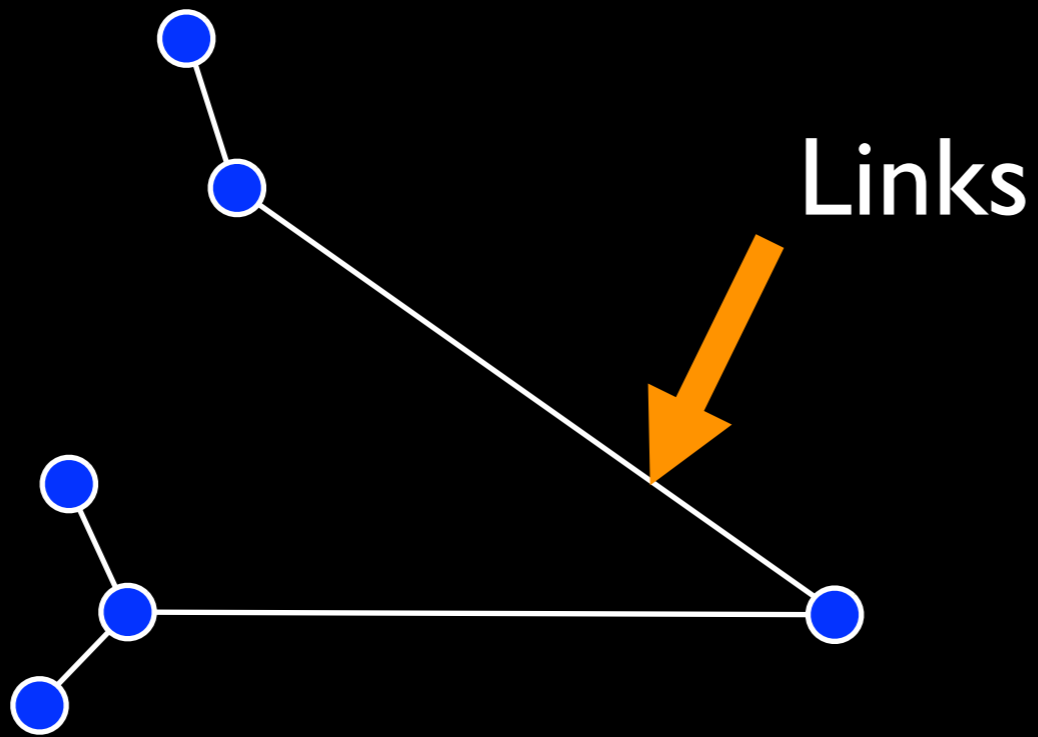
Network Analysis

Assignment Schedule

- Web Crawlers Assignment
 - Due 25-Nov
- SNA Assignment
 - Out: Tomorrow | Due: 2-Dec
- Archiving Assignment
 - Out: 3-Dec | Due: 10-Dec
- Social Assignment
 - Out: 10-Dec | Due: 16-Dec

- Focuses on the relationship between actors
- Not attributes of individual actors
- More holistic

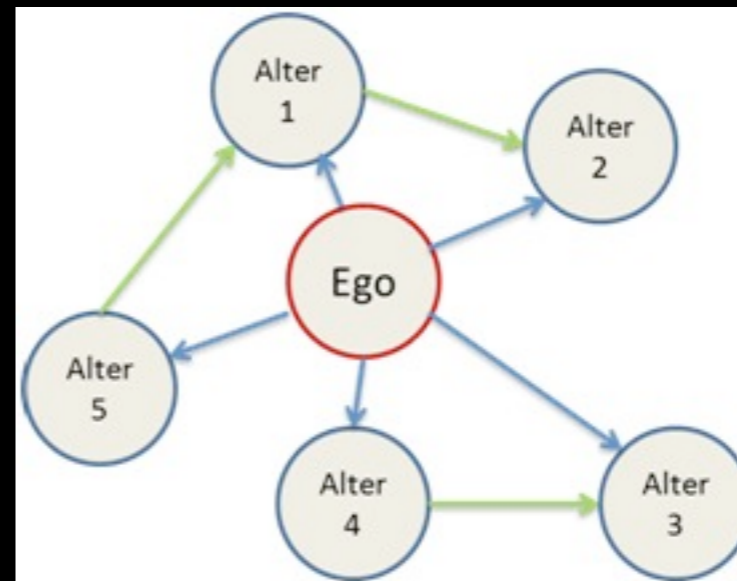




Sampling

- Full network
 - Collect information about the ties of all of the members of the network
 - Easier for small groups
- Snowball
 - Actors are asked to identify ties to other actors
 - Often used for “special” populations

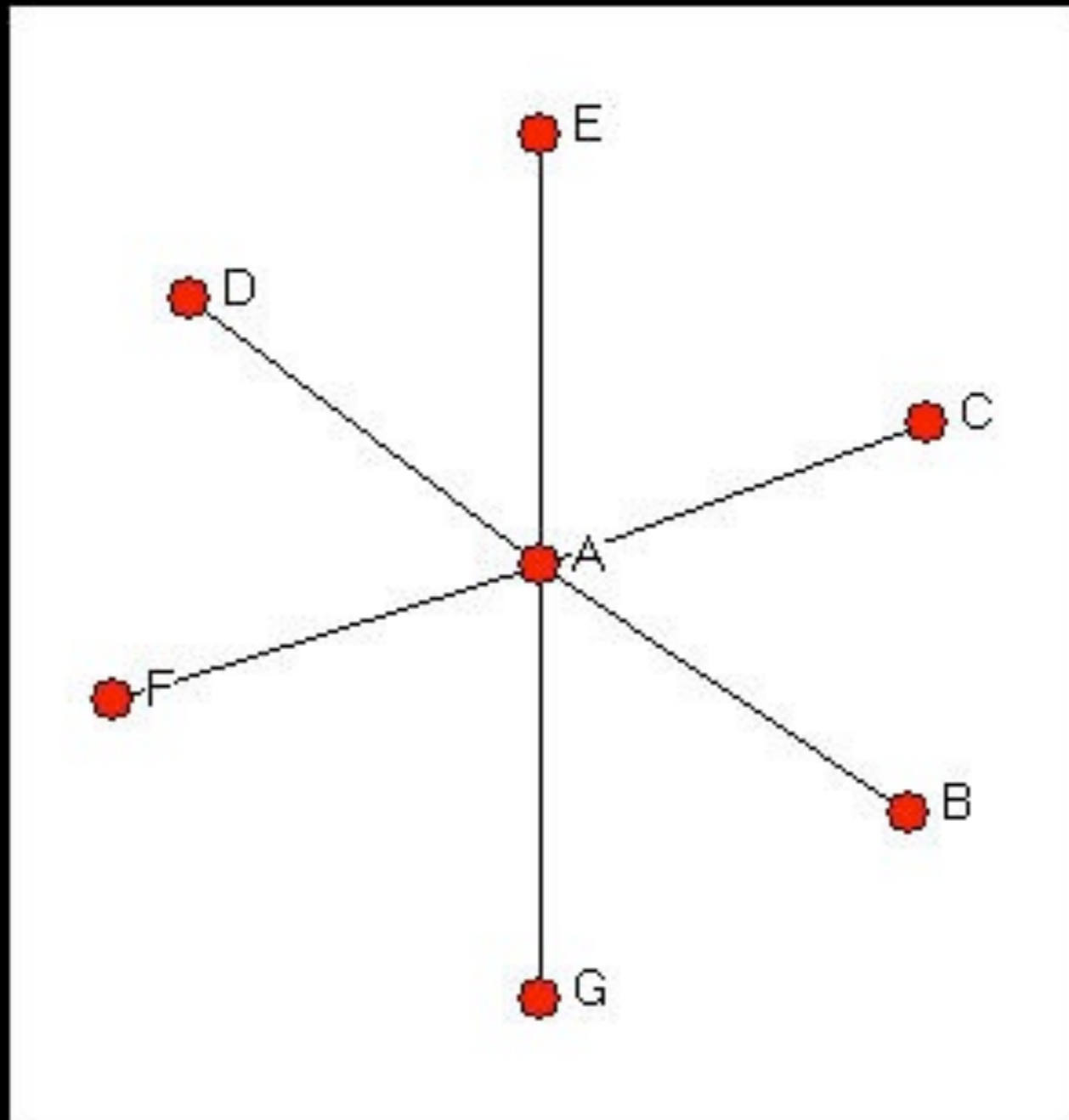
- Ego-centric



Star Network

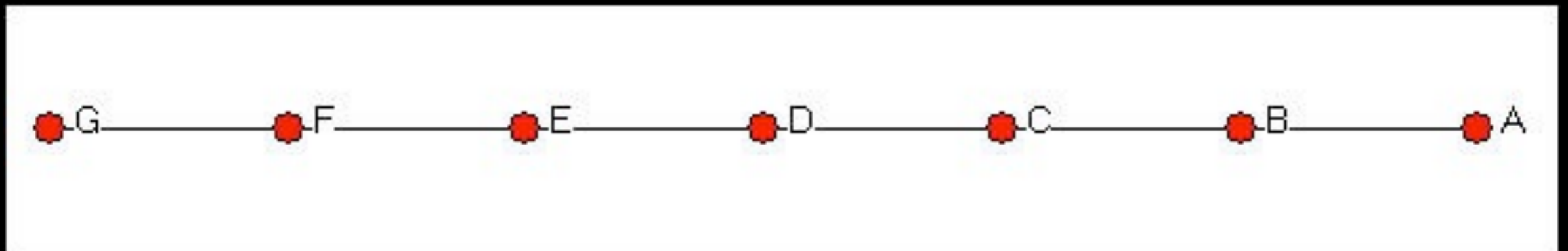
A: degree 6

degree 1



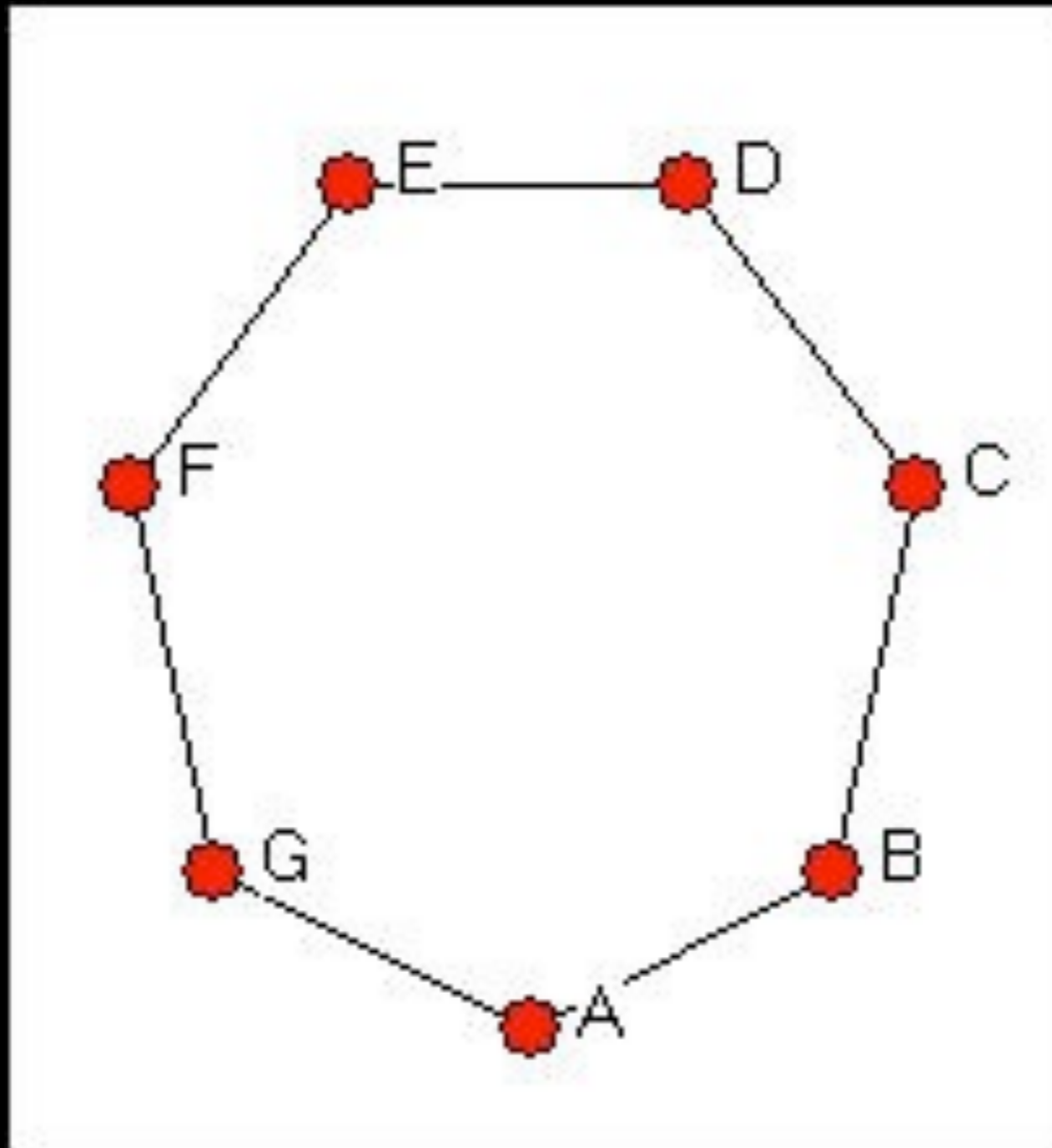
What's the best position

Line Network



What's the
best position

Circle Network



What's the
best position

Measures of Relation

- Binary
 - Connected/not-connected
- Multiple Categories
 - Friend, lover, student, advisor
- Ordinal (strength of ties)
 - Liked, disliked, neutral

Matrices

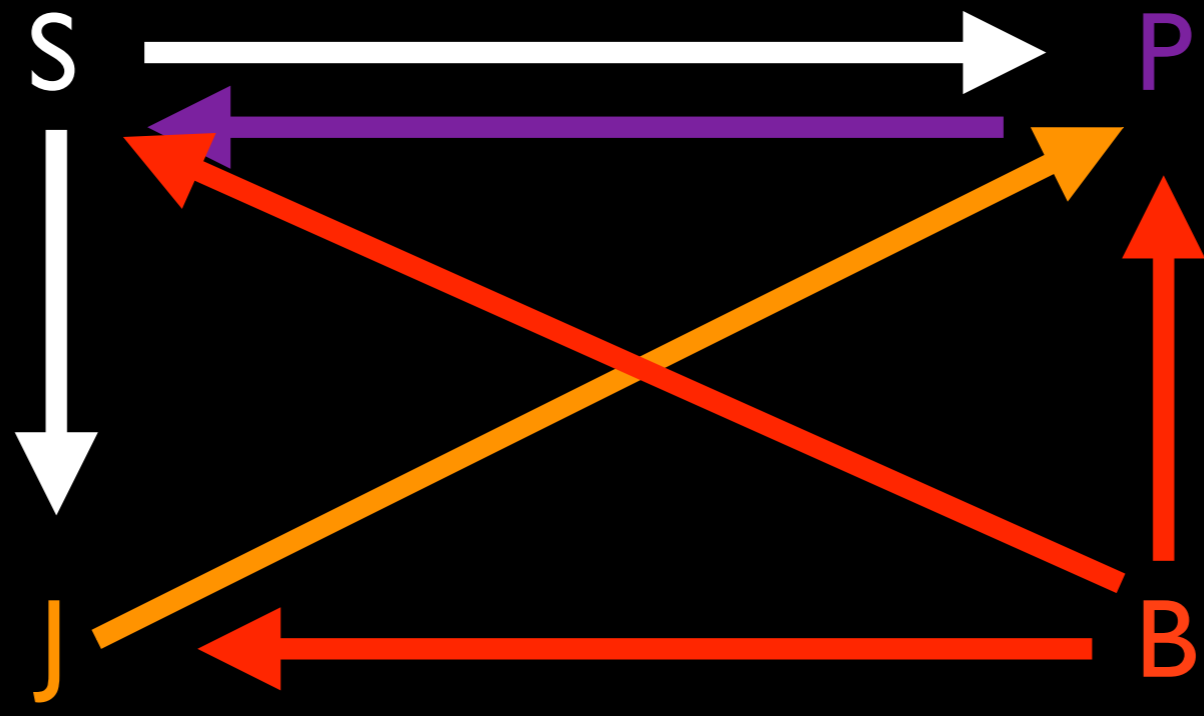
- Friends Shawn, Jeff, Patricia, and Beth
- Shawn likes Jeff and Patricia, but not Beth
- Jeff likes Patricia, but not Beth or Shawn
- Beth likes everyone
- Patricia only likes Shawn

S

P

J

B



Binary and Directed

	Shawn	Jeff	Patricia	Beth
Shawn				
Jeff				
Patricia				
Beth				

	Shawn	Jeff	Patricia	Beth
Shawn	-	1	1	0
Jeff	0	-	1	0
Patricia	1	0	-	0
Beth	1	1	1	-

- Binary (0/1)
- Signed (-1/0/+1)
- Valued Graph (1-100)

- Friends Shawn, Jeff, Patricia, and Beth
- Shawn:
Jeff (BFF) and Patricia (girlfriend), but not Beth (hates)
- Jeff
Patricia (acquaintance), but not Beth (dislikes) or Shawn (secretly hates)
- Beth likes everyone
- Patricia only likes Shawn (boyfriend)

Fill In

	Shawn	Jeff	Patricia	Beth
Shawn				
Jeff				
Patricia				
Beth				

Draw

S

P

J

B

Size

- Size: number of nodes in network
- What's the size?
- 4

Number of Possible Connections

- $k = \text{number of notes}$
- $k * (k-1)$
- $4 * (4-1) = 12$

Mean Strength

- Mean strength of ties
- Average of all of the ties

Degree Centrality

$$d(i) = \frac{\text{number of links to node } i}{k-1}$$

The higher the value, the more connected a node is

Closeness Centrality

$$c(i) = \sum \textit{Distance to each node}$$

The lower the value, the more connected a node is

Betweenness Centrality

$$b(i) = \sum \frac{\text{\# of shortest paths from } j \text{ to } k}{\text{\# of shortest paths from } j \text{ to } k \text{ that pass through node } i}$$

The higher the number, the more this node facilitates information flow in the network

Group Presentation if I
forgot.
:-)

How does SNA Apply to IR?

How can SNA change
our search results
(personalization)?