

Last Time

- Embeddedness
- Structural Holes
- Homophily
 - Homophily test
 - Drivers of homophily
 1. Selection
 2. Social influence

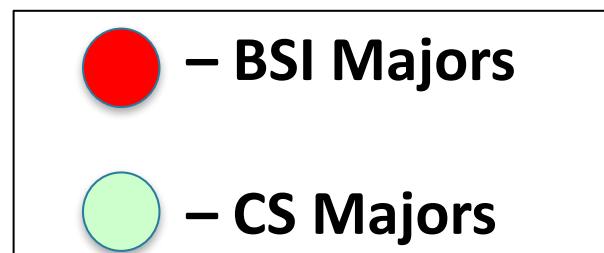
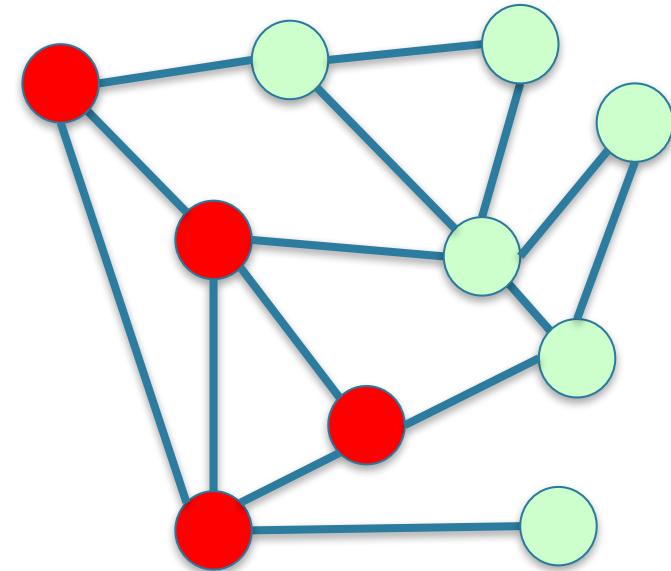
Distinguishing Between Selection and Social Influence

Is the homophily in this network due to selection, social influence, or something else?

Impossible to tell.

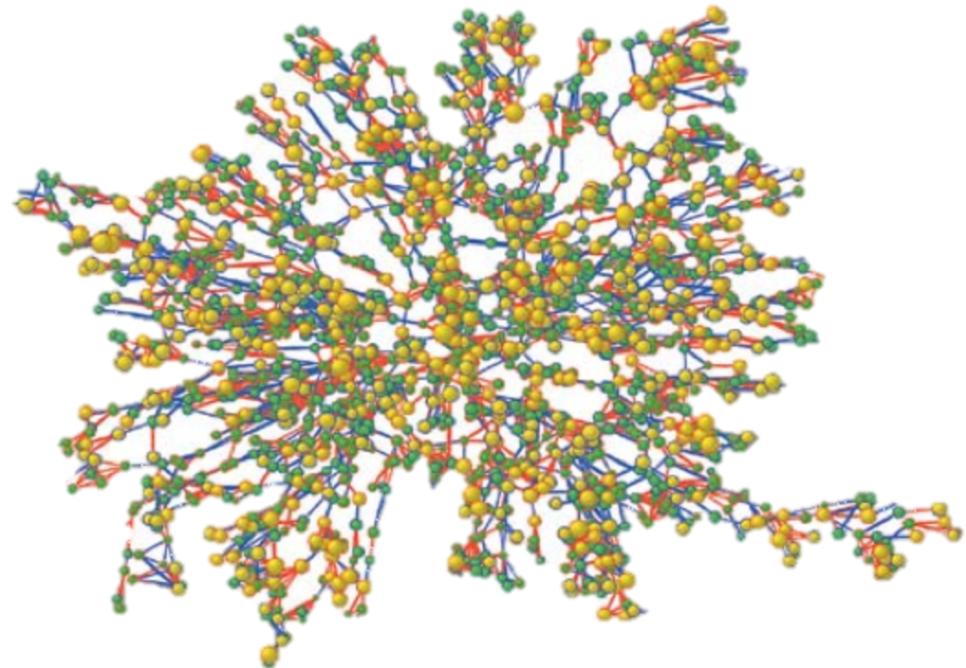
We would need information about the temporal dynamics of the network.

Specific mechanisms are important to identify if we want to make interventions to encourage or discourage a certain behavior.



The Spread of Obesity in Social Networks

- Social network of 12,067 people assessed repeatedly for 32 years (1971 to 2003).
- Yellow node denotes an obese person (body-mass index ≥ 30) and green denotes a nonobese person (sample network from 2000).



[Christakis – Fowler 2007]

- Homophily was observed in all 32 years.
- Longitudinal statistical models suggest that **social influence** and **not selection** was the driver of homophily.

The Spread of Obesity Media Coverage



©THE COLBERT REPORT

MENU ⌂

TIME

2013 2012 2011 2010 2009

The 2009 TIME 100 fblike Facebook Social Plugin
In our annual TIME 100 issue, we do the impossible: name the people who most affect our world

Like 408 Tweet 1 G+1 Share 4

SCIENTISTS & THINKERS
Nicholas Christakis
By Dan Ariely | Thursday, Apr. 30, 2009

96 of 109 ◀ ▶
VIEW ALL

Social scientists used to have a straightforward, if tongue-in-cheek, answer to the question of how to become happy: Surround yourself with people who are uglier, poorer and shorter than you are — and who are unhappily married and have annoying kids. You will compare yourself with these people, and the contrast will cheer you up.

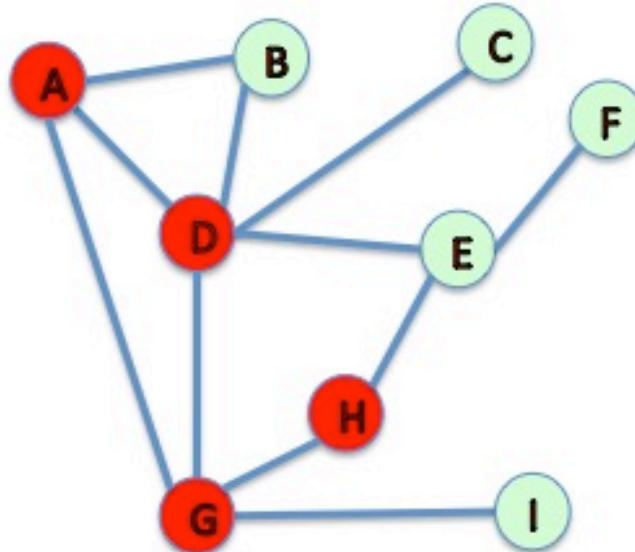
Nicholas Christakis, 47, a physician and sociologist at Harvard University, challenges this idea. Using data from a study that tracked about 5,000 people over 20 years, he suggests that happiness, like the flu, can spread from person to person.



<http://www.cc.com/video-clips/c3suh9/the-colbert-report-james-fowler>

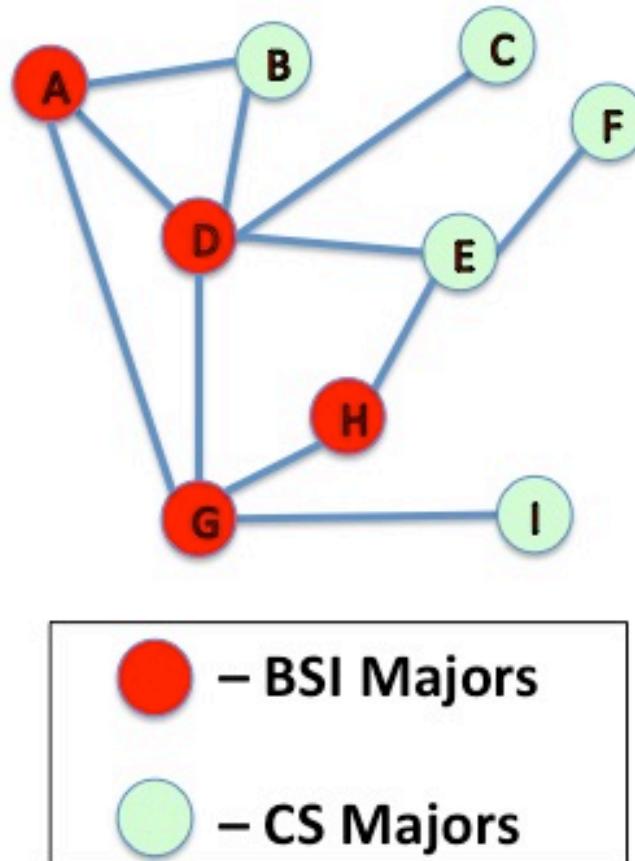
Warm up

Run the homophily test on
this network



Warm up

1. If edges were assigned randomly, approximately what percentage of edges would be connecting a BSI student to a CS students?
 - A. 55.3%
 - B. 30.5%
 - C. 49.4%
 - D. 63.6%



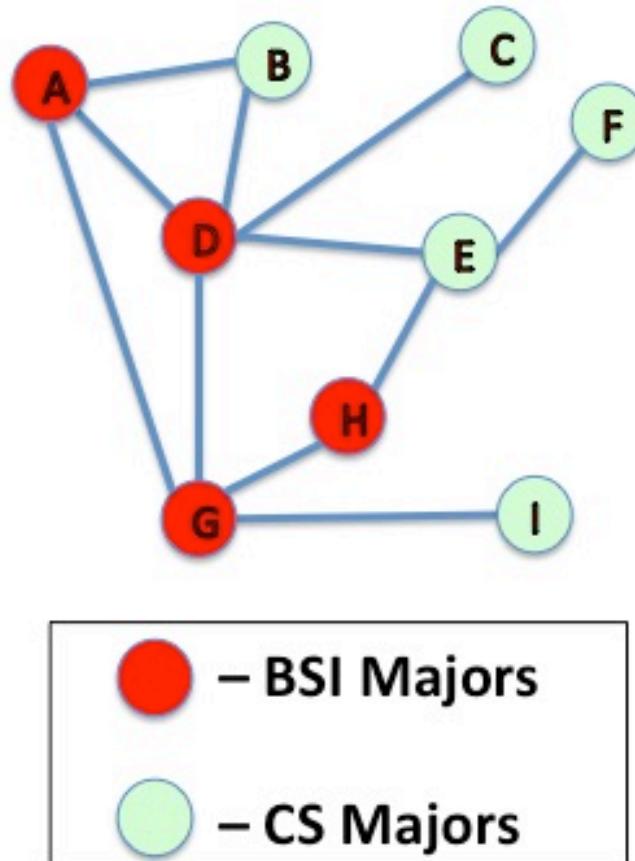
Warm up

1. If edges were assigned randomly, approximately what percentage of edges would be connecting a BSI student to a CS students?

$$p = \% \text{ BSI majors} = 4/9$$

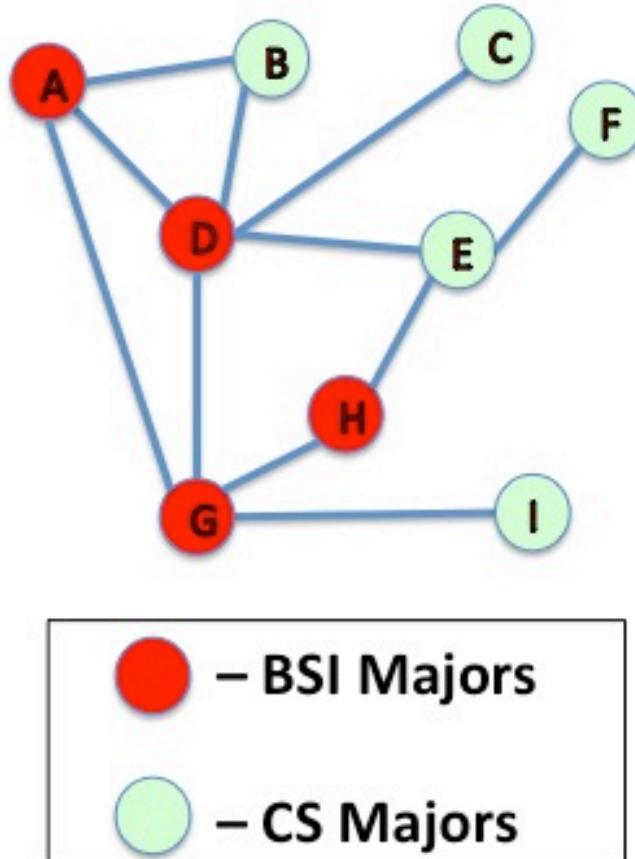
$$q = \% \text{ CS majors} = 5/9$$

$$2*p*q = 2*(4/9)*(5/9) = 0.494$$



Warm up

2. Does this network exhibit homophily?
 - A. Yes
 - B. No
 - C. Test is inconclusive
 - D. None of the above



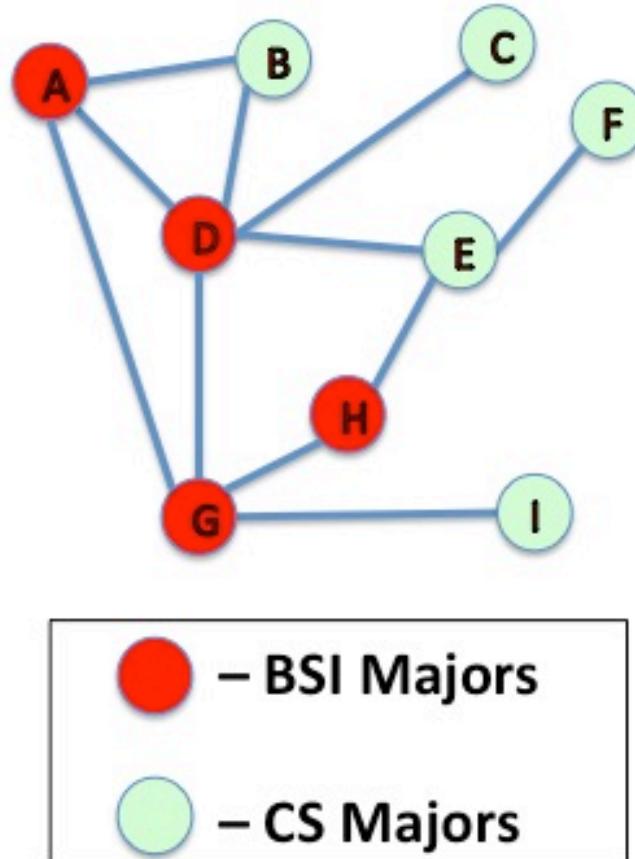
Warm up

2. Does this network exhibit homophily?

Expected fraction of cross-type edges with no homophily = .494

Actual percentage of cross-type edges = $6/11 = .545$

No, we do not have evidence of homophily.



Affiliation Networks

An **affiliation network** is a network that indicates the affiliation of people in groups or foci.

Affiliation Networks

An **affiliation network** is a network that indicates the affiliation of people in groups or foci.

Two kinds of nodes:

- i. People
- ii. Groups

Edges represent affiliation of people in the groups.

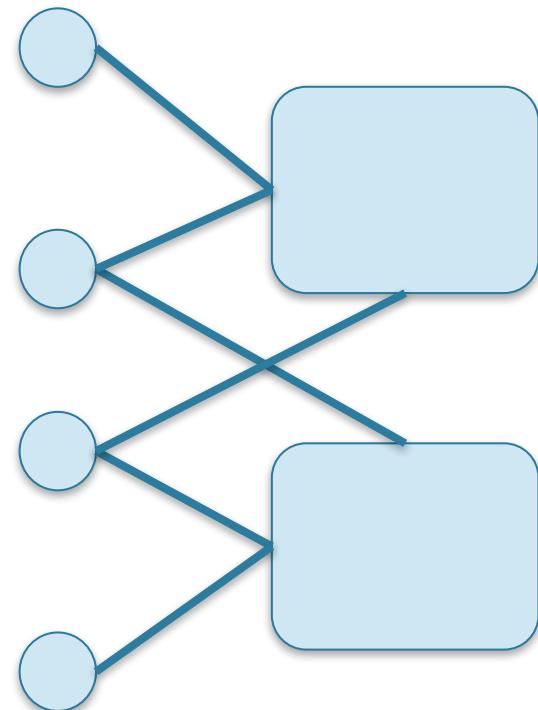
Affiliation Networks

An **affiliation network** is a network that indicates the affiliation of people in groups or foci.

Two kinds of nodes:

- i. People
- ii. Groups

Edges represent affiliation of people in the groups.



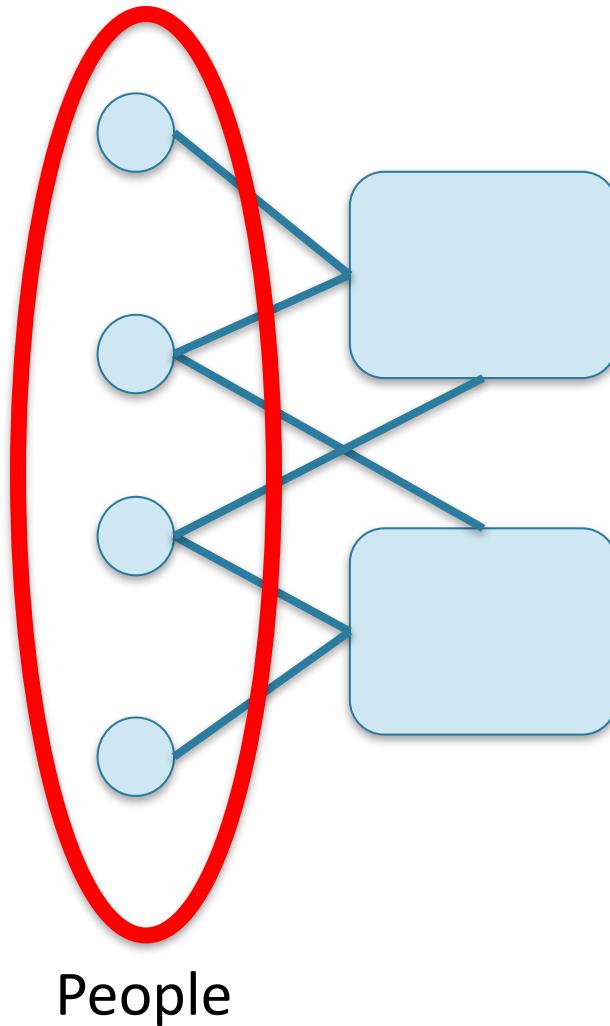
Affiliation Networks

An **affiliation network** is a network that indicates the affiliation of people in groups or foci.

Two kinds of nodes:

- i. People
- ii. Groups

Edges represent affiliation of people in the groups.



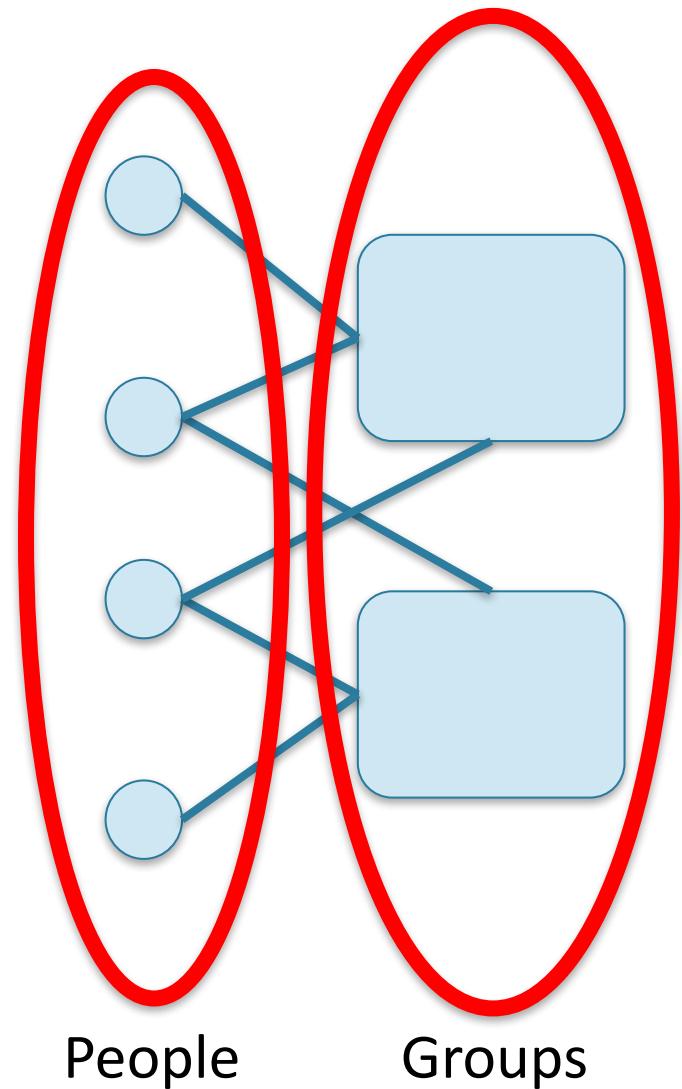
Affiliation Networks

An **affiliation network** is a network that indicates the affiliation of people in groups or foci.

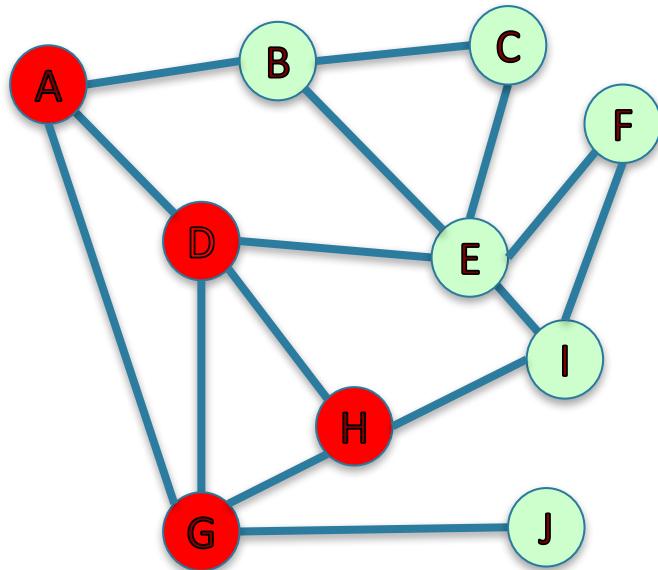
Two kinds of nodes:

- i. People
- ii. Groups

Edges represent affiliation of people in the groups.



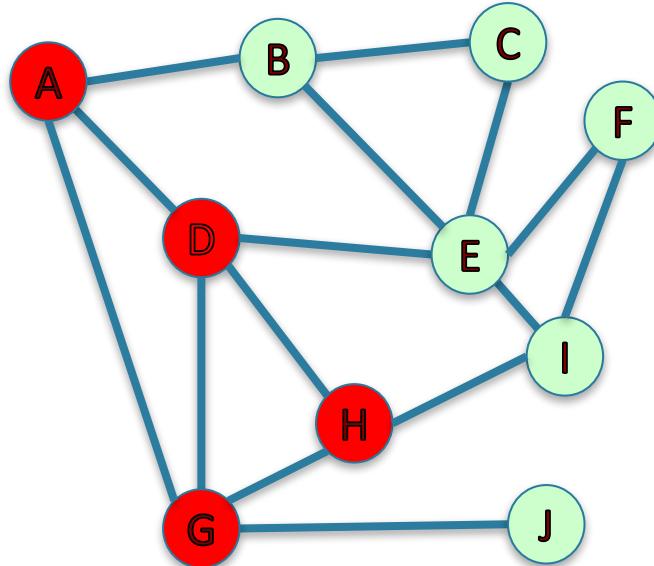
Affiliation Networks



● – BSI Majors

● – CS Majors

Affiliation Networks

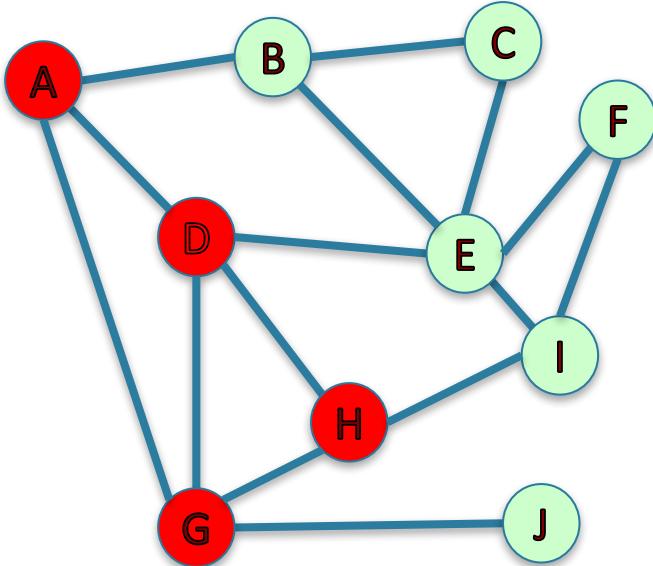


● – BSI Majors

● – CS Majors



Affiliation Networks



● – BSI Majors

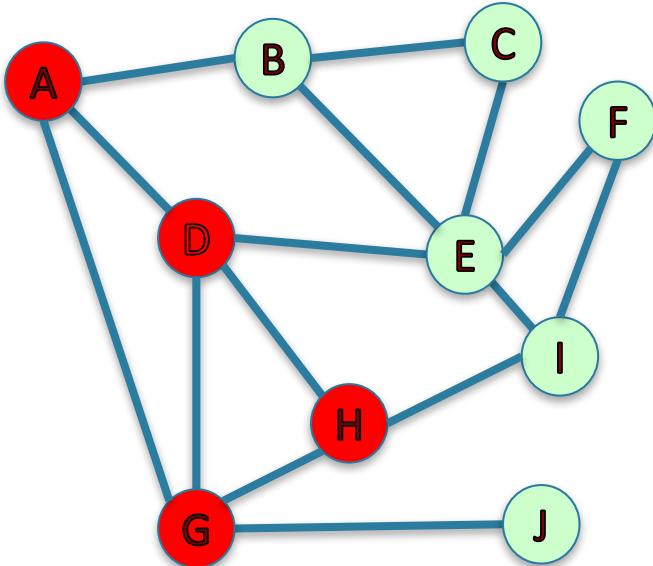
● – CS Majors



BSI

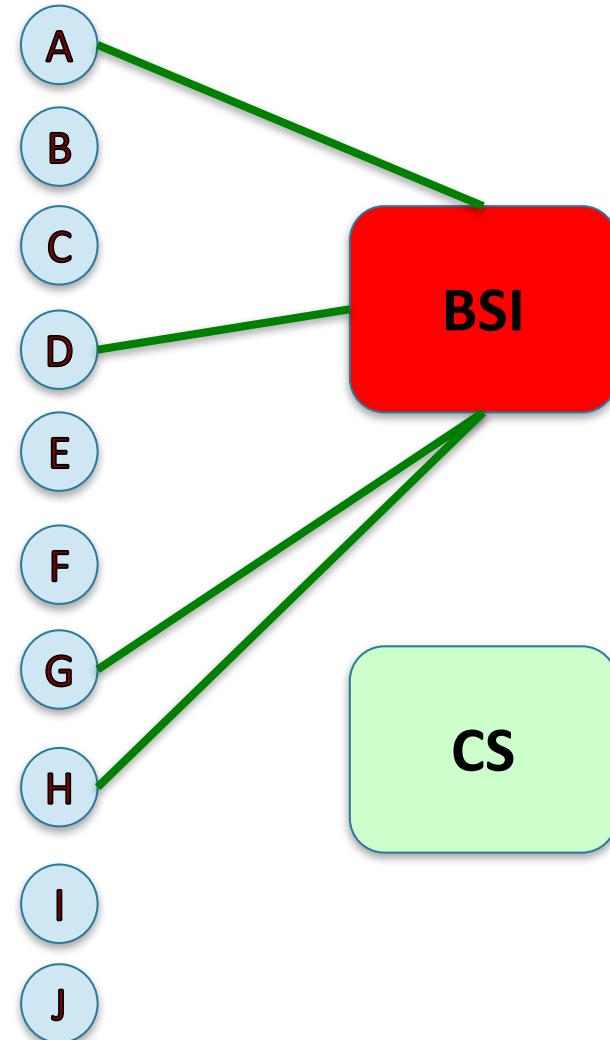
CS

Affiliation Networks

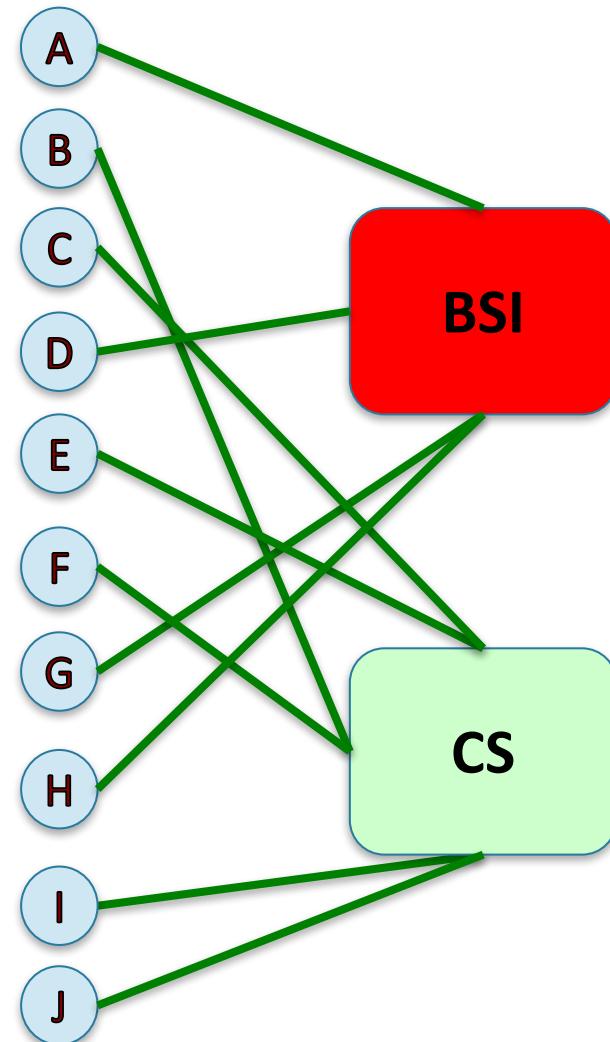
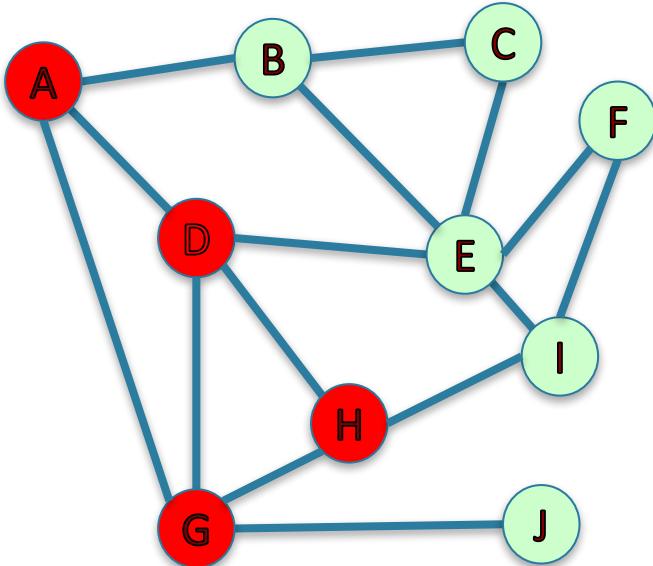


— BSI Majors

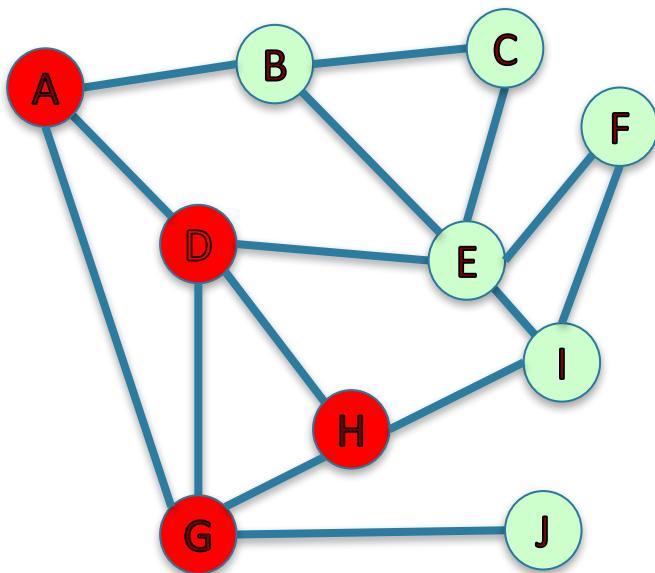
— CS Majors



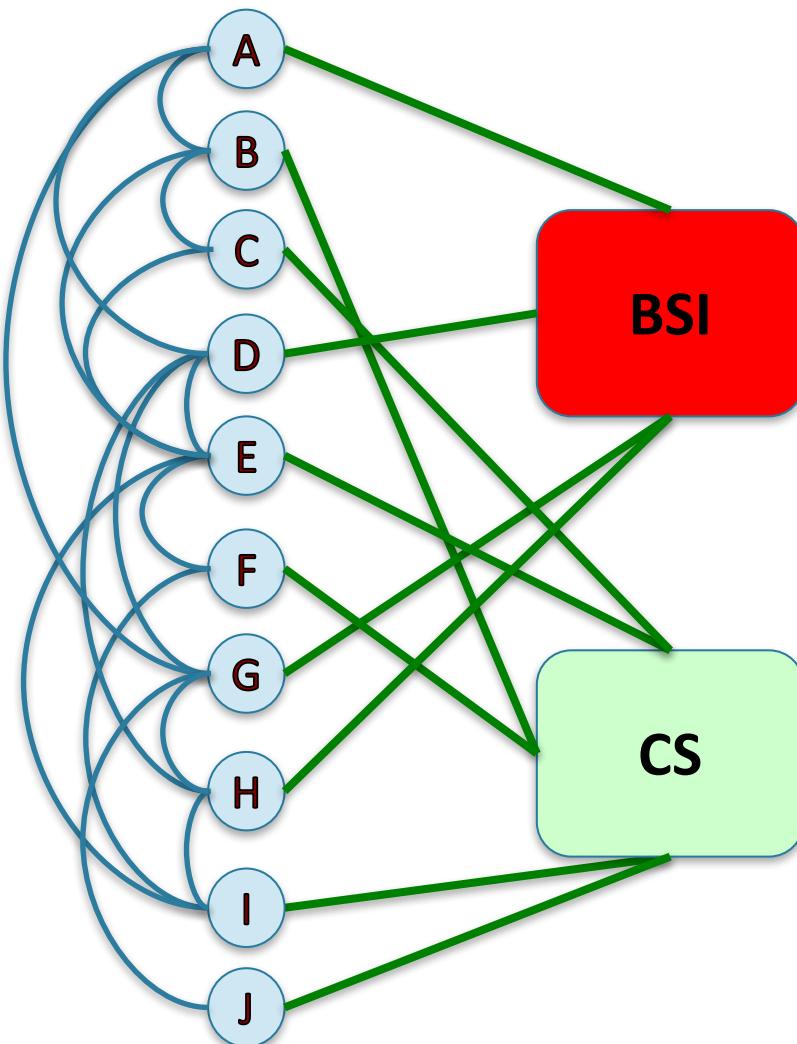
Affiliation Networks



Social-Affiliation Networks



- – BSI Majors
- – CS Majors



Social-Affiliation Networks

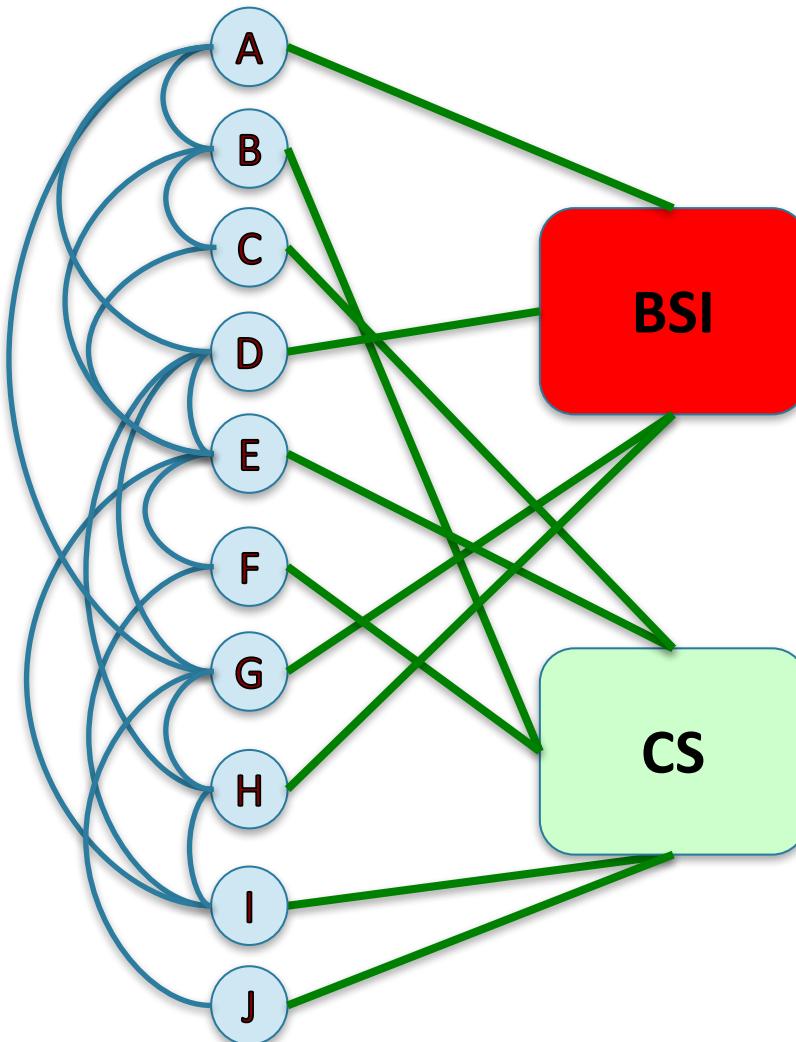
An **social-affiliation network** indicates affiliation to groups as well as social connections among people.

Two kinds of nodes:

- i. People
- ii. Groups

Two kind of edges:

- i. People to groups
- ii. People to people

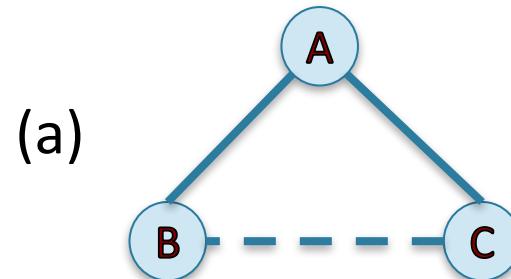


Closure in Social-Affiliation Networks

Types of closure in social-affiliation networks:

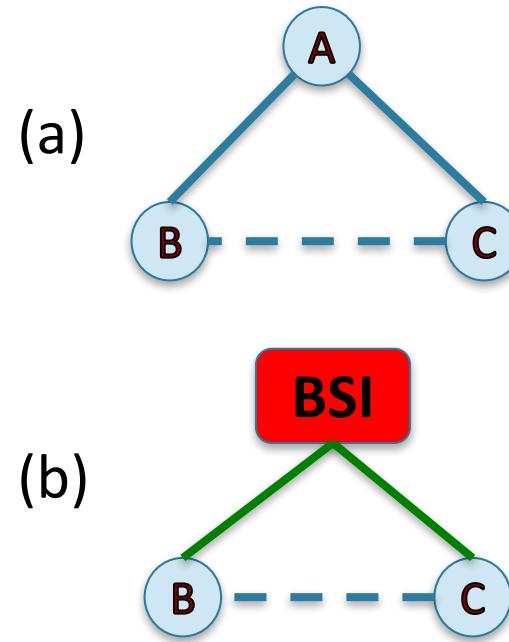
Closure in Social-Affiliation Networks

Types of closure in social-affiliation networks:



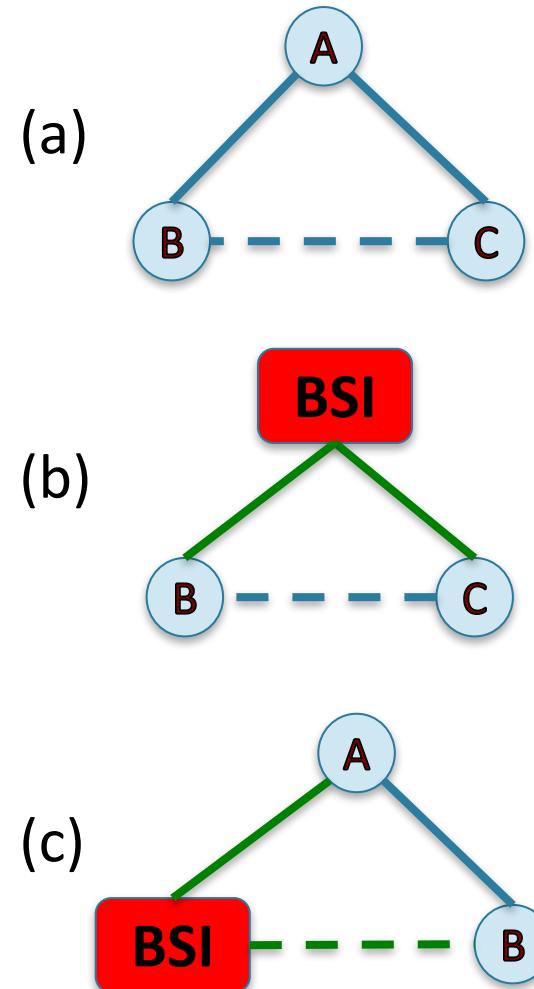
Closure in Social-Affiliation Networks

Types of closure in social-affiliation networks:



Closure in Social-Affiliation Networks

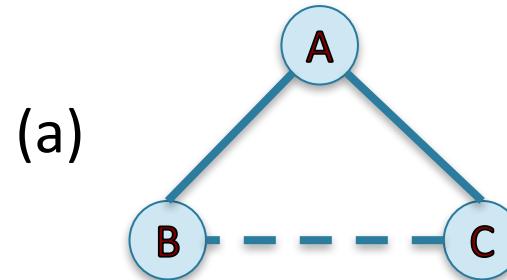
Types of closure in social-affiliation networks:



Closure in Social-Affiliation Networks

Types of closure in social-affiliation networks:

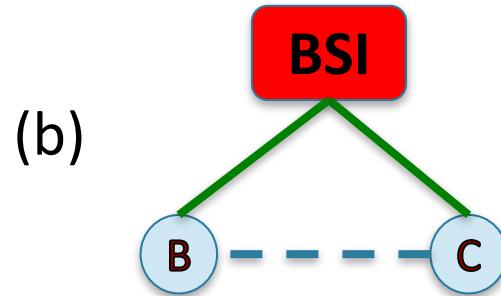
- (a) ***Triadic closure:*** People with a friend in common are more likely to become friends.



Closure in Social-Affiliation Networks

Types of closure in social-affiliation networks:

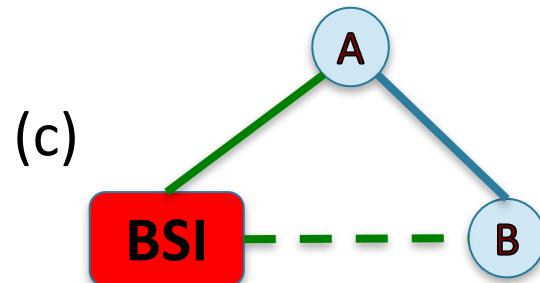
- (a) ***Triadic closure***: People with a friend in common are more likely to become friends.
- (b) ***Group or focal closure***: People who belong to the same group are more likely to become friends.



Closure in Social-Affiliation Networks

Types of closure in social-affiliation networks:

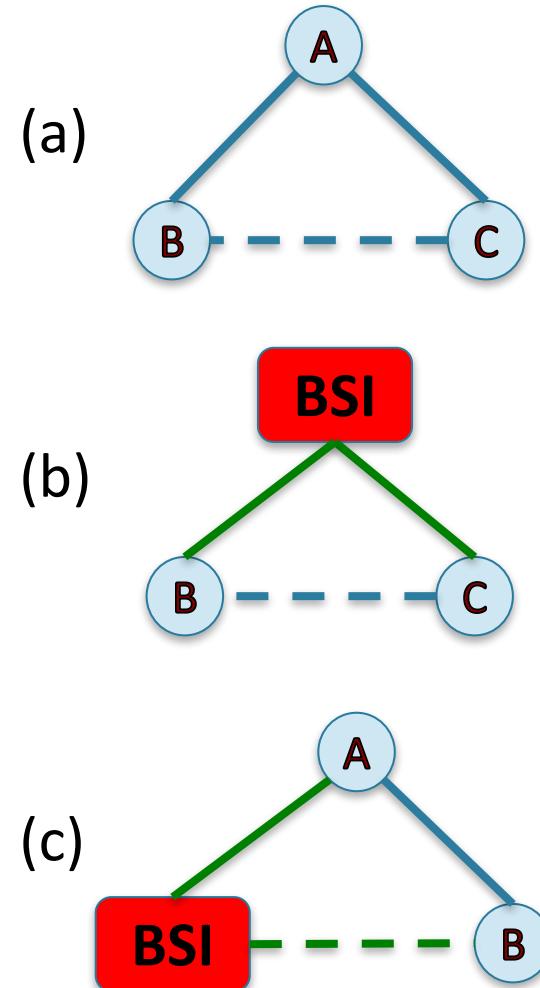
- (a) ***Triadic closure***: People with a friend in common are more likely to become friends.
- (b) ***Group or focal closure***: People who belong to the same group are more likely to become friends.
- (c) ***Membership closure***: People are more likely to join groups where they already have friends.



Closure in Social-Affiliation Networks

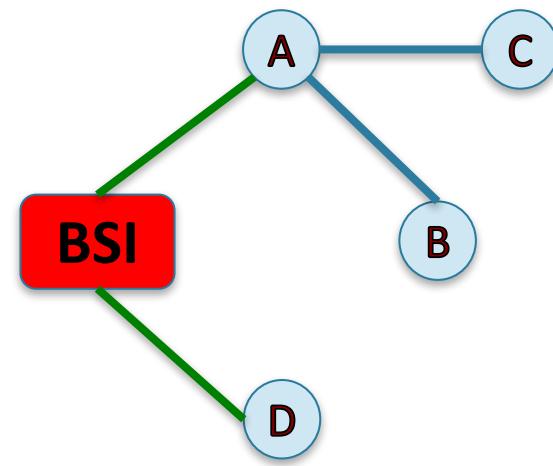
Types of closure in social-affiliation networks:

- (a) **Triadic closure:** People with a friend in common are more likely to become friends.
- (b) **Group or focal closure:** People who belong to the same group are more likely to become friends.
- (c) **Membership closure:** People are more likely to join groups where they already have friends.



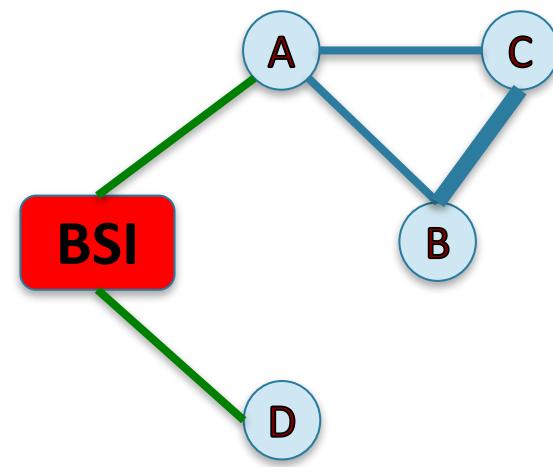
Closure in Social-Affiliation Networks

Possible interpretations of the different types of closure:



Closure in Social-Affiliation Networks

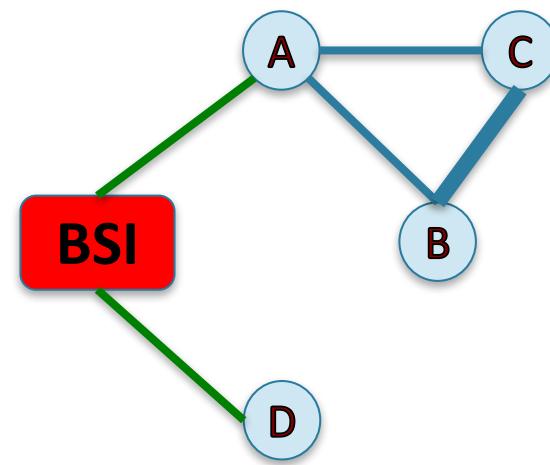
Possible interpretations of the different types of closure:



Closure in Social-Affiliation Networks

Possible interpretations of the different types of closure:

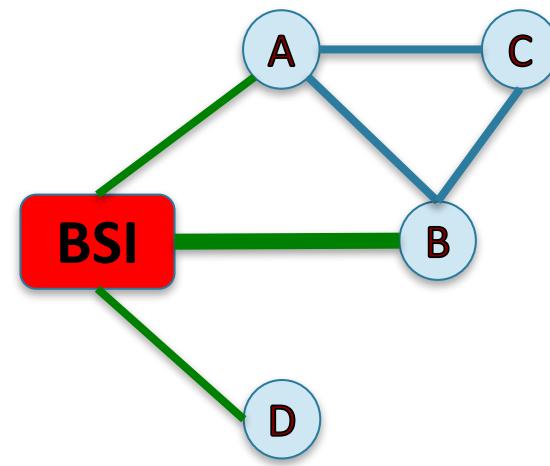
B – C: Triadic closure. A introduced B to C.



Closure in Social-Affiliation Networks

Possible interpretations of the different types of closure:

B – C: Triadic closure. A introduced B to C.

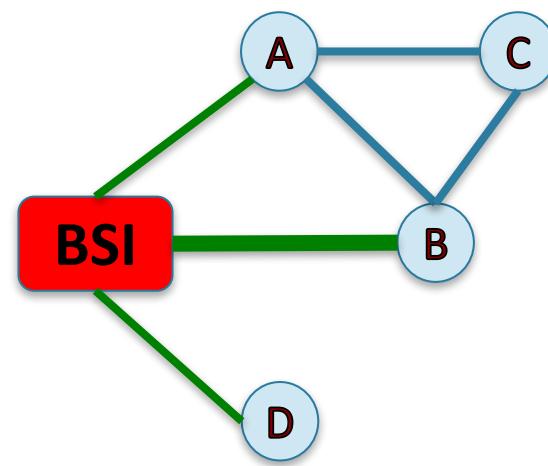


Closure in Social-Affiliation Networks

Possible interpretations of the different types of closure:

B – C: Triadic closure. A introduced B to C.

BSI – B: Membership closure. B chose to become a BSI student because her best friend, A, really enjoyed being a BSI student (social influence).

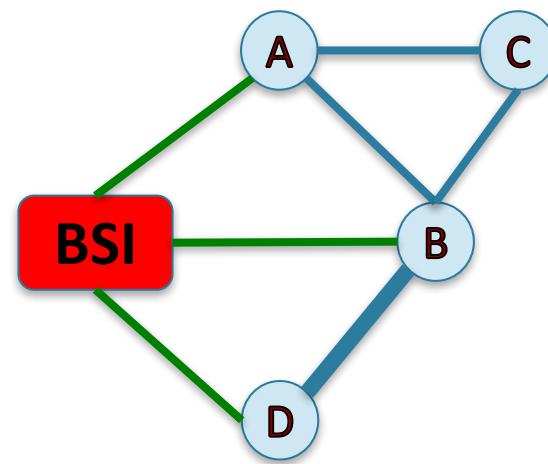


Closure in Social-Affiliation Networks

Possible interpretations of the different types of closure:

B – C: Triadic closure. A introduced B to C.

BSI – B: Membership closure. B chose to become a BSI student because her best friend, A, really enjoyed being a BSI student (social influence).



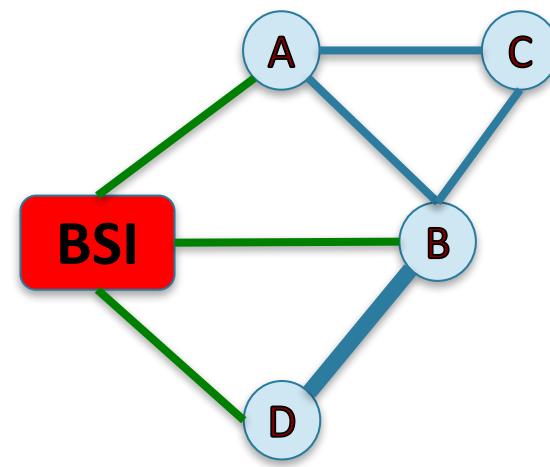
Closure in Social-Affiliation Networks

Possible interpretations of the different types of closure:

B – C: Triadic closure. A introduced B to C.

BSI – B: Membership closure. B chose to become a BSI student because her best friend, A, really enjoyed being a BSI student (social influence).

B – D: Group or focal closure: B and D became friends because they were both BSI students (selection).



How prevalent is closure in online
social-affiliation networks?

Triadic Closure in Email Communication

Nodes: 22K students at a U.S. university.

Edges: Reciprocal e-mail communication.

Triadic Closure in Email Communication

Nodes: 22K students at a U.S. university.

Edges: Reciprocal e-mail communication.

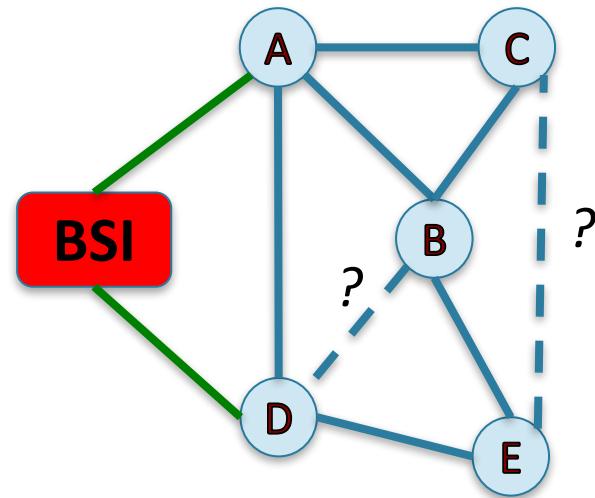
Authors take snapshots of the network at various times, and compute *probability of a link given the number of friends in common*.

Triadic Closure in Email Communication

Nodes: 22K students at a U.S. university.

Edges: Reciprocal e-mail communication.

Authors take snapshots of the network at various times, and compute *probability of a link given the number of friends in common*.

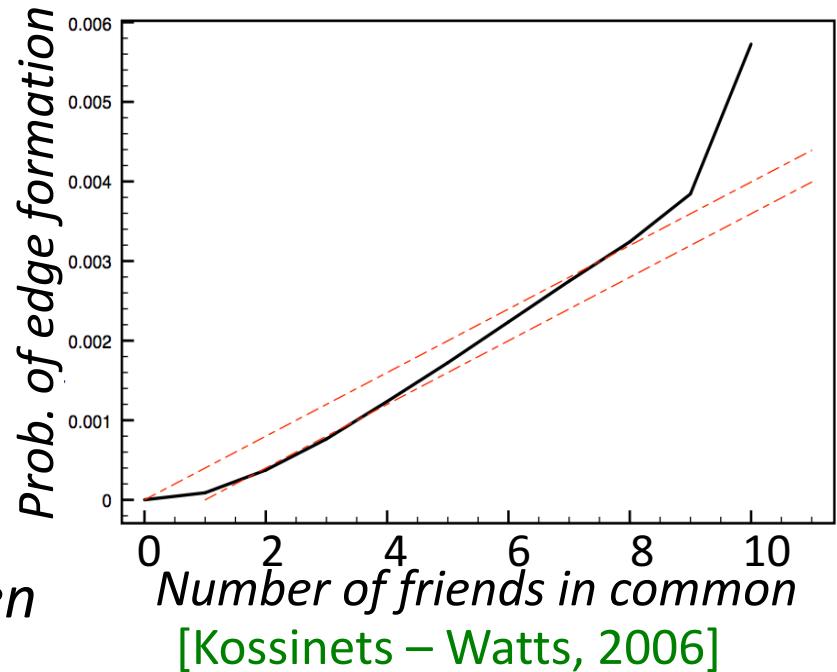


Triadic Closure in Email Communication

Nodes: 22K students at a U.S. university.

Edges: Reciprocal e-mail communication.

Authors take snapshots of the network at various times, and compute *probability of a link given the number of friends in common*.

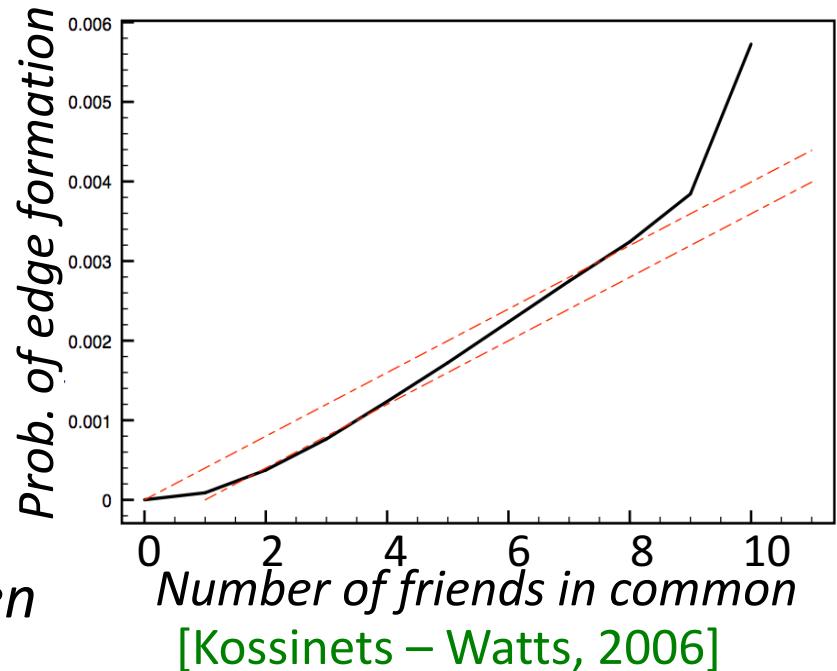


Triadic Closure in Email Communication

Nodes: 22K students at a U.S. university.

Edges: Reciprocal e-mail communication.

Authors take snapshots of the network at various times, and compute *probability of a link given the number of friends in common*.



Results:

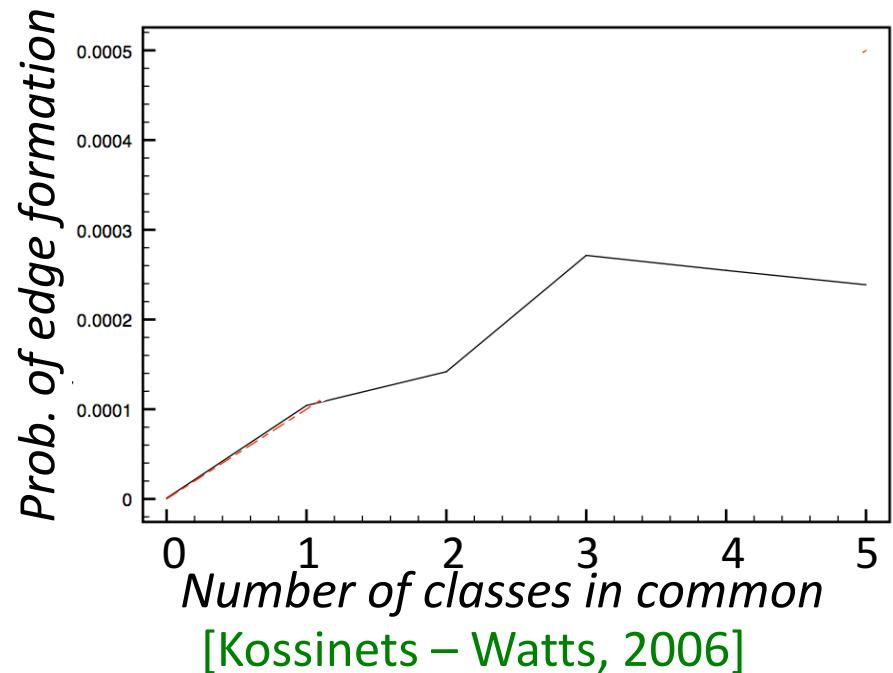
- Prob. of link increases with number of friend in common.
- Evidence of **triadic closure** in the formation of the network.

Group Closure in Email Communication

Authors take snapshots of the network at various times, and compute *probability of a link given the number of classes in common.*

Group Closure in Email Communication

Authors take snapshots of the network at various times, and compute *probability of a link given the number of classes in common*.

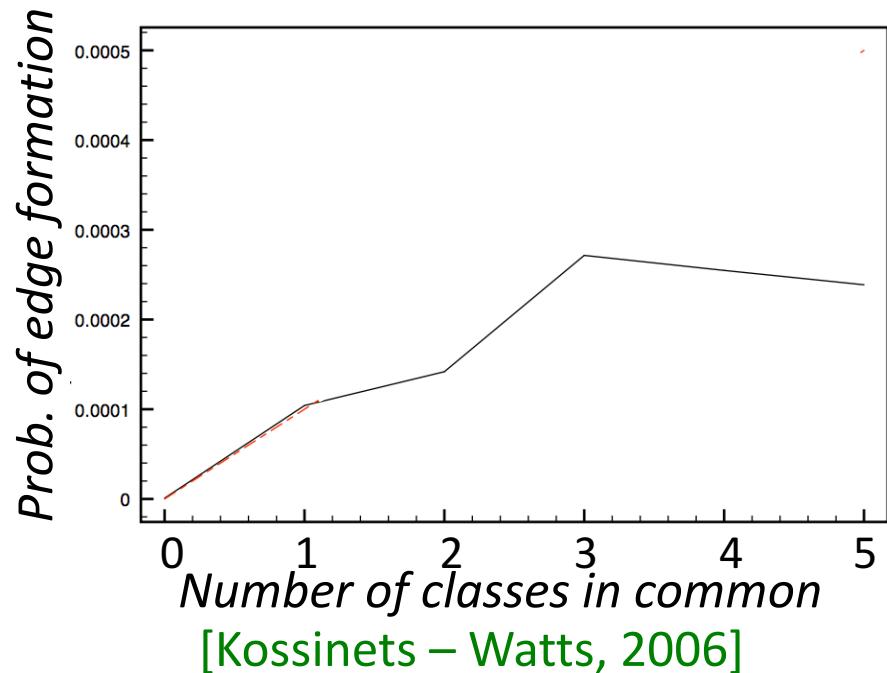


Group Closure in Email Communication

Authors take snapshots of the network at various times, and compute *probability of a link given the number of classes in common*.

Results:

- Prob. of link initially increases with number of classes in common, but stops increasing after 3 classes.
- Evidence of **group closure** in the formation of the network, but it's not as strong and persistent as triadic closure.



Membership Closure in Wikipedia

Nodes: Wikipedia editors.

Edges: Reciprocal communication
on Wikipedia.

Membership Closure in Wikipedia

Nodes: Wikipedia editors.

Edges: Reciprocal communication
on Wikipedia.

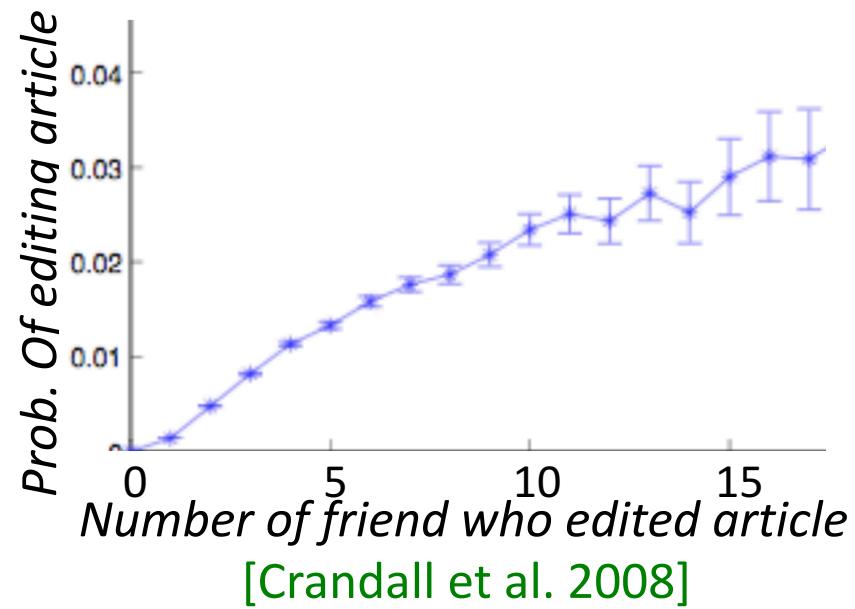
Authors take snapshots of the network at various times, and compute *probability of editing an article given the number of friends who have edited the article.*

Membership Closure in Wikipedia

Nodes: Wikipedia editors.

Edges: Reciprocal communication
on Wikipedia.

Authors take snapshots of the network at various times, and compute *probability of editing an article given the number of friends who have edited the article.*



[Crandall et al. 2008]

Membership Closure in Wikipedia

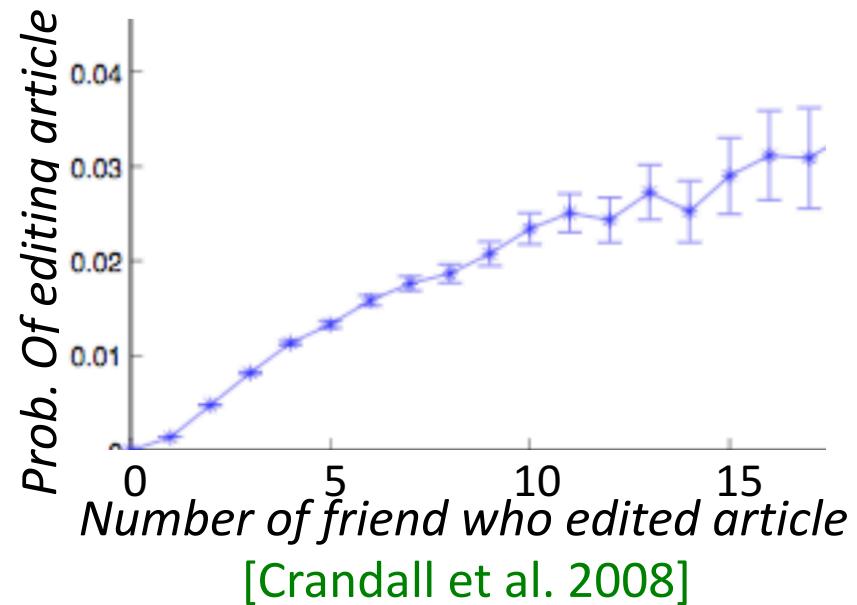
Nodes: Wikipedia editors.

Edges: Reciprocal communication on Wikipedia.

Authors take snapshots of the network at various times, and compute *probability of editing an article given the number of friends who have edited the article*.

Results:

Prob. of editing article increases with number of friends who have edited article. Evidence of **membership closure**.



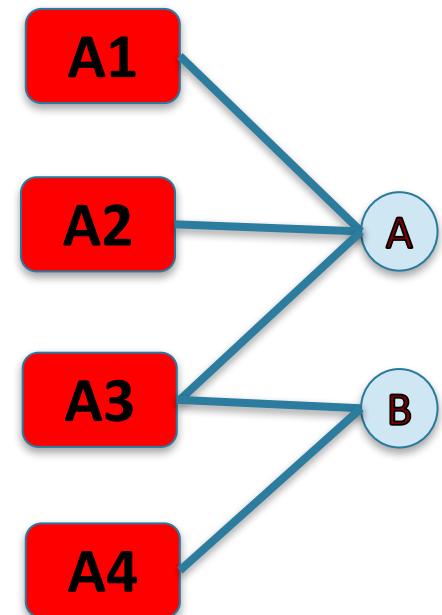
[Crandall et al. 2008]

Selection and Social Influence in Wikipedia

Similarity between two editors A and B:

$$\frac{\text{Number of articles edited by both A and B}}{\text{Number of articles edited by at least one of A and B}}$$

In this example A and B have similarity 1/4

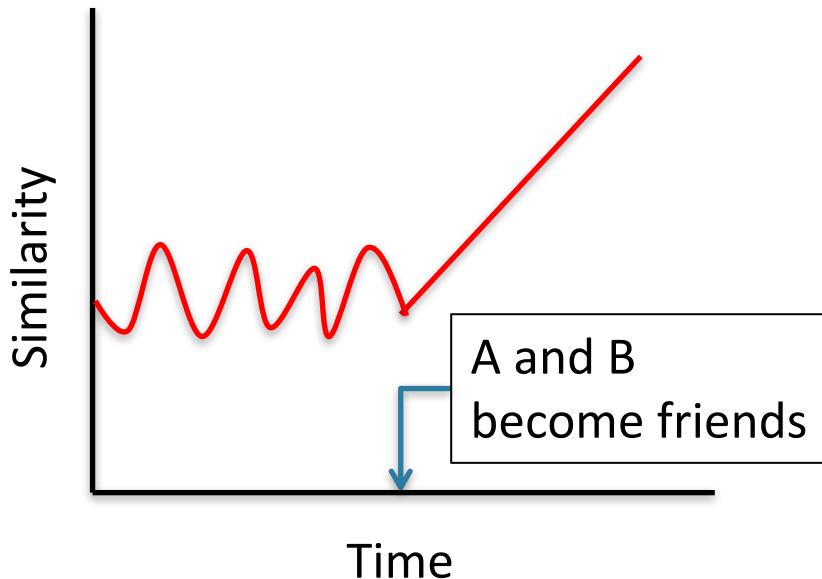


Selection and Social Influence in Wikipedia

Assume editors A and B become friends on Wikipedia.
How does similarity of A and B change before and after A and B
become friends?

Selection and Social Influence in Wikipedia

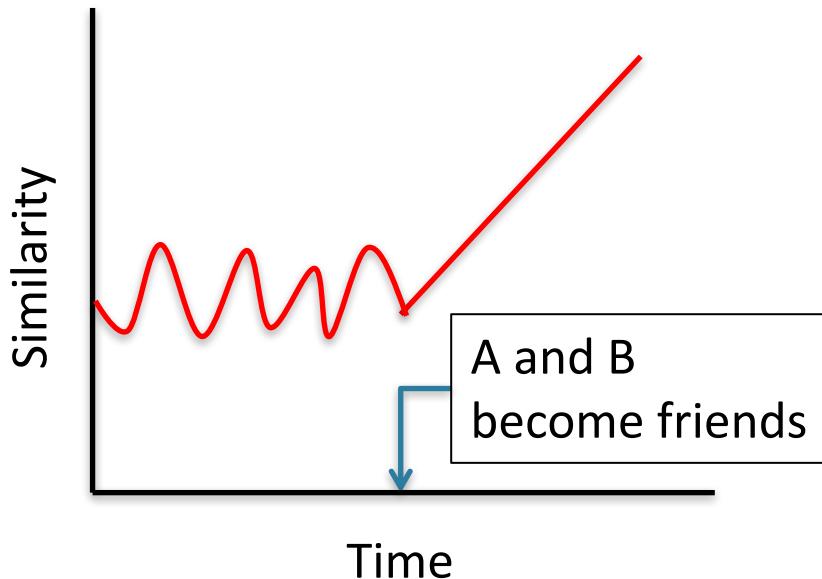
Assume editors A and B become friends on Wikipedia.
How does similarity of A and B change before and after A and B become friends?



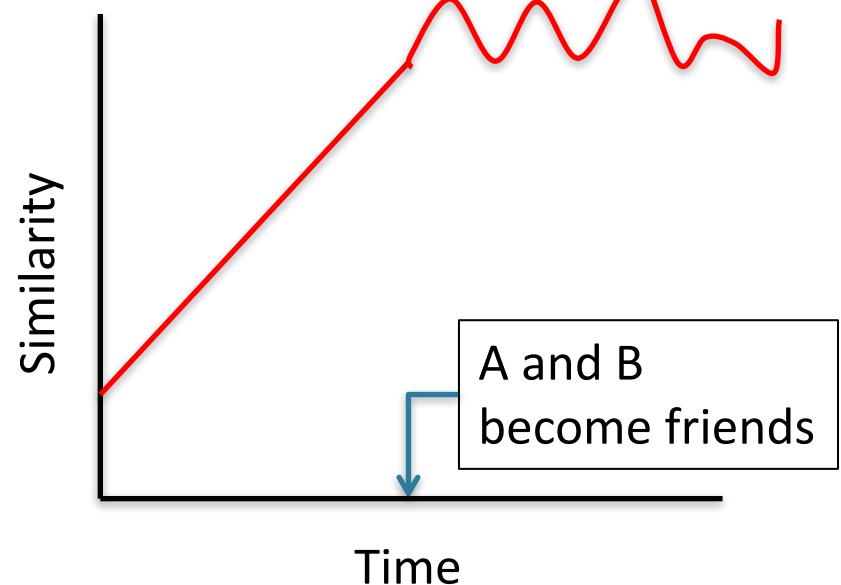
Social Influence

Selection and Social Influence in Wikipedia

Assume editors A and B become friends on Wikipedia.
How does similarity of A and B change before and after A and B become friends?

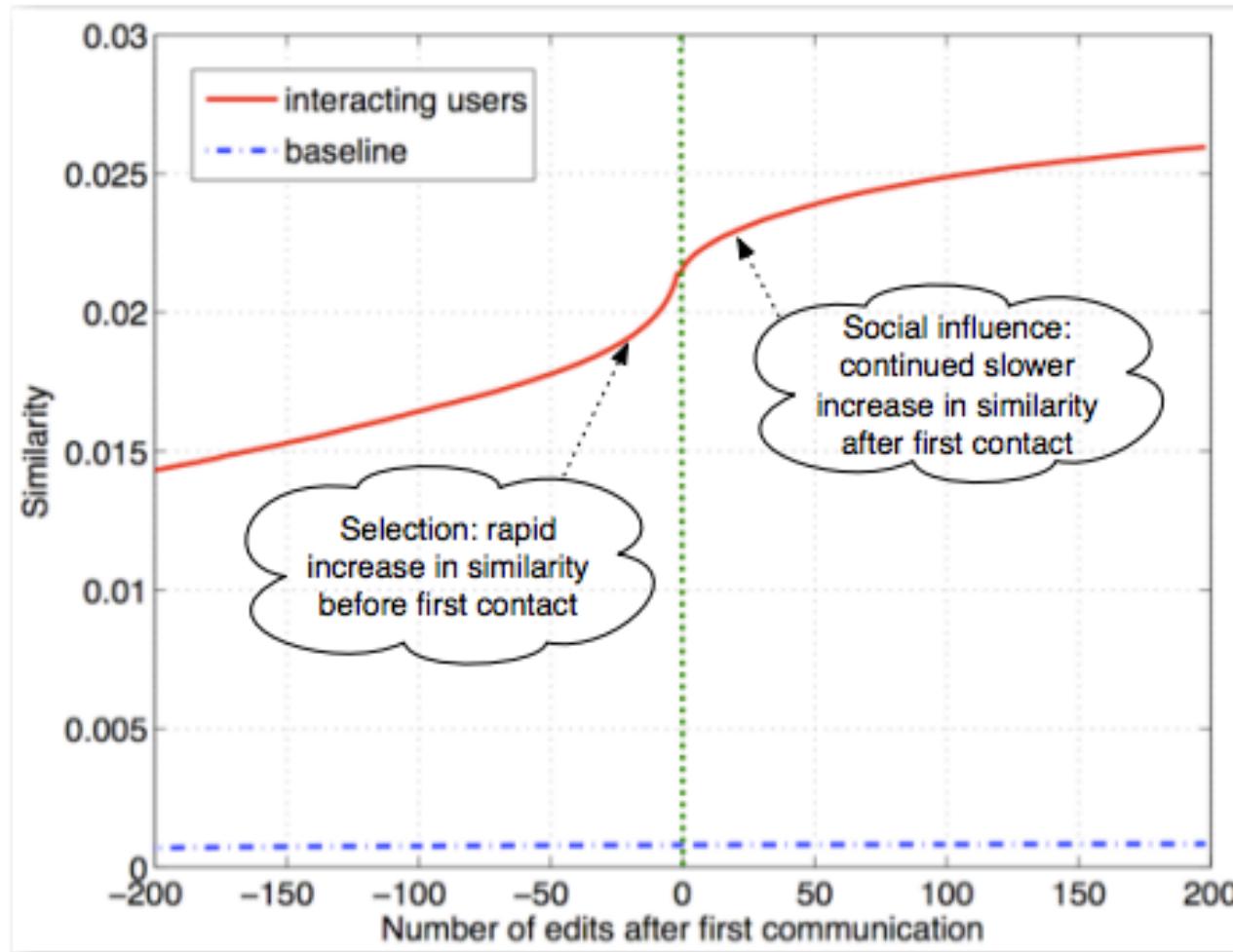


Social Influence



Selection

Selection and Social Influence in Wikipedia



[Crandall et al. 2008]

Positive/Negative Edges

- Relationships in networks discussed so far have been mostly positive.

Positive/Negative Edges

- Relationships in networks discussed so far have been mostly positive.
 - friendship

Positive/Negative Edges

- Relationships in networks discussed so far have been mostly positive.
 - friendship
 - collaboration

Positive/Negative Edges

- Relationships in networks discussed so far have been mostly positive.
 - friendship
 - collaboration
 - following on Twitter

Positive/Negative Edges

- Relationships in networks discussed so far have been mostly positive.
 - friendship
 - collaboration
 - following on Twitter
 - membership within a group

Positive/Negative Edges

- are there networks containing the edges represent negative relationships?

Positive/Negative Edges

- are there networks containing the edges represent negative relationships?
- Controversies, disagreements, and outright conflicts are important components of interactions between people or groups.

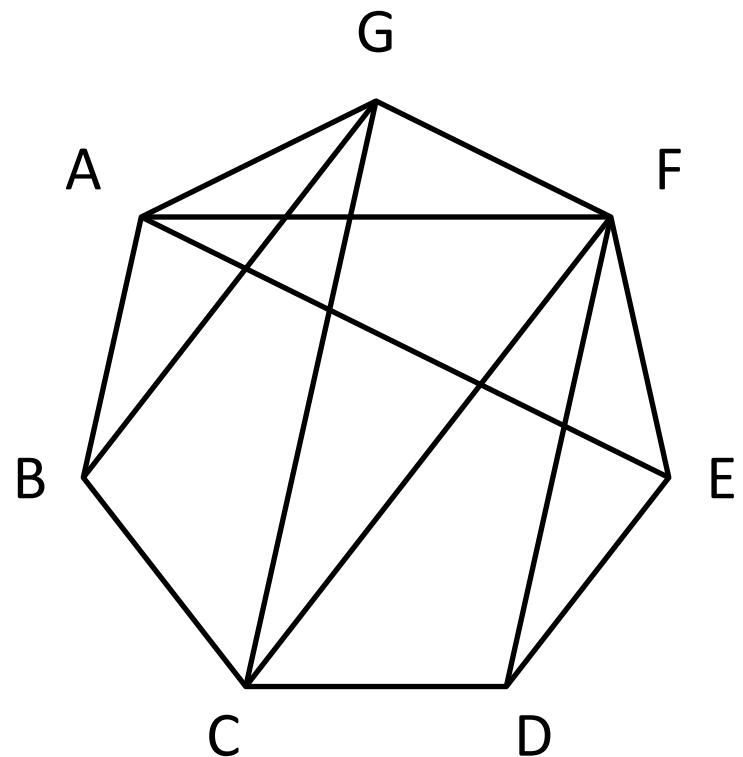
Positive/Negative Edges

- are there networks containing the edges represent negative relationships?
- Controversies, disagreements, and outright conflicts are important components of interactions between people or groups.
- Interpersonal relationships can be friendship (+) or antagonism (-).

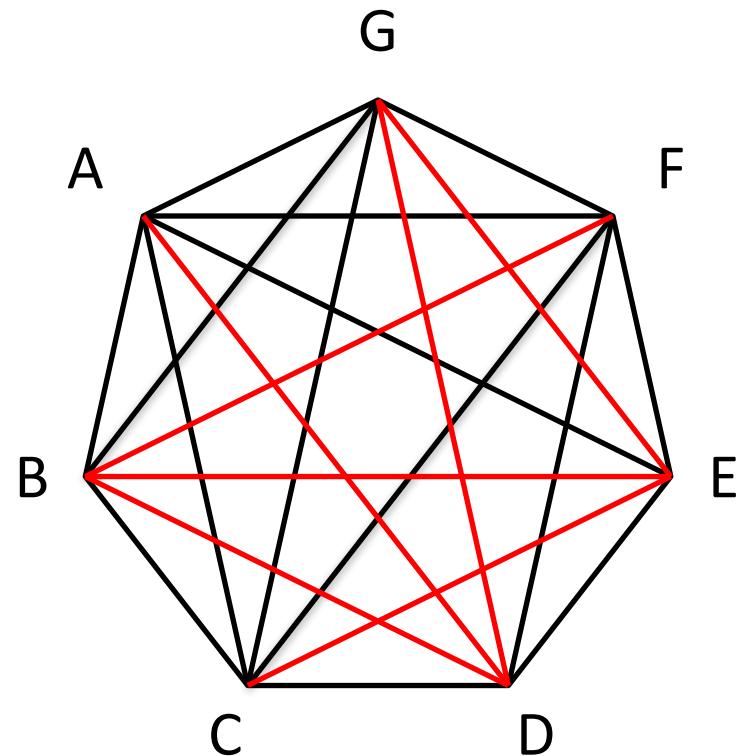
Positive/Negative Edges

- are there networks containing the edges represent negative relationships?
- Controversies, disagreements, and outright conflicts are important components of interactions between people or groups.
- Interpersonal relationships can be friendship (+) or antagonism (-).
- How do nodes interact within such a network?

Complete graph

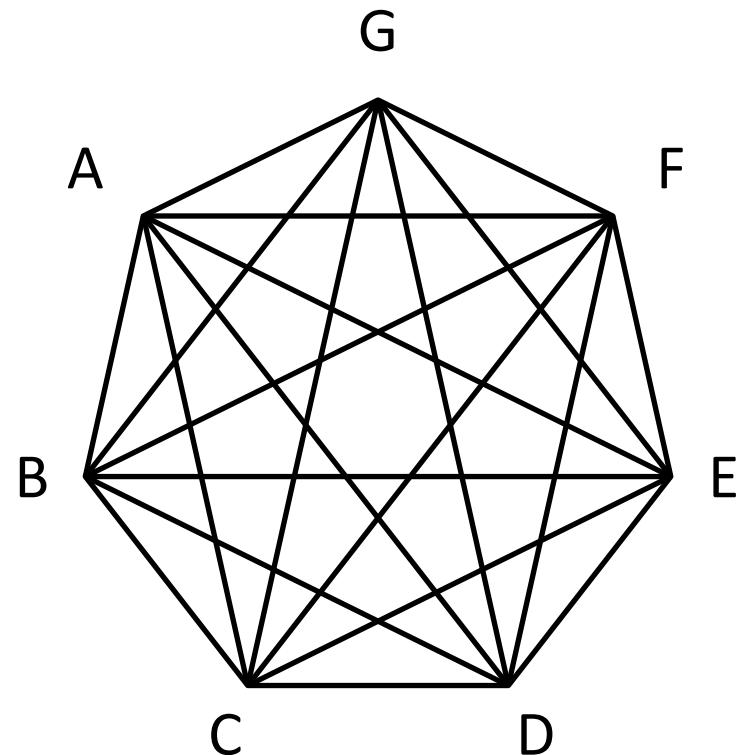


Complete graph



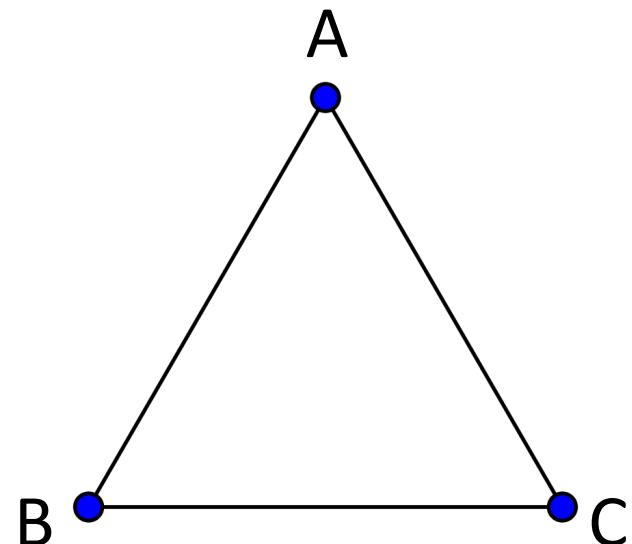
Complete graph

A network where there is an edge connecting each pair of nodes is called a **complete graph**.



Complete graph

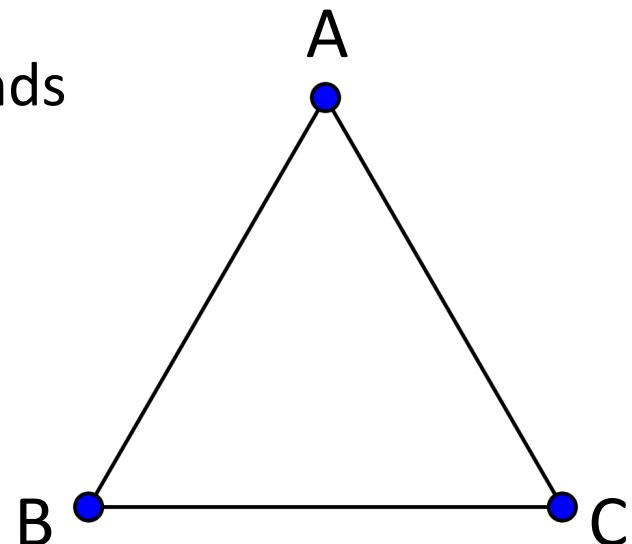
A network where there is an edge connecting each pair of nodes is called a **complete graph**.



Complete graph

A network where there is an edge connecting each pair of nodes is called a **complete graph**.

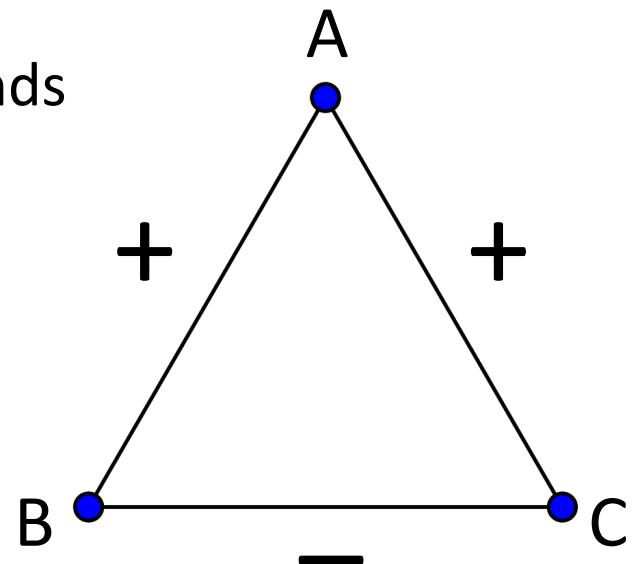
Label each edge with either + or –
A ‘+’ stands for friendship and a ‘–’ stands for enemy.



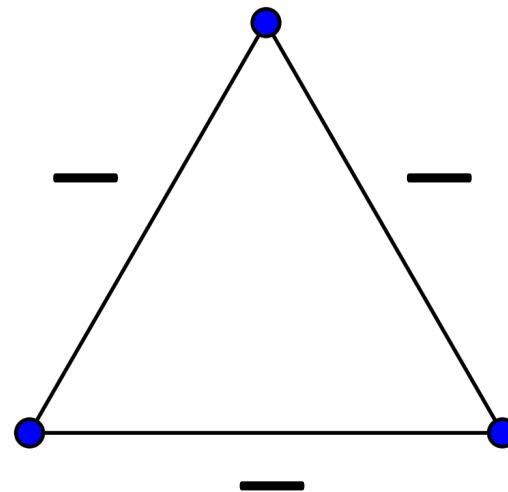
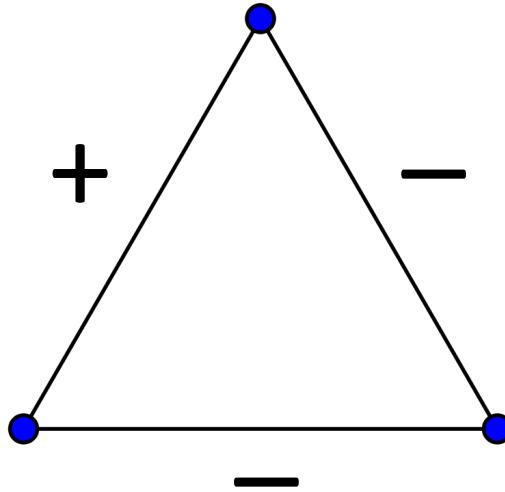
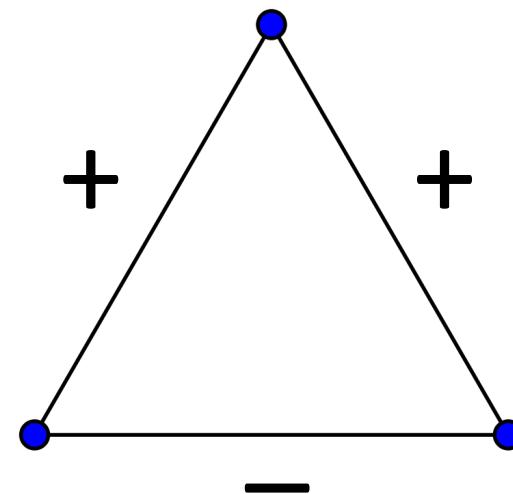
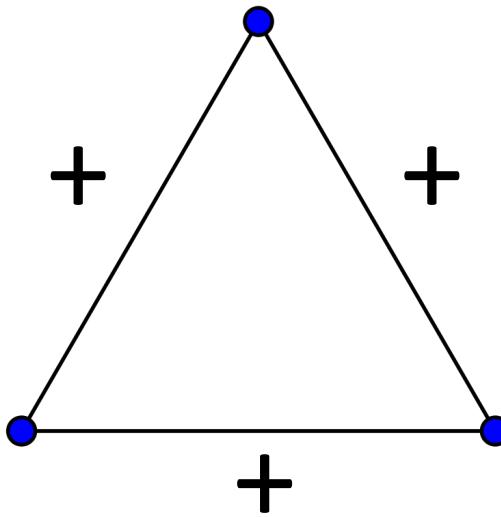
Complete graph

A network where there is an edge connecting each pair of nodes is called a **complete graph**.

Label each edge with either + or –
A ‘+’ stands for friendship and a ‘–’ stands for enemy.



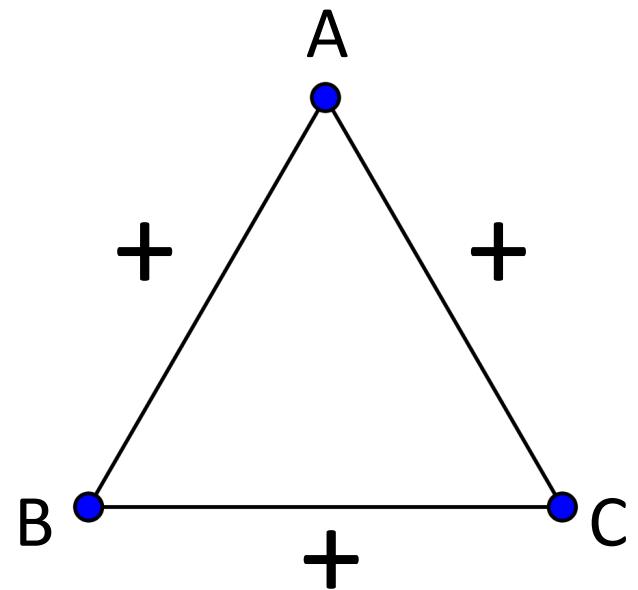
Possible ways to label a triangle



Possible ways to label a triangle

Three pluses:

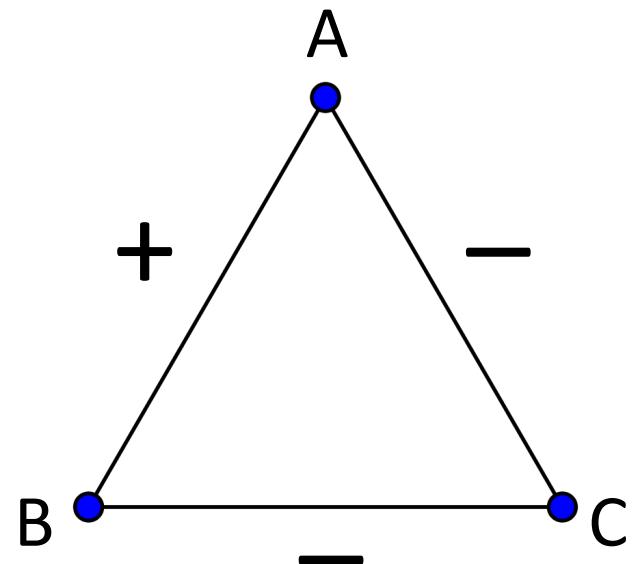
Three people who like each other.



Possible ways to label a triangle

One plus & two minuses:

Two among the three are friends, and they have a mutual enemy in the third.

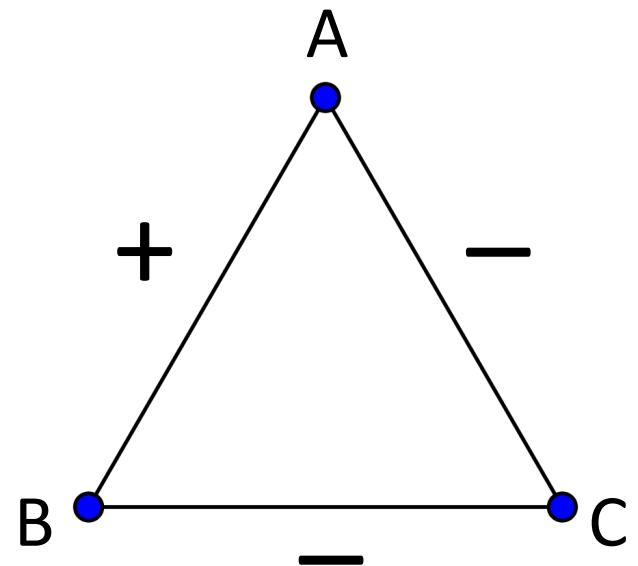


Possible ways to label a triangle

One plus & two minuses:

Two among the three are friends, and they have a mutual enemy in the third.

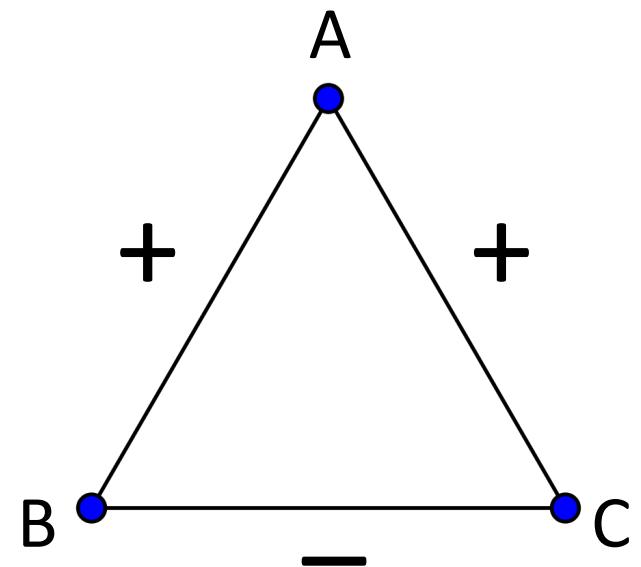
“The enemy of my enemy is my friend.”



Possible ways to label a triangle

Two pluses & one minus:

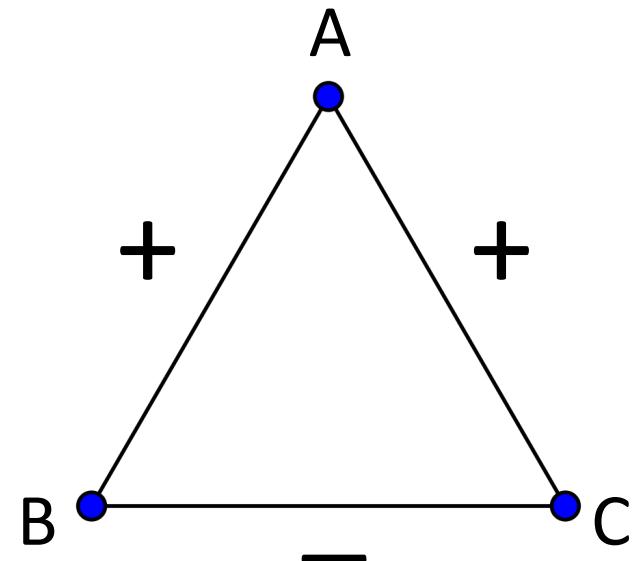
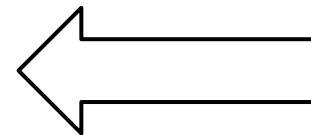
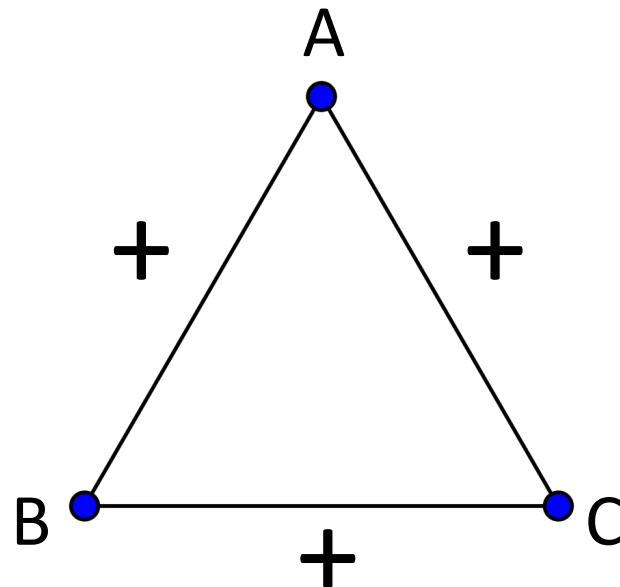
I have two friends who hate each other.



Possible ways to label a triangle

Two pluses & one minus:

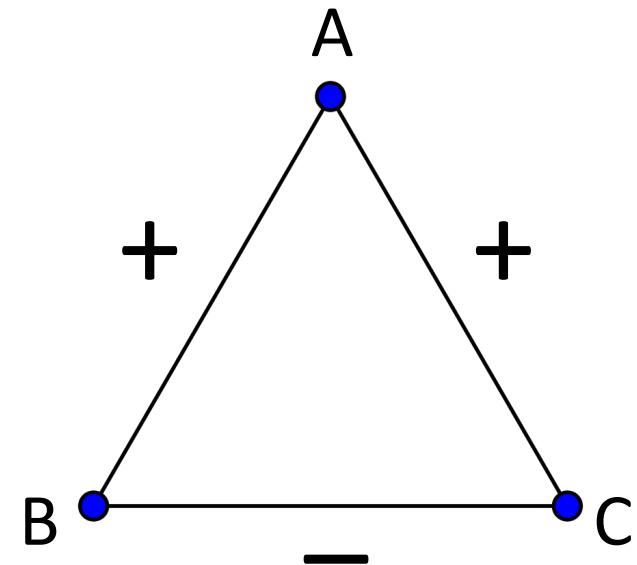
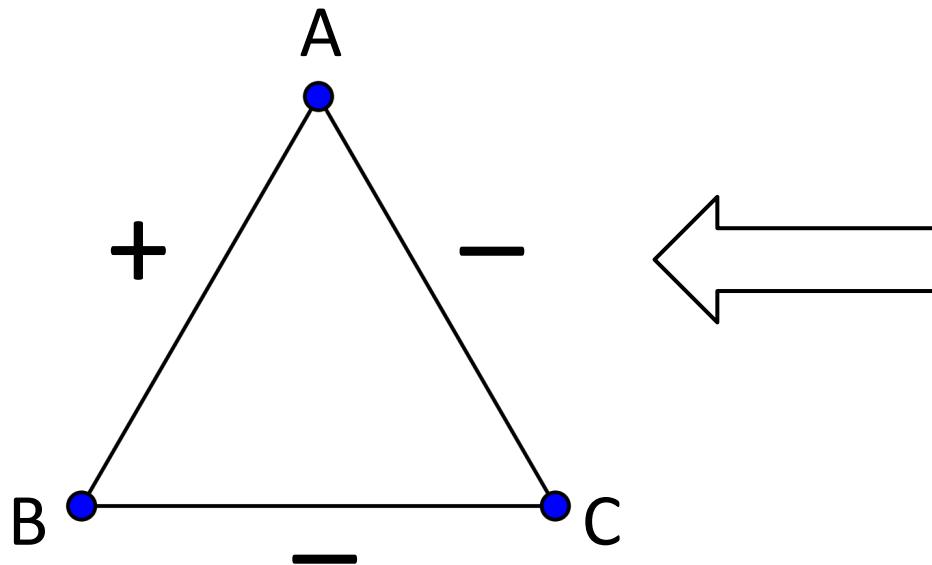
I have two friends who hate each other.



Possible ways to label a triangle

Two pluses & one minus:

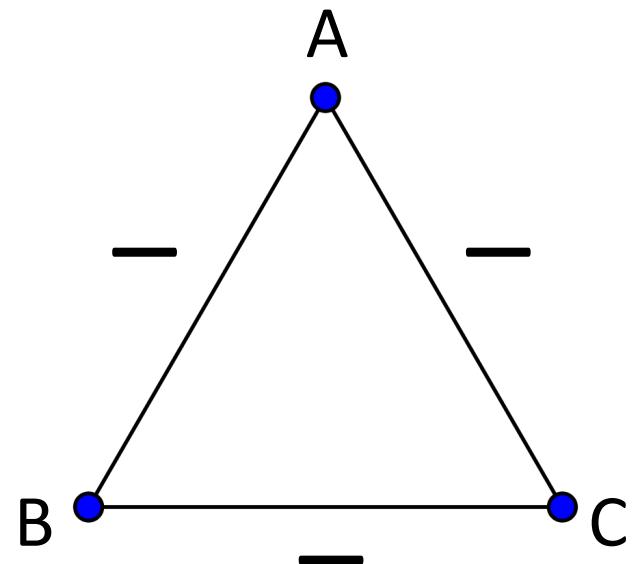
I have two friends who hate each other.



Possible ways to label a triangle

Three minuses:

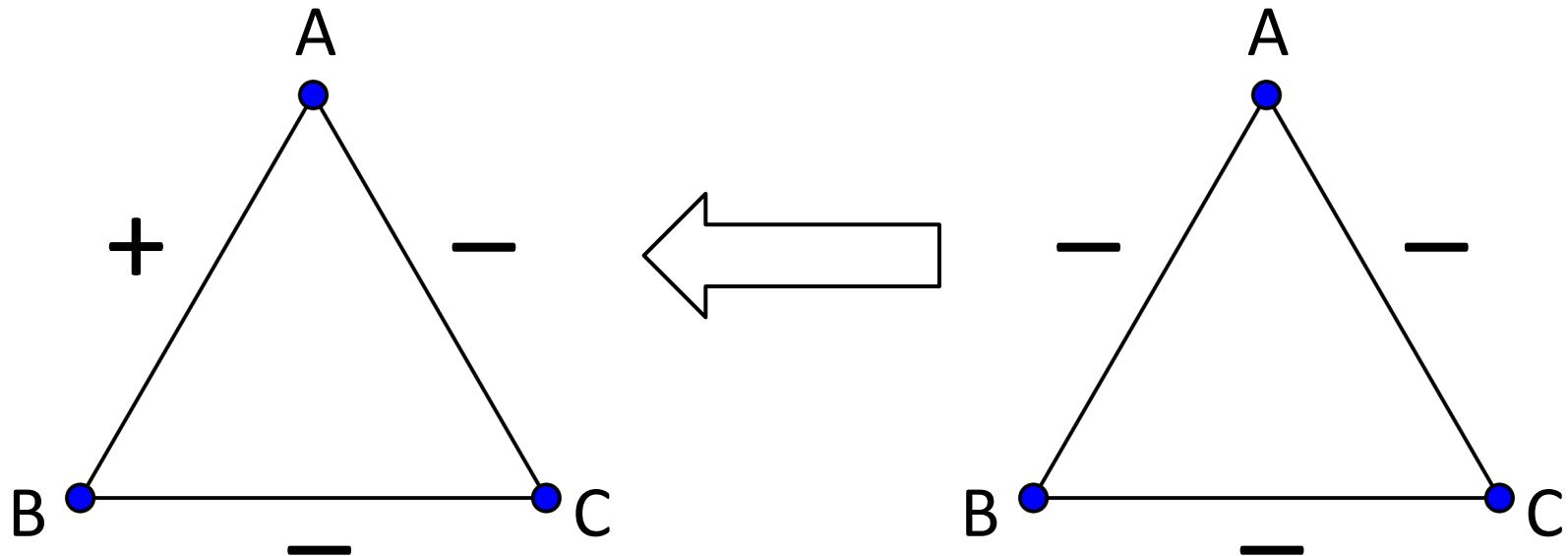
*Three people who hate each other
pairwisely...*



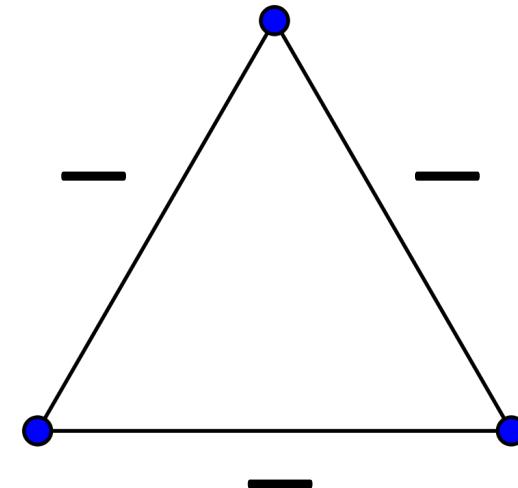
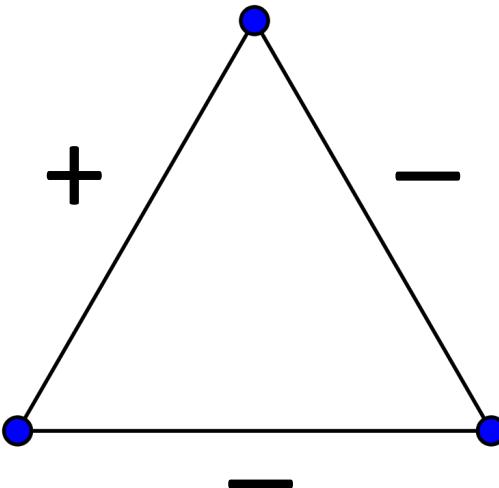
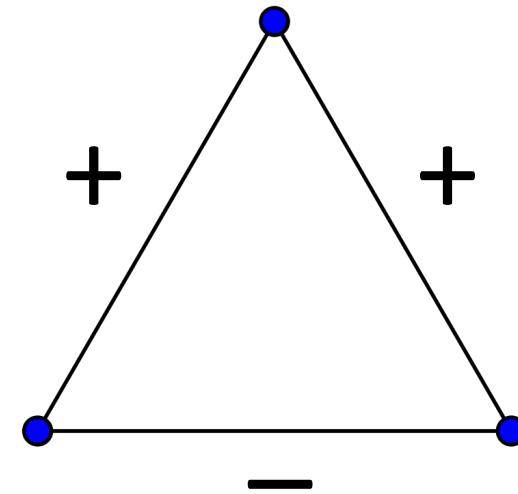
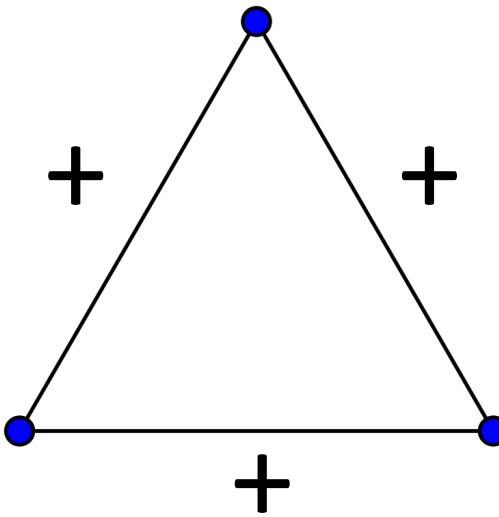
Possible ways to label a triangle

Three minuses:

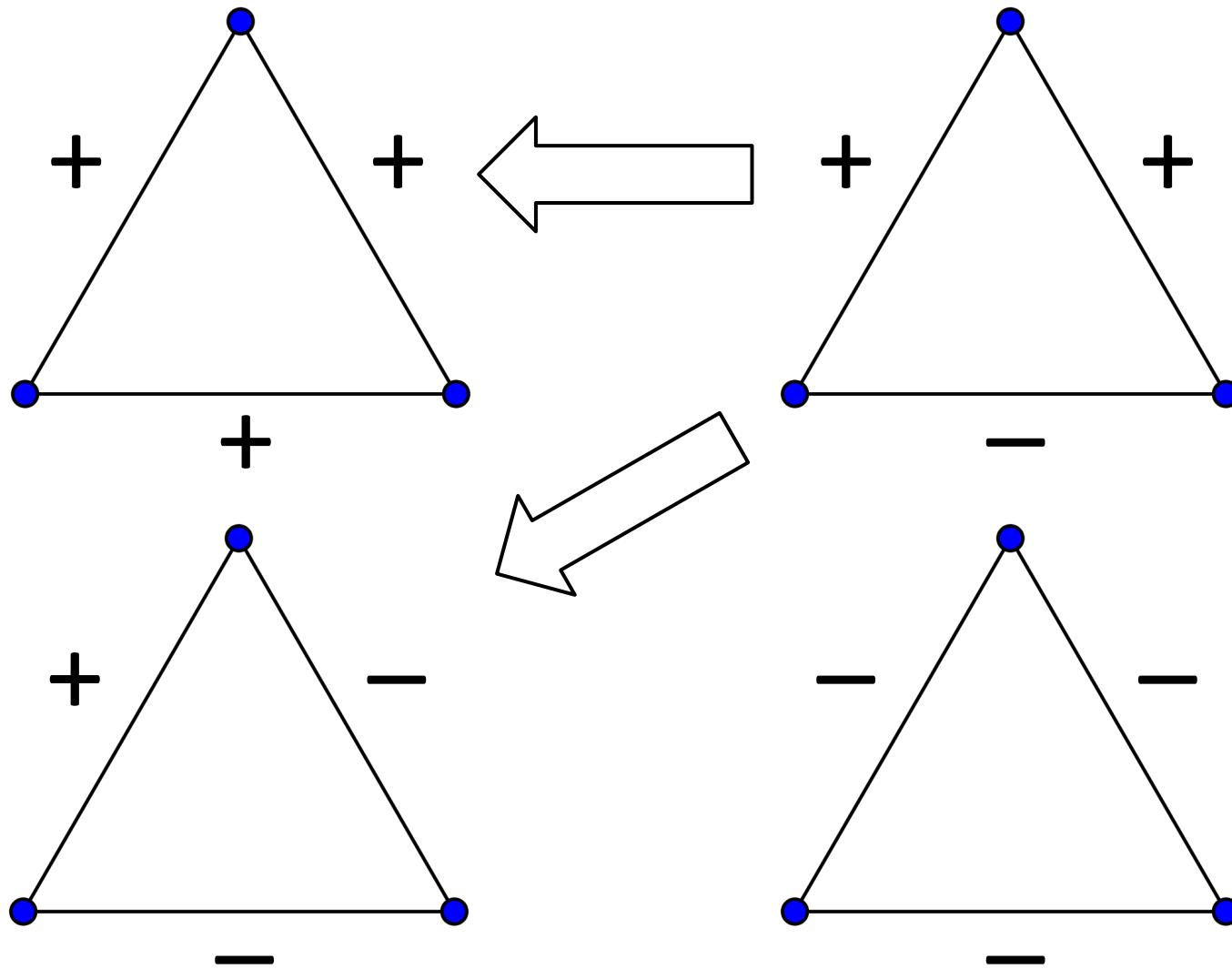
*Three people who hate each other
pairwisely...*



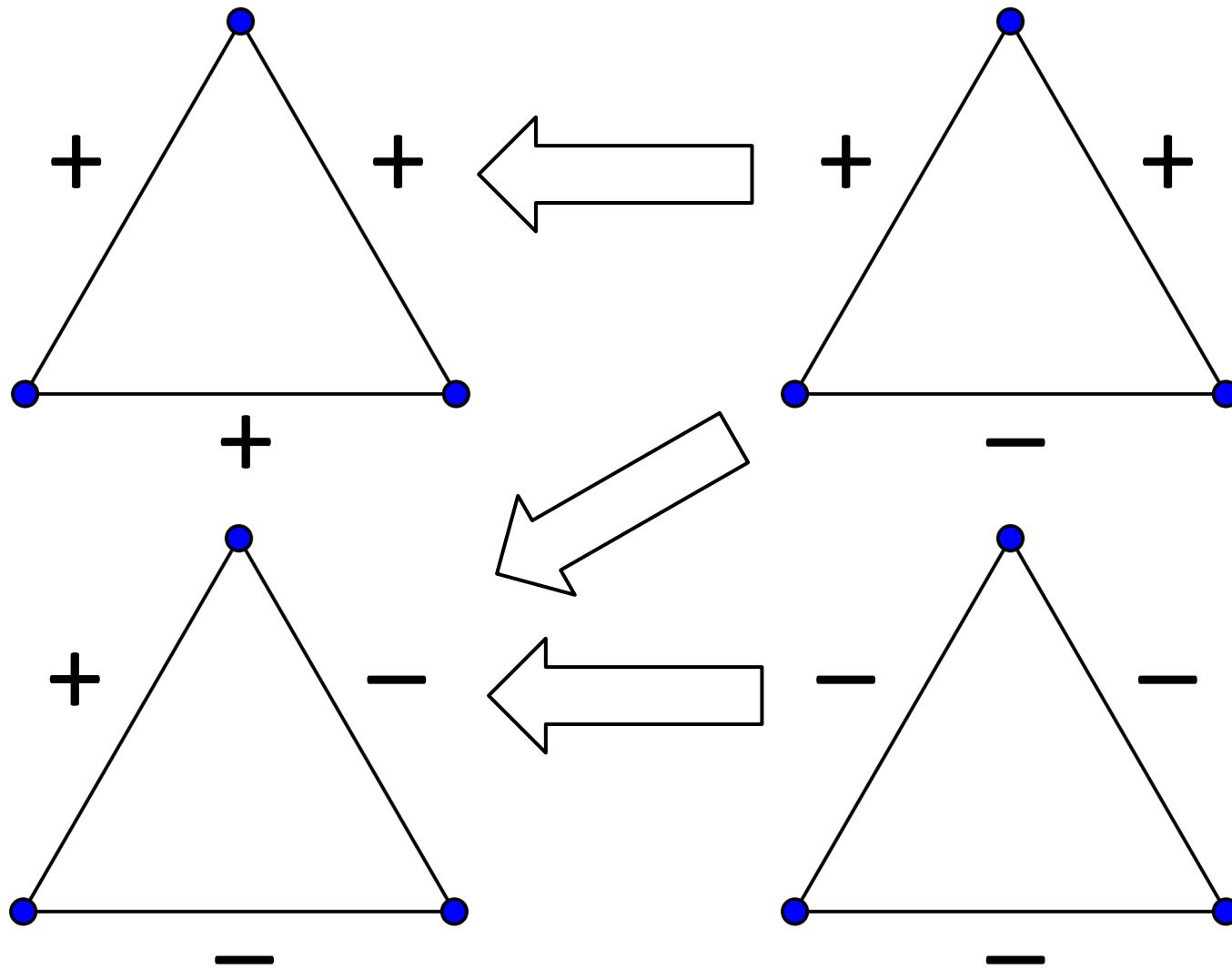
Possible ways to label a triangle



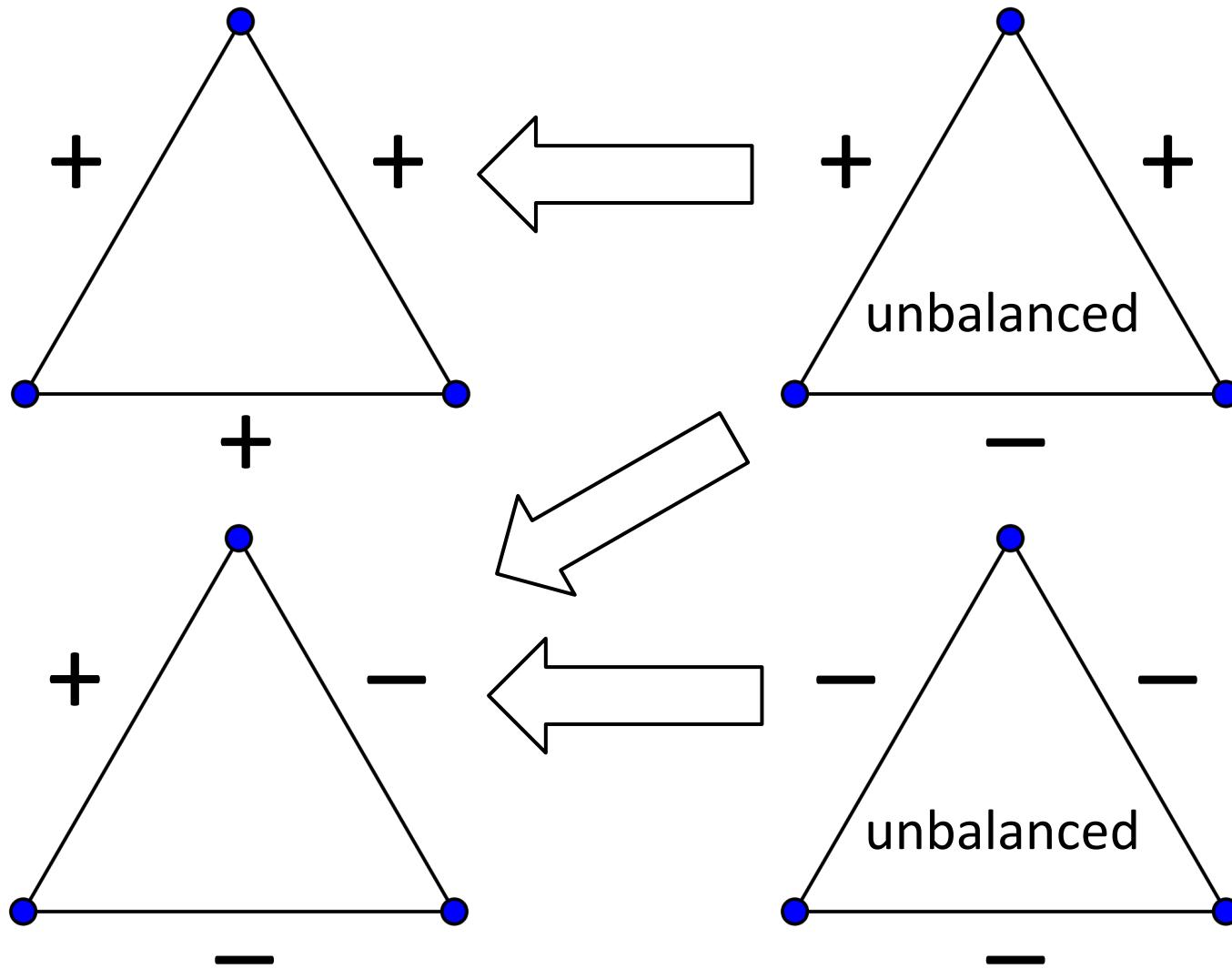
Possible ways to label a triangle



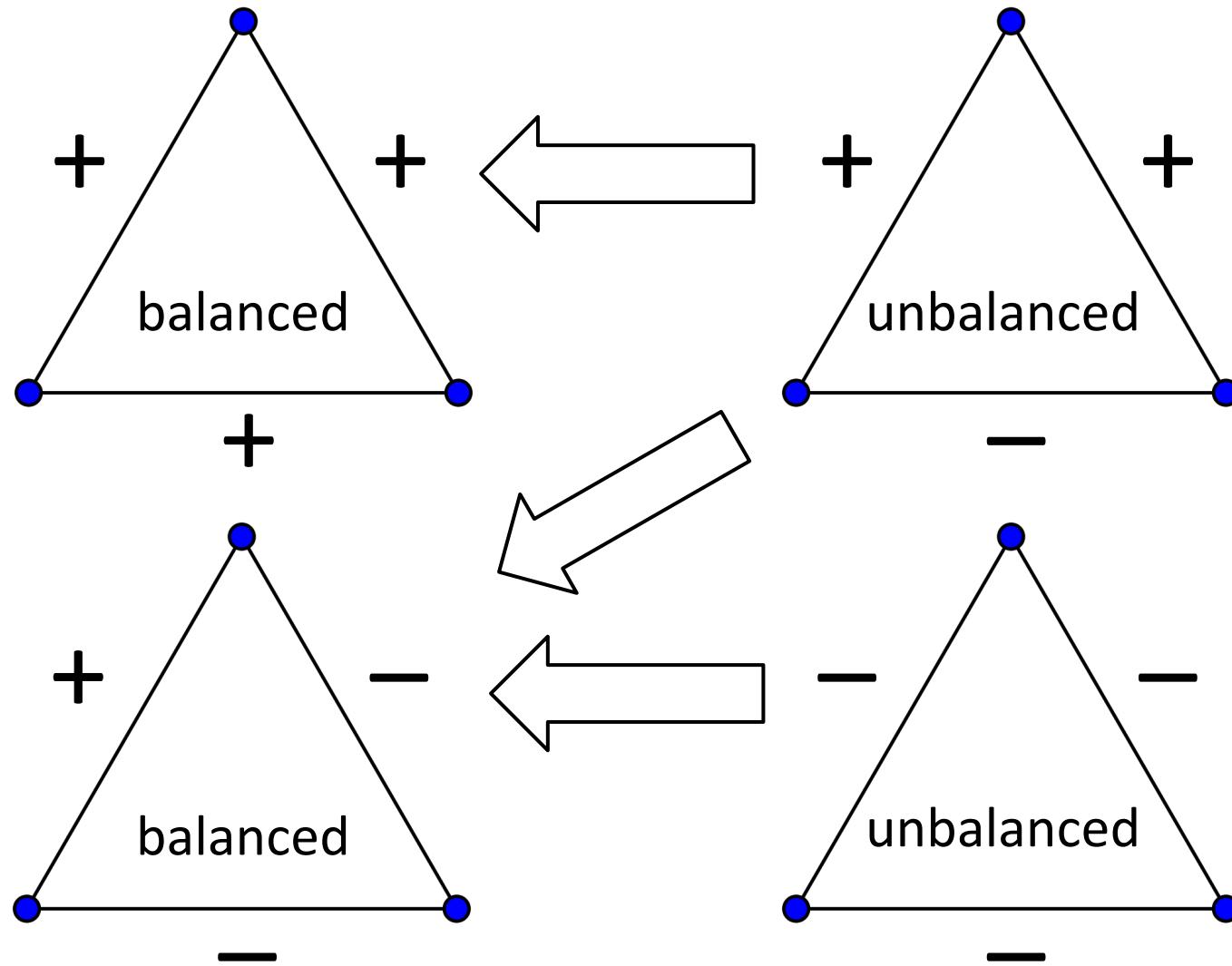
Possible ways to label a triangle



Possible ways to label a triangle



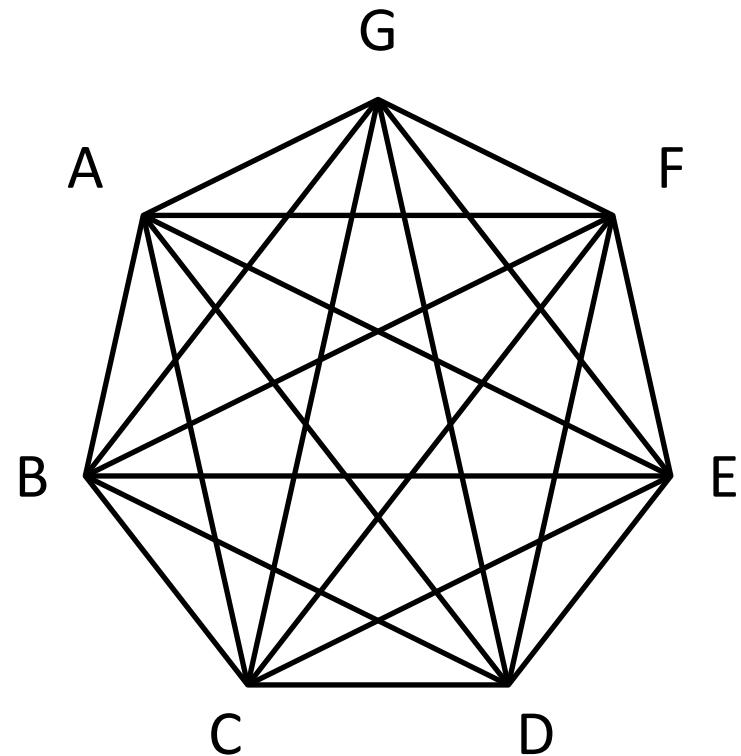
Possible ways to label a triangle



Structural Balance

Structural Balance Property:

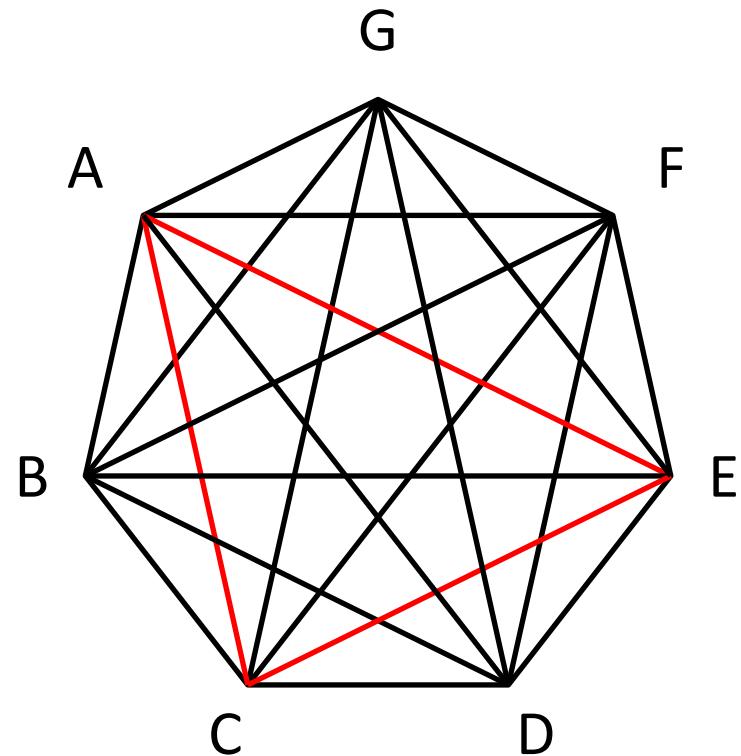
For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



Structural Balance

Structural Balance Property:

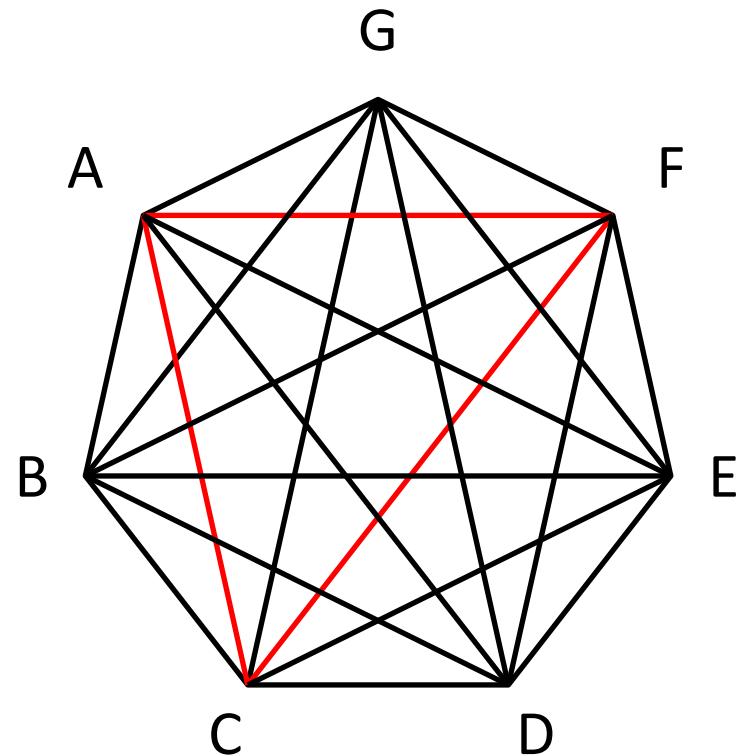
For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



Structural Balance

Structural Balance Property:

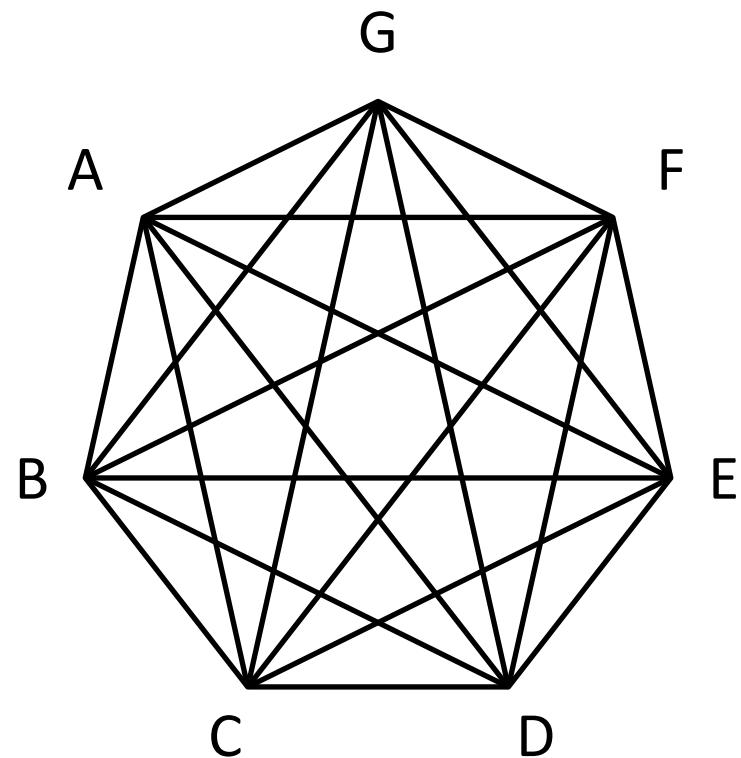
For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



Structural Balance

Structural Balance Property:

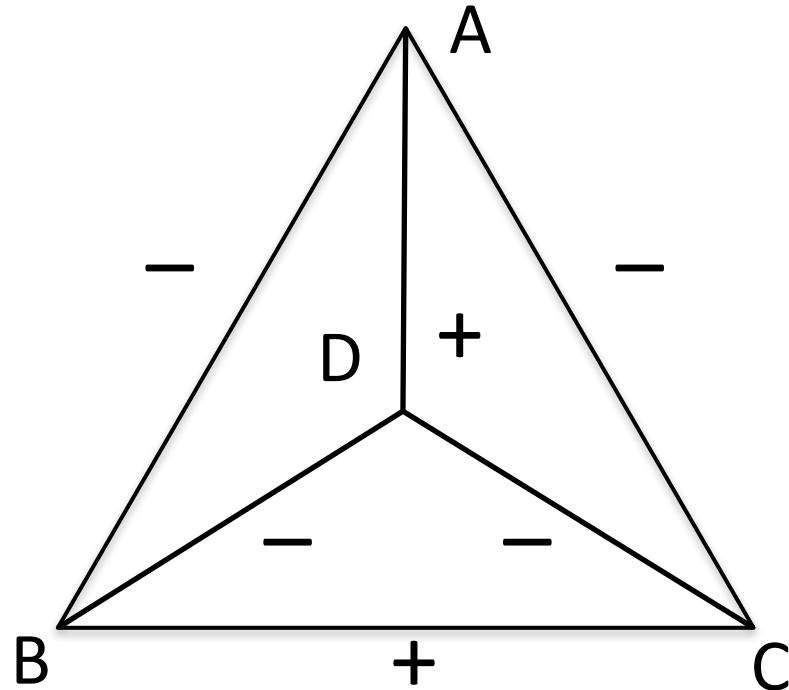
For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



Structural Balance

Structural Balance Property:

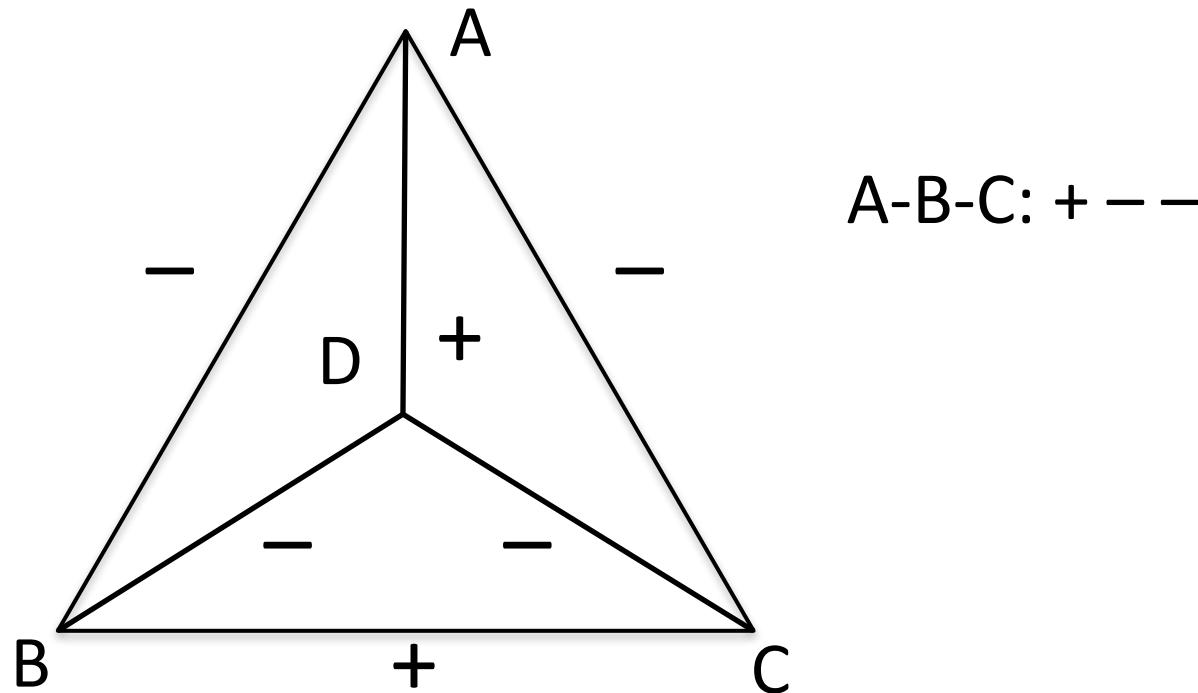
For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



Structural Balance

Structural Balance Property:

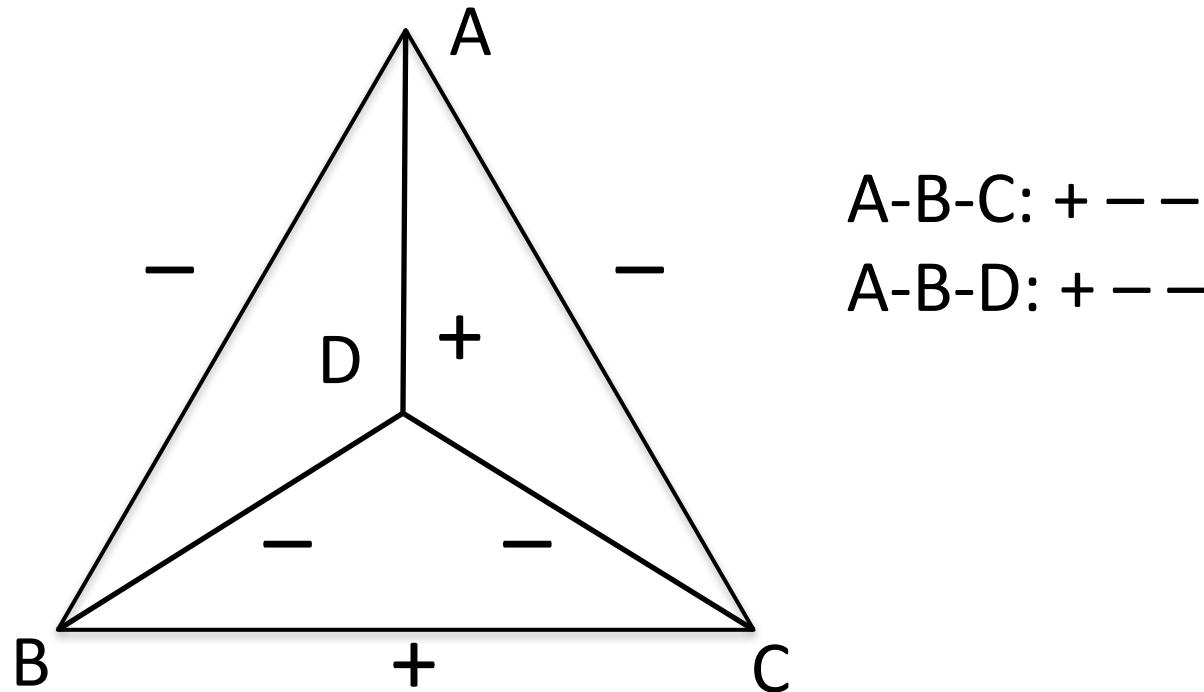
For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



Structural Balance

Structural Balance Property:

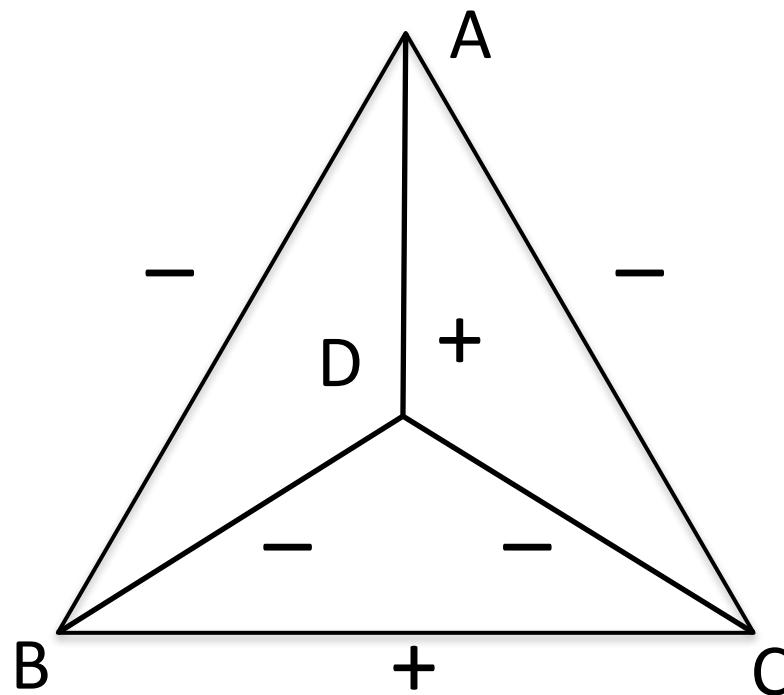
For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



Structural Balance

Structural Balance Property:

For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



A-B-C: +---

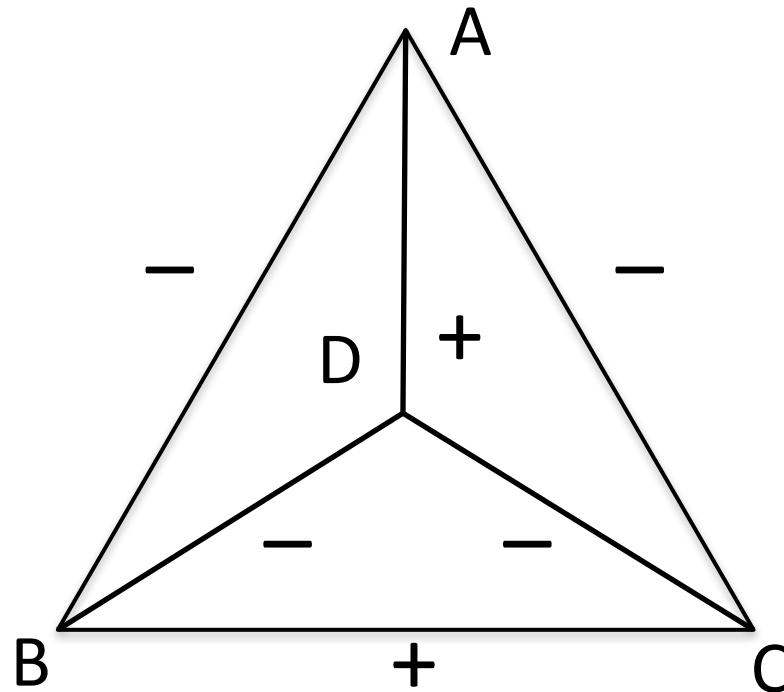
A-B-D: +---

A-C-D: +---

Structural Balance

Structural Balance Property:

For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



A-B-C: + - -

A-B-D: + - -

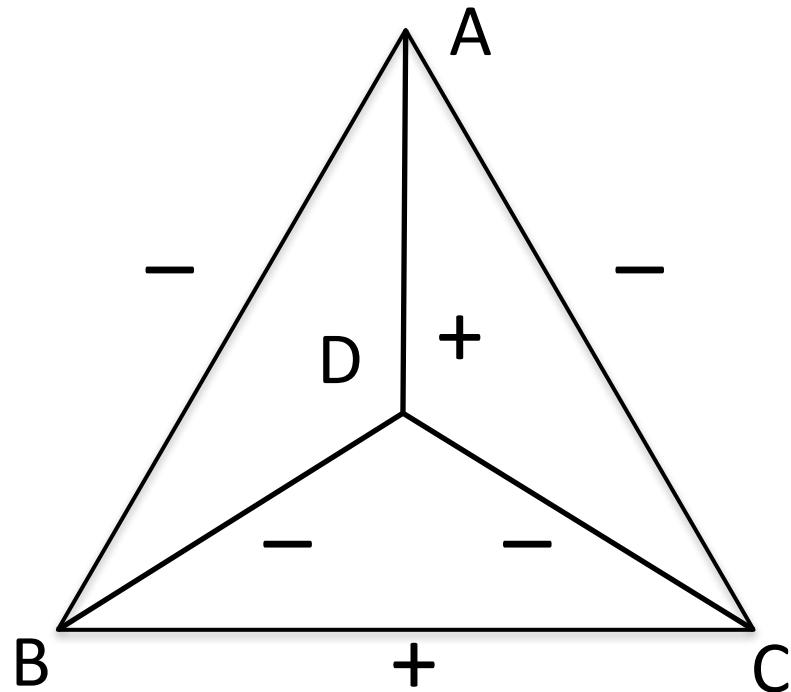
A-C-D: + - -

B-C-D: + - -

Structural Balance

Structural Balance Property:

For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.



A-B-C: + - -

A-B-D: + - -

A-C-D: + - -

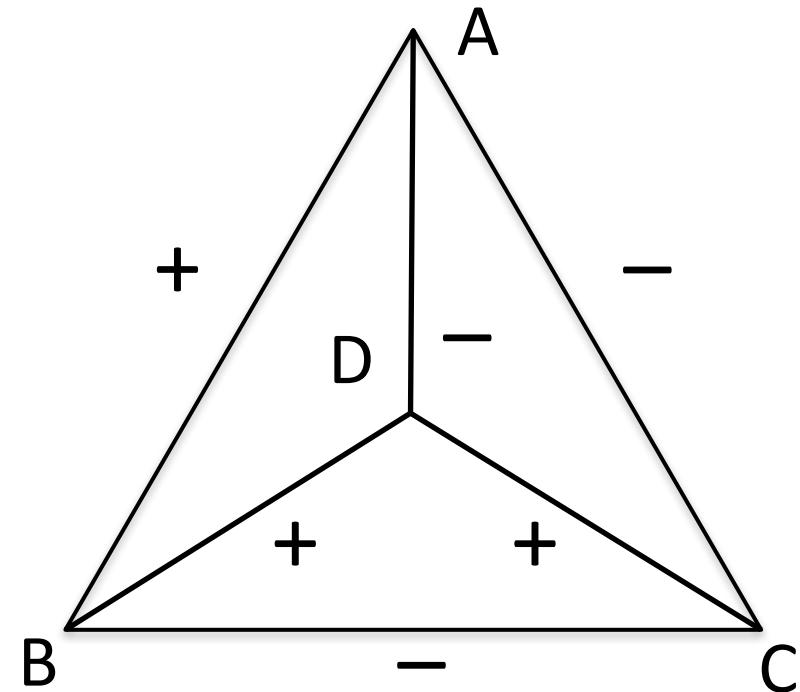
B-C-D: + - -

Network is balanced.

Structural Balance

Structural Balance Property:

For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.

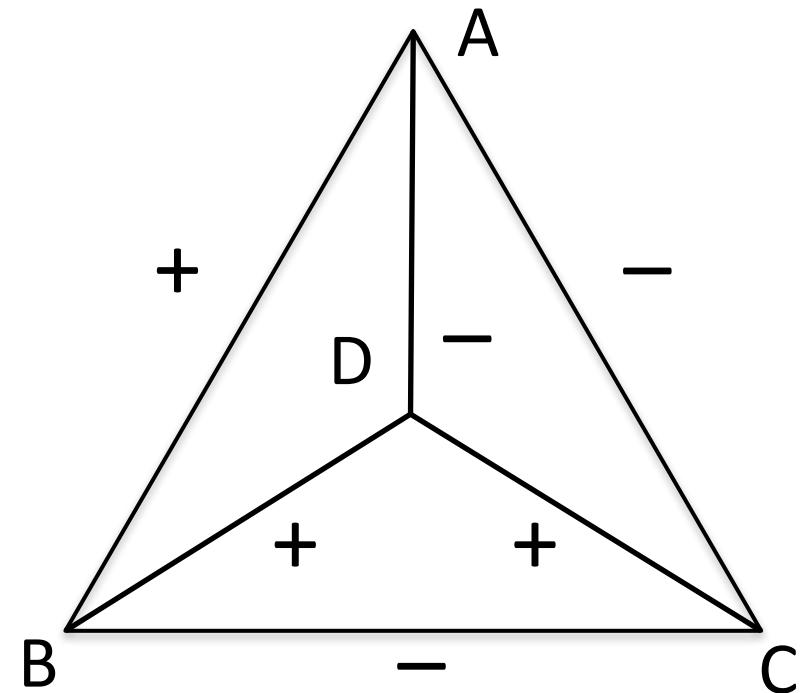


Structural Balance

Structural Balance Property:

For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.

A-B-C: + - -



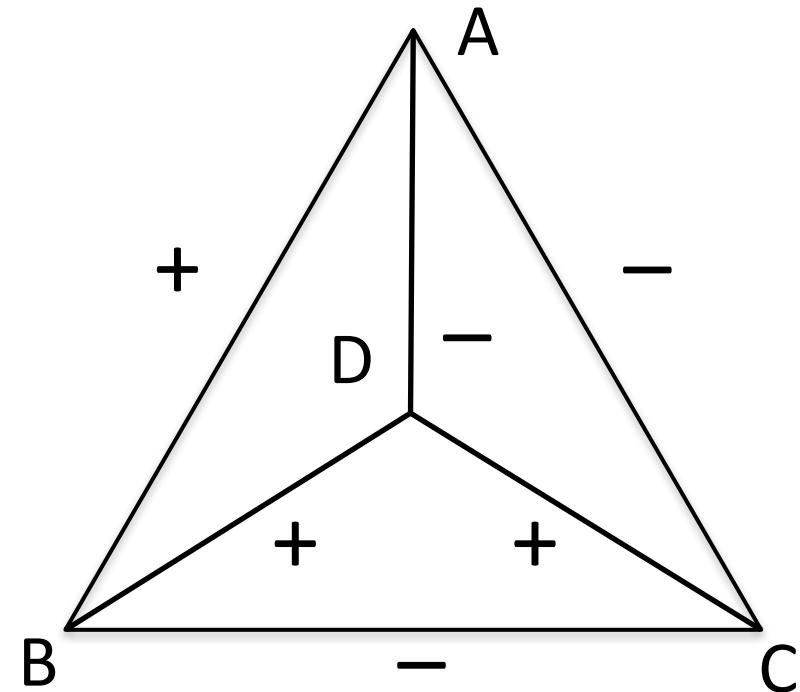
Structural Balance

Structural Balance Property:

For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.

A-B-C: + - -

A-B-D: + + -



Structural Balance

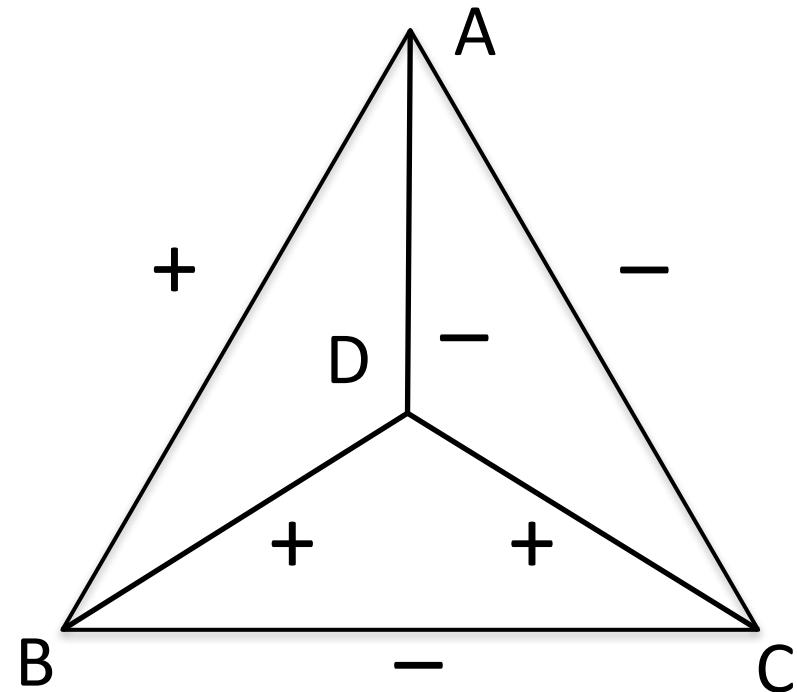
Structural Balance Property:

For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.

A-B-C: + - -

A-B-D: + + -

A-C-D: + - -



Structural Balance

Structural Balance Property:

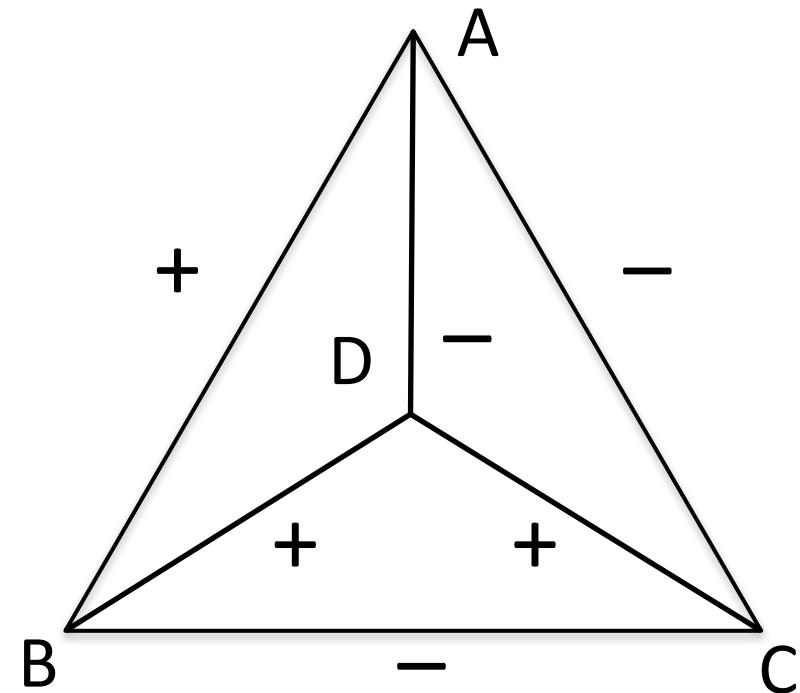
For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.

A-B-C: + - -

A-B-D: + + -

A-C-D: + - -

B-C-D: + + -



Structural Balance

Structural Balance Property:

For *every* set of three nodes within a labeled complete graph, either all three of their edges are labeled +, or exactly one of them is labeled +.

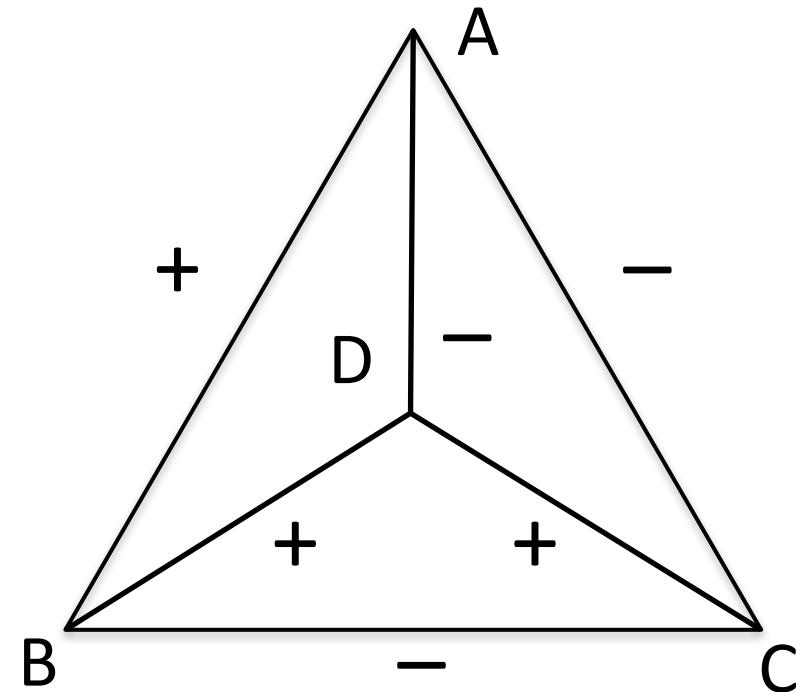
A-B-C: + - -

A-B-D: + + -

A-C-D: + - -

B-C-D: + + -

Network is not balanced.

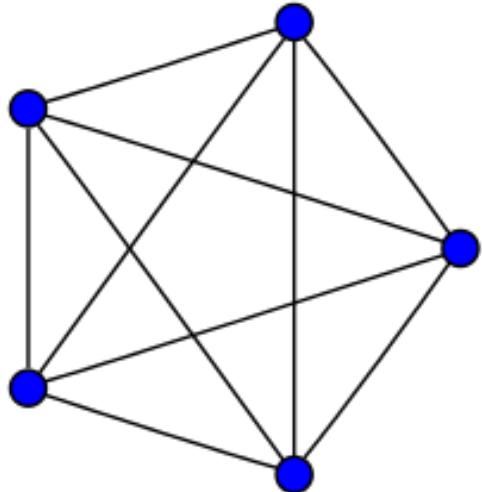


Structural Balance

- To check whether a complete graph satisfies structural balance, we need to go through *every* triangle.

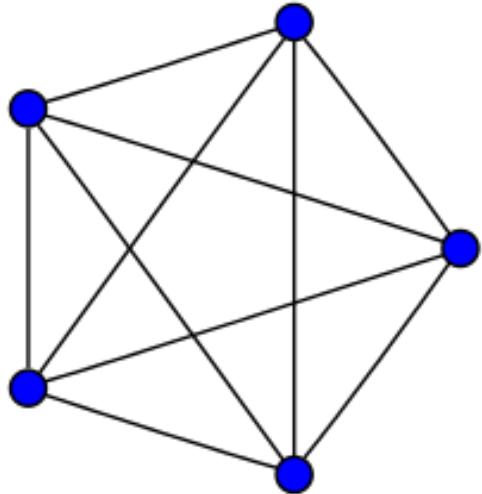
Structural Balance

- To check whether a complete graph satisfies structural balance, we need to go through *every* triangle.



Structural Balance

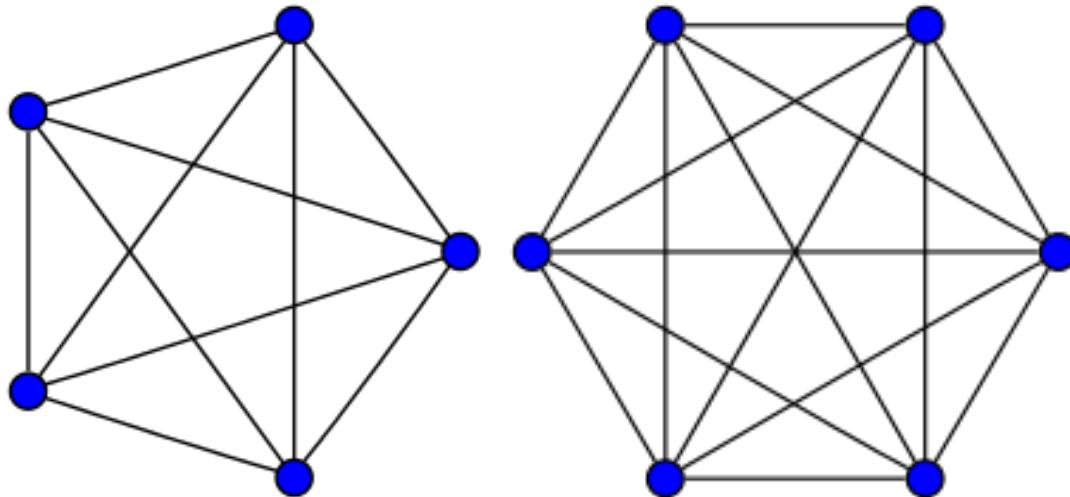
- To check whether a complete graph satisfies structural balance, we need to go through *every* triangle.



5 nodes
10 triangles

Structural Balance

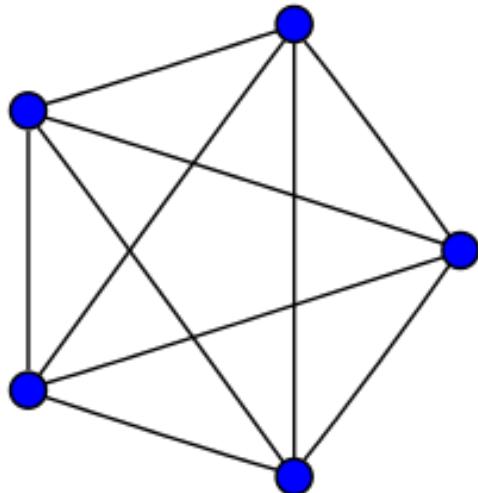
- To check whether a complete graph satisfies structural balance, we need to go through *every* triangle.



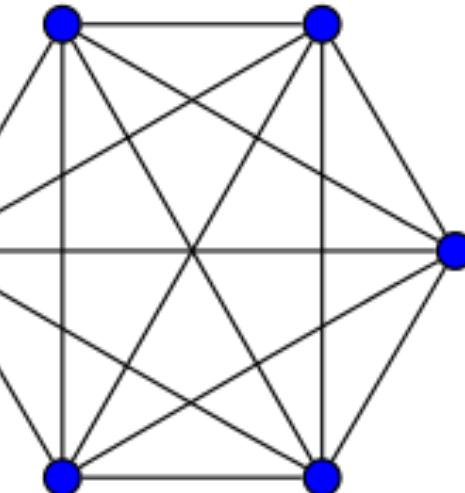
5 nodes
10 triangles

Structural Balance

- To check whether a complete graph satisfies structural balance, we need to go through *every* triangle.



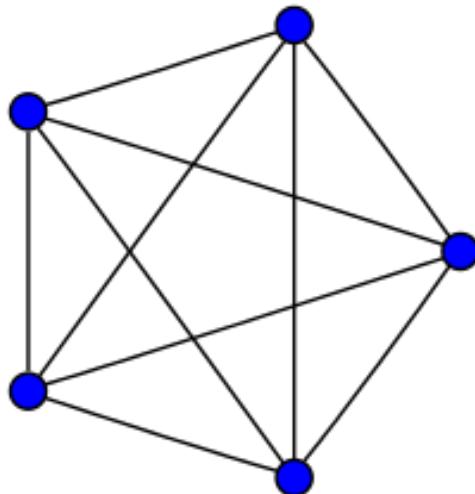
5 nodes
10 triangles



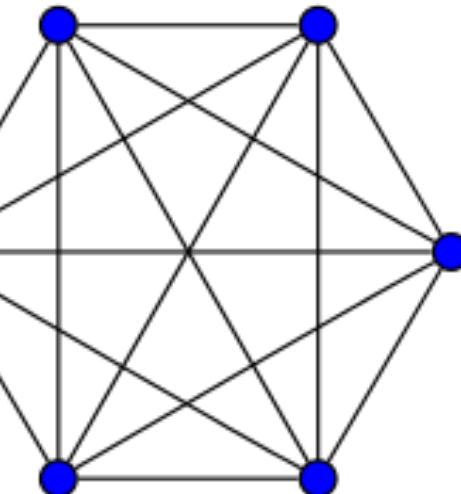
6 nodes
20 triangles

Structural Balance

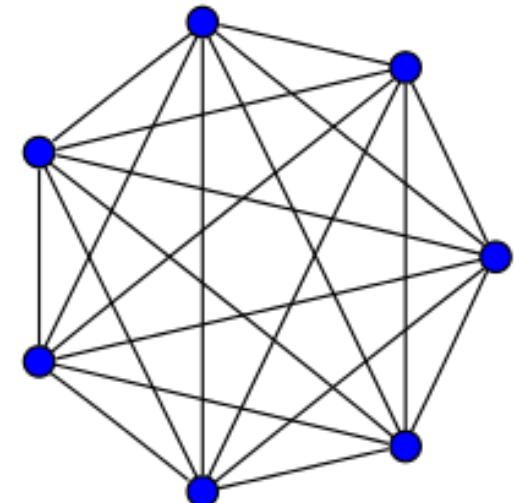
- To check whether a complete graph satisfies structural balance, we need to go through *every* triangle.



5 nodes
10 triangles

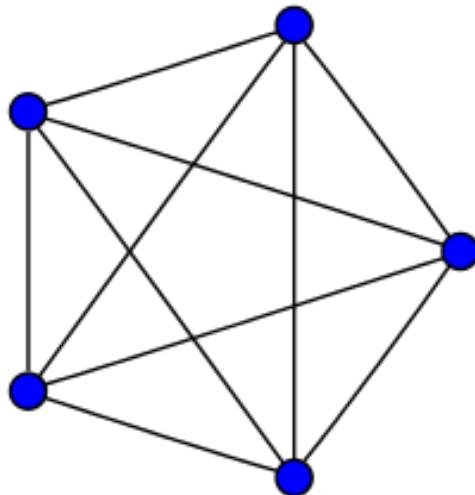


6 nodes
20 triangles

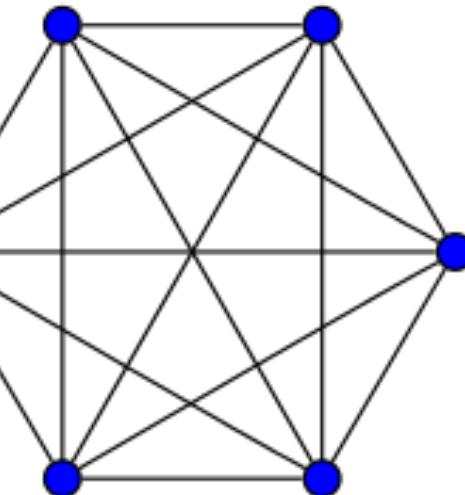


Structural Balance

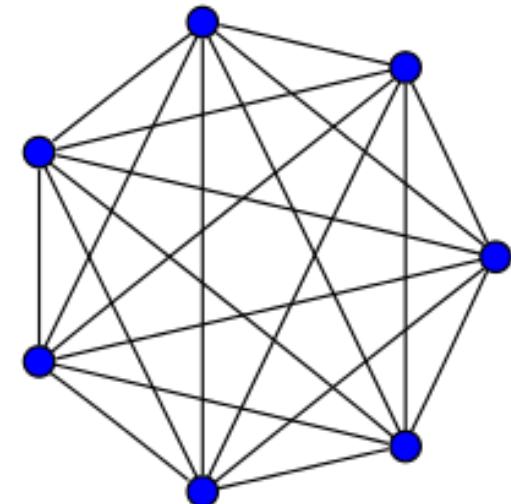
- To check whether a complete graph satisfies structural balance, we need to go through *every* triangle.



5 nodes
10 triangles



6 nodes
20 triangles



7 nodes
35 triangles

Structural Balance

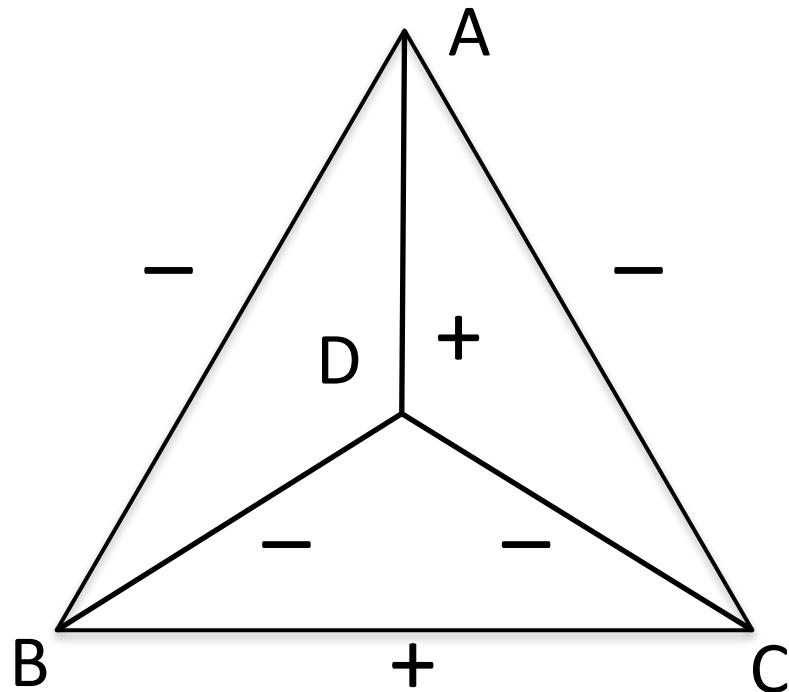
- As the number of nodes grow, the number of triangles grow much faster.

Structural Balance

- As the number of nodes grow, the number of triangles grow much faster.
- Check structural balance manually (or even computationally for very large networks) is difficult if not impossible.

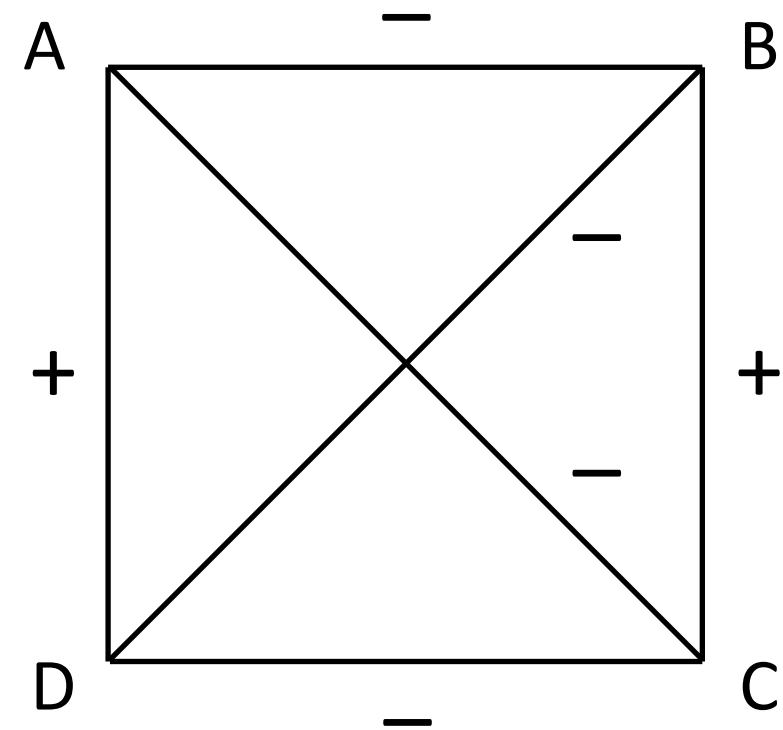
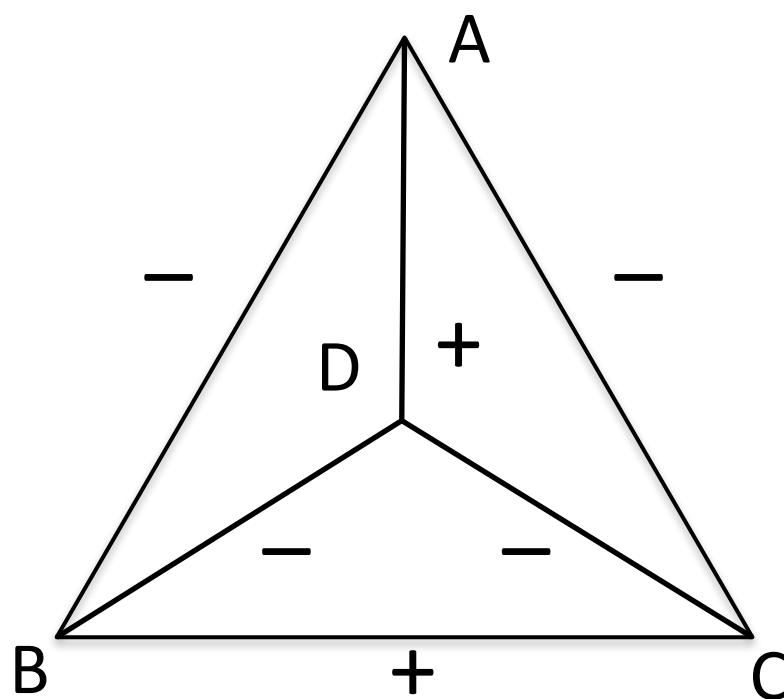
Structural Balance

- Fortunately, any balanced network, *no matter how complicated it is*, has a simple feature that characterizes its structure.



Structural Balance

- Fortunately, any balanced network, *no matter how complicated it is*, has a simple feature that characterizes its structure.



Structural Balance

- Fortunately, any balanced network, *no matter how complicated it is*, has a simple feature that characterizes its structure.

