# Human Evaluations and Signed Networks

# Announcement: Course Project

- **Project is a substantial part of the class**
  - Students put significant effort and great things have been done
- **Types of projects:**
  - **(1)** Analysis of an interesting dataset with the goal to develop a (new) model or an algorithm
  - **(2)** A test of a model or algorithm (that you have read about or your own) on real & simulated data.
    - Fast algorithms for big graphs.
- **Other points:**
  - The project should contain some mathematical analysis, and some experimentation on real or synthetic data
  - The result of the project will typically be an 8 page paper, describing the approach, the results, and related work.
  - **Come to us if you need help with a project idea!**

UNC CHARLOTTE

# Announcement: Project Proposal

**Project proposal: 3-5 pages, teams of up to 2 students**

- **Project proposal has 3 parts:**
  - **(0) Quick 200 word abstract**
  - **(1) Reaction paper/Related work (2-3 pages):**
    - **Read 3 papers related to the project/class (ask me for sugggestions)**
    - Do reading beyond what was covered in class
    - Think beyond what you read. Don't take other's work for granted!
    - **2-3 pages:** Summary (~1 page), Critique (~1 page)
  - **(2) Proposal (1-2 pages):**
    - Clearly define the problem you are solving.
    - **How does it relate to what you read for the Reaction paper?**
    - What data will you use? (**make sure you already have it!**)
    - Which algorithm/model will you use/develop? **Be specific!**
    - How will you evaluate/test your method?

UNC CHARLOTTE

# People Express Opinions

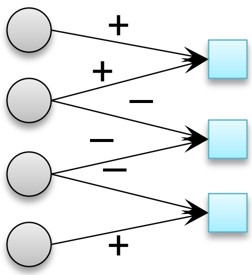
**In many online applications users express positive and negative attitudes/opinions:**

- Through **actions**:
  - Rating a product/person
  - Pressing a "like" button
- Through **text**:
  - Writing a comment, a review
- **Success of these online applications is built on people expressing opinions**
  - Recommender systems
  - Wisdom of the Crowds
  - Sharing economy

# People & Evaluations

- **About items:**
  - Movie and product reviews

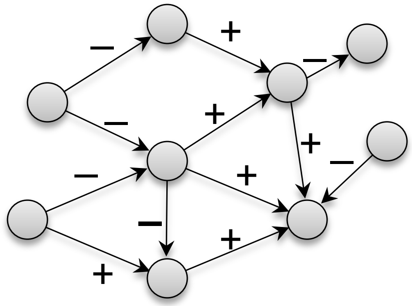
IMDb Earth's Biggest Movie Database™ amazon.com



- **About other users:**
  - Online communities

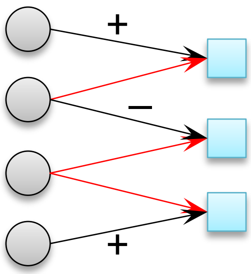
UBER

Epinions.com WIKIPEDIA airbnb



- **About items created by others:**
  - Q&A websites

stackoverflow YAHOO! ANSWERS



# User-User Evaluations

- **Many on-line settings where one person expresses an opinion about another** (or about another's content)
  - **I trust you** [Kamvar-Schlosser-Garcia-Molina '03]
  - **I agree with you** [Adamic-Glance '04]
  - **I vote in favor of admitting you into the community** [Cosley et al. '05, Burke-Kraut '08]
  - **I find your answer/opinion helpful** [Danescu-Niculescu-Mizil et al. '09, Borgs-Chayes-Kalai-Malekian-Tennenholtz '10]

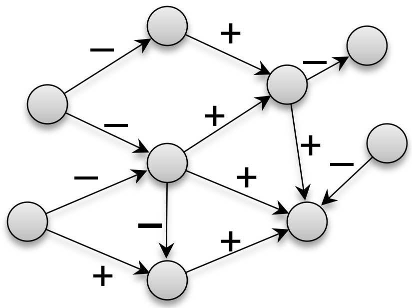
# Evaluations: Some Issues

**Some of the central issues:**

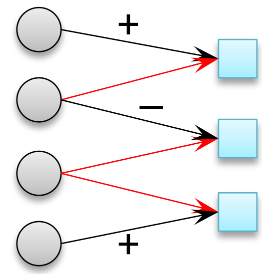
- **Factors:**
  **What factors drive one's evaluations?**
- **Synthesis:**
  **How do we create a composite description that accurately reflects cumulative opinion of the community?**

UNC CHARLOTTE

# Evaluations: the Setting



Direct

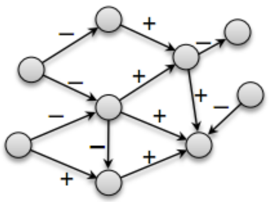


Indirect

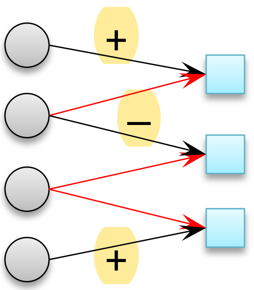
- **Direct:** User to user
- **Indirect:** User to content (created by another member of a community)
- **Where online does this explicitly occur on a large scale?**

# Evaluations: the Data

- **Wikipedia adminship elections**
  - Support/Oppose (120k votes in English)
  - 4 languages: EN, GER, FR, SP



- **Stack Overflow Q&A community**
  - Upvote/Downvote (7.5M votes)

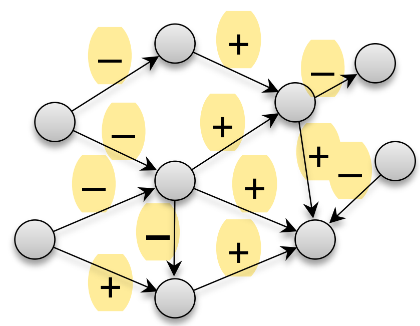



- **Epinions product reviews**
  - Ratings of others' product reviews (13M)
    - 5 = positive, 1-4 = negative

# Two ways to look at this

- There are two ways to look at this: One person evaluates the other via a positive/negative evaluation


First we focus on a single evaluation (without the context of a network)



Then we will focus on evaluations in the context of a network

# Human Evaluations

- **What drives human evaluations?**

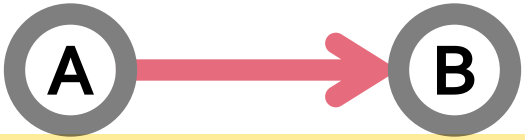

- How do properties of **evaluator A** and **target B** affect A's vote?
  - **Status** and **Similarity** are two fundamental drivers behind human evaluations

# Definitions

- **Status:** Level of recognition, merit, achievement, reputation in the community
  - Wikipedia: # edits, # barnstars
  - Stack Overflow: # answers
- **User-user Similarity:**
  - Overlapping topical interests of **A** and **B**
    - **Wikipedia:** Similarity of the articles edited
    - **Stack Overflow:** Similarity of users evaluated

# Relative vs. Absolute Assessment

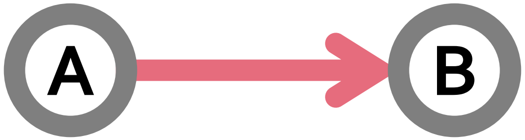
- How do properties of **evaluator A** and **target B** affect A's vote?




- **Two natural (but competing) hypotheses:**
  - **(1)** Prob. that B receives a positive evaluation depends primarily on the characteristics of B
    - There is some objective criteria for user B to receive a positive evaluation

# Relative vs. Absolute Assessment

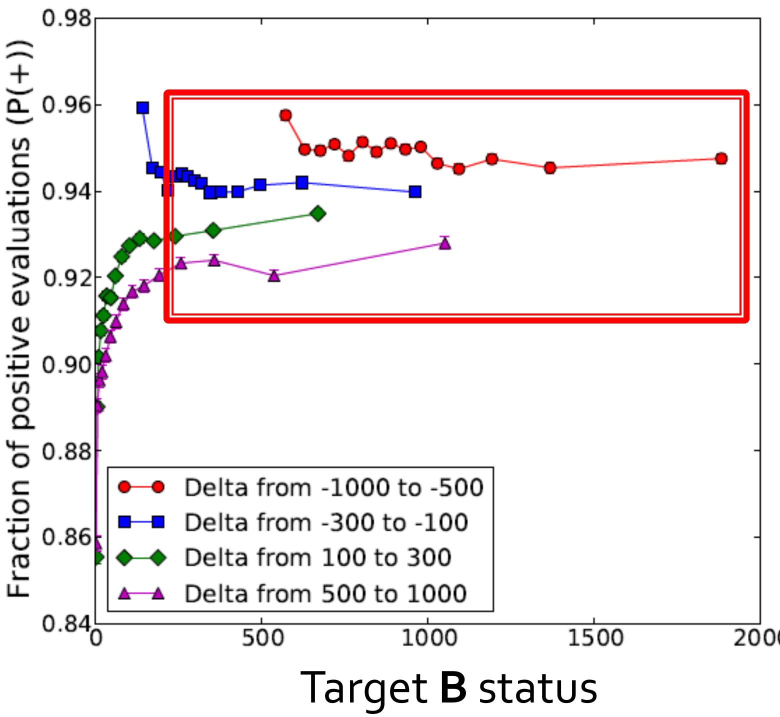
- How do properties of **evaluator A** and **target B** affect A's vote?

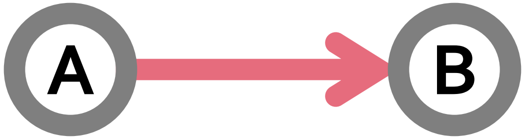



- **Two natural (but competing) hypotheses:**
  - **(2)** Prob. that B receives a positive evaluation depends on relationship between the characteristics of A and B
    - User A compares herself to user B and then makes the evaluation

# Effects of Status

- **How does status of B affect A's evaluation?**
  - Each curve is fixed status difference: $\Delta = S_A - S_B$
- **Observations:**
  - **Flat curves:** Prob. of positive eval. P(+) doesn't depend on B's status
  - **Different levels:** Different values of $\Delta$ result in different behavior



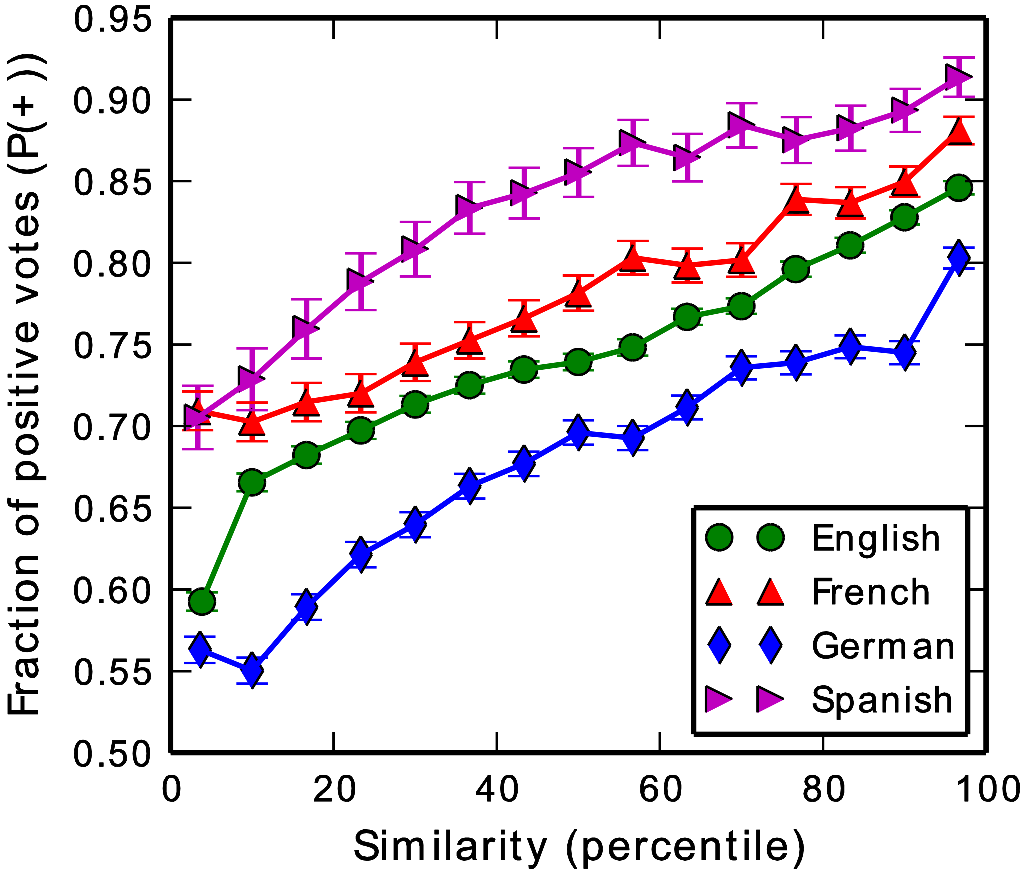




We keep increasing status of B, while keeping the status difference ($S_A$-$S_B$) fixed

# Effects of Similarity

- **How does prior interaction shape evaluations? 2 hypotheses:**
  - **(1)** Evaluators are more supportive of targets in their area
    - "The more similar you are, the more I like you"
  - **(2)** More familiar evaluators know weaknesses and are more harsh
    - "The more similar you are, the better I can understand your weaknesses"

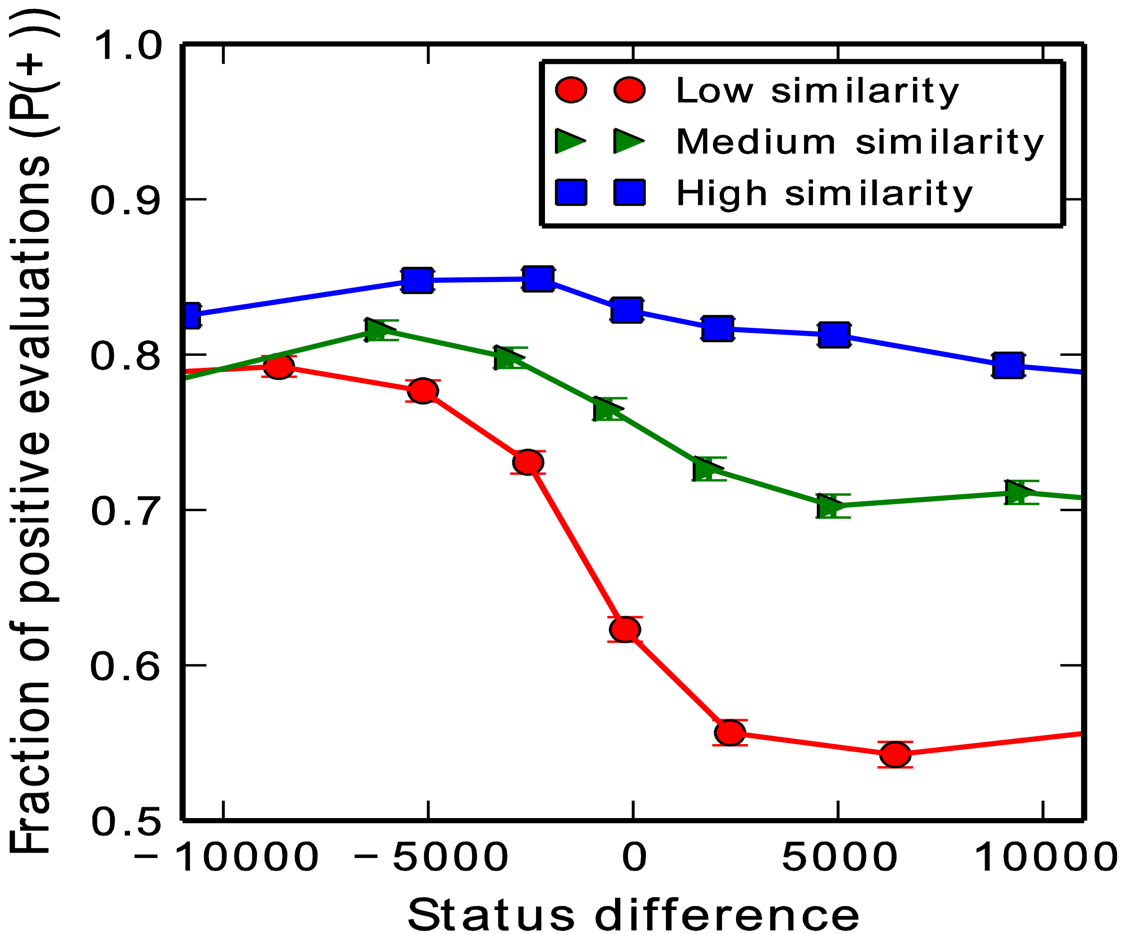
# Effects of Similarity




Prior interaction/ similarity boosts positive evaluations

**Similarity:** For each user create a set of words of all articles she edited. The similarity is then the Jaccard similarity between the two sets of words. Then sort the user pairs by similarity and bucket them into percentile.

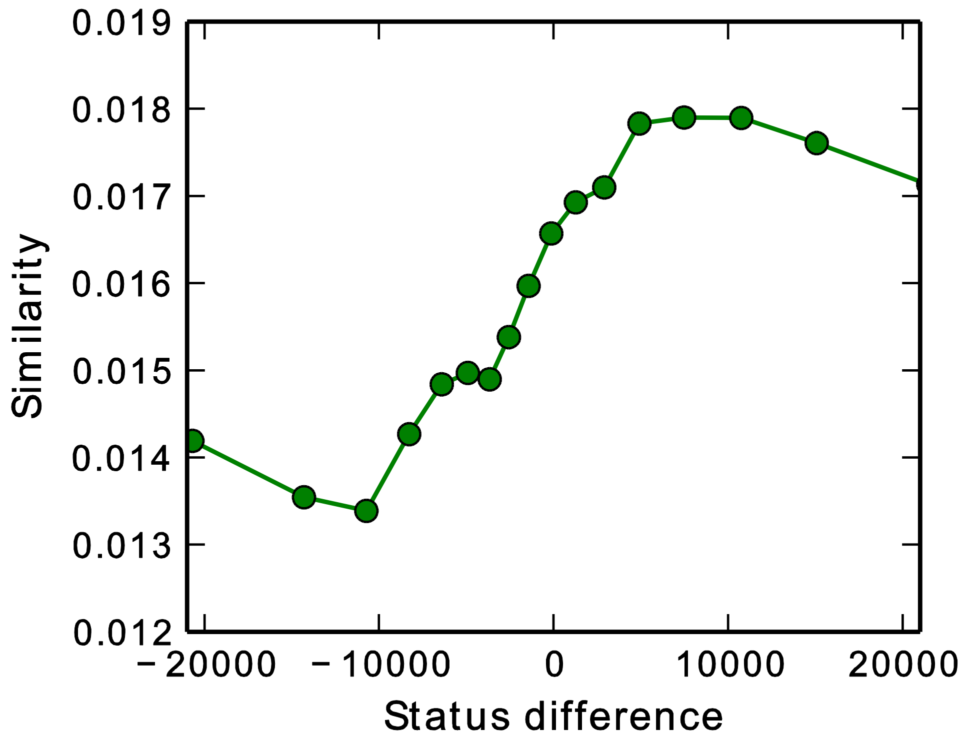
# Status & Similarity




Status is a proxy for quality when evaluator does not know the target

# Status & Similarity

- **Who shows up to evaluate?**



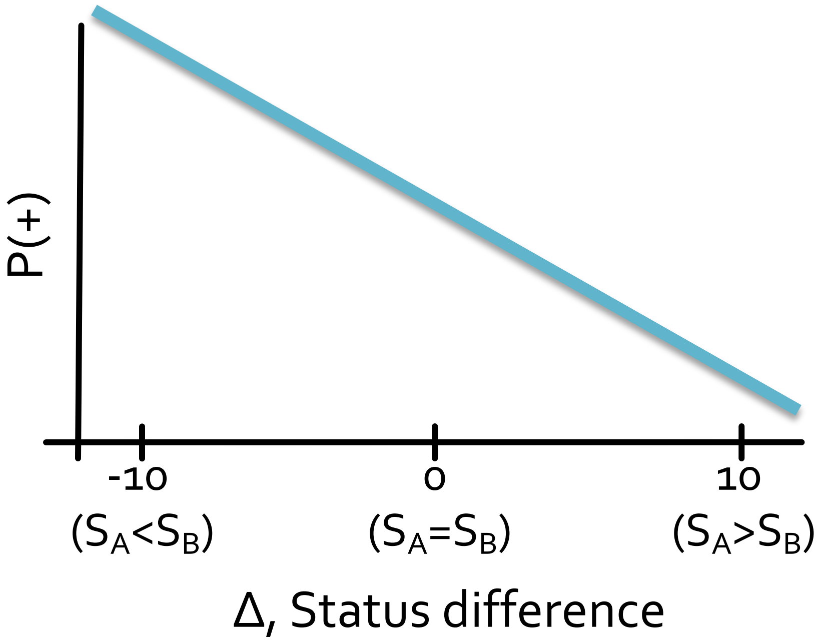

Elite evaluators vote on targets in their area of expertise

- Selection effect in who gives the evaluation
  - If $S_A > S_B$ then A and B are more likely to be similar

UNC CHARLOTTE

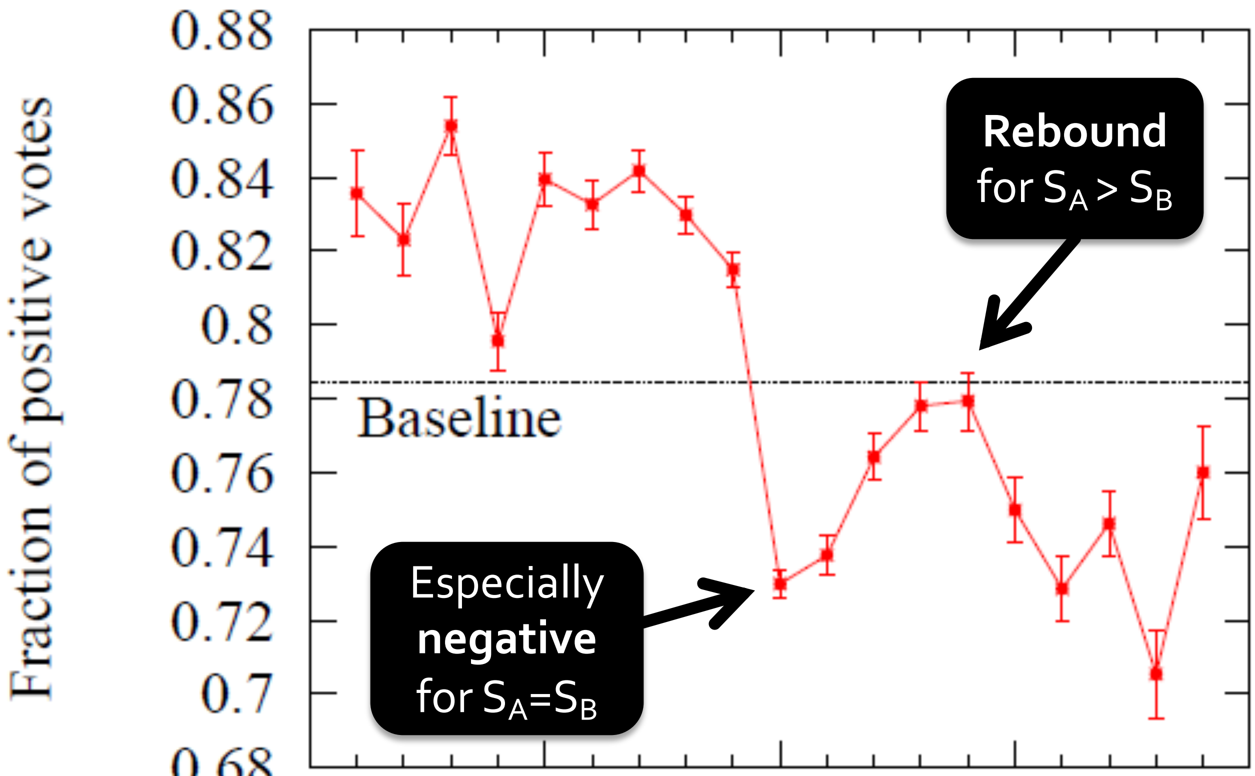
# A Puzzle

- **What is P(+) as a function of $\Delta = S_A - S_B$?**
  - Based on findings so far: **Monotonically decreasing**



# A Puzzle: The Mercy Bounce

- **What is P(+) as a function of $\Delta = S_A - S_B$?**

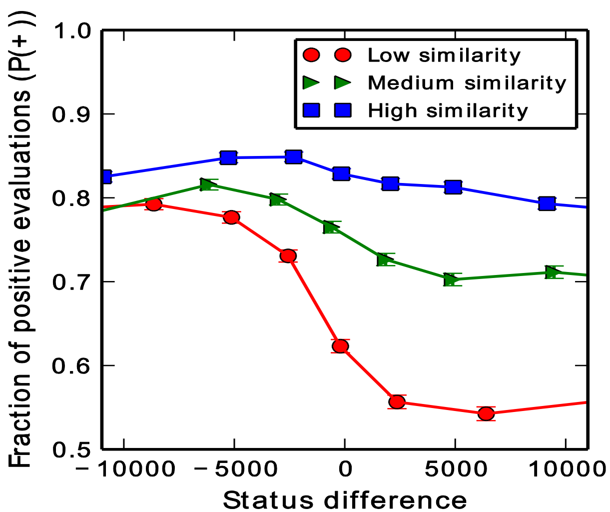


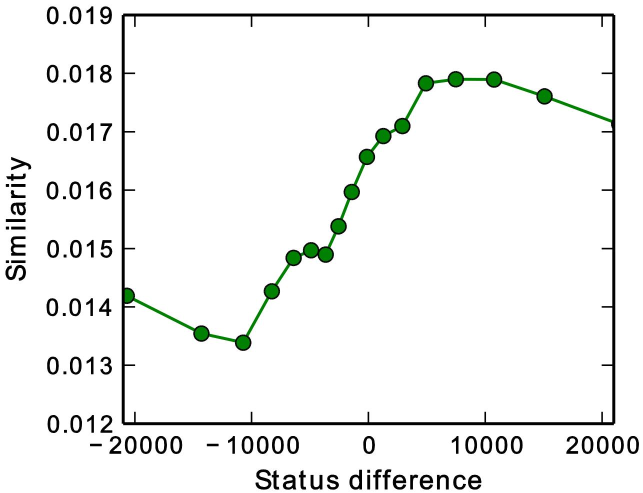

Computed over 120k votes

**How can we explain this?**

# The Mercy Bounce

- **Why low evals. of users of same status?**
  - Not due to users being tough on each other
  - **But due to the effects of similarity**




**Explanation:** For **negative status difference** we have low similarity people which behave according to the red curve on the left plot. As status difference increases the similarity also increases (green curve). For **positive status difference**, similarity is high, and evaluations follow the blue curve (left). By having a particularly weighted combination of red, green, and blue curve we observe the "mercy bounce" from the previous slide.

- **So we get the "mercy" bounce due to uneven mixing of votes.**

# Aggregating Evaluations

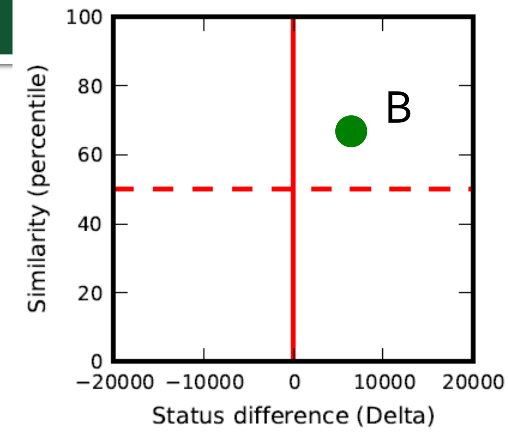
- **So far:** Properties of individual evaluations
- **But:** Evaluations need to be “summarized”
  - Determining rankings of users or items
  - Multiple evaluations lead to a group decision
- **How to aggregate user evaluations to obtain the opinion of the community?**
  - Can we guess community’s opinion from a small fraction of the makeup of the community?

# Ballot-blind Prediction

- **Predict Wikipedia adminship election results without seeing the votes**
  - Observe identities of the first *k* (=5) people voting (but *not* how they voted)
  - Want to predict the election outcome
    - Promotion vs. no promotion
- **Why is it hard?**
  - Don't see the votes (just voters)
  - Only see first 5 voters (out of ~50)

UNC CHARLOTTE

# Ballot-blind: The Model




- **Want to model prob. user *A* votes + in election of user *B***
- **Our model:**

$$P(A = +|B) = P_A + d(S_A - S_B, sim(A, B))$$

  - $P_A$ ... empirical fraction of **+**votes of A
  - *d(status,similarity)* ... avg. deviation in frac. of **+**votes
    - When *A* evaluates *B* from a particular *(status, similarity)* quadrant, how does this change their behavior on average?
      - Note: *d(status,similarity)* only takes 4 different values (based on the quadrant in the *(status,similarity)* space)
- Predict 'elected' if: $\sum_{i=1}^{k} P(A_i = +|B) > w$

# Ballot-blind Prediction

- **Based on only who showed to vote predict the outcome of the election**

| Number of voters seen | Accuracy |
|---|---|
| 5 | 71.4% |
| 10 | 75.0% |
| all | 75.6% |

- **Other methods:**
  - Guessing gives 52% accuracy
  - Logistic Regression on status and similarity features: 67%
  - If we see the first *k=5* votes 85% (gold standard)

**Theme:** Learning from implicit feedback

**Audience composition tells us something about their reaction**

CHARLOTTE

# Summary

- **Social media sites are governed by** (often implicit) **user evaluations**
- Wikipedia voting process has an explicit, public and recorded process of evaluation
- **Main characteristics:**
  - Importance of relative assessment: **Status**
  - Importance of prior interaction: **Similarity**
  - Diversity of individuals' response functions

- **Application:** Ballot-blind prediction

UNC CHARLOTTE

# Important Points

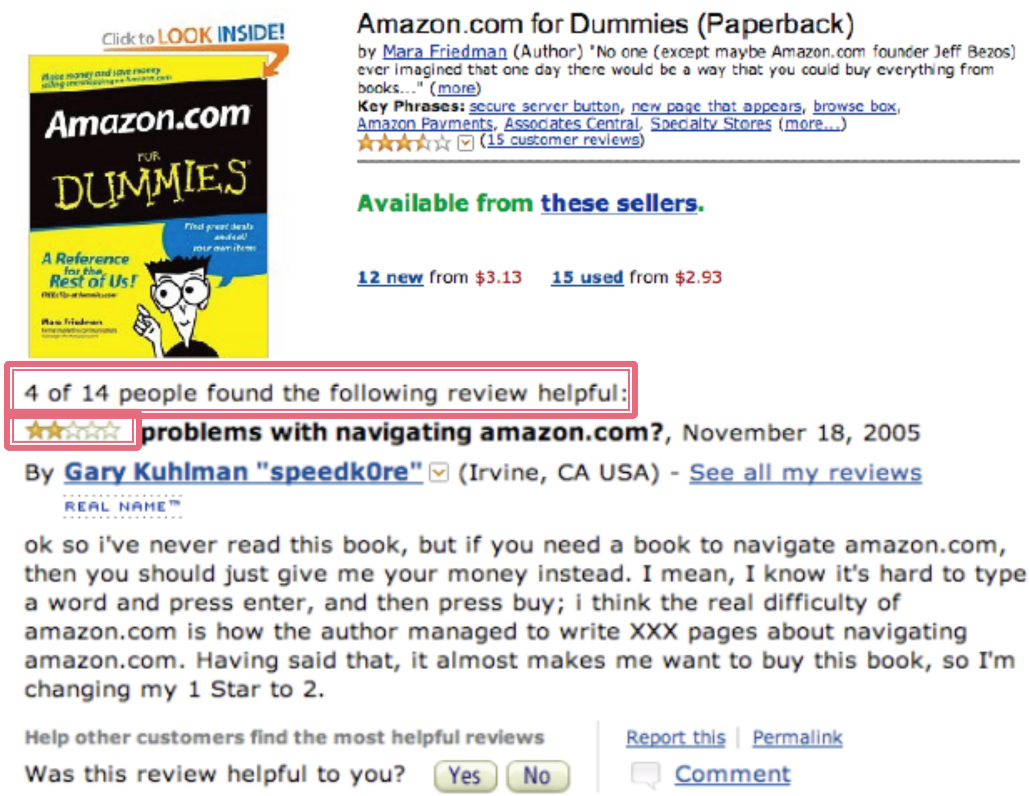
- **Status seems to be salient feature**
- **Similarity** also plays important role
- Audience composition helps predict audience's reaction
- **What kinds of opinions do people find helpful?**

UNC CHARLOTTE

[Danescu et al., 2009]

# What do People Find Helpful?

- **What do people think about our recommendations and opinions?**

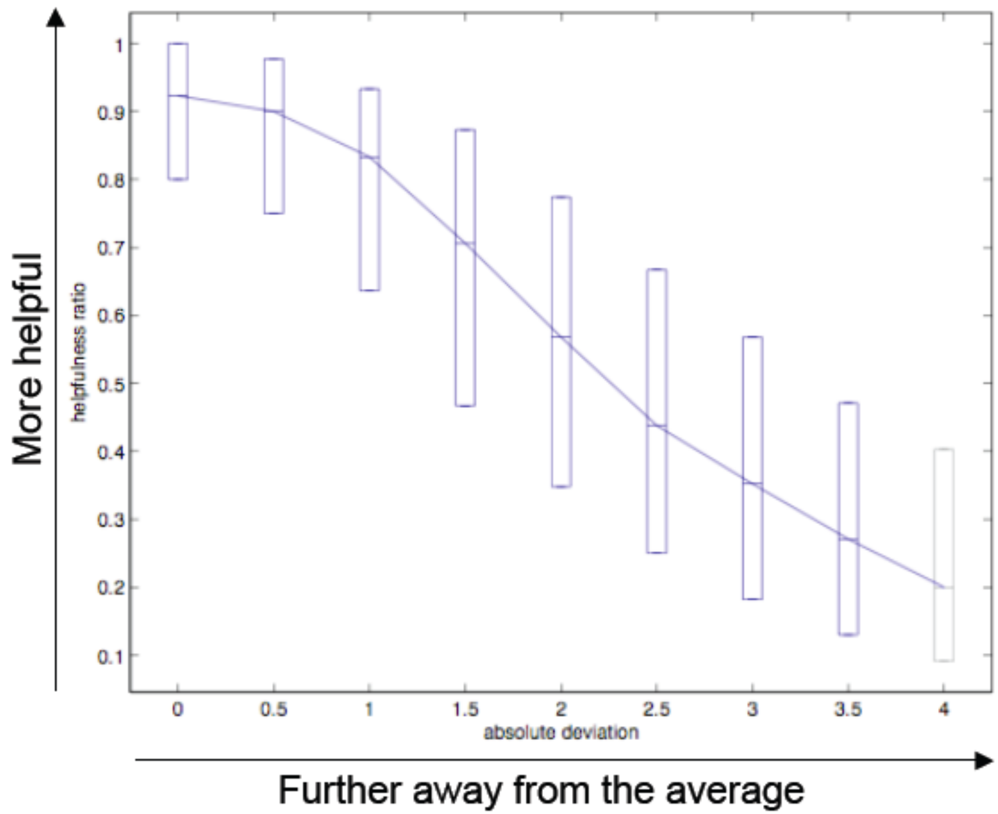



UNC CHARLOTTE

[Danescu et al., 2009]

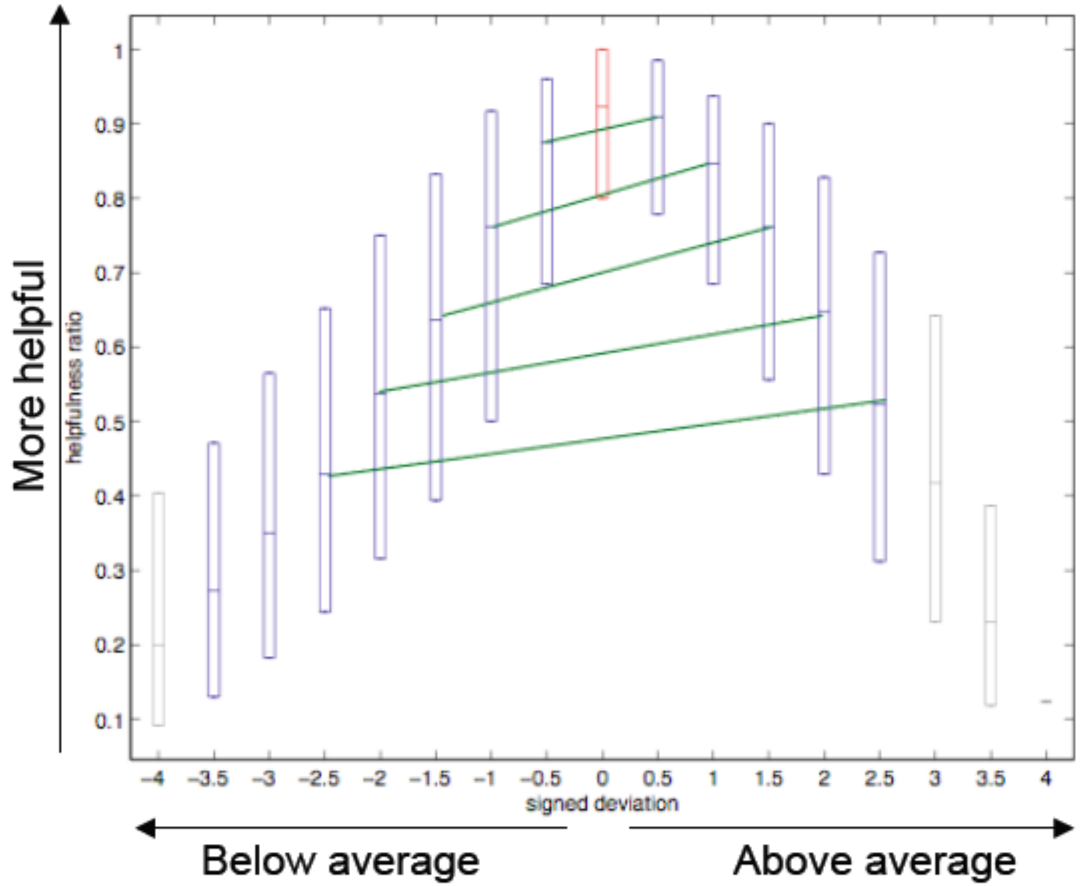
# Review Helpfulness: Conformity

- **Are conforming opinions more helpful?**



# Review Helpfulness: Deviation

- Are Positive reviews are more helpful?



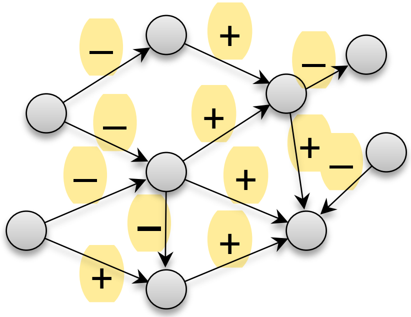

**Slight bias towards positive reviews**

# Evaluations happen in a context of a network!

# Two ways to look at this

- There are two ways to look at this: **One person evaluates the other via a positive/negative evaluation**


So far we focused on a single evaluation (without the context of a network)

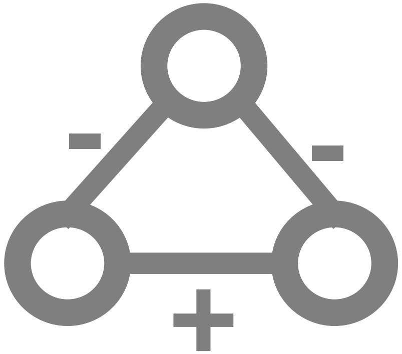


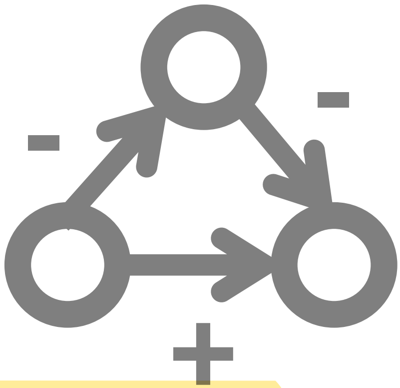

Now we will focus on evaluations in the context of a network

UNC CHARLOTTE

# Signed Networks

- **Networks with positive and negative relationships**
- Our basic unit of investigation will be **signed triangles**
- First we talk about **undirected** networks then **directed**
- **Plan:**
  - **Model:** Consider two social theories of signed nets
  - **Data:** Reason about them in large online networks
  - **Application:** Predict if A and B are linked with + or -




UNC CHARLOTTE

# Signed Networks

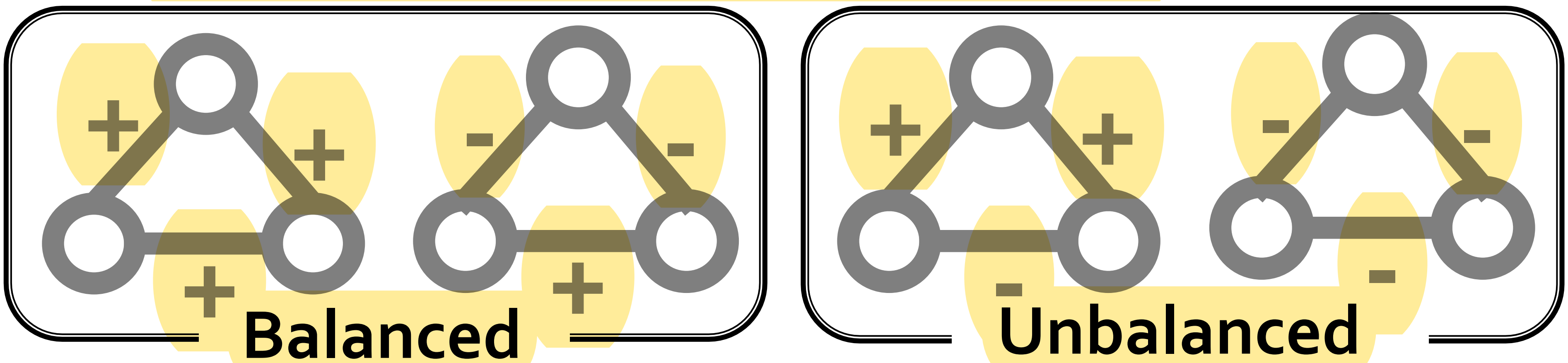
- **Networks with positive and negative relationships**

- Consider an **undirected complete graph**
- Label each edge as either:
  - **Positive:** friendship, trust, positive sentiment, …
  - **Negative:** enemy, distrust, negative sentiment, …

- Examine triples of connected nodes A, B, C

# Theory of Structural Balance

- **Start with the intuition** [Heider '46]:
  - Friend of my friend is my friend
  - Enemy of enemy is my friend
  - Enemy of friend is my enemy
- Look at connected triples of nodes:

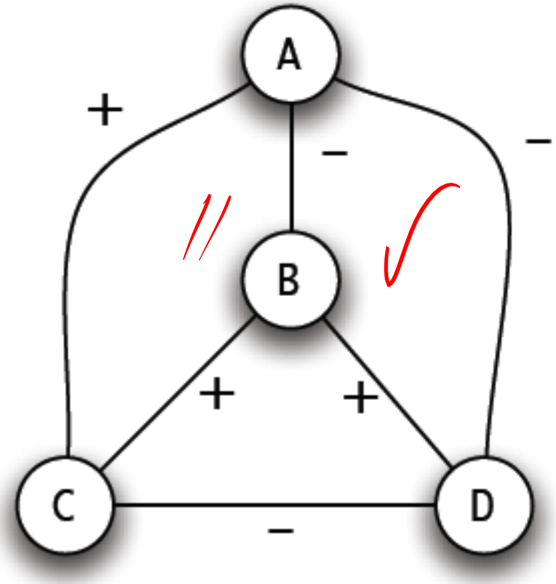



**Consistent** with "friend of a friend" or "enemy of the enemy" intuition

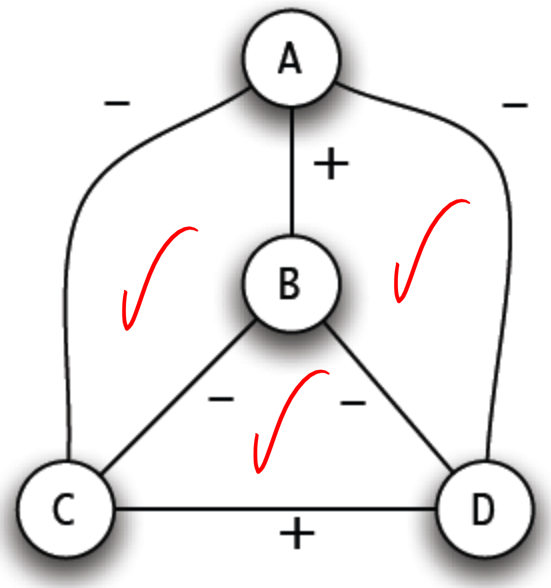
**Inconsistent** with the "friend of a friend" or "enemy of the enemy" intuition

# Balanced/Unbalanced Networks

- Graph is **balanced** if every connected triple of nodes has:
  - All 3 edges labeled +, or
  - Exactly 1 edge labeled +



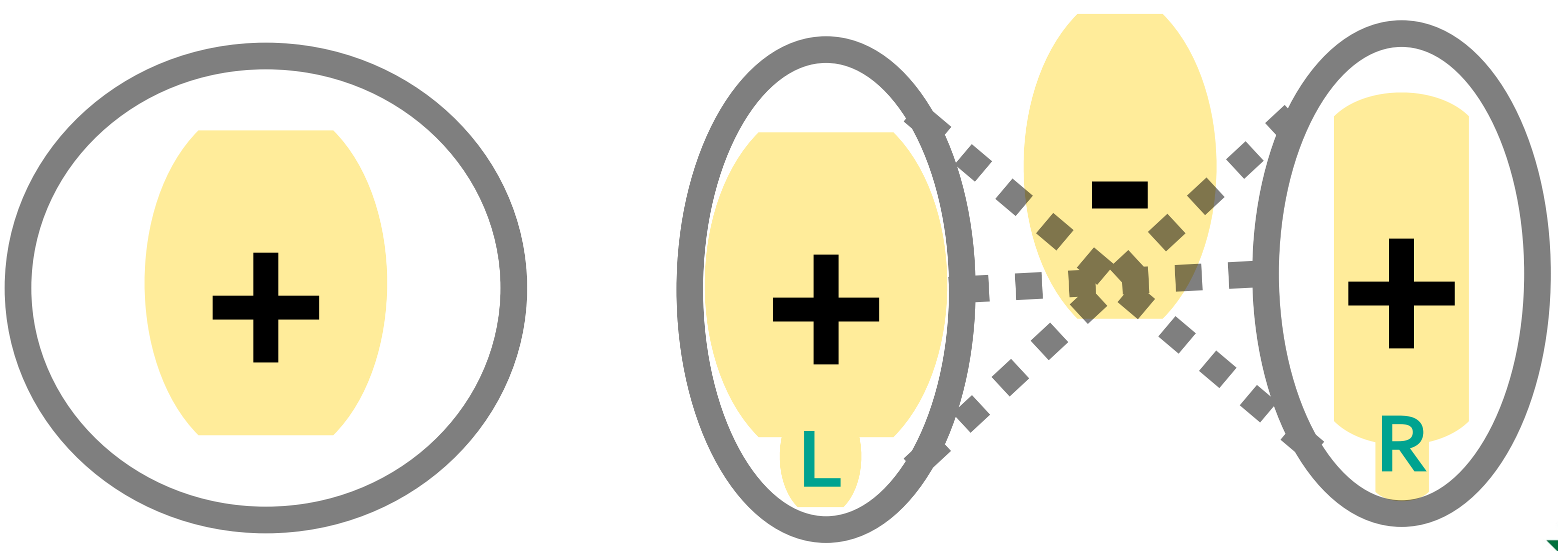

Unbalanced




Balanced

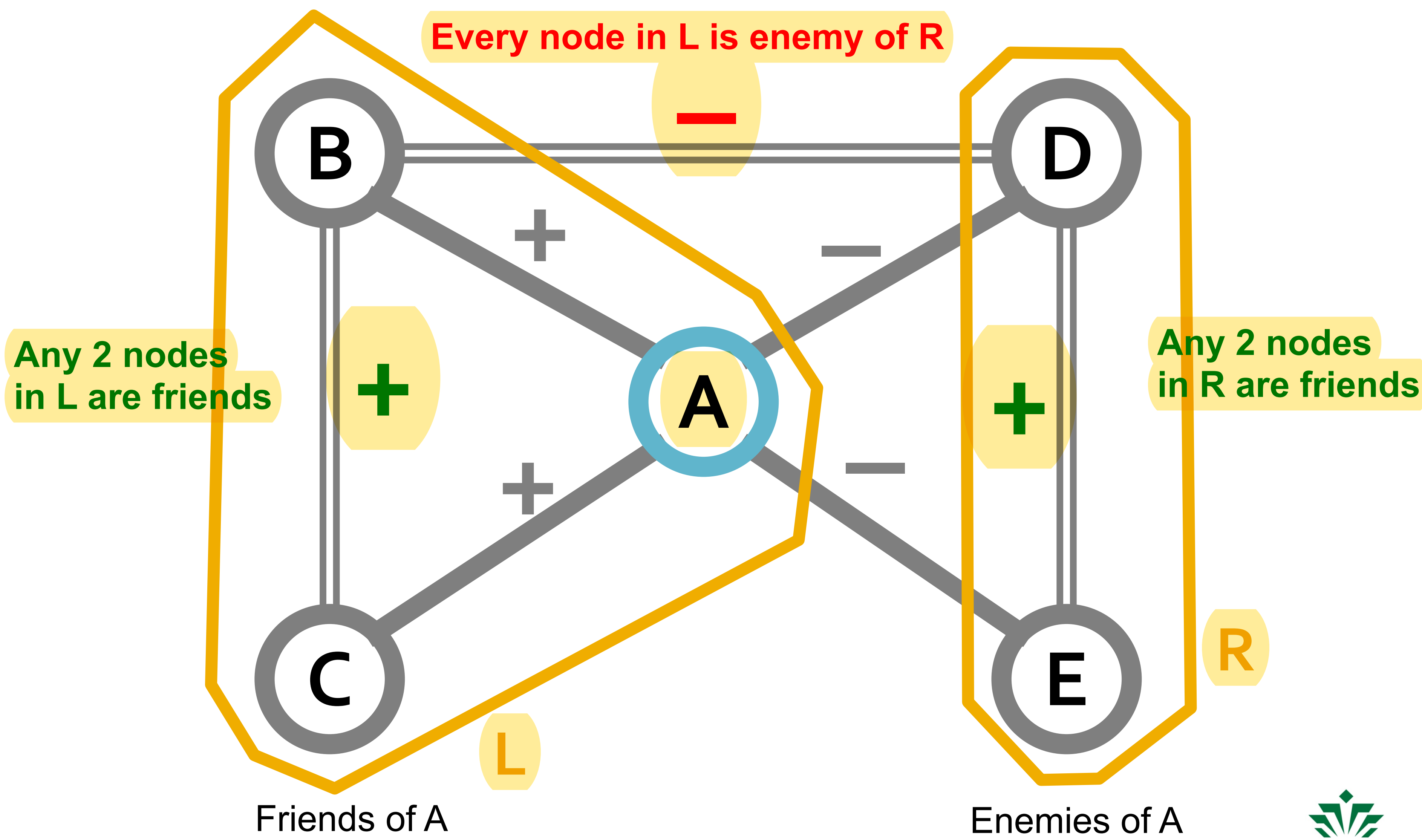
# Local Balance → Global Factions

- **Balance implies global coalitions** [Cartwright-Harary]
- If **all triangles are balanced**, then either:
  - The network contains only positive edges, or
  - Nodes can be split into 2 sets where negative edges only point between the sets




UNC CHARLOTTE

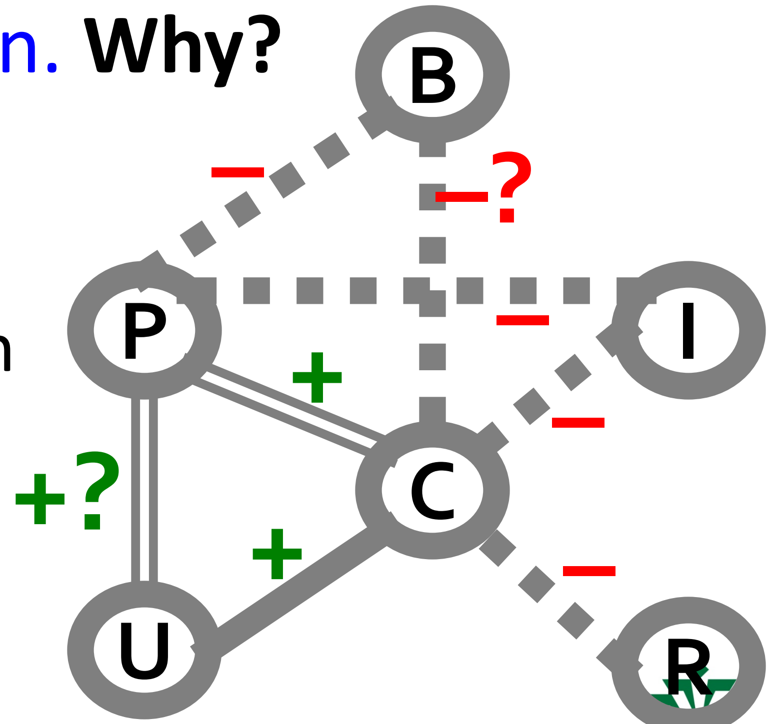
# Analysis of Balance




UNC CHARLOTTE

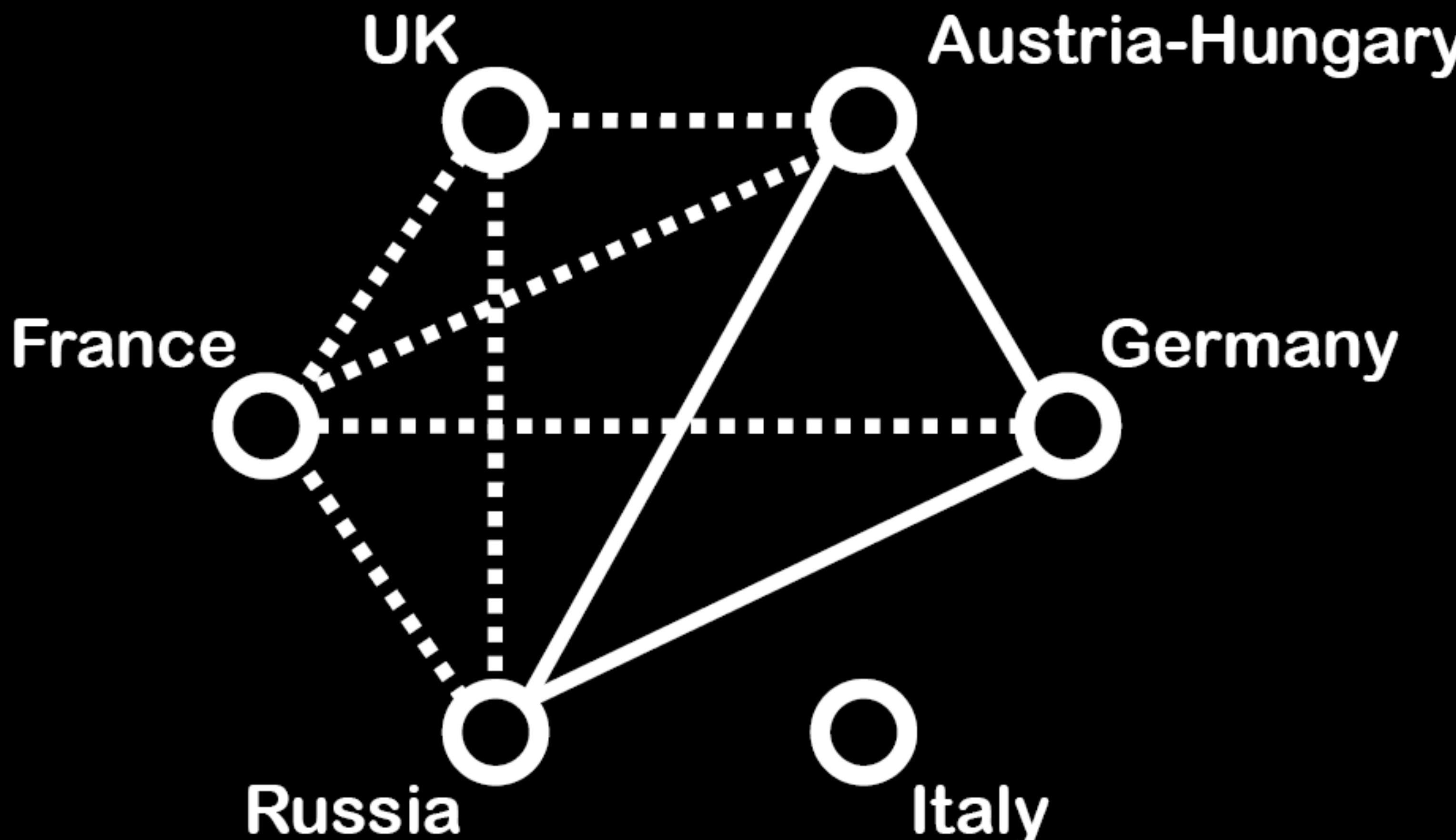
# Example: International Relations

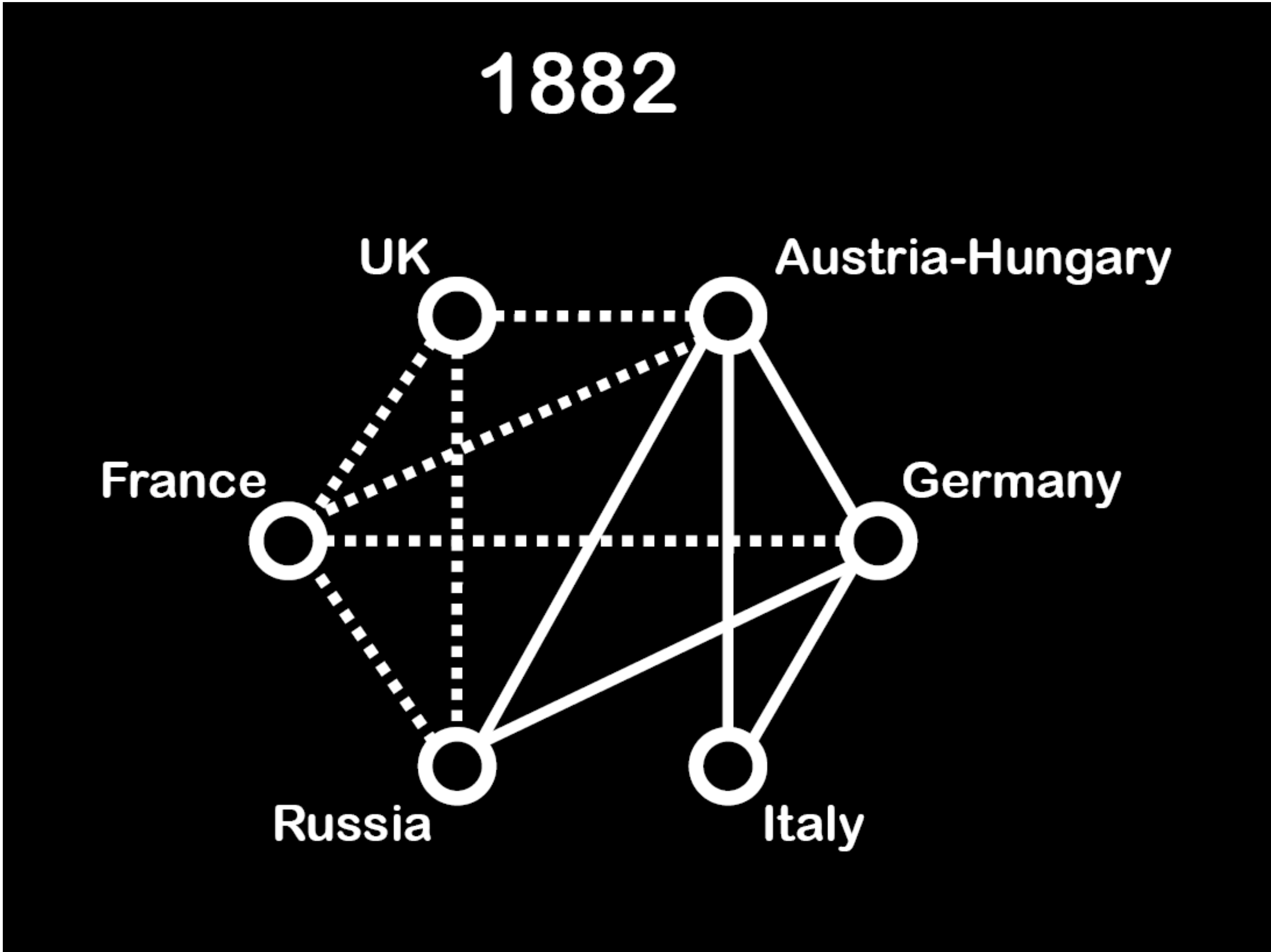
- **International relations:**
  - Positive edge: alliance
  - Negative edge: animosity
- Separation of Bangladesh from Pakistan in 1971: US supports Pakistan. **Why?**
  - USSR was enemy of China
  - China was enemy of India
  - India was enemy of Pakistan
  - US was friendly with China
  - China vetoed Bangladesh from U.N.

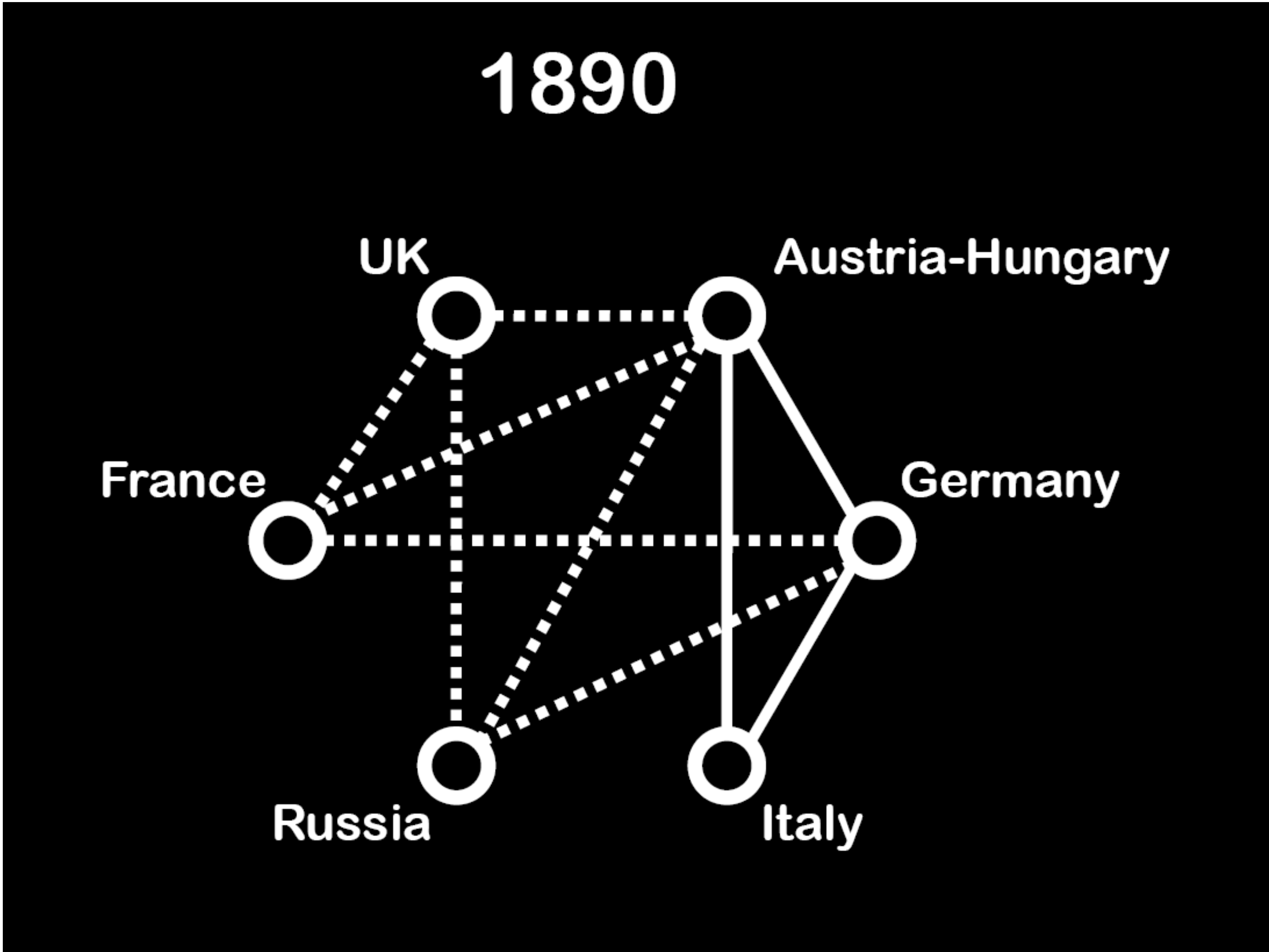



UNC CHARLOTTE

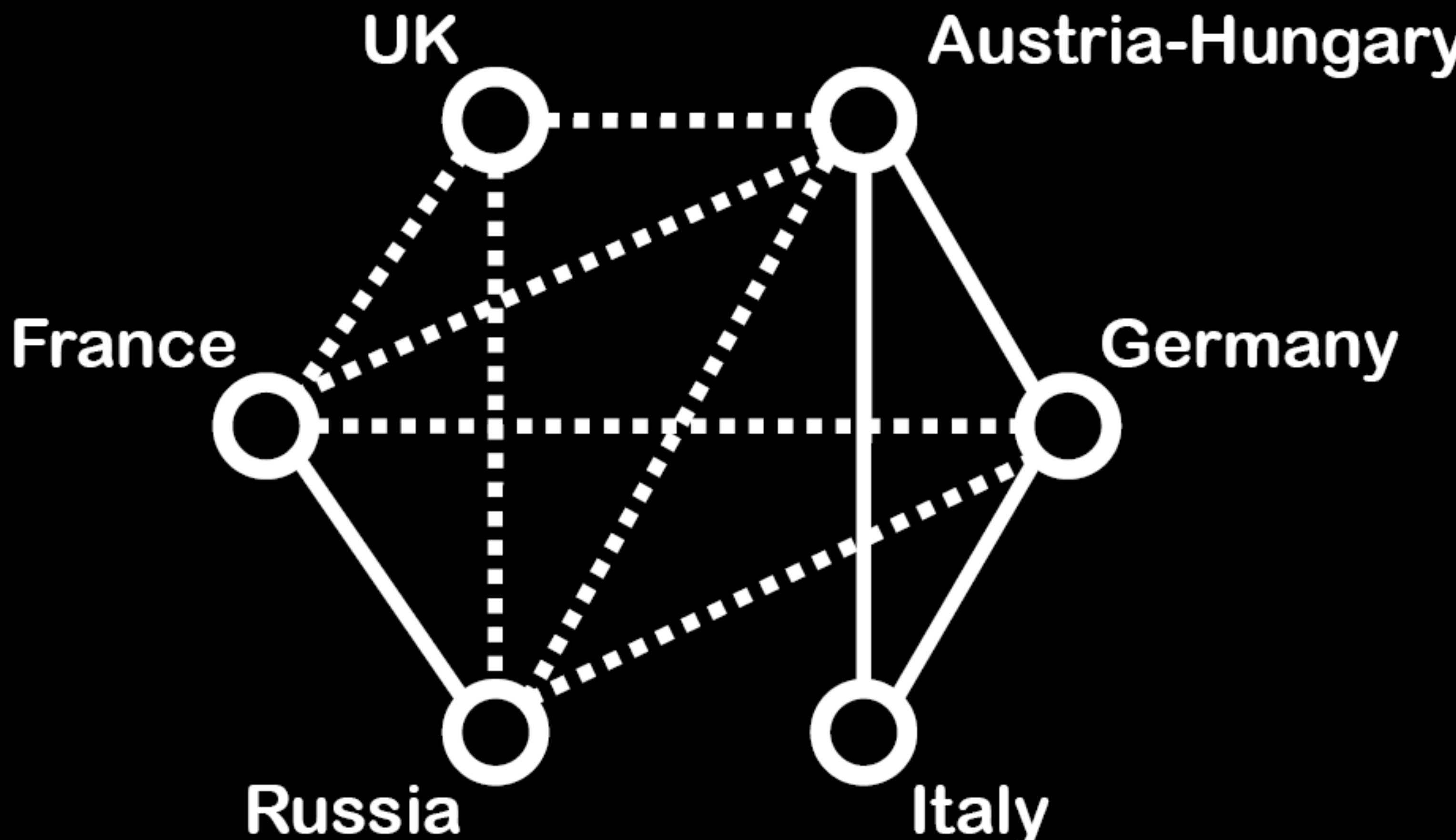
# 1872-1881



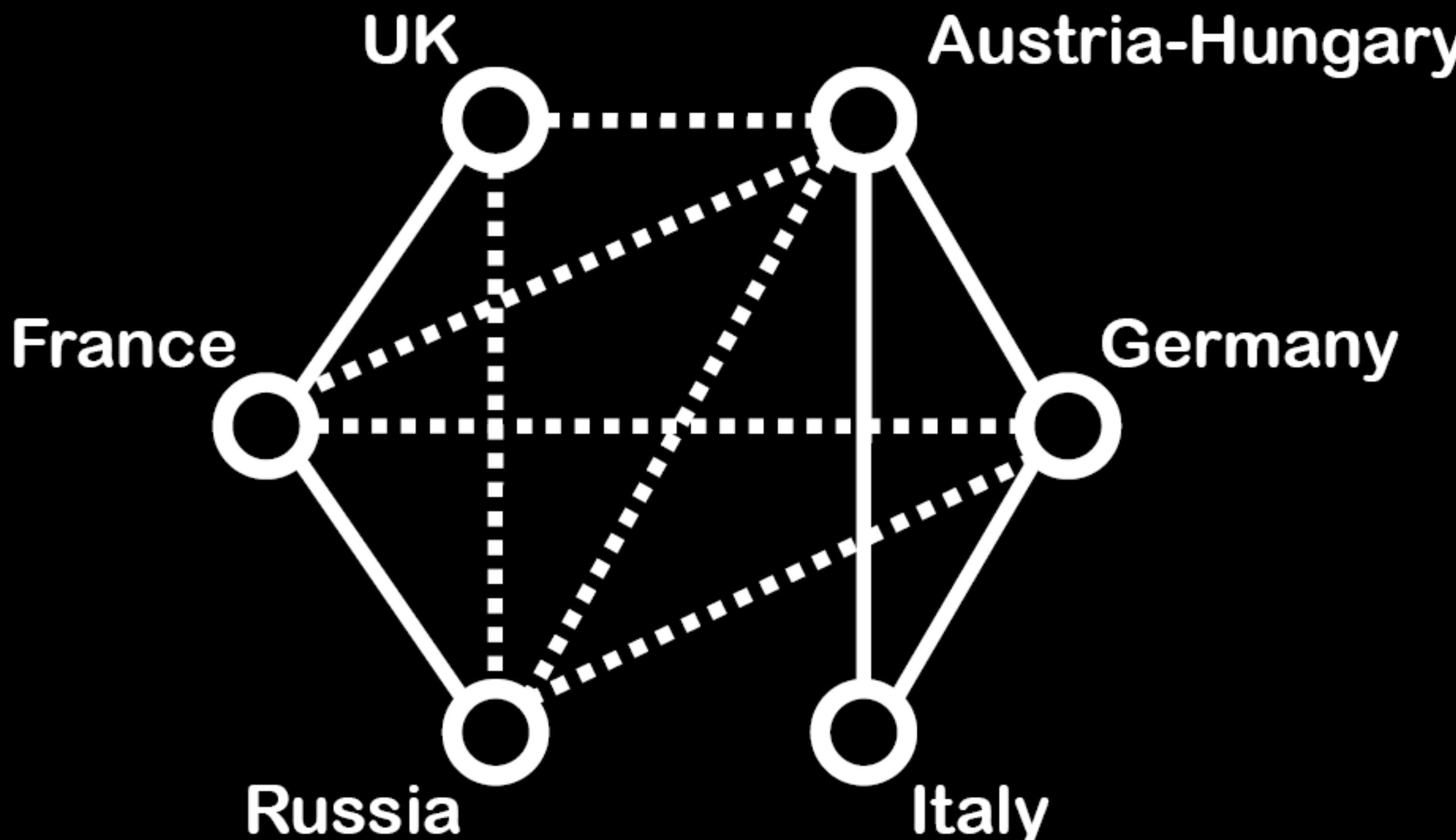

1882
UK
Austria-Hungary
France
Germany
Russia
Italy


1890
UK
Austria-Hungary
France
Germany
Russia
Italy

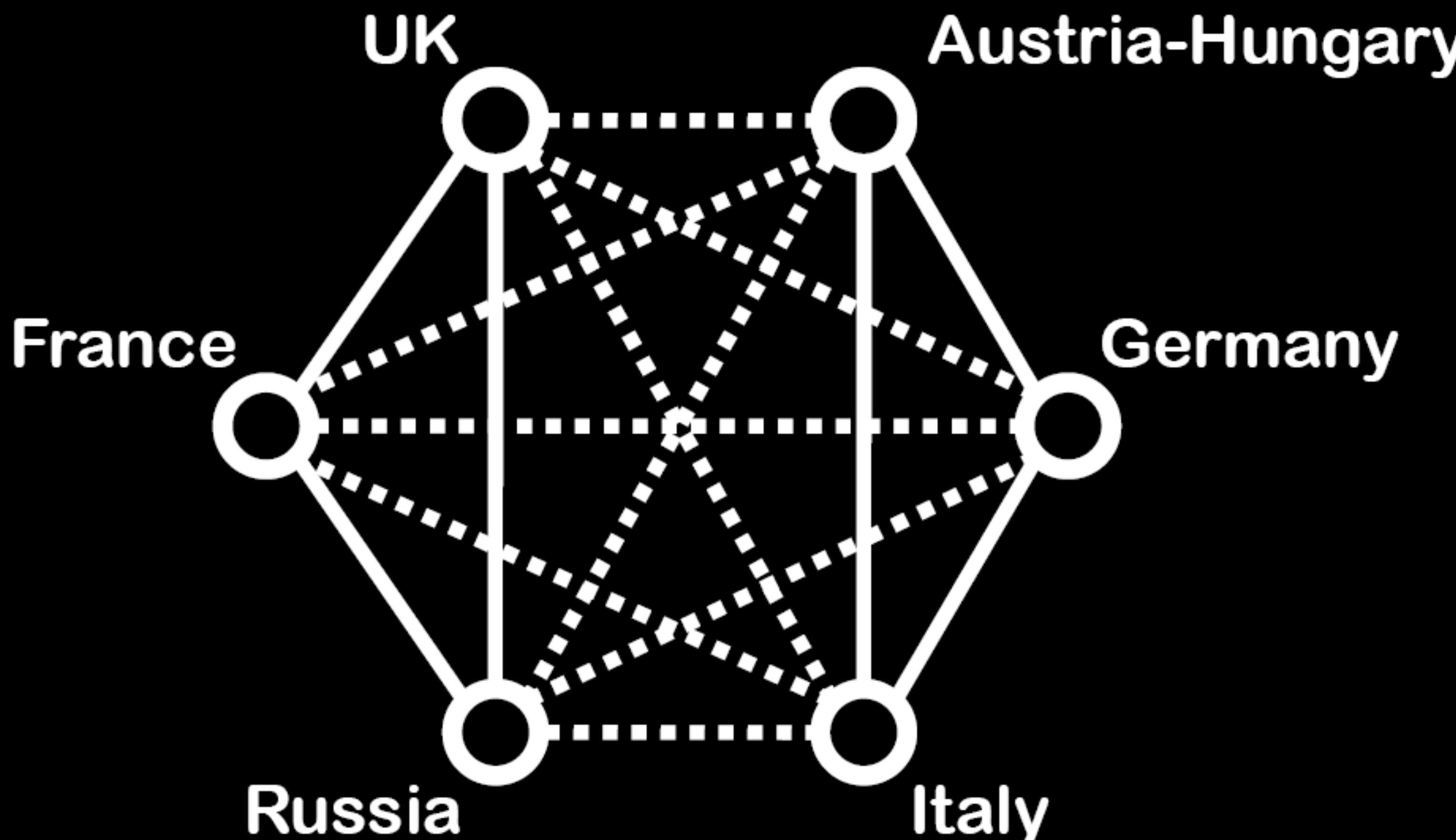
# 1891-1894



# 1904

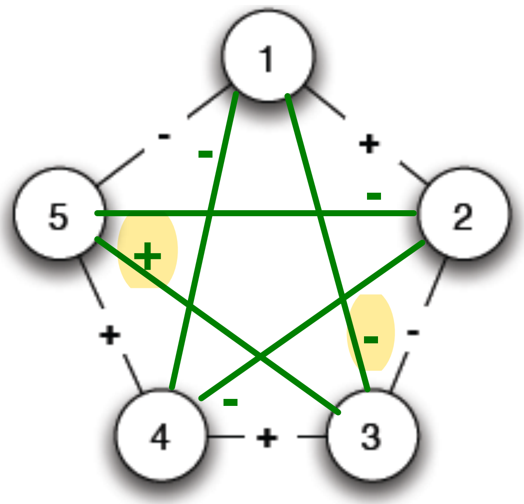


# 1907



# Balance in General Networks

- **So far we talked about complete graphs**




Balanced?

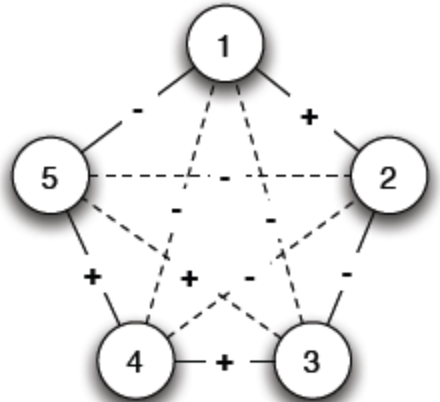
**Def 1: Local view**

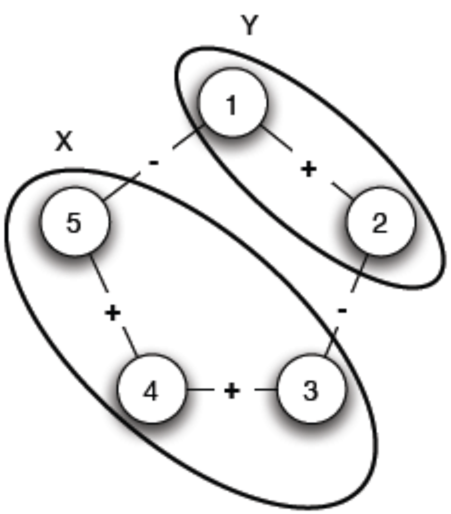
Fill in the missing edges to achieve balance

**Def 2: Global view**

Divide the graph into two coalitions

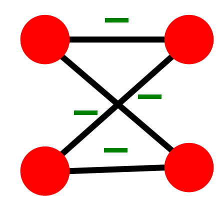
**The 2 definitions are equivalent!**



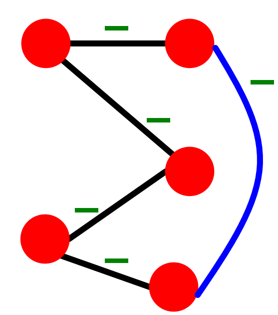



# Is a Signed Network Balanced?

- Graph is **balanced** if and only if it contains **no cycle with an odd number of negative** edges
- **How to compute this?**
  - Find connected components on + edges
    - If we find a component of nodes on +edges that contains a – edge ⇒ **Unbalanced**
  - For each component create a super-node
  - Connect components A and B if there is a negative edge between the members
  - Assign super-nodes to sides using BFS

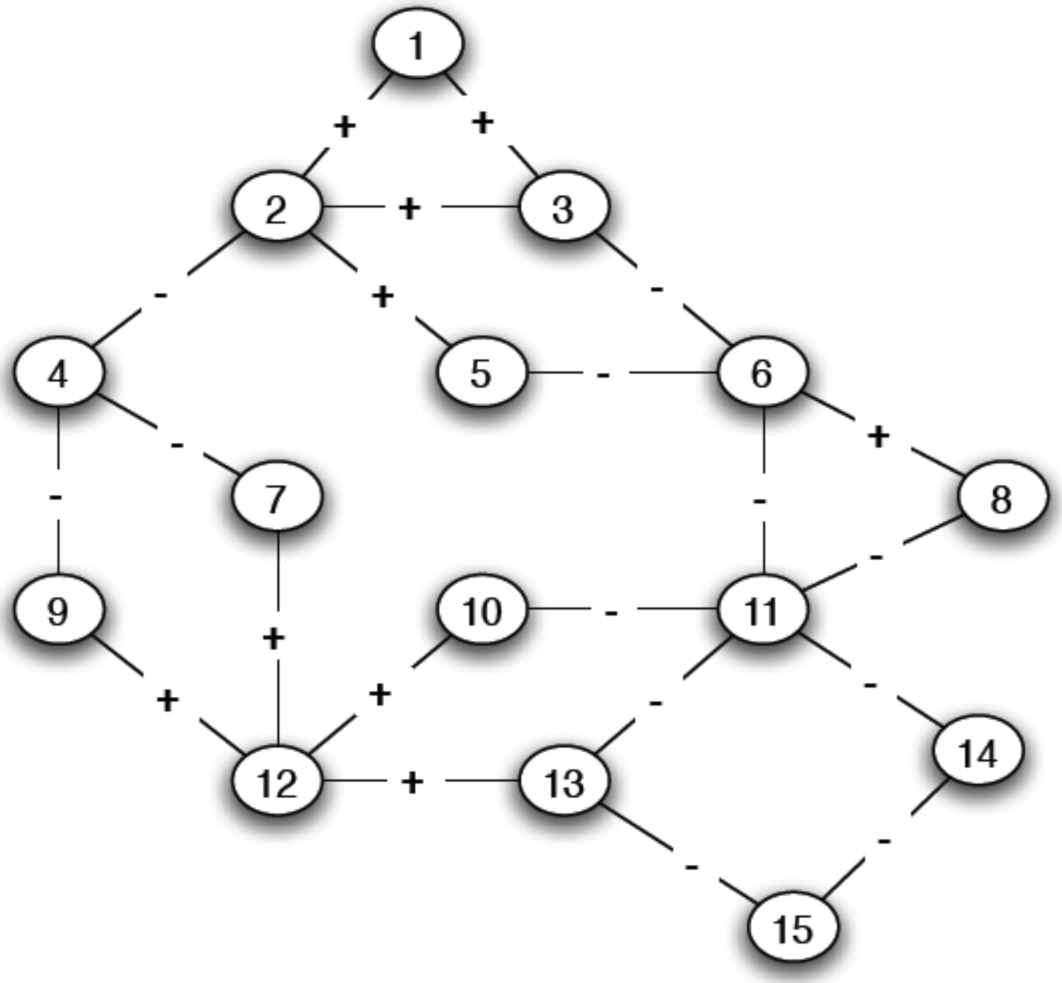



Even length cycle

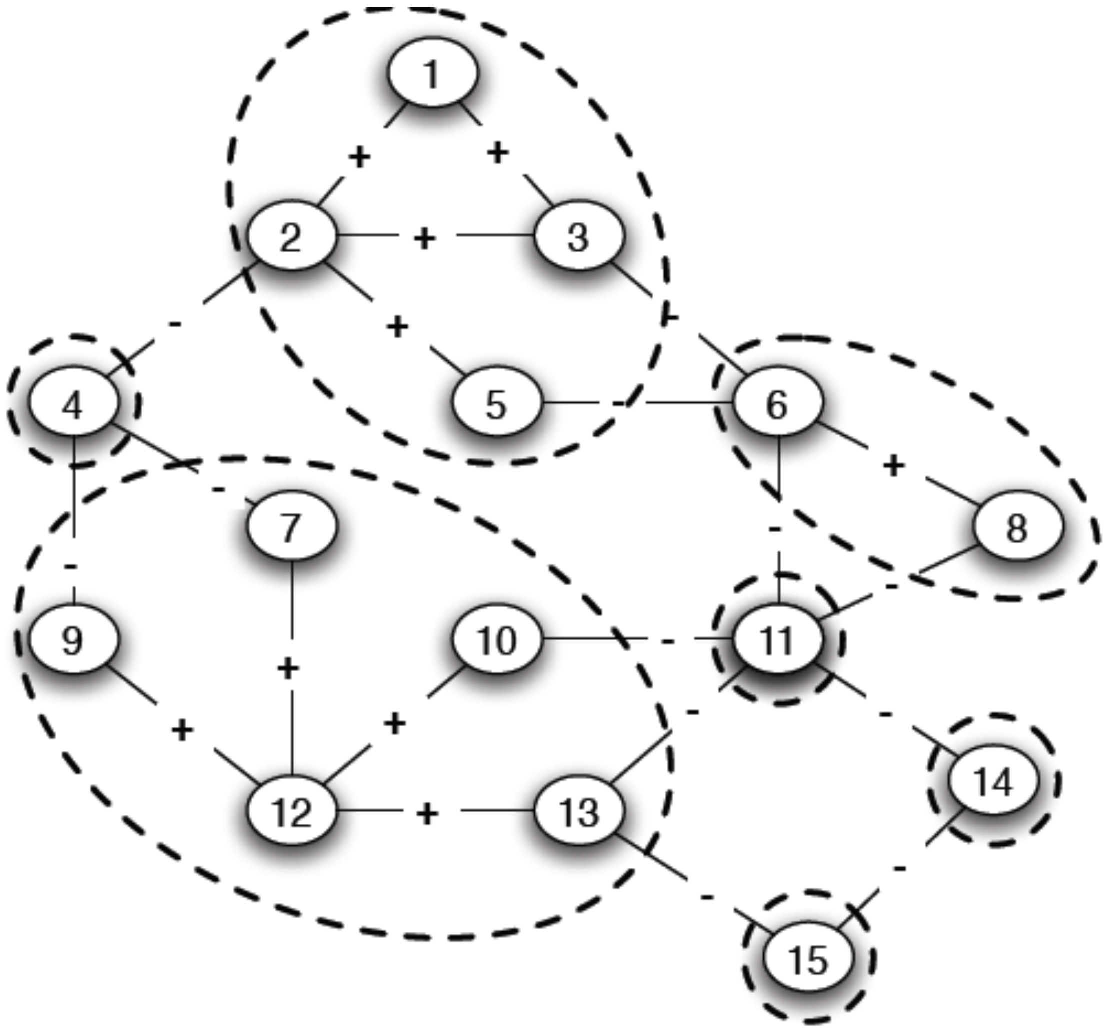



Odd length cycle

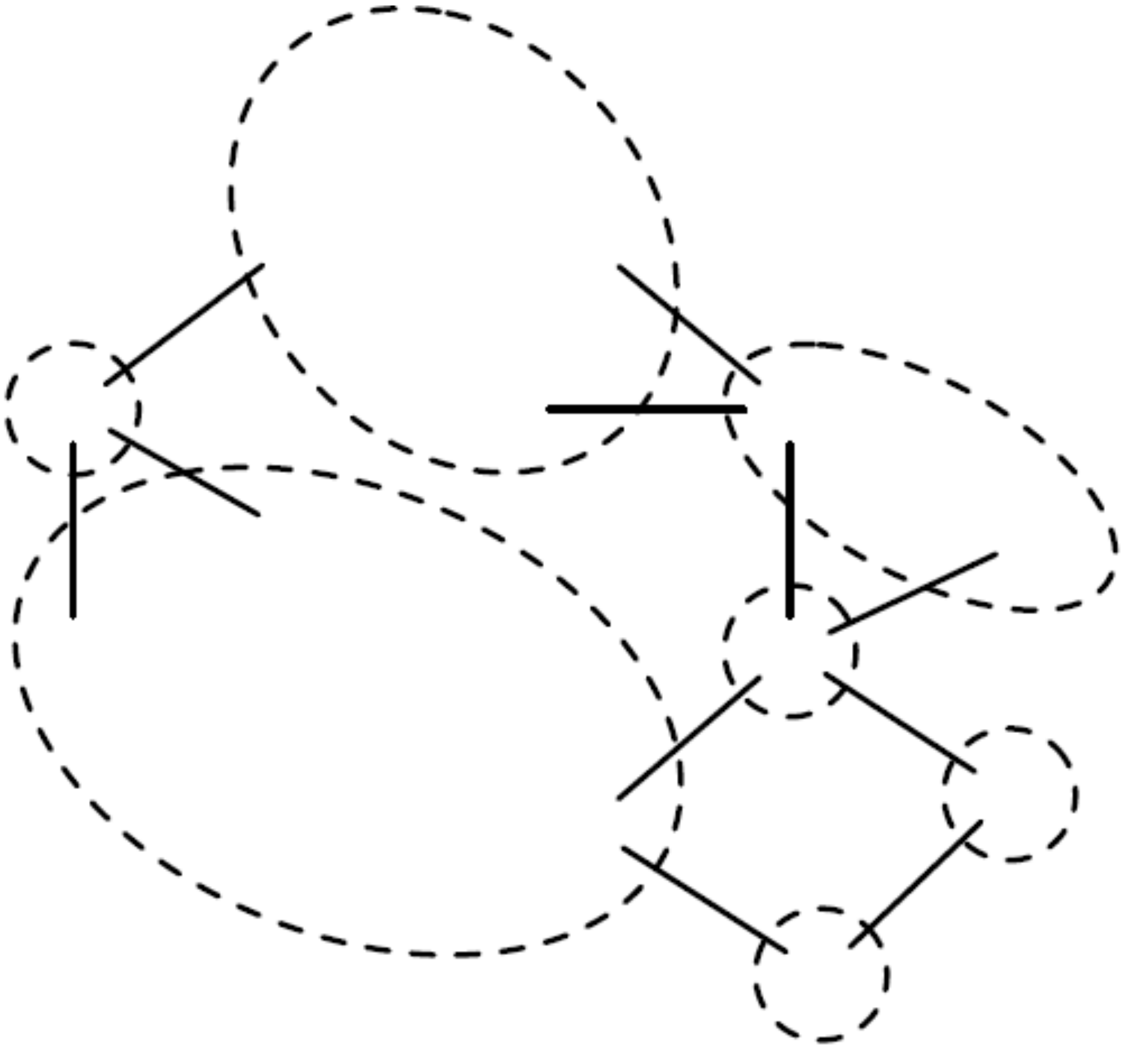
# Signed Graph: Is it Balanced?



# Positive Connected Components

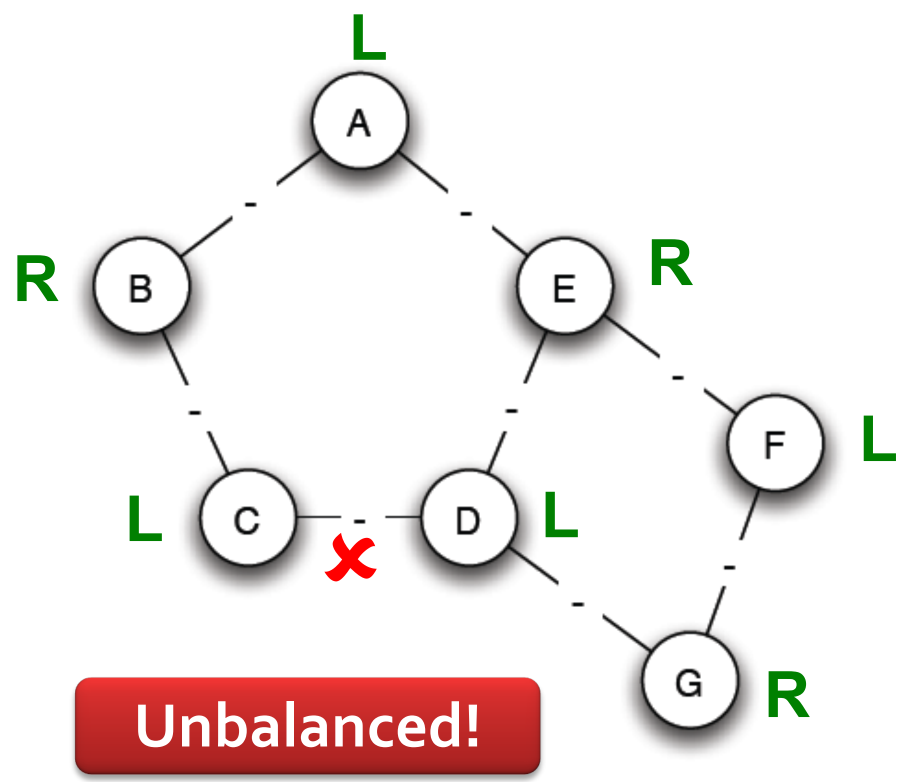


# Reduced Graph on Super-Nodes



# BFS on Reduced Graph

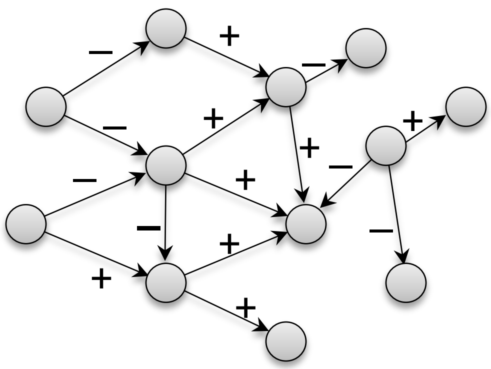
- Using BFS assign each node a **side**
- Graph is **unbalanced** if any two connected super-nodes are assigned the **same side**



# Exploring Real Data

# Real Large Signed Networks

- Each link **A→B** is **explicitly** tagged with a sign:
  - Epinions: Trust/Distrust
    - Does A trust B's product reviews?
      (only positive links are visible to users)
  - Wikipedia: Support/Oppose
    - Does A support B to become Wikipedia administrator?
  - Slashdot: Friend/Foe
    - Does A like B's comments?
  - **Other examples:**
    - Online multiplayer games




| | Epinions | Slashdot | Wikipedia |
|---|---|---|---|
| Nodes | 119,217 | 82,144 | 7,118 |
| Edges | 841,200 | 549,202 | 103,747 |
| + edges | 85.0% | 77.4% | 78.7% |
| − edges | 15.0% | 22.6% | 21.2% |

UNC CHARLOTTE

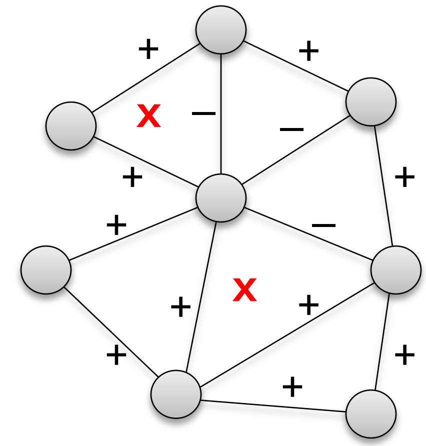
# Balance in Our Network Data

- **Does structural balance hold?**
  - Compare frequencies of signed triads in real and "shuffled" signs

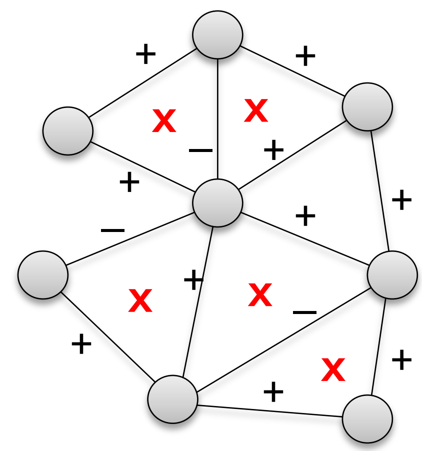
| | Triad | Epinions | | Wikipedia | | Consistent with Balance? |
|---|---|---|---|---|---|---|
| | | P(T) | $P_0(T)$ | P(T) | $P_0(T)$ | |
| Balanced | + + + | 0.87 | 0.62 | 0.70 | 0.49 | ✓ |
| Balanced | - - + | 0.07 | 0.05 | 0.21 | 0.10 | ✓ |
| Unbalanced | + + - | 0.05 | 0.32 | 0.08 | 0.49 | ✓ |
| Unbalanced | - - - | 0.007 | 0.003 | 0.011 | 0.010 | ✗ |

P(T) … fraction of a triads

$P_0(T)$… triad fraction if the signs would appear at random



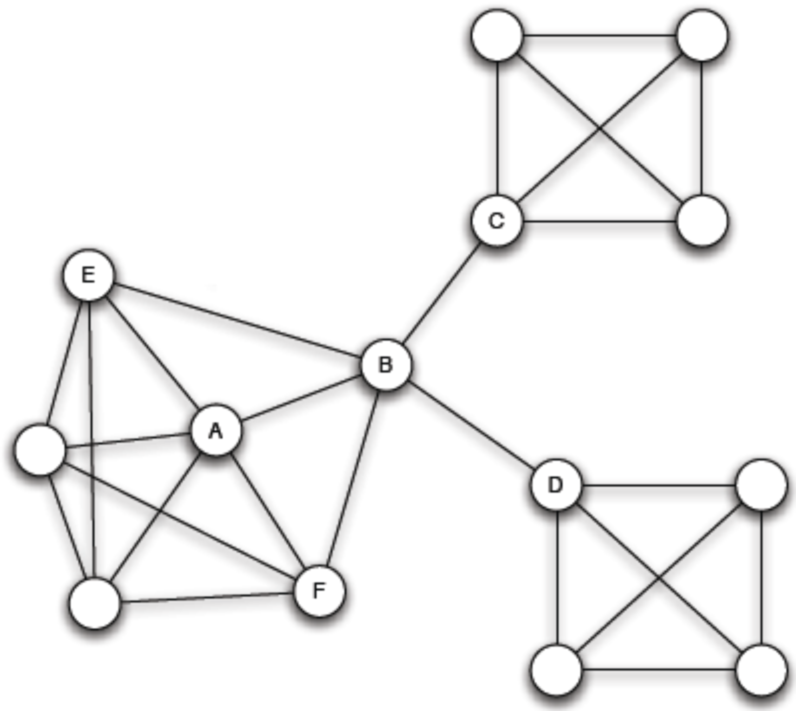
Real data



Shuffled data

# Global Structure of Signed Nets

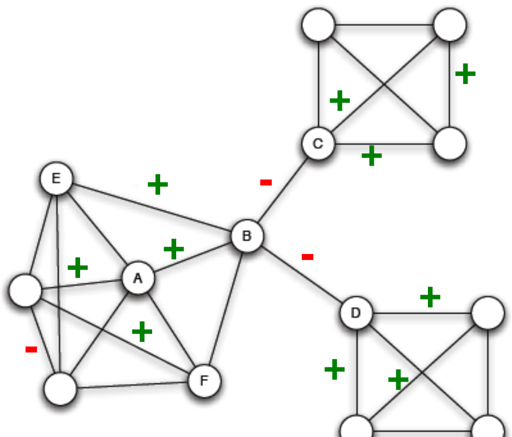
- Intuitive picture of social network in terms of densely linked clusters
- How does structure interact with links?
- Embeddedness of link (A,B): Number of shared neighbors

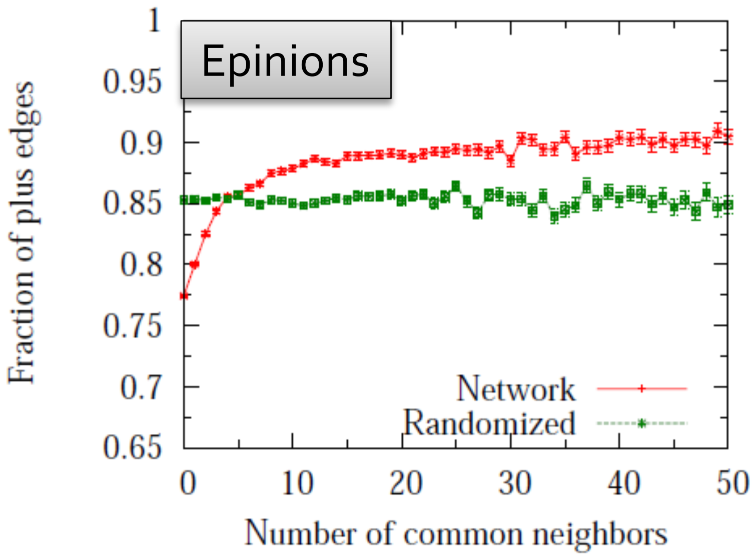


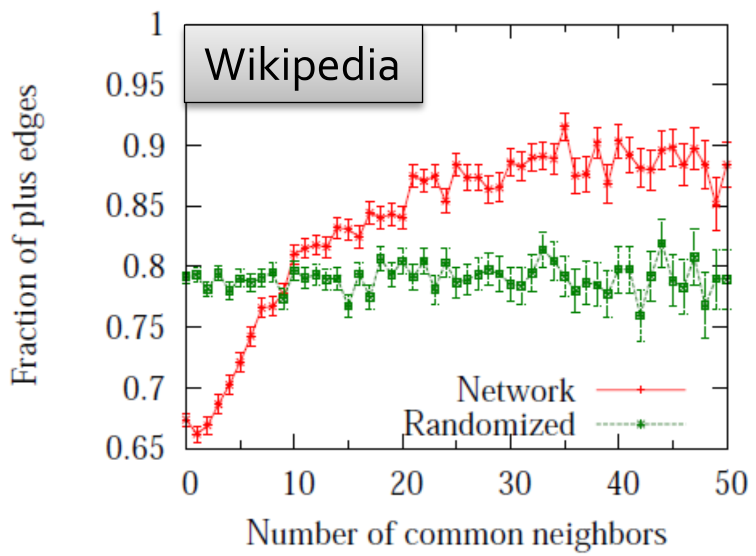

UNC CHARLOTTE

# Global Factions: Embeddedness

- **Embeddedness of ties:**
  - Positive ties tend to be **more** embedded



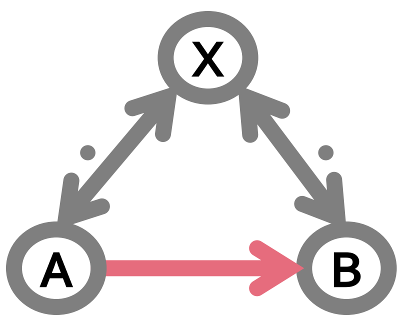



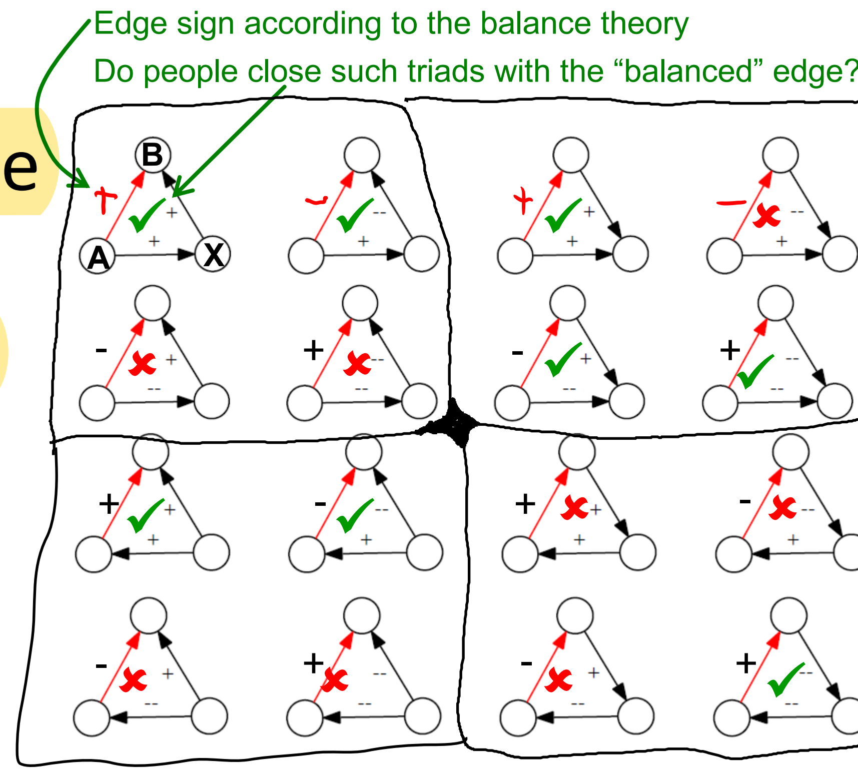
[CHI '10]

# Evolving Directed Networks

- **New setting:** Links are **directed**, created over time
  - Node **A** links to **B**
  - Directions and signs of links from/to X provide context




- How many △ are now explained by balance?
  - Only half (8 out of 16)




16 signed directed triads

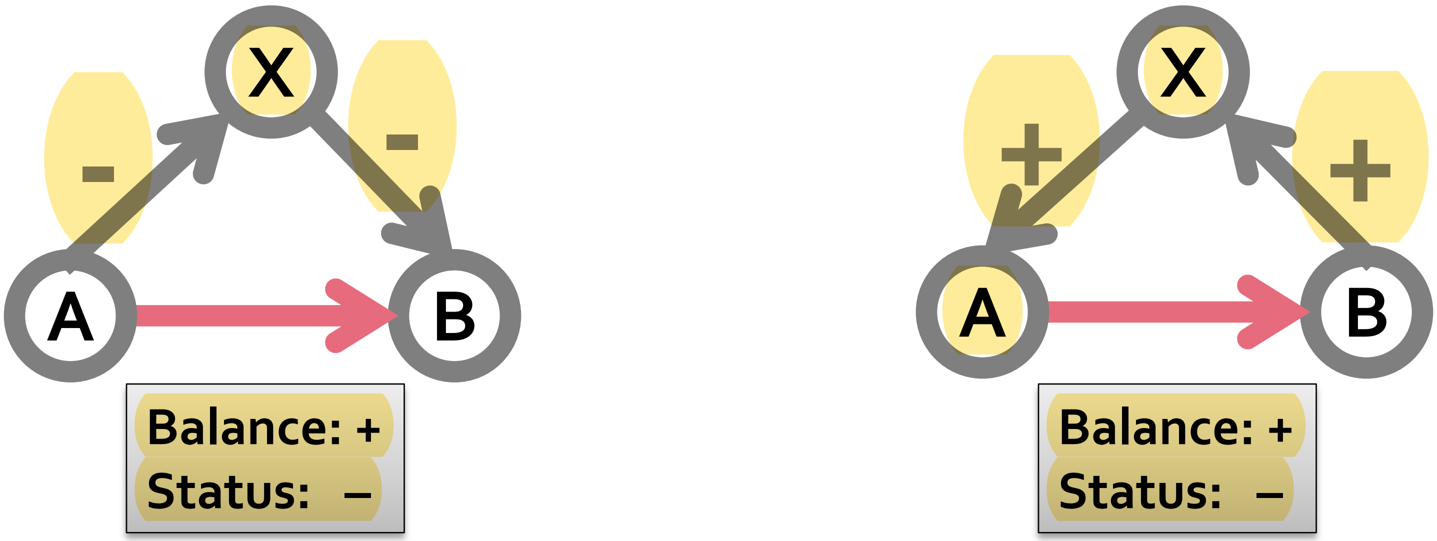
(in directed networks people traditionally applied balance by ignoring edge directions)

# Alternate Theory: Status

- **Status** in a network [Davis-Leinhardt '68]
  - **A** $\xrightarrow{+}$ **B** :: **B** has higher status than **A**
  - **A** $\xrightarrow{-}$ **B** :: **B** has lower status than **A**
    - Note the notion of status is now implicit and governed by the network (rather than the number of edits)
  - **Apply this principle transitively over paths**
    - Can replace each **A** $\xrightarrow{-}$ **B** with **A** $\xleftarrow{+}$ **B**
    - Obtain an all-positive network with same status interpretation

# Status vs. Balance

[CHI '10]



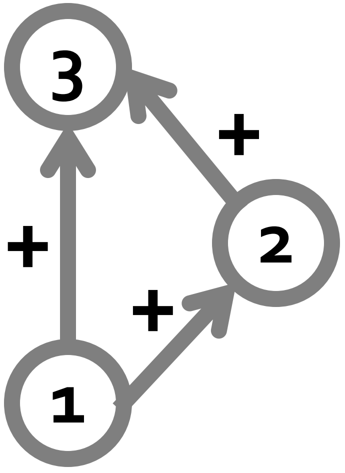

**Status and balance give different predictions!**

UNC CHARLOTTE

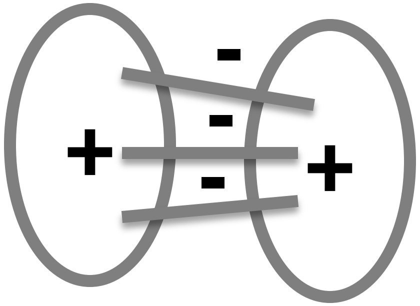
# Status vs. Balance

**At a global level (in the ideal case):**

- **Status ⇒ Hierarchy**
  - All-positive directed network should be approximately **acyclic**

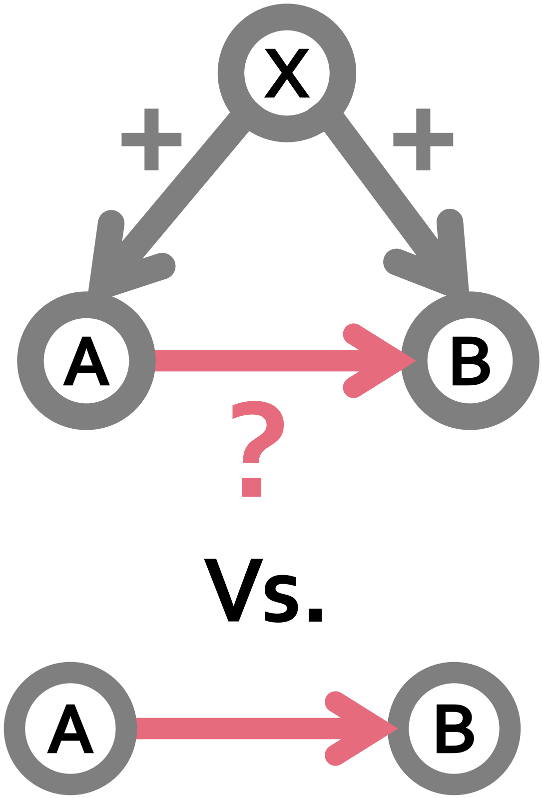



- **Balance ⇒ Coalitions**
  - Balance ignores directions and implies that subgraph of negative edges should be approximately **bipartite**



# Theory of Status

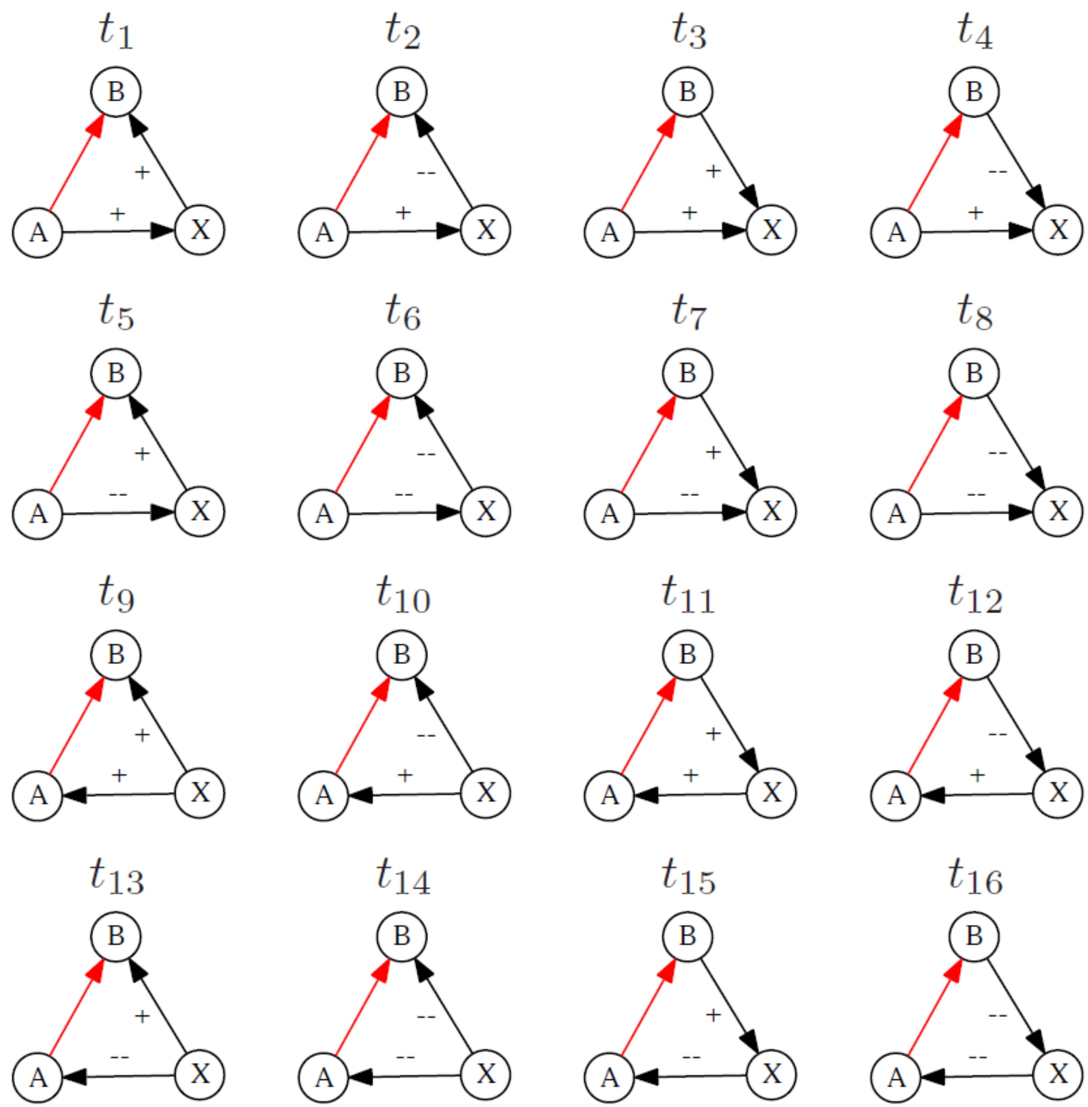
- **Edges are directed:**
  - X has links to A and B
  - Now, A links to B (triad A-B-X)
  - **How does sign of A⟶B depend signs from/to X?**
    **$P(A \xrightarrow{+} B \mid X)$ vs. $P(A \xrightarrow{+} B)$**
- **We need to formalize:**
  - **1) Links are embedded in triads:** Triads provide context for signs
  - **2) Users are heterogeneous** in their linking behavior




UNC CHARLOTTE

# 1) Context: 16 Types

- **Link A⟶B appears in context X: A⟶B | X**
- **16 possible contexts:**




**Note:** Context of a link is uniquely determined by the directions and signs of links from/to **X**

# 2) Heterogeneity in linking behavior

- Users differ in frac. of + links they give/receive
- **For a user U:**
  - Generative **baseline:** Frac. of **+** given by **U**
  - Receptive **baseline:** Frac. of **+** received by **U**

**Basic question:**

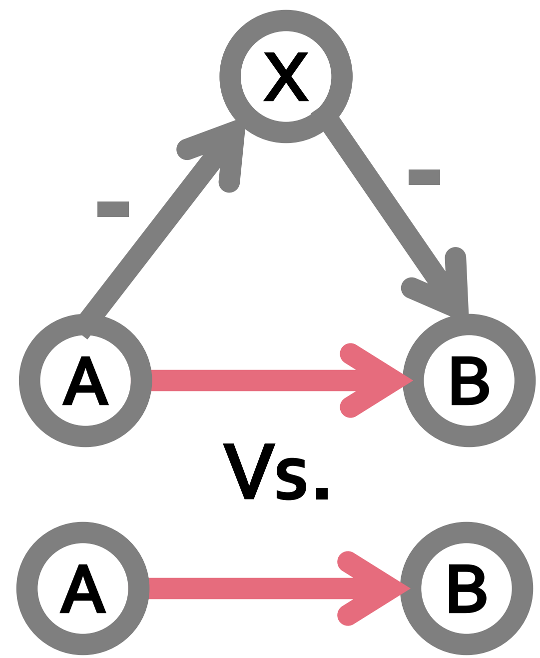
- How do **different link contexts** cause users to **deviate from their baselines**?
  - Link contexts as modifiers on a person's predicted behavior
  - **Def:** **Surprise**: How much behavior of A/B **deviates** from his/her **baseline** when A/B is in **context X**

UNC CHARLOTTE

# Computing Surprise

- **Surprise:** How much behavior of user **deviates** from **baseline** in **context X**
  - **Baseline:** For every user $A_i$ :
    $p_g(A_i)$… **generative baseline** of $A_i$
    - Fraction of times $A_i$ gives a plus
  - **Context:** $(A_1, B_1 | X_1), \ldots, (A_n, B_n | X_n)$
    … all instances of triad context $X$
    - $(\mathbf{A_i}, \mathbf{B_i}, \mathbf{X_i})$ … an instance where when user $\mathbf{A_i}$ links to user $\mathbf{B_i}$ the triad of type **X** is created.
    - Say $k$ of those triads closed with a plus
      - $k$ out of $n$ times: $\mathbf{A_i} \xrightarrow{+} \mathbf{B_i}$

Context X:



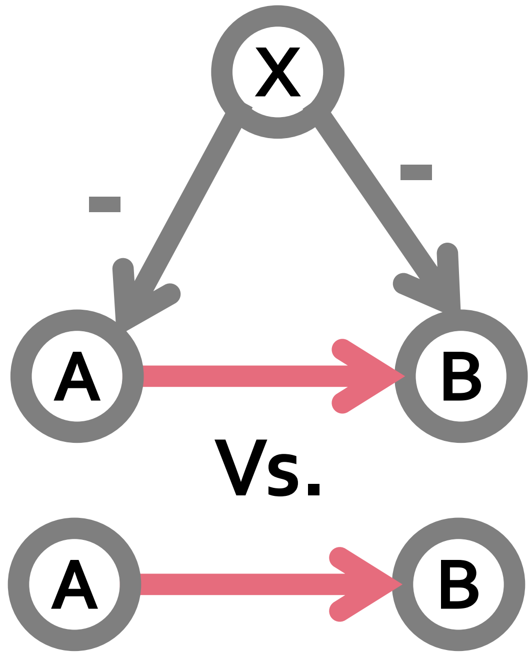
# Computing Surprise

- **Surprise:** How much behavior of users **deviates** from **baseline** when in **context X**
  - **Generative surprise of context *X*:**

$$s_g(X) = \frac{k - \sum_{i=1}^{n} p_g(A_i)}{\sqrt{\sum_{i=1}^{n} p_g(A_i)(1 - p_g(A_i))}}$$

    - **$p_g(A_i)$ … generative baseline** of **$A_i$**
    - **Context X:** $(A_1, B_1 | X_1), \ldots, (A_n, B_n | X_n)$
    - ***k*** of instances of triad ***X*** closed with a plus edges
  - Receptive surprise is similar, just use **$p_r(A_i)$**

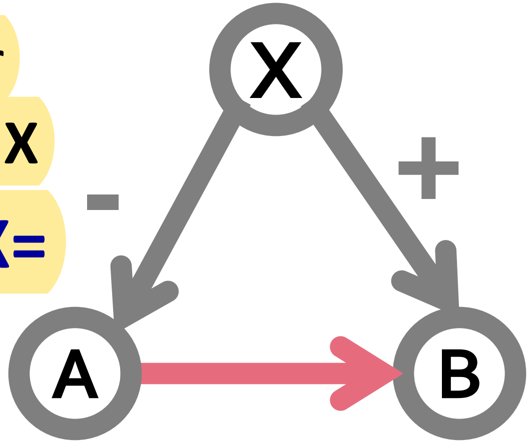
Context X:




UNC CHARLOTTE

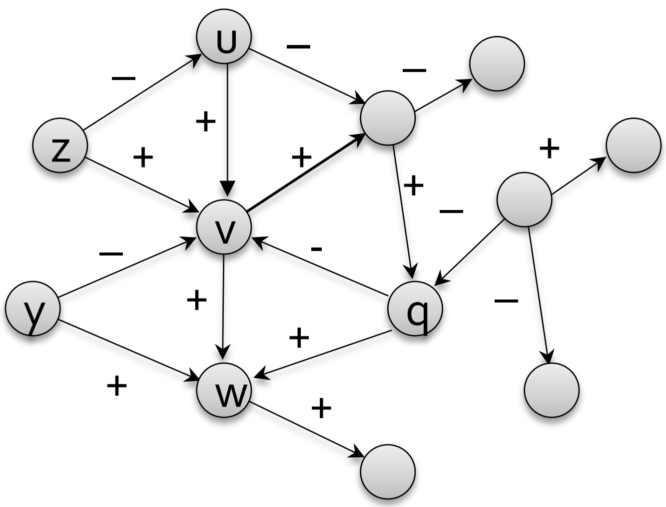
# Example: Computing Surprise

- **Surprise:** How much behavior of user **deviates** from **baseline** when in **context X**
  - **Generative surprise of context X=**




n = 3

$$s_g(X) = \frac{k - \sum_{i=1}^{n} p_g(A_i)}{\sqrt{\sum_{i=1}^{n} p_g(A_i)(1 - p_g(A_i))}}$$




**We have 3 triads of context X:** (z,u,v), (y,v,w), (q,v,w)

They all close with a plus: So *k*=3

$P_g(u)=1/2=0.5$ $P_g(v)=2/2=1$

$S_g(X)=(3-2.5)/\sqrt{(0.5*0.5+1*0+1*0)}$ **= 1**

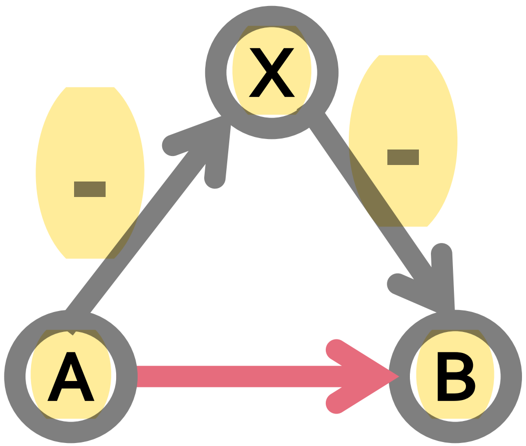
u v v

$P_g(u)$ $P_g(v)$ $P_g(v)$

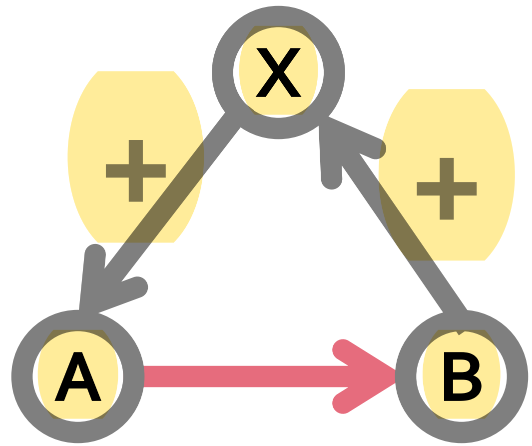
$P_g(v)$ : generating nod

# Status: Two Examples

- Assume status is at work
- What happens?



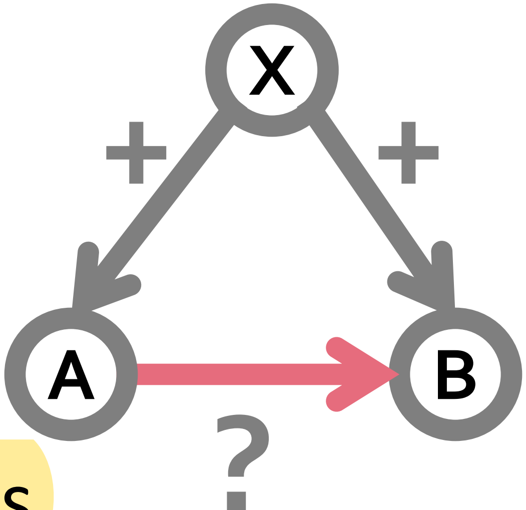

Gen. surprise of A: –
Rec. surprise of B: –




Gen. surprise of A: –
Rec. surprise of B: –

# Joint Positive Endorsement

- **X** positively endorses **A** and **B**
- Now **A** links to **B**




**A puzzle:**

- In our data we observe:
  Fraction of positive links deviates
  - **Above generative baseline of A: $S_g(X) > 0$**
  - **Below receptive baseline of B: $S_r(X) < 0$**

- **Why?**

UNC CHARLOTTE

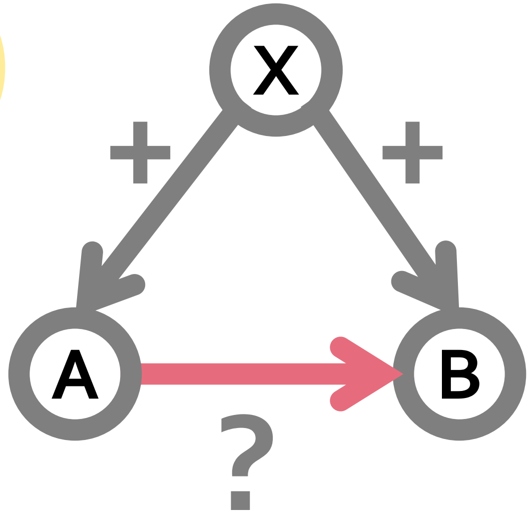
# A Story: Soccer Team

[CHI '10]

- Ask every node: How does skill of B compare to yours?
  - Build a signed directed network

- We haven't asked **A** about **B**
- But we know that **X** thinks **A** and **B** are both better than him

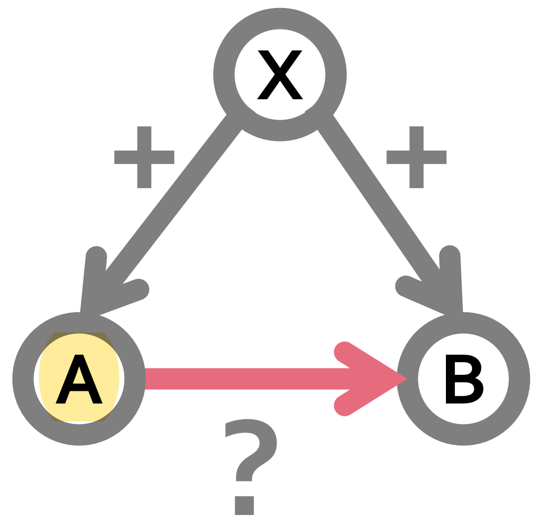
- What can we infer about A's answer?




UNC CHARLOTTE

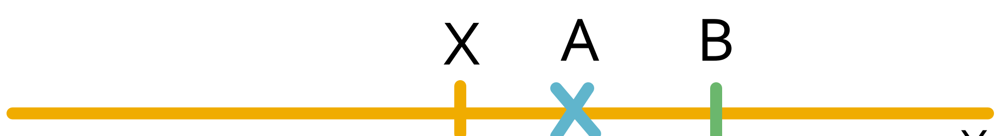
# A Story: Soccer Team

- **A’s viewpoint:**
  - Since **B** has positive evaluation, **B** is high status
  - Thus, evaluation **A** gives is **more likely to be positive** than the baseline




**How does A evaluate B?**

**A** is evaluating someone who is better than avg.
→ **A** is more positive than average

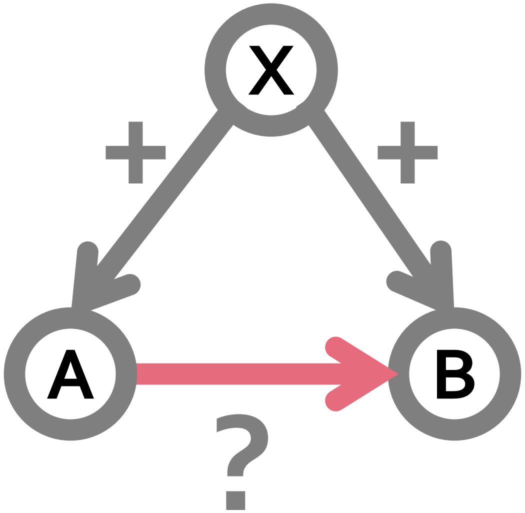



X… “average” node

UNC CHARLOTTE

# A Story: Soccer Team

- **B's viewpoint:**
  - Since **A** has positive evaluation, **A** is high status
  - Thus, evaluation **B** receives is **less likely to be positive** than the baseline




How is B evaluated by A?

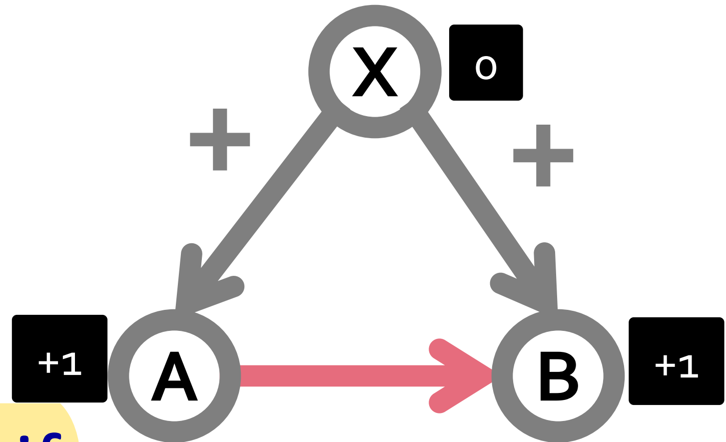
**B** is evaluated by someone better than average.

→ They will be **more negative to B than average**

**Surprise of A→B deviates in different directions depending on the viewpoint!**

# Consistency with Status

- **Determine node status:**
  - Assign ***X*** status 0
  - Based on signs and directions of edges set status of **A** and **B**
- Surprise is **status**-consistent, if:
  - Gen. surprise is status-consistent if it has **same** sign as status of **B**
  - Rec. surprise is status-consistent if it has the **opposite** sign from the status of **A**
- Surprise is **balance**-consistent, if:
  - If it completes a balanced triad




**Status-consistent if:**
Gen. surprise > 0
Rec. surprise < 0

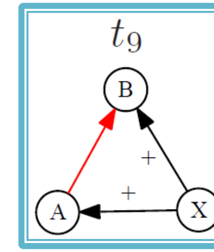
UNC CHARLOTTE

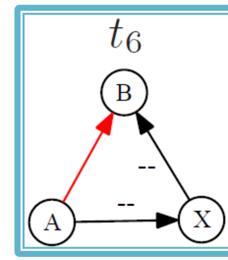
[CHI '10]

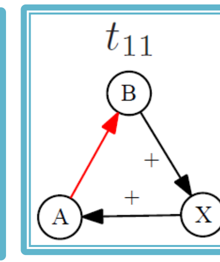
# Status vs. Balance (Epinions)

## Predictions by status and balance:

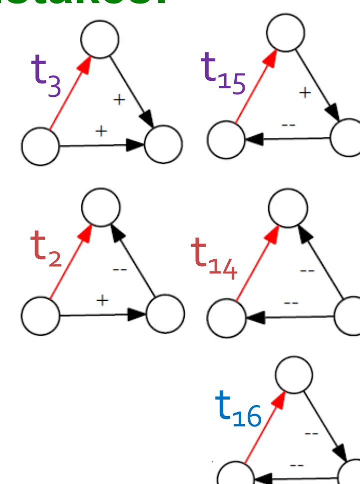
| $t_i$ | count | $P(+)$ | $S_g(t_i)$ | $S_r(t_i)$ | $B_g$ | $B_r$ | $S_g$ | $S_r$ |
|---|---|---|---|---|---|---|---|---|
| $t_1$ | 178,051 | 0.97 | 95.9 | 197.8 | ✓ | ✓ | ✓ | ✓ |
| $t_2$ | 45,797 | 0.54 | -151.3 | -229.9 | ✓ | ✓ | ✓ | ● |
| $t_3$ | 246,371 | 0.94 | 89.9 | 195.9 | ✓ | ✓ | ● | ✓ |
| $t_4$ | 25,384 | 0.89 | 1.8 | 44.9 | ○ | ○ | ✓ | ✓ |
| $t_5$ | 45,925 | 0.30 | 18.1 | -333.7 | ○ | ✓ | ✓ | ✓ |
| $t_6$ | 11,215 | 0.23 | -15.5 | -193.6 | ○ | ○ | ✓ | ✓ |
| $t_7$ | 36,184 | 0.14 | -53.1 | -357.3 | ✓ | ✓ | ✓ | ✓ |
| $t_8$ | 61,519 | 0.63 | 124.1 | -225.6 | ✓ | ○ | ✓ | ✓ |
| $t_9$ | 338,238 | 0.82 | 207.0 | -239.5 | ✓ | ○ | ✓ | ✓ |
| $t_{10}$ | 27,089 | 0.20 | -110.7 | -449.6 | ✓ | ✓ | ✓ | ✓ |
| $t_{11}$ | 35,093 | 0.53 | -7.4 | -260.1 | ○ | ○ | ✓ | ✓ |
| $t_{12}$ | 20,933 | 0.71 | 17.2 | -113.4 | ○ | ✓ | ✓ | ✓ |
| $t_{13}$ | 14,305 | 0.79 | 23.5 | 24.0 | ○ | ○ | ✓ | ✓ |
| $t_{14}$ | 30,235 | 0.69 | -12.8 | -53.6 | ○ | ○ | ✓ | ● |
| $t_{15}$ | 17,189 | 0.76 | 6.4 | 24.0 | ○ | ○ | ● | ✓ |
| $t_{16}$ | 4,133 | 0.77 | 11.9 | -2.6 | ✓ | ○ | ✓ | ● |
| Number of correct predictions | | | | | 8 | 7 | 14 | 13 |



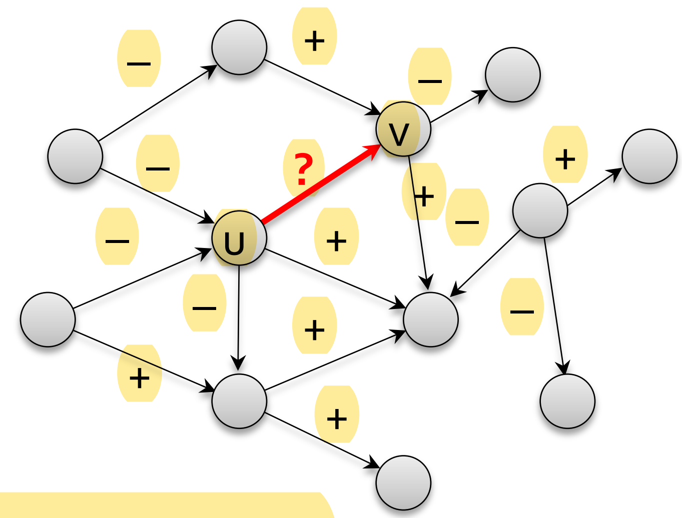

### Mistakes:

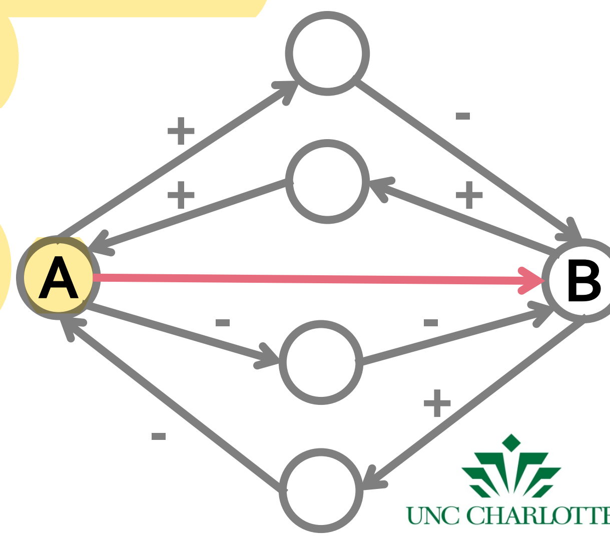


# Predicting Edge Signs

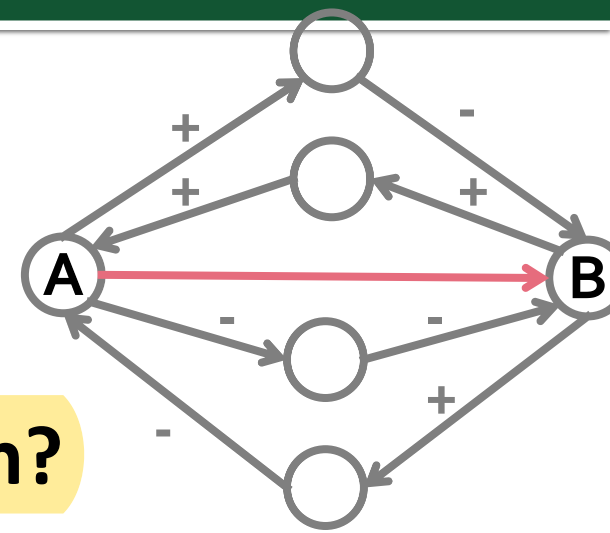
## Edge sign prediction problem

- Given a network and signs on all but one edge, predict the missing sign
- **Friend recommendation:**
  - Predicting whether you know someone vs. Predicting what you think of them
- **Setting:**
  - **Given edge *(A,B)*, predict its sign:**
  - **Let's look at signed triads *(A,B)* belongs to:**



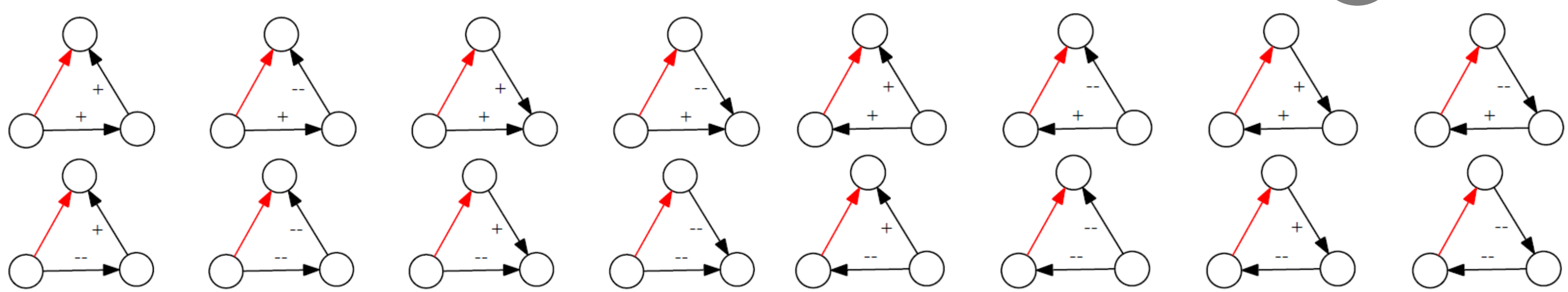



# Features for Learning




**For the edge (A,B) we examine Its network context:**

- **In what types of triads does our red-edge participate in?**



- Each triad then "votes" and we determine the sign

# Balance and Status: Complete Model

| Triad | Bal |
|---|---|
| const | |
| $\bullet \xrightarrow{+} \bullet \xrightarrow{+} \bullet$ | 1 |
| $\bullet \xrightarrow{+} \bullet \xrightarrow{-} \bullet$ | -1 |
| $\bullet \xrightarrow{-} \bullet \xrightarrow{+} \bullet$ | -1 |
| $\bullet \xrightarrow{-} \bullet \xrightarrow{-} \bullet$ | 1 |
| $\circ \xrightarrow{+} \circ \xleftarrow{+} \circ$ | 1 |
| $\circ \xrightarrow{+} \circ \xleftarrow{-} \circ$ | -1 |
| $\bullet \xrightarrow{-} \bullet \xleftarrow{+} \bullet$ | -1 |
| $\circ \xrightarrow{-} \circ \xleftarrow{-} \circ$ | 1 |
| $\circ \xleftarrow{+} \circ \xrightarrow{+} \circ$ | 1 |
| $\bullet \xleftarrow{+} \bullet \xrightarrow{-} \bullet$ | -1 |
| $\circ \xleftarrow{-} \circ \xrightarrow{+} \circ$ | -1 |
| $\circ \xleftarrow{-} \circ \xrightarrow{-} \circ$ | 1 |
| $\circ \xleftarrow{+} \circ \xleftarrow{+} \circ$ | 1 |
| $\circ \xleftarrow{+} \circ \xleftarrow{-} \circ$ | -1 |
| $\circ \xleftarrow{-} \circ \xleftarrow{+} \circ$ | -1 |
| $\circ \xleftarrow{-} \circ \xleftarrow{-} \circ$ | 1 |

# Balance and Status: Complete Model

| Triad | Bal | Stat |
|---|---|---|
| const | | |
| ● $\xrightarrow{+}$ ● $\xrightarrow{+}$ ● | 1 | 1 |
| ● $\xrightarrow{+}$ ● $\xrightarrow{-}$ ● | -1 | 0 |
| ● $\xrightarrow{-}$ ● $\xrightarrow{+}$ ● | -1 | 0 |
| ● $\xrightarrow{-}$ ● $\xrightarrow{-}$ ● | 1 | -1 |
| ○ $\xrightarrow{+}$ ○ $\xleftarrow{+}$ ○ | 1 | 0 |
| ○ $\xrightarrow{+}$ ○ $\xleftarrow{-}$ ○ | -1 | 1 |
| ● $\xrightarrow{-}$ ● $\xleftarrow{+}$ ● | -1 | -1 |
| ○ $\xrightarrow{-}$ ○ $\xleftarrow{-}$ ○ | 1 | 0 |
| ○ $\xleftarrow{+}$ ○ $\xrightarrow{+}$ ○ | 1 | 0 |
| ● $\xleftarrow{+}$ ● $\xrightarrow{-}$ ● | -1 | -1 |
| ○ $\xleftarrow{-}$ ○ $\xrightarrow{+}$ ○ | -1 | 1 |
| ○ $\xleftarrow{-}$ ○ $\xrightarrow{-}$ ○ | 1 | 0 |
| ○ $\xleftarrow{+}$ ○ $\xleftarrow{+}$ ○ | 1 | -1 |
| ○ $\xleftarrow{+}$ ○ $\xleftarrow{-}$ ○ | -1 | 0 |
| ○ $\xleftarrow{-}$ ○ $\xleftarrow{+}$ ○ | -1 | 0 |
| ○ $\xleftarrow{-}$ ○ $\xleftarrow{-}$ ○ | 1 | 1 |

# Balance and Status: Complete Model

| Triad | Bal | Stat | **Epin** | **Slashd** | **Wikip** |
|---|---|---|---|---|---|
| const | | | -0.2 | 0.02 | -0.2 |
| ● $\xrightarrow{+}$ ● $\xrightarrow{+}$ ● | 1 | 1 | **0.5** | **0.9** | **0.3** |
| ● $\xrightarrow{+}$ ● $\xrightarrow{-}$ ● | -1 | 0 | **-0.5** | **-0.9** | **-0.4** |
| ● $\xrightarrow{-}$ ● $\xrightarrow{+}$ ● | -1 | 0 | **-0.4** | **-1.1** | **-0.3** |
| ● $\xrightarrow{-}$ ● $\xrightarrow{-}$ ● | 1 | -1 | **-0.7** | **-0.6** | **-0.8** |
| ○ $\xrightarrow{+}$ ○ $\xleftarrow{+}$ ○ | 1 | 0 | 0.3 | 0.4 | 0.05 |
| ○ $\xrightarrow{+}$ ○ $\xleftarrow{-}$ ○ | -1 | 1 | -0.01 | -0.1 | -0.01 |
| ● $\xrightarrow{-}$ ● $\xleftarrow{+}$ ● | -1 | -1 | **-0.9** | **-1.2** | **-0.2** |
| ○ $\xrightarrow{-}$ ○ $\xleftarrow{-}$ ○ | 1 | 0 | 0.04 | -0.07 | -0.03 |
| ○ $\xleftarrow{+}$ ○ $\xrightarrow{+}$ ○ | 1 | 0 | 0.08 | 0.4 | 0.1 |
| ● $\xleftarrow{+}$ ● $\xrightarrow{-}$ ● | -1 | -1 | **-1.3** | **-1.1** | **-0.4** |
| ○ $\xleftarrow{-}$ ○ $\xrightarrow{+}$ ○ | -1 | 1 | -0.1 | -0.2 | 0.05 |
| ○ $\xleftarrow{-}$ ○ $\xrightarrow{-}$ ○ | 1 | 0 | 0.08 | -0.02 | -0.1 |
| ○ $\xleftarrow{+}$ ○ $\xleftarrow{+}$ ○ | 1 | -1 | -0.09 | -0.09 | -0.01 |
| ○ $\xleftarrow{+}$ ○ $\xleftarrow{-}$ ○ | -1 | 0 | -0.05 | -0.3 | -0.02 |
| ○ $\xleftarrow{-}$ ○ $\xleftarrow{+}$ ○ | -1 | 0 | -0.04 | -0.3 | 0.05 |
| ○ $\xleftarrow{-}$ ○ $\xleftarrow{-}$ ○ | 1 | 1 | -0.02 | 0.2 | -0.2 |

# Edge Sign Prediction

- **Prediction accuracy:**

| | Balance | Status | Triads |
|---|---|---|---|
| Epinions | 80% | 82% | 93.5% |
| Slashdot | 84% | 72% | 94.4% |
| Wikipedia | 64% | 70% | 81% |

- **Observations:**
  - **Signs can be modeled from local network structure alone!**
  - **Triad counts perform less well for less embedded edges ($E_m$)**
    - Wikipedia is harder to model:
      - Votes are publicly visible

# Generalization

- Do people use these very different linking systems by obeying the same principles?
  - How generalizable are the results across the datasets?

| Train on row, test on column | Epinions | Slashdot | Wikipedia |
|---|---|---|---|
| Epinions | 0.9342 | 0.9289 | 0.7722 |
| Slashdot | 0.9249 | 0.9351 | 0.7717 |
| Wikipedia | 0.9272 | 0.9260 | 0.8021 |

- Nearly **perfect generalization** of the models even though networks come from very different applications!

# Summary: Signed Networks

- Signed networks provide insight into how social computing systems are used:
  - Status vs. Balance
  - Role of embeddedness and public display
  - More evidence that networks are globally organized based on status
- Sign of relationship can be reliably predicted from the local network context
  - ~90% accuracy sign of the edge
  - People use signed edges consistently regardless of particular application
    - Near perfect generalization of models across datasets