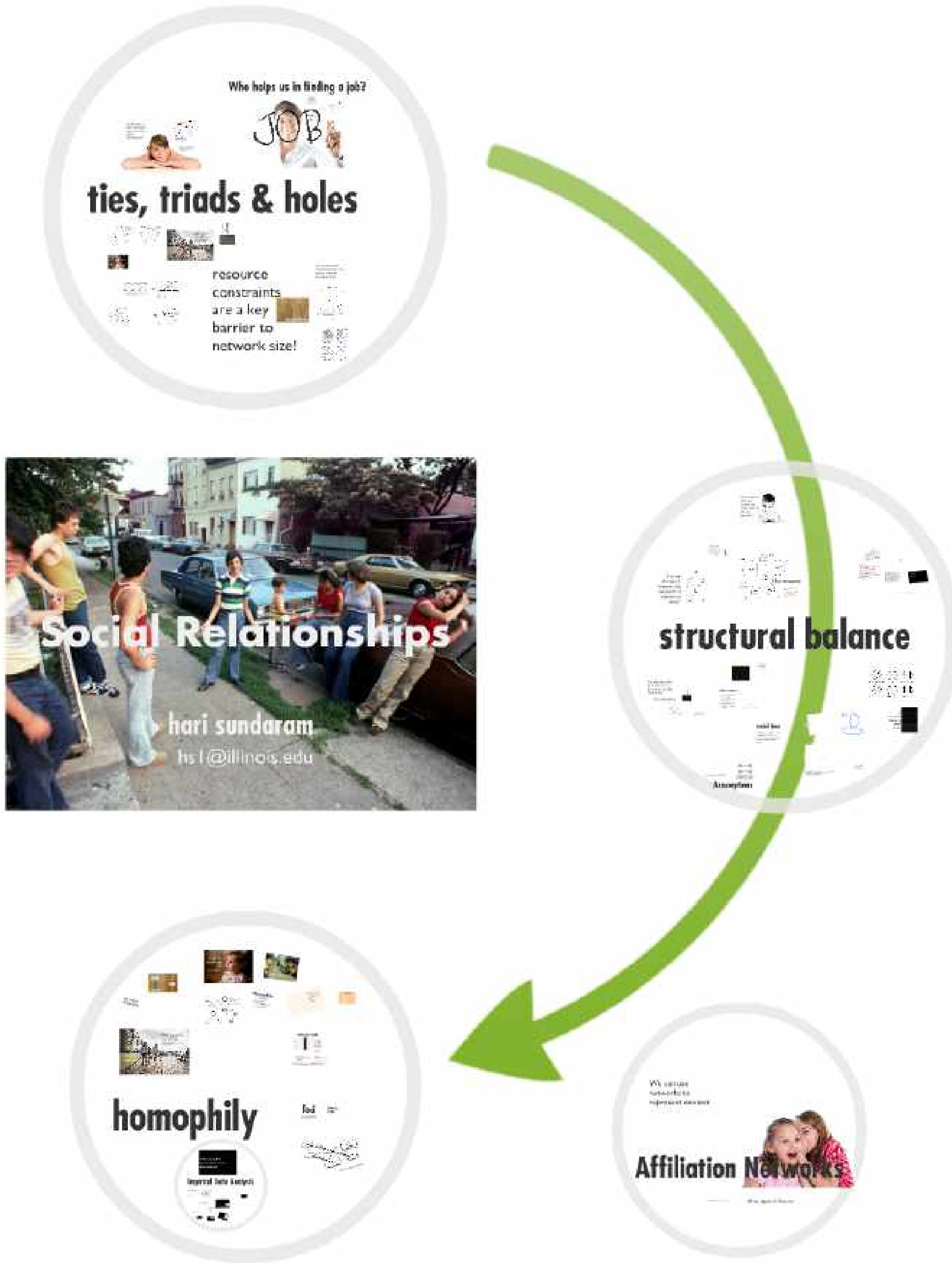
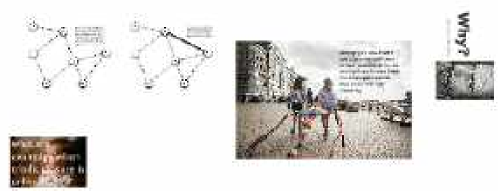

Who helps us in finding a job?
JOB
ties, triads & holes
resource constraints are a key barrier to network size!
Social Relationships
hari sundaram
hs1@illinois.edu
structural balance
homophily
Affiliation Networks

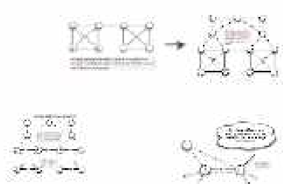

Social Relationships
hari sundaram
hs1@illinois.edu


Who helps us in finding a job?
JOB
ties, triads & holes
resource constraints are a key barrier to network size!
Social Relationships
hari sundaram
hs1@illinois.edu
structural balance
homophily
Affiliation Networks

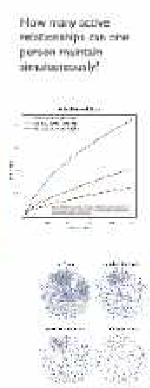
Who helps us in finding a job?

ties, triads & holes





resource constraints are a key barrier to network size!



Who helps us in finding a job?

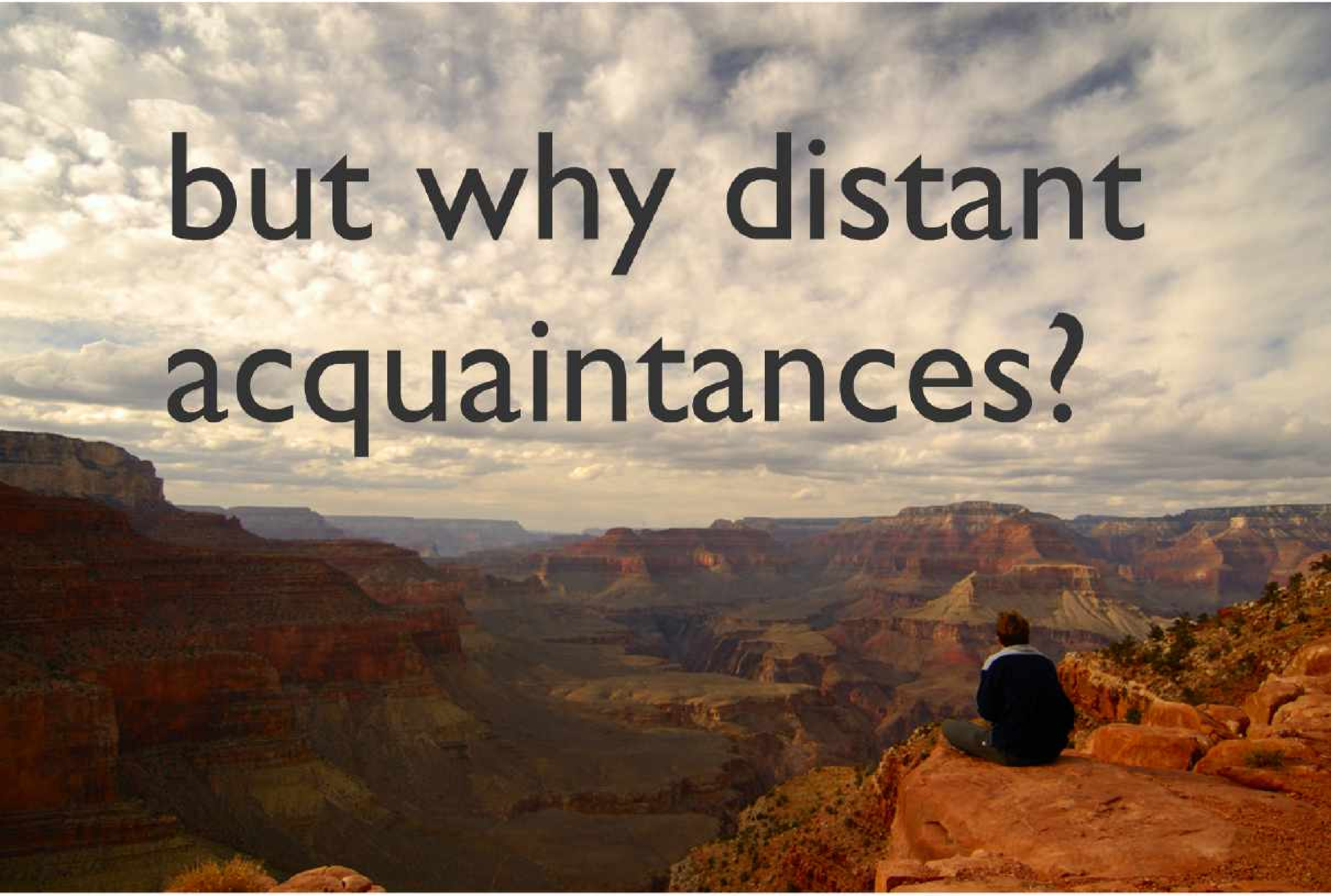
Granovetter made an interesting observation

M. S. Granovetter. The strength of weak ties. American journal of sociology, pages 1360–1380, 1973.

1 distant acquaintances were more helpful

2

tie strength and network structure were at play


but why distant
acquaintances?

THE STRENGTH OF TIES

Most intuitive notions of the "strength" of an interpersonal tie should be satisfied by the following definition: the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie.[2] Each of these is somewhat independent of the other, though the set is obviously highly intracorrelated. Discussion of operational measures of and weights attaching to each of the four elements is postponed to future empirical studies.[3] It is sufficient for the present purpose if most of us can agree, on a rough intuitive basis, whether a given tie is strong, weak, or absent.[4]

M. S. Granovetter. The strength of weak ties. American journal of sociology, pages 1360–1380, 1973.

At what point is a tie weak?

The existing literature suggests seven dimensions of tie strength: Intensity, Intimacy, Duration, Reciprocal Services, Structural, Emotional Support and Social Distance. As manifested in social media, can these dimensions predict tie strength? In what combination?

E. Gilbert and K. Karahalios. Predicting tie strength with social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 211–220. ACM, 2009.

With these dimensions as a guide, they identified 74 Facebook variables as potential predictors of tie strength.

n=35

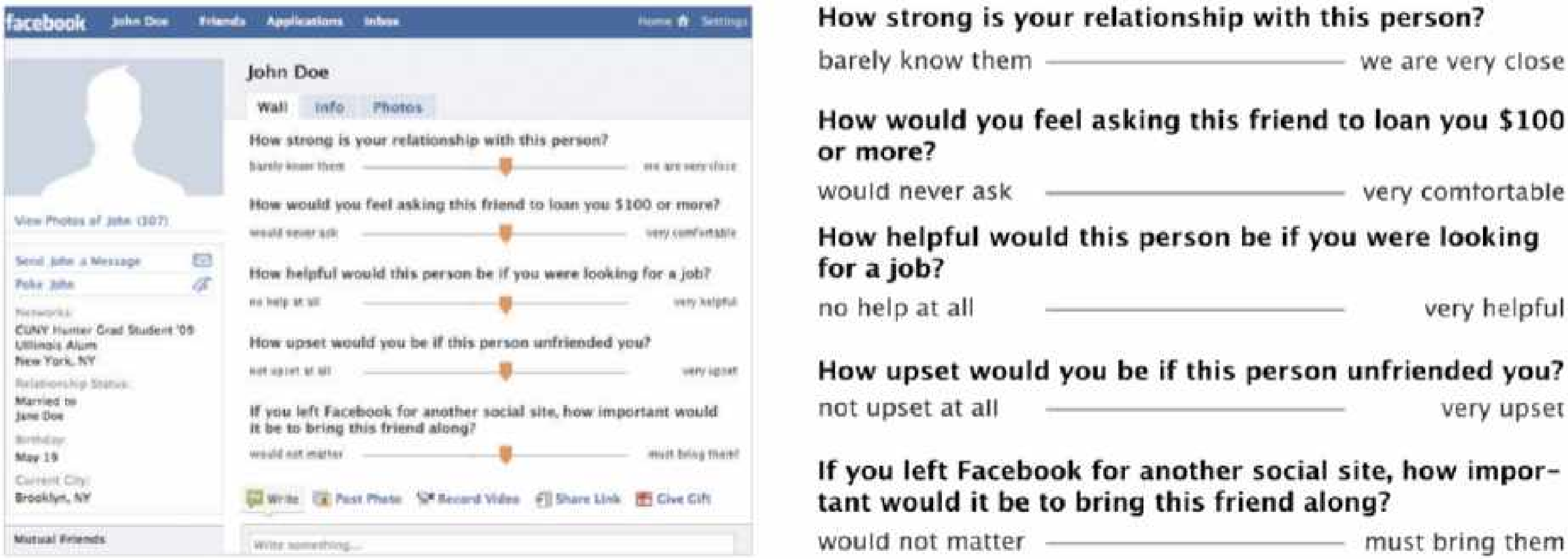


Figure 1. The questions used to assess tie strength, embedded into a friend's profile as participants experienced them. An automated script guided participants through a random subset of their Facebook friends. As participants answered each question by dragging a slider, the script collected data describing the friendship. The questions reflect a diversity of views on tie strength.

E. Gilbert and K. Karahalios. Predicting tie strength with social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 211–220. ACM, 2009.

Used a linear combination of the predictive variables, interaction effects as well as network structure

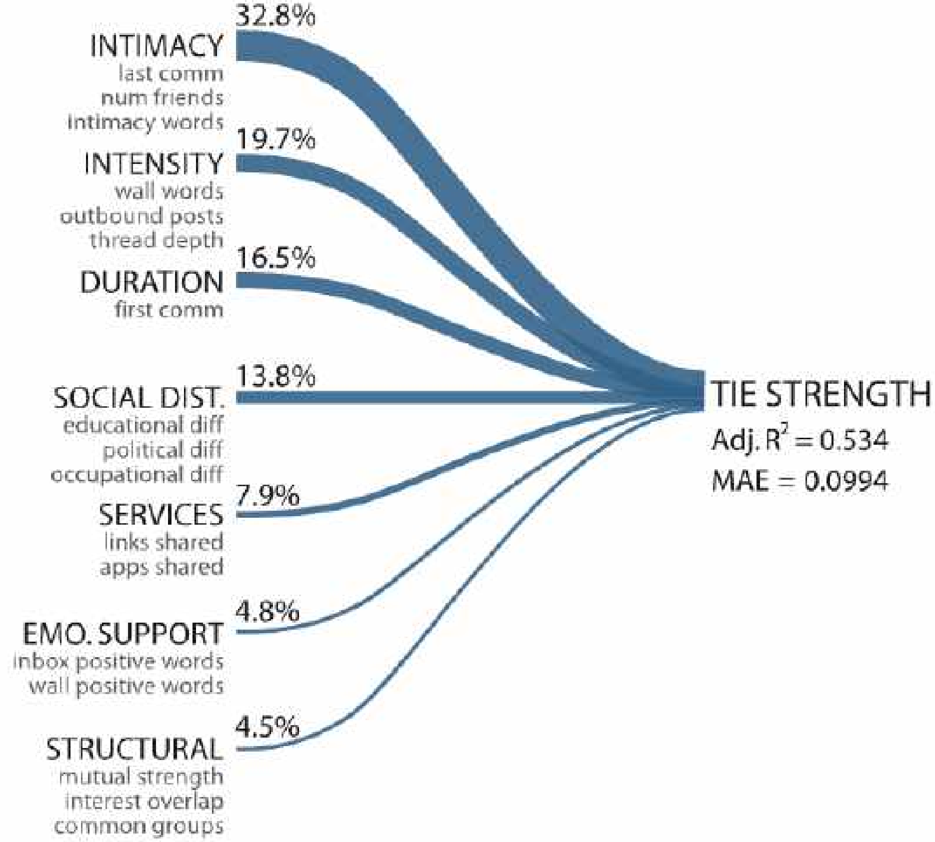


Figure 3. The predictive power of the seven tie strength dimensions, presented here as part of the *How strong?* model. A dimension's weight is computed by summing the absolute values of the coefficients belonging to it. The diagram also lists the top three predictive variables for each dimension. On average, the model predicts tie strength within one-tenth of its true value on a continuous 0–1 scale.

cause predictions are capped between 0 and 1. In addition, we found strong evidence of four dimension interactions (p < 0.001): *Intimacy* × *Structural*, $F_{1,911}$ = 12.37; *Social Distance* × *Structural*, $F_{1,911}$ = 34; *Reciprocal Services* × *Reciprocal Services*, $F_{1,911}$ = 14.4; *Structural* × *Structural*, $F_{1,911}$ = 12.41. As we demonstrate shortly, the *Structural* dimension plays a minor role as a linear factor. However, it has an important modulating role via these interactions. One way to read this result is that individual relationships matter, but they get filtered through a friend's clique before impacting tie strength.

cause predictions are capped between 0 and 1. In addition, we found strong evidence of four dimension interactions ($p < 0.001$): *Intimacy* × *Structural*, $F_{1,971} = 12.37$; *Social Distance* × *Structural*, $F_{1,971} = 34$; *Reciprocal Services* × *Reciprocal Services*, $F_{1,971} = 14.4$; *Structural* × *Structural*, $F_{1,971} = 12.41$. As we demonstrate shortly, the *Structural* dimension plays a minor role as a linear factor. However, it has an important modulating role via these interactions. One way to read this result is that individual relationships matter, but they get filtered through a friend's clique before impacting tie strength.

asymmetric friendships

e.g. ex-girlfriend / ex-boyfriend

Failure cases:

use channels other than Facebook to stay in touch

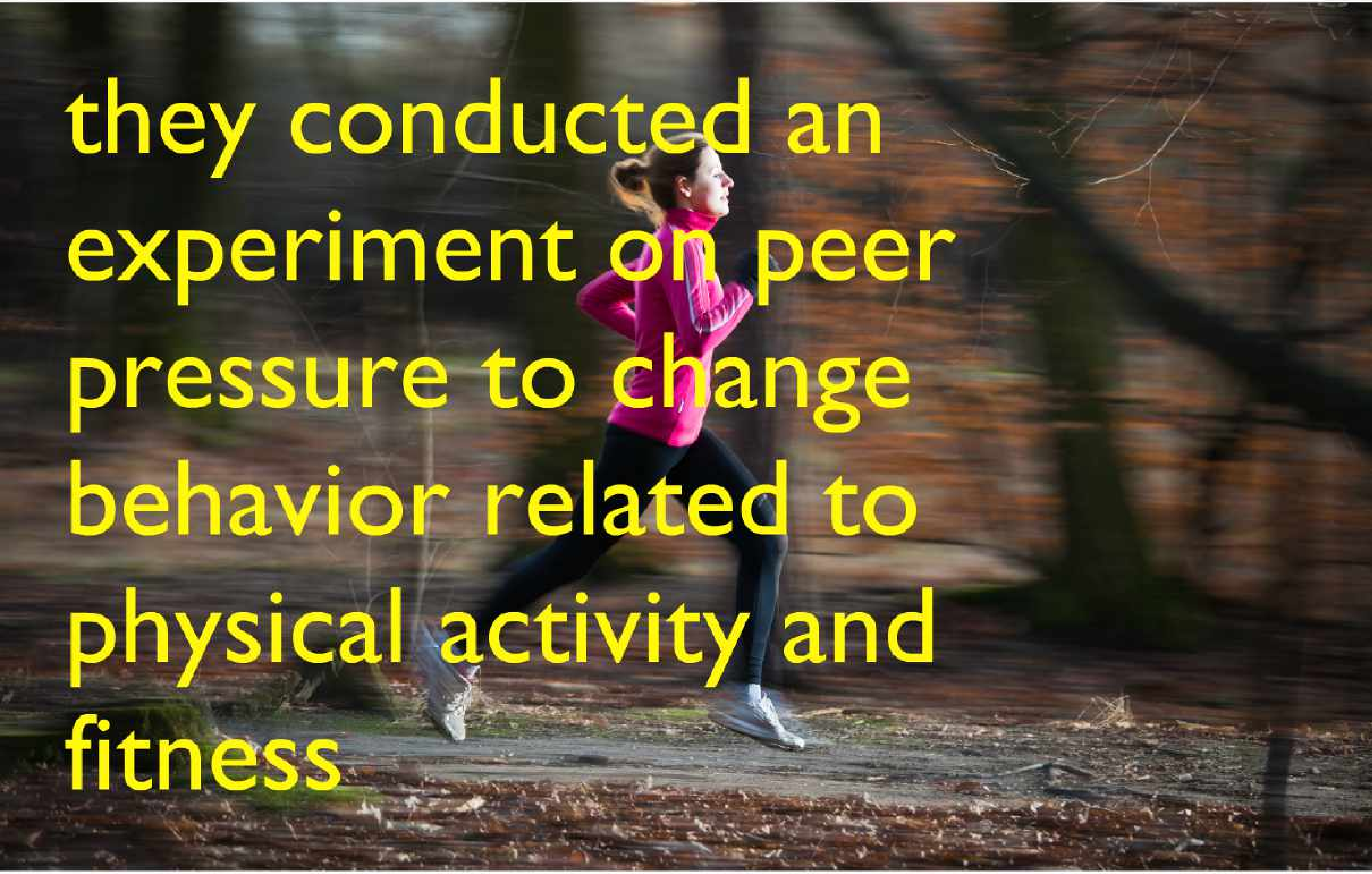
educational differences

e.g. professor / student

Tie Asymmetry and Persuasion

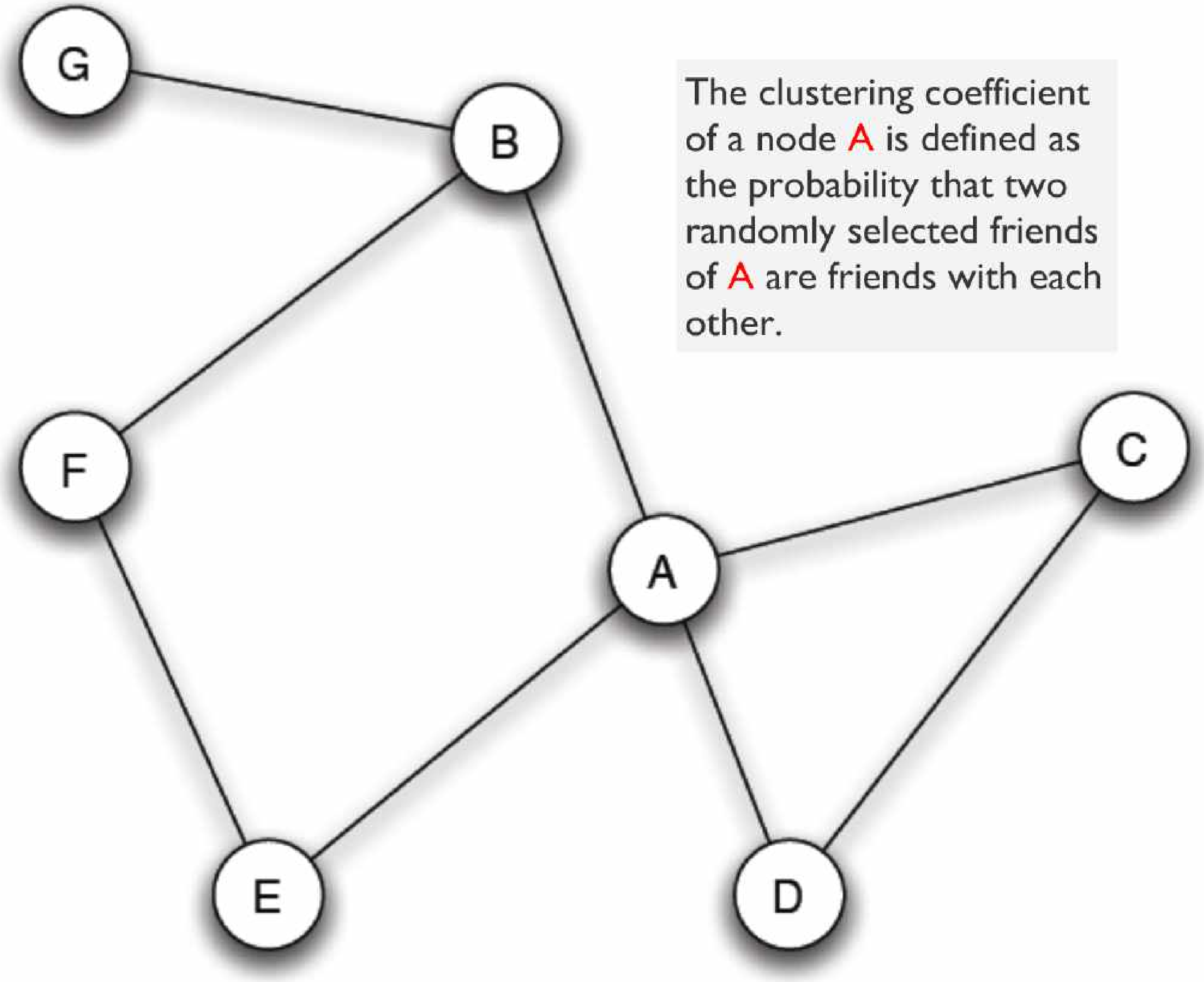
A. Almaatouq, L. Radaelli, A. Pentland, and E. Shmueli. Are you your friends' friend? poor perception of friendship ties limits the ability to promote behavioral change. PLoS ONE, 11(3):1–13, 03 2016.

"Moreover, we show that the effect of directionality is larger than the effect of the self-reported strength of a friendship tie [16] and thus of the implied 'social capital' of a relationship."

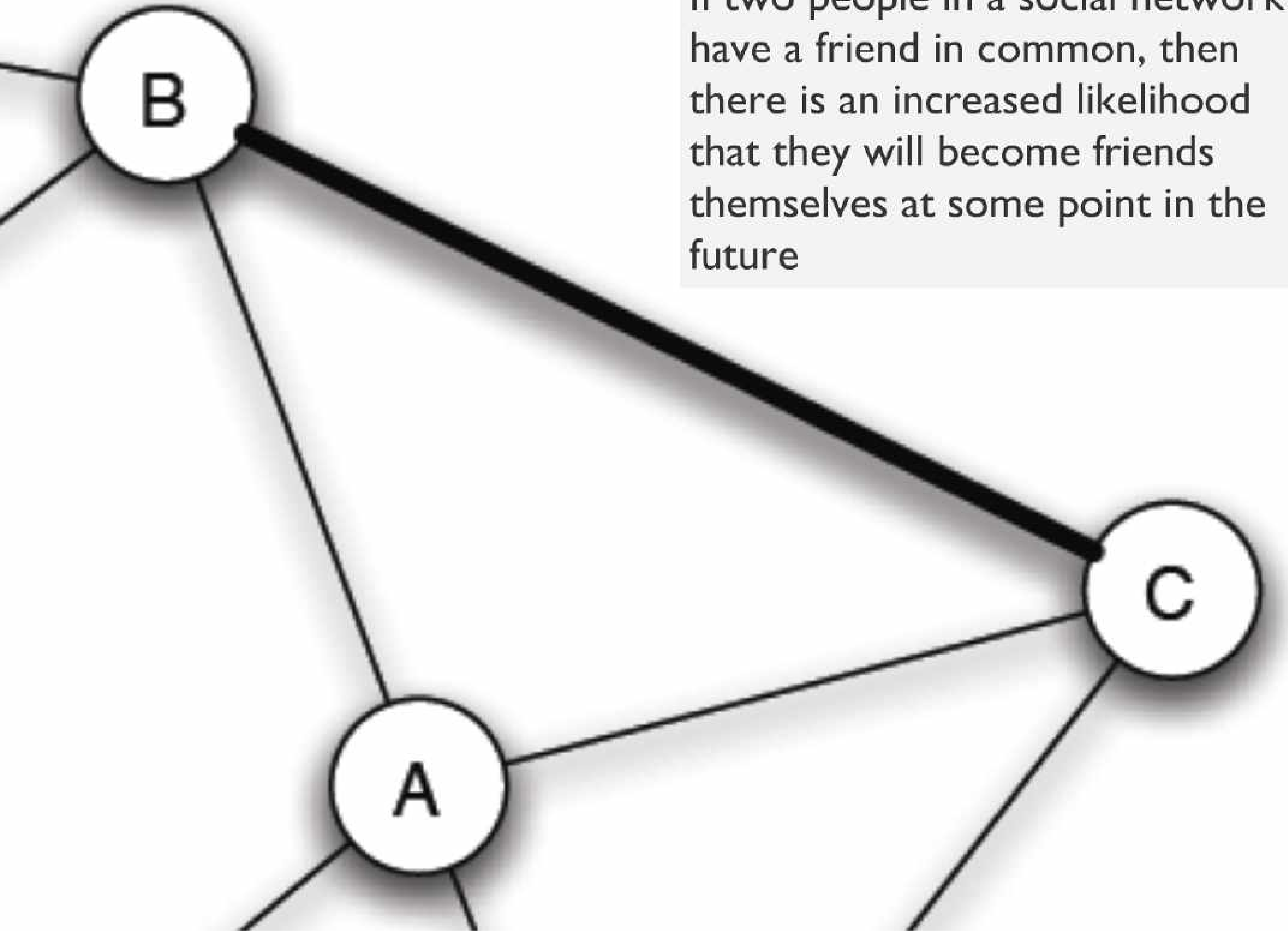


they conducted an experiment on peer pressure to change behavior related to physical activity and fitness

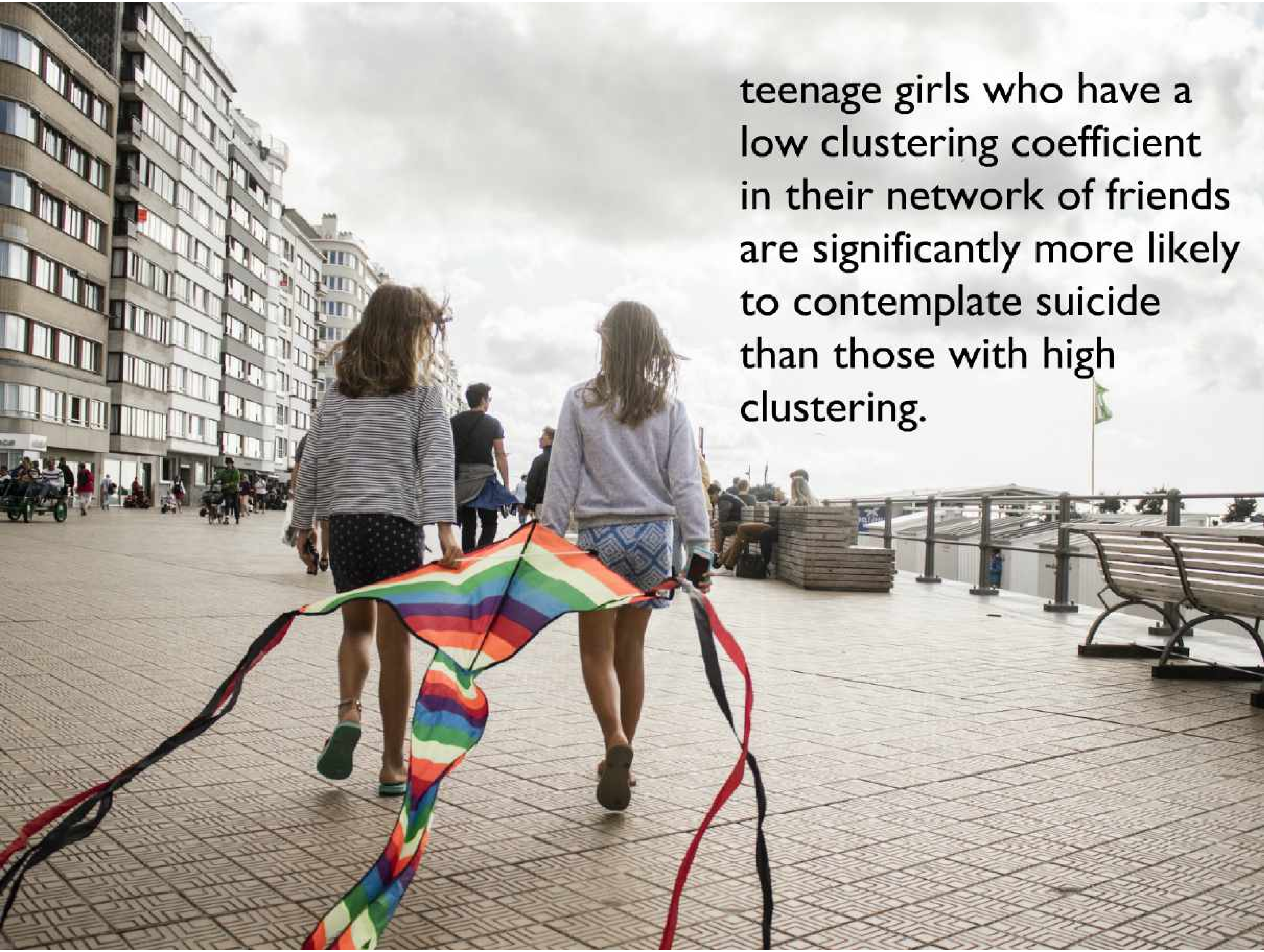
"In this experiment, we find that effective behavioral change occurs when subjects share reciprocal ties, or when a unilateral friendship tie exists from the person applying the peer pressure to the subject receiving the pressure, but not when the friendship tie is from the subject to the person applying peer pressure."

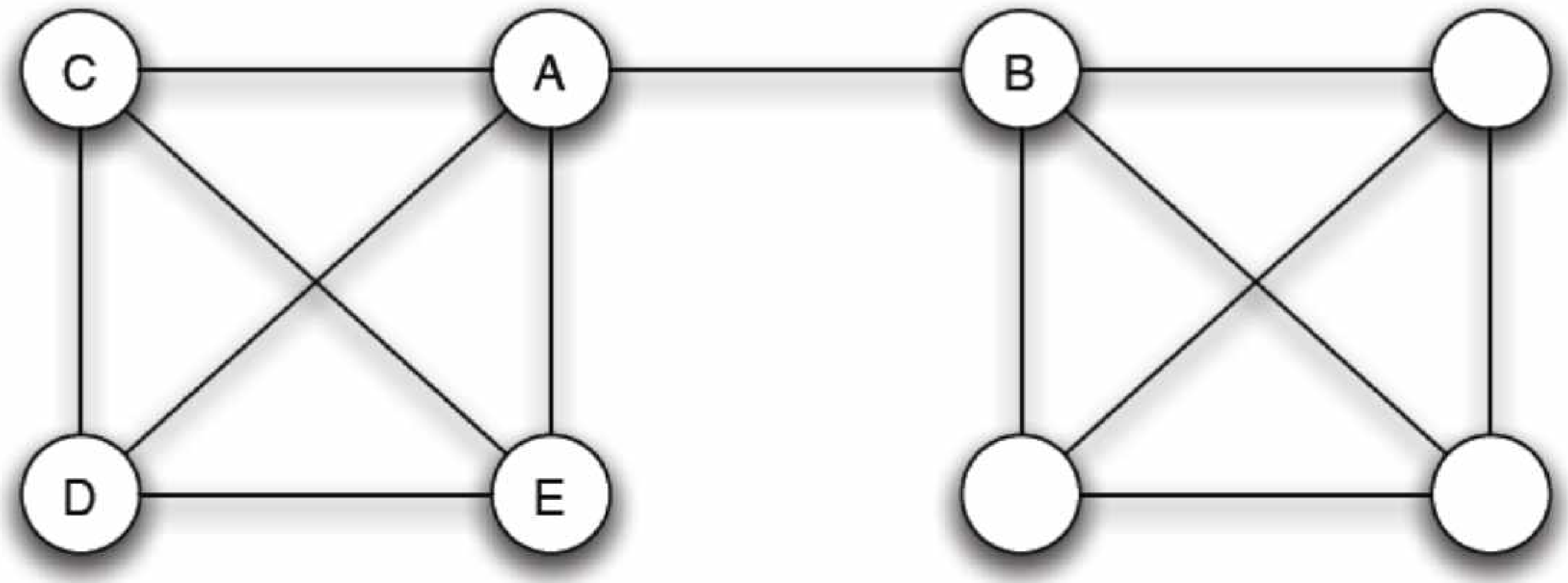


The clustering coefficient of a node A is defined as the probability that two randomly selected friends of A are friends with each other.



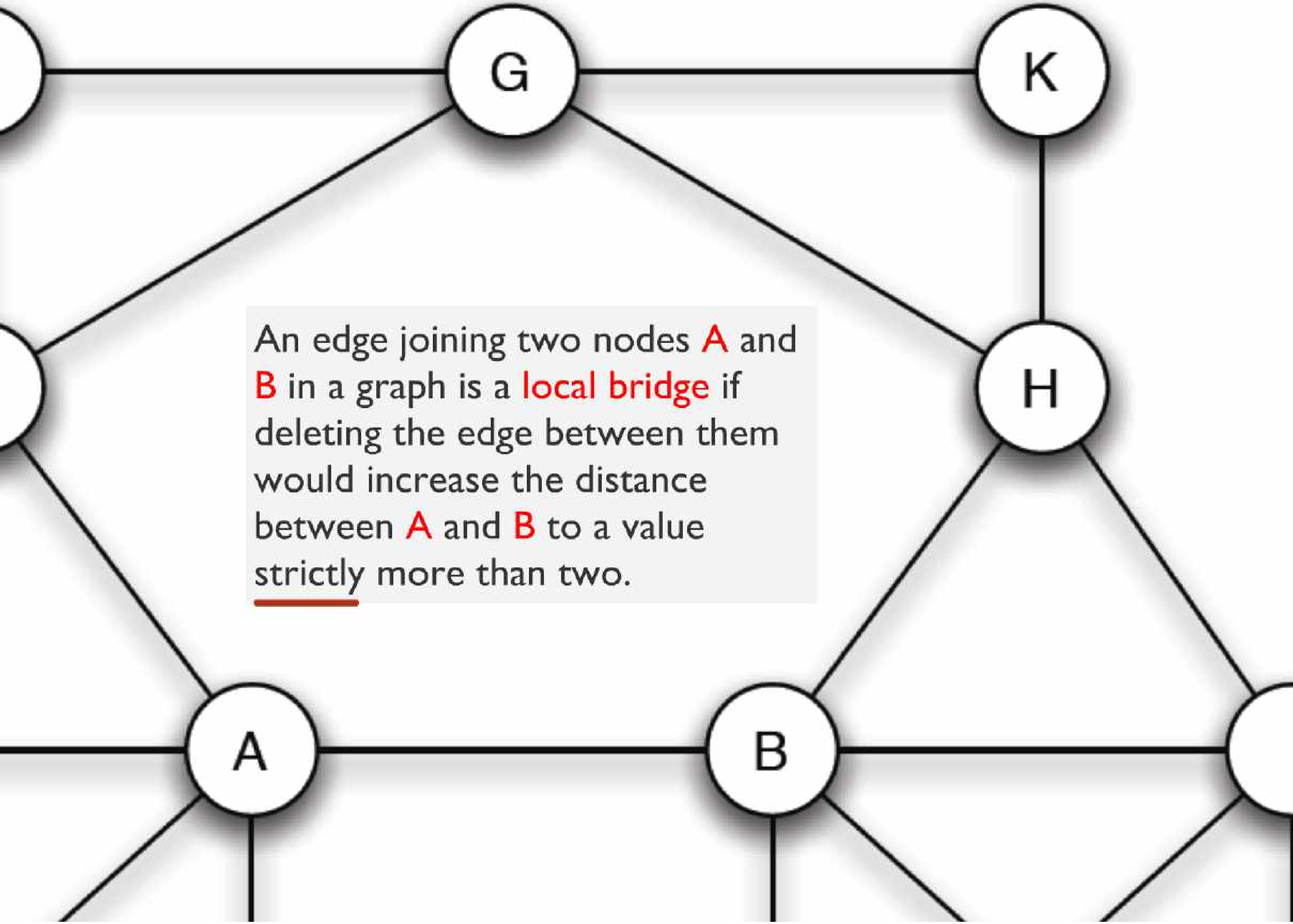
If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future


teenage girls who have a low clustering coefficient in their network of friends are significantly more likely to contemplate suicide than those with high clustering.



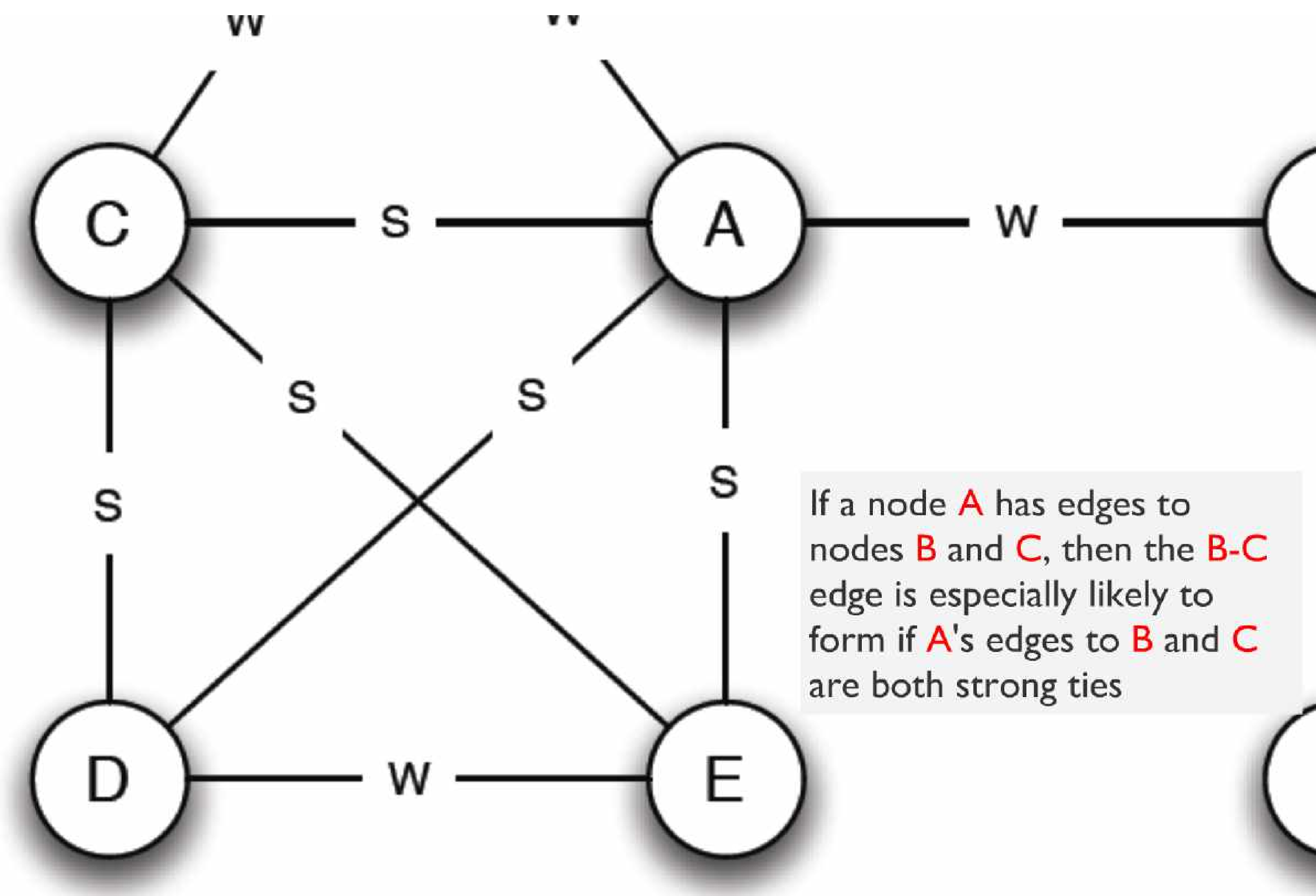
An edge joining two nodes A and B in a graph is a **bridge** if deleting the edge would cause A and B to lie in two different components

An edge joining two nodes A and B in a graph is a local bridge if deleting the edge between them would increase the distance between A and B to a value strictly more than two.

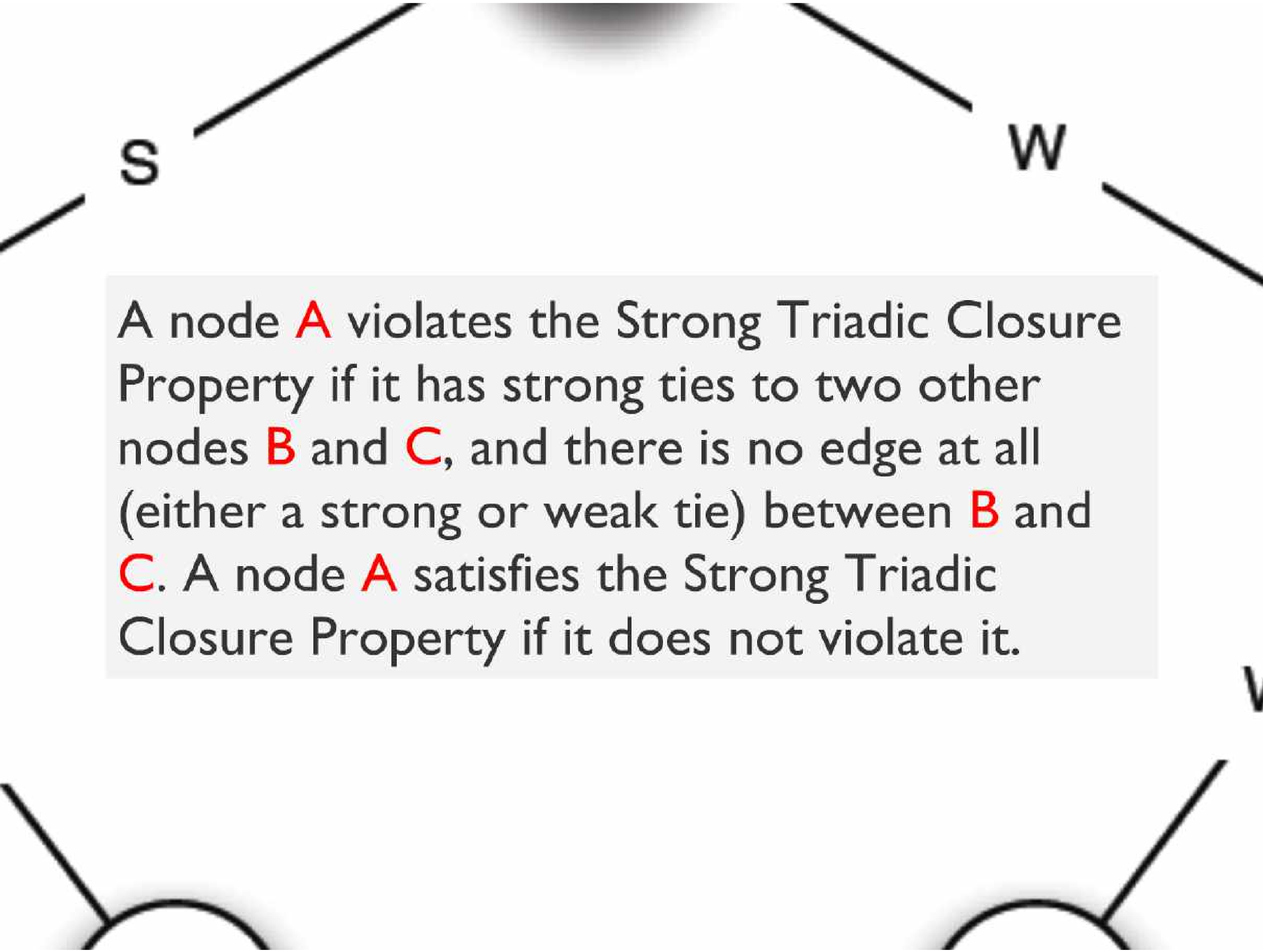


An edge joining two nodes A and B in a graph is a **local bridge** if deleting the edge between them would increase the distance between A and B to a value strictly more than two.

If a node A has edges to nodes B and C, then the B-C edge is especially likely to form if A's edges to B and C are both strong ties

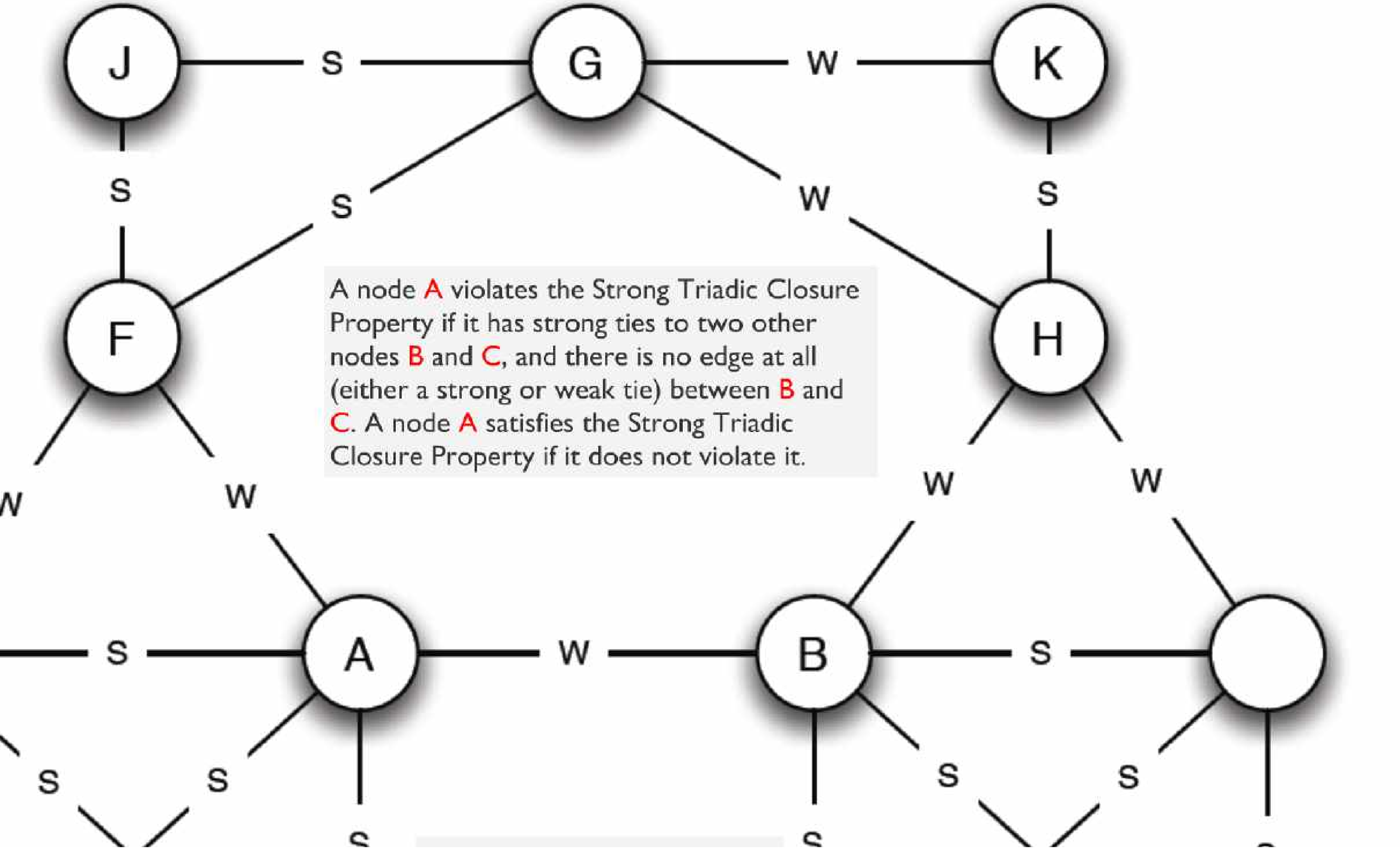


If a node A has edges to nodes B and C, then the B-C edge is especially likely to form if A's edges to B and C are both strong ties



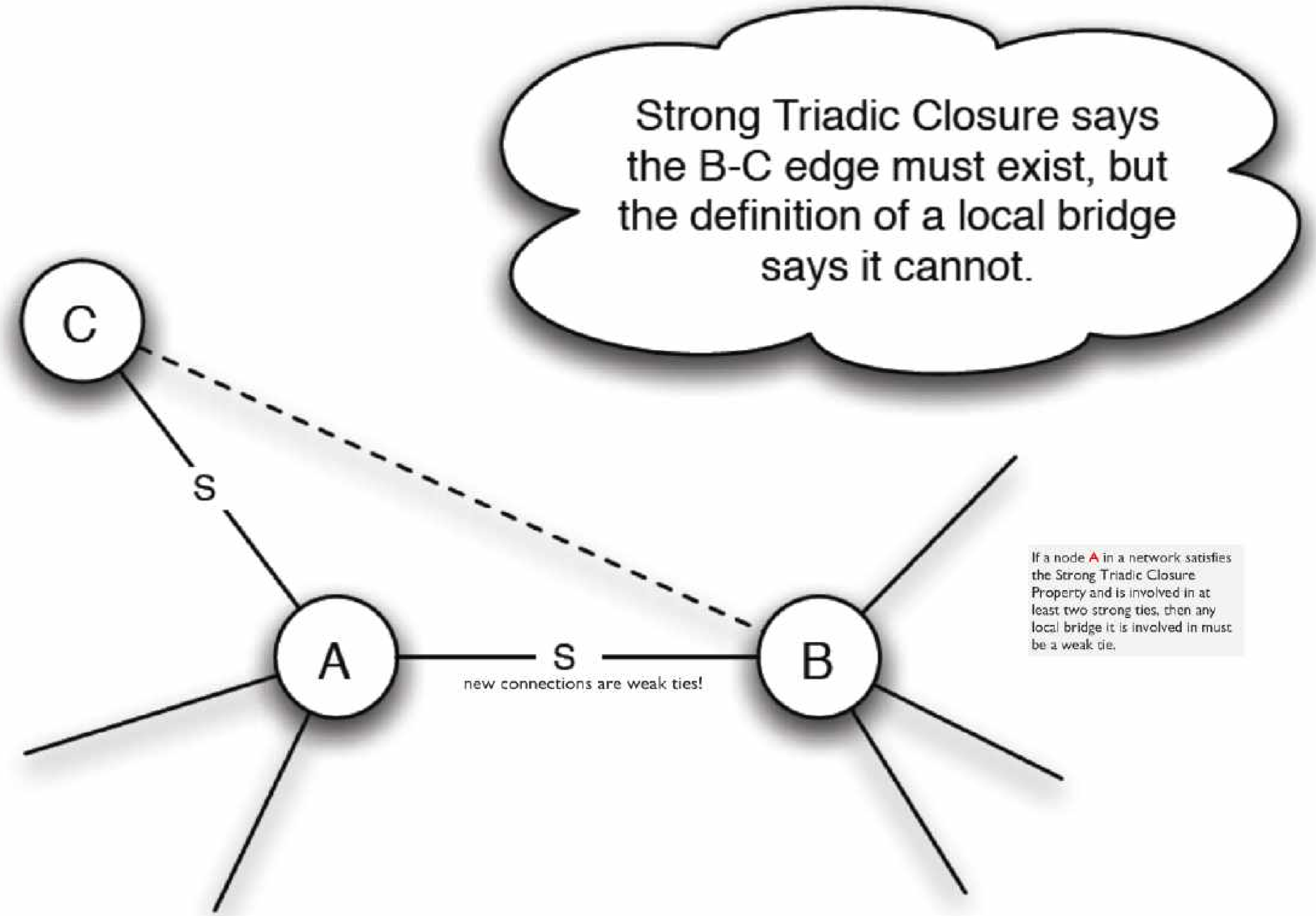
A node A violates the Strong Triadic Closure Property if it has strong ties to two other nodes B and C, and there is no edge at all (either a strong or weak tie) between B and C. A node A satisfies the Strong Triadic Closure Property if it does not violate it.

strong triadic closure property

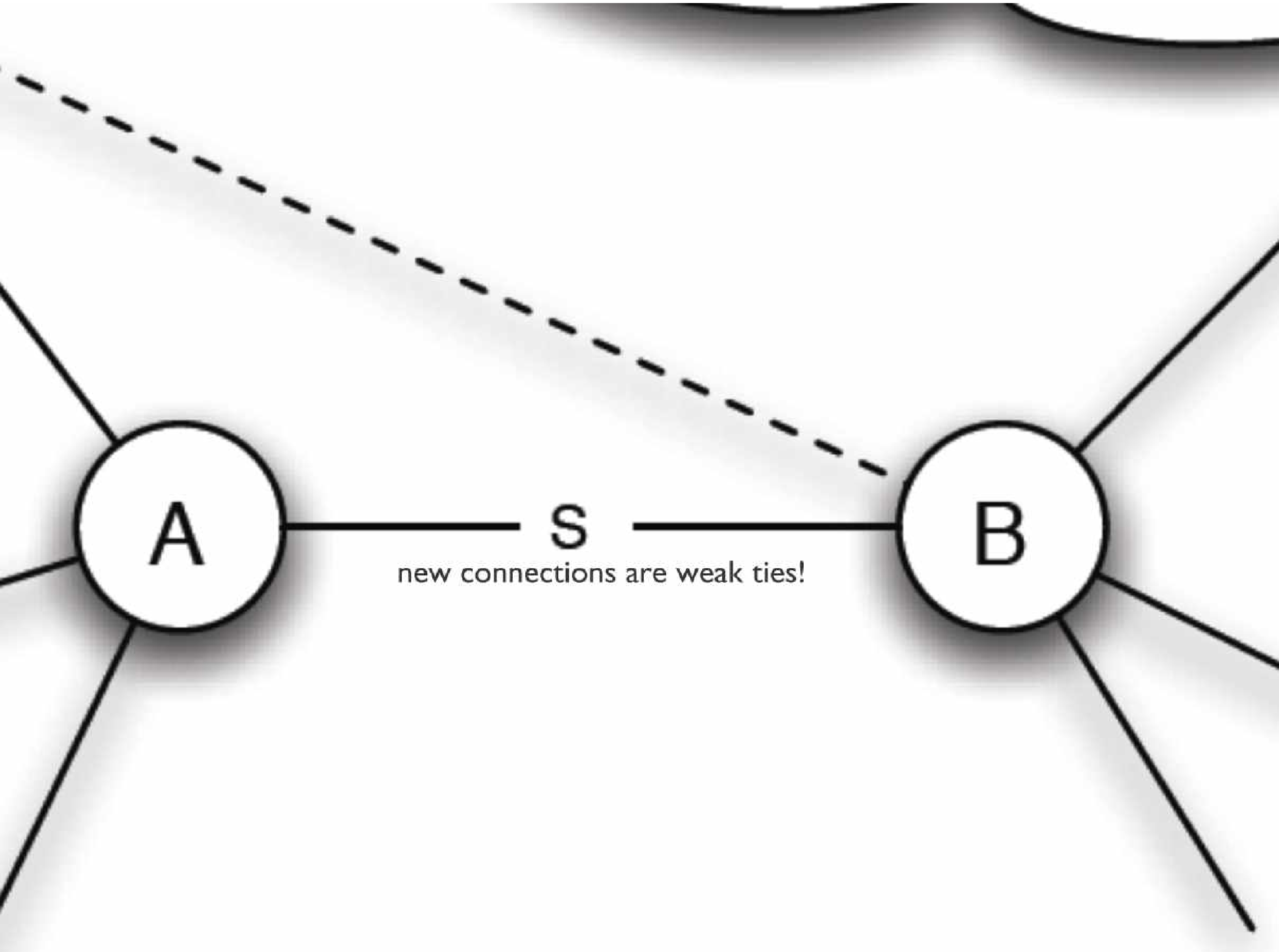


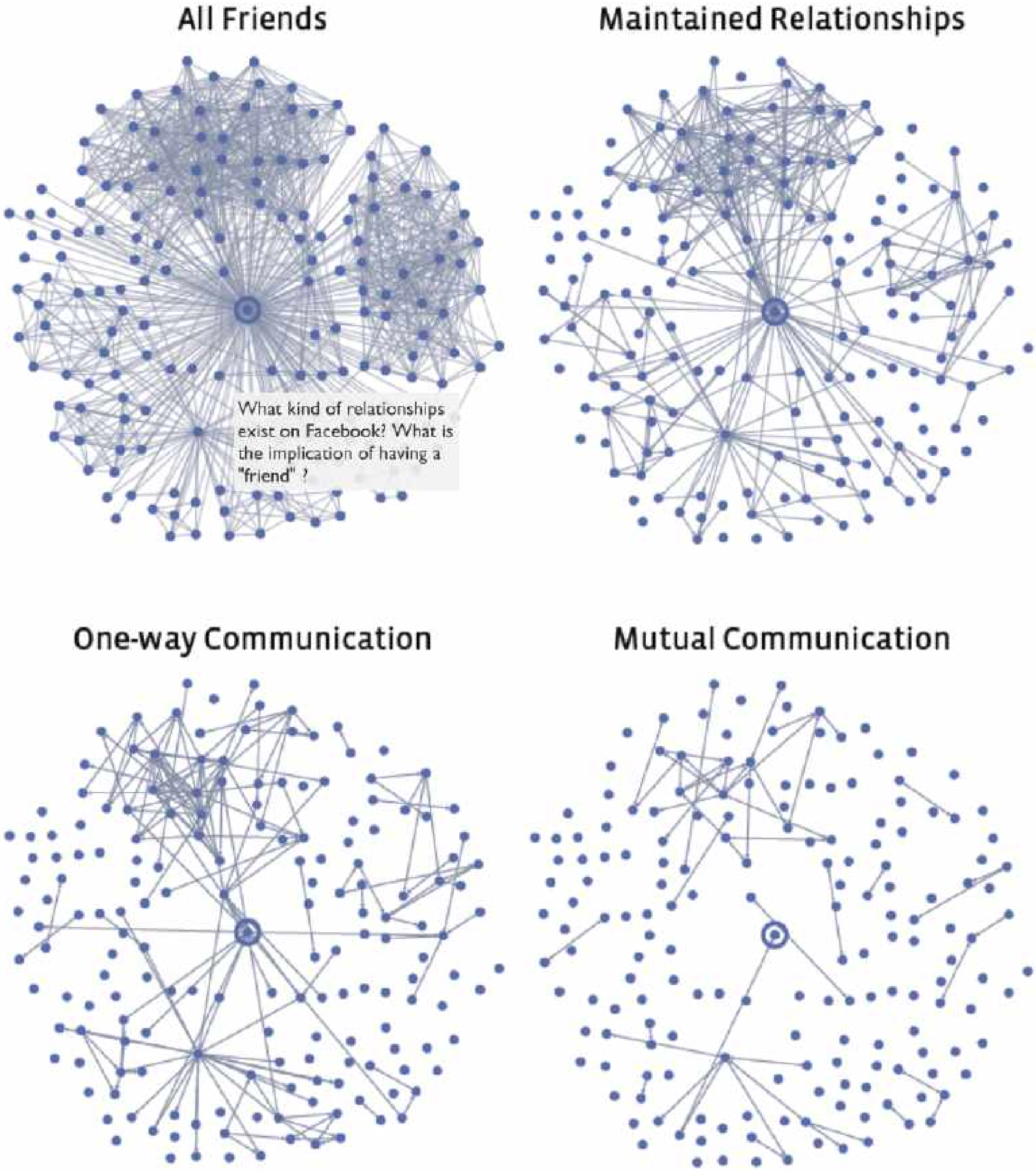
A node A violates the Strong Triadic Closure Property if it has strong ties to two other nodes B and C, and there is no edge at all (either a strong or weak tie) between B and C. A node A satisfies the Strong Triadic Closure Property if it does not violate it.

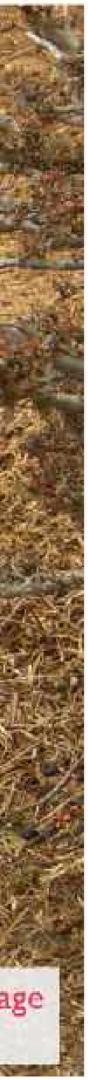
If a node A in a network satisfies the Strong Triadic Closure Property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie.

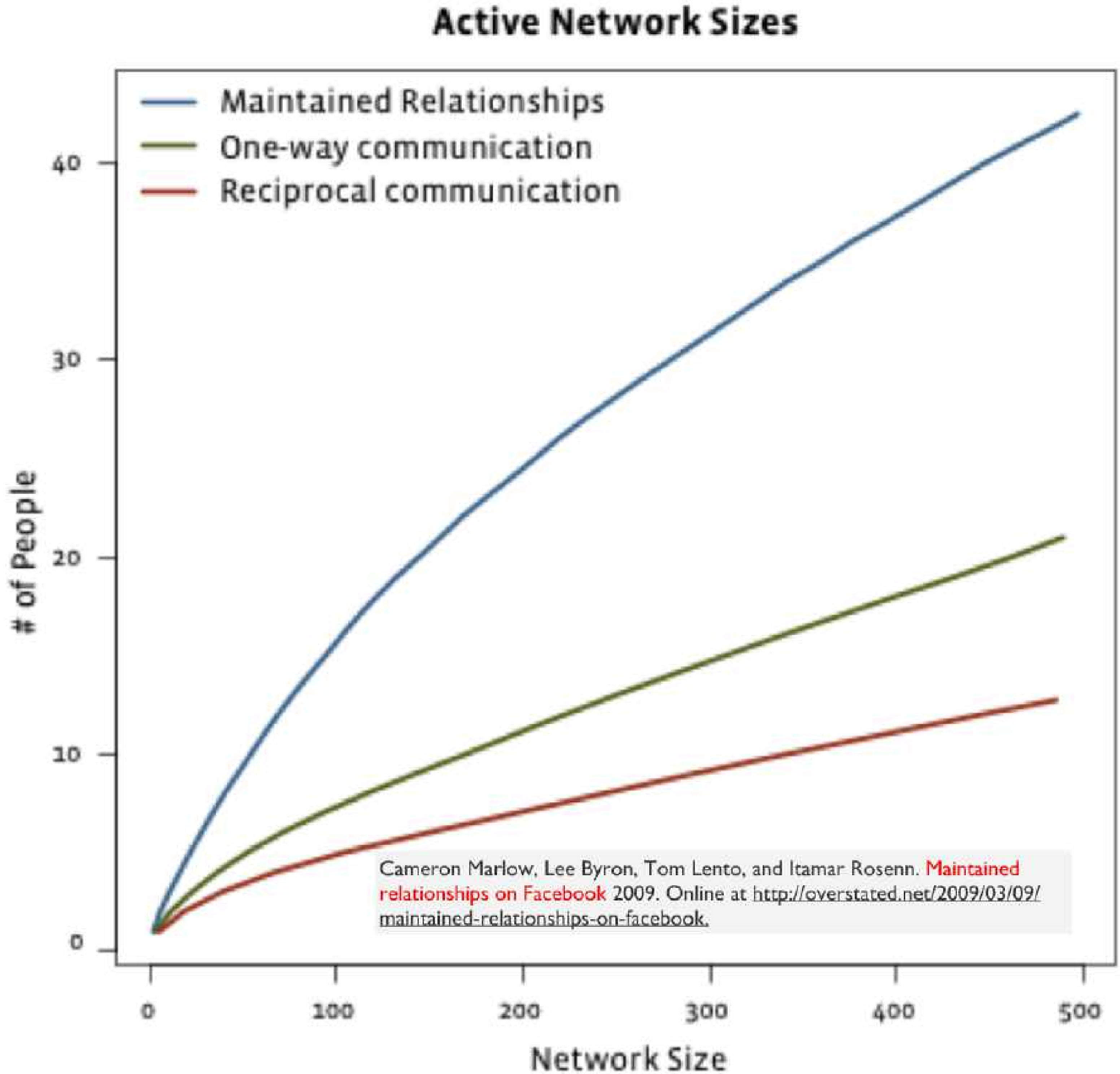


If a node A in a network satisfies the Strong Triadic Closure Property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie.


A
S
B
new connections are weak ties!


All Friends
Maintained Relationships
What kind of relationships exist on Facebook? What is the implication of having a "friend" ?
One-way Communication
Mutual Communication





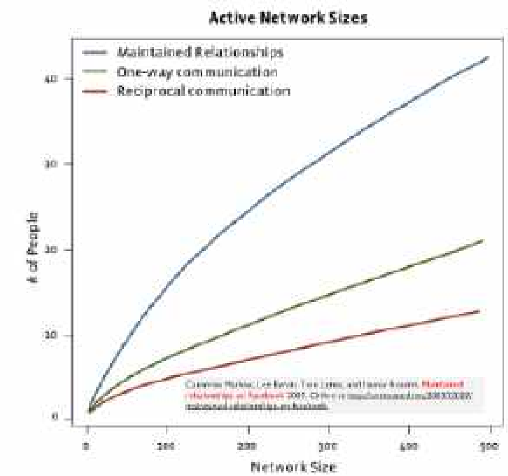
Cameron Marlow, Lee Byron, Tom Lento, and Itamar Rosenn. Maintained relationships on Facebook 2009. Online at http://overstated.net/2009/03/09/maintained-relationships-on-facebook.

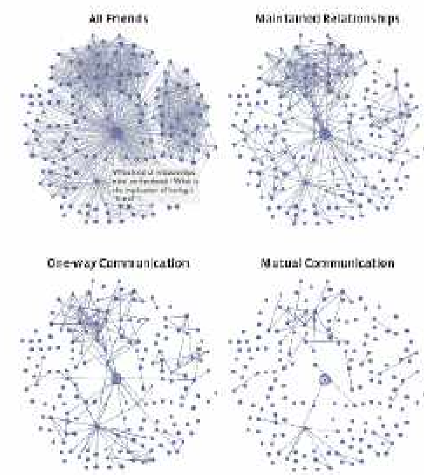
How many active relationships can one person maintain simultaneously?

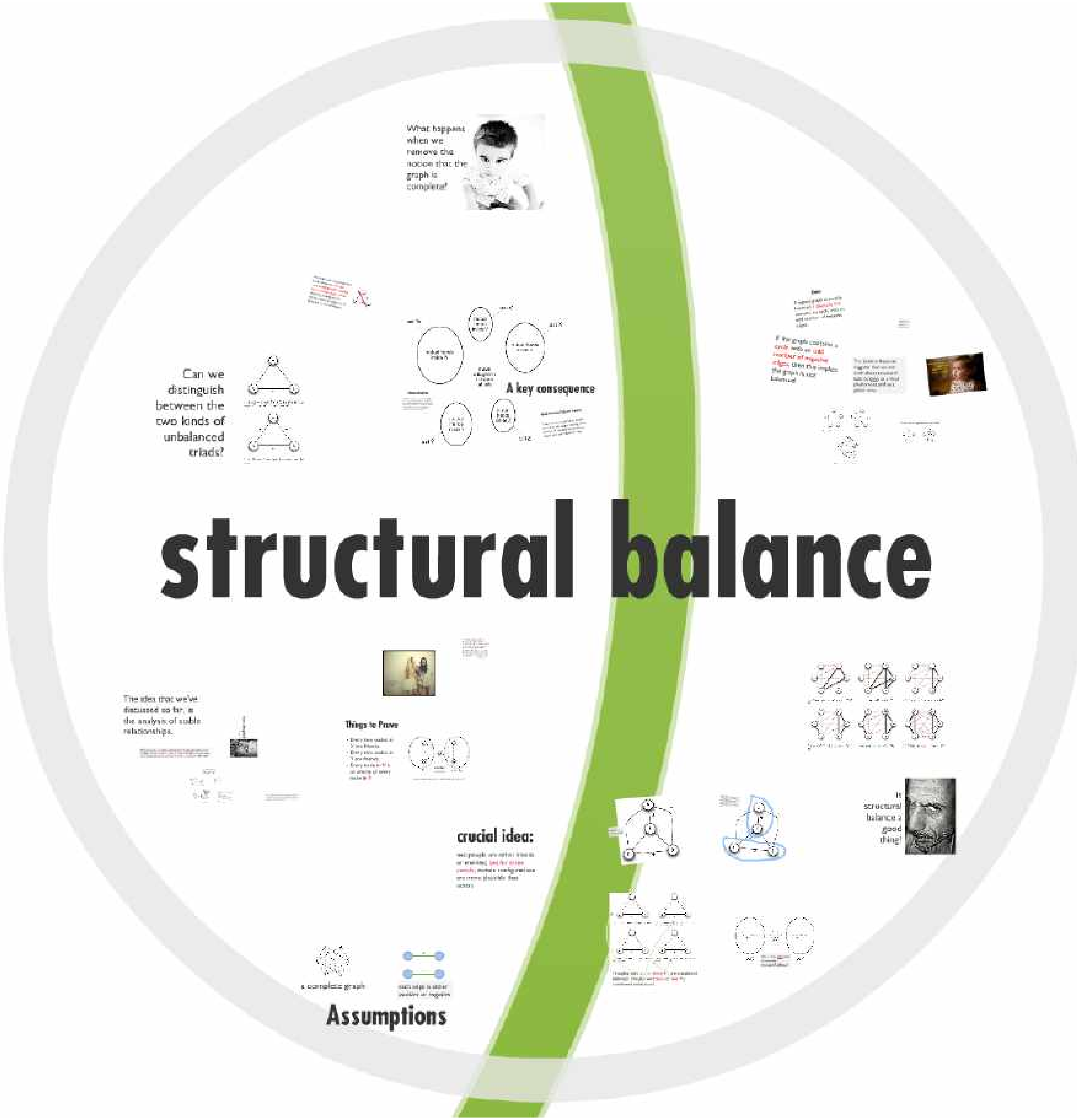
R. I. M. Dunbar. Coevolution of neocortical size, group size and language in humans. Behavioral and Brain Sciences, 16:681–694, 11 1993.

resource constraints are a key barrier to network size!

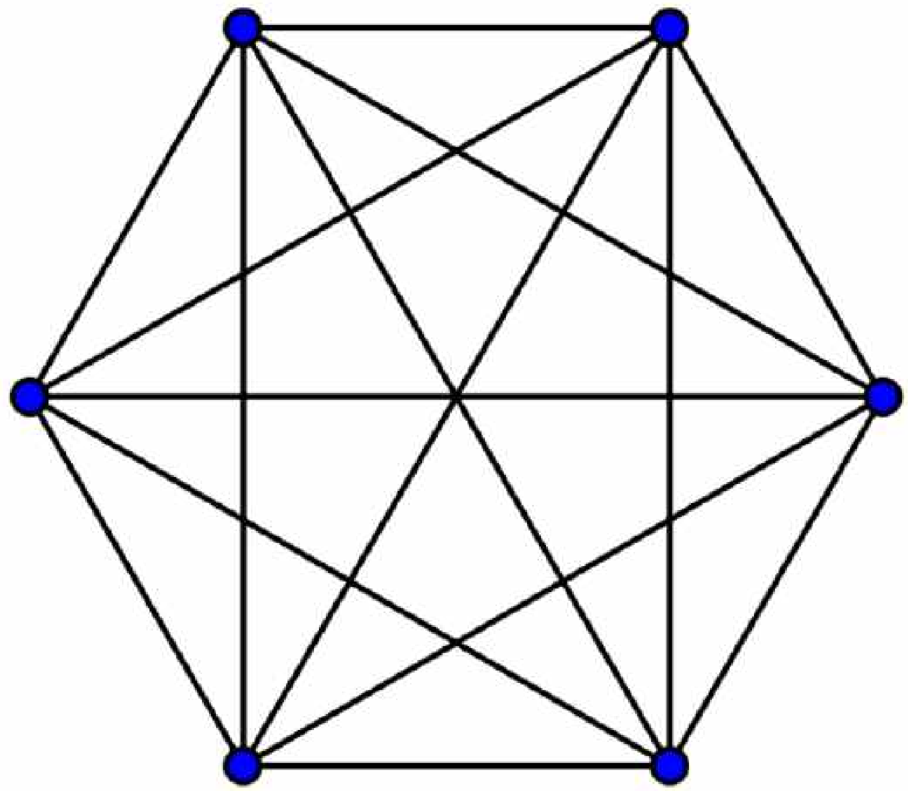
How many active relationships can one person maintain simultaneously?



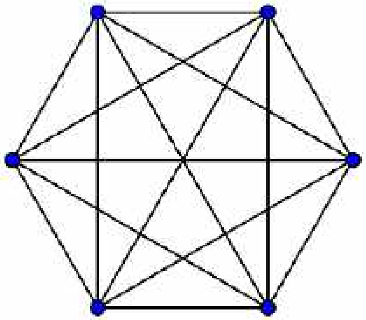



What happens when we remove the notion that the graph is complete?
Can we distinguish between the two kinds of unbalanced triads?
A key consequence
structural balance
Things to Prove
crucial idea:
Is structural balance a good thing?
a complete graph
each edge is either positive or negative
Assumptions

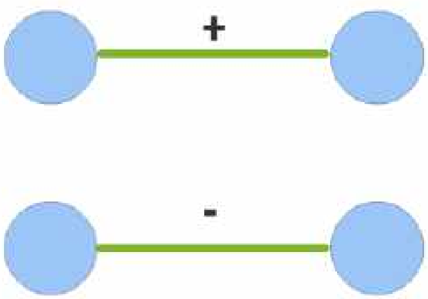

it is optimistic to
assume that all ties
are friendly



a complete graph



a complete graph

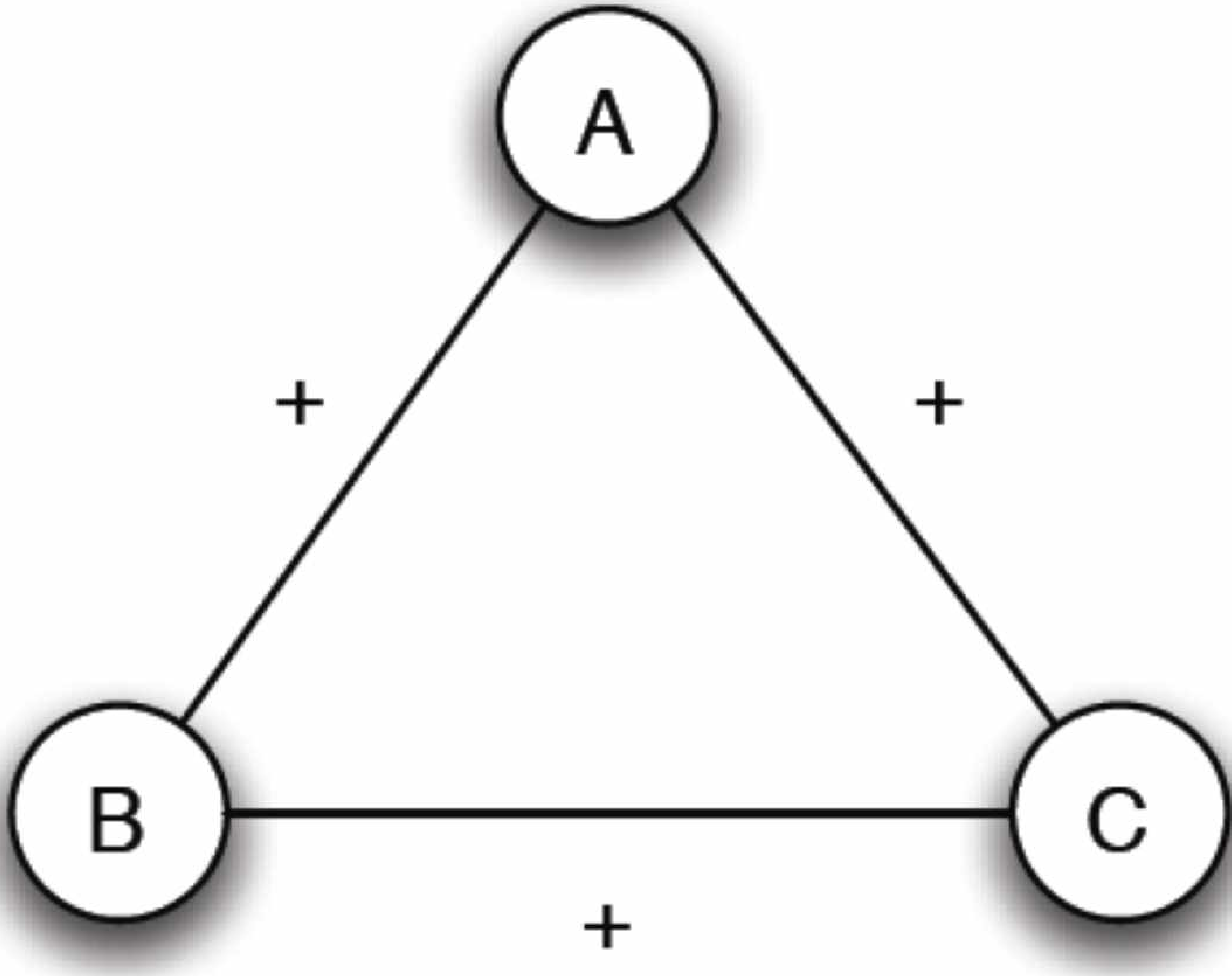


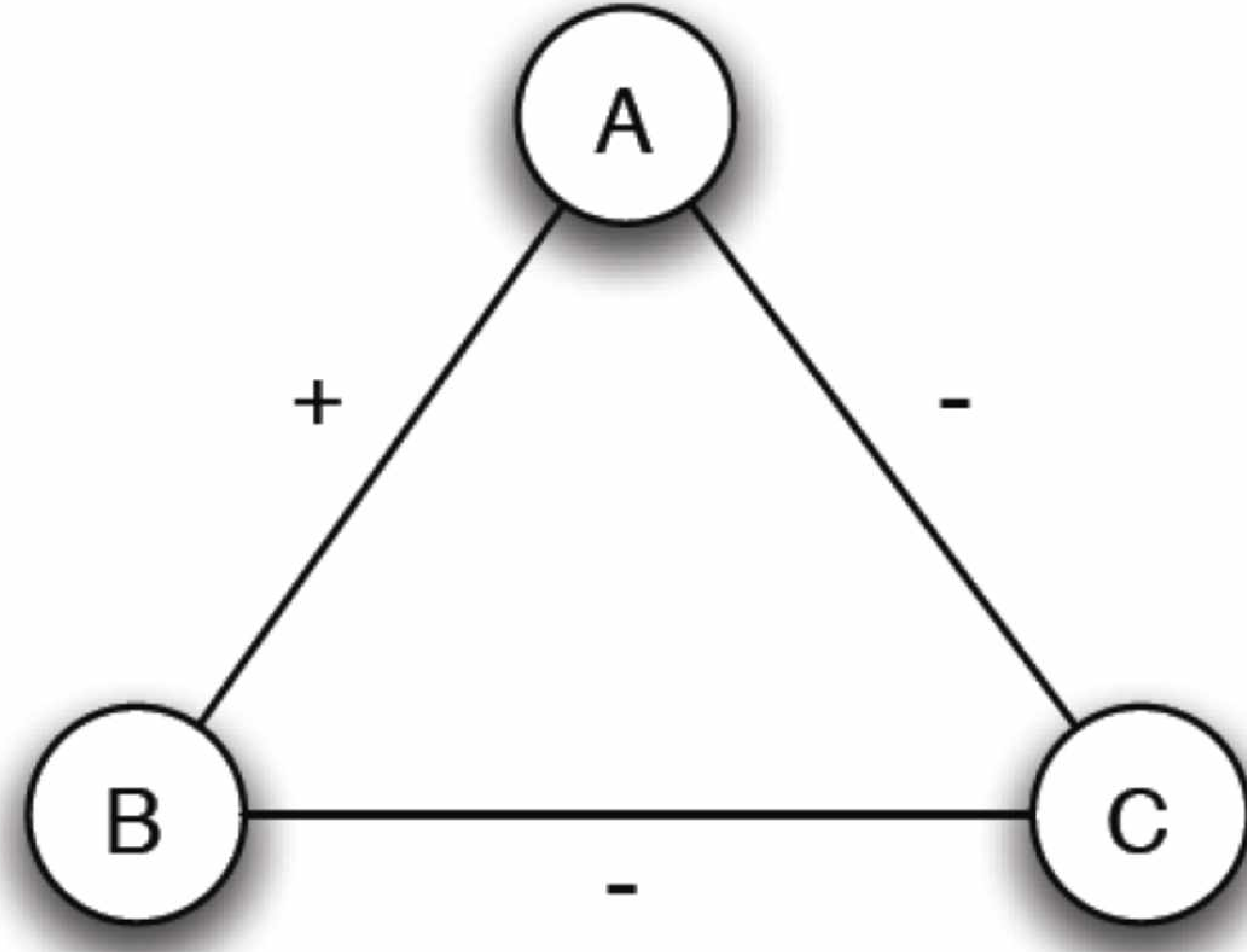
each edge is either positive or negative

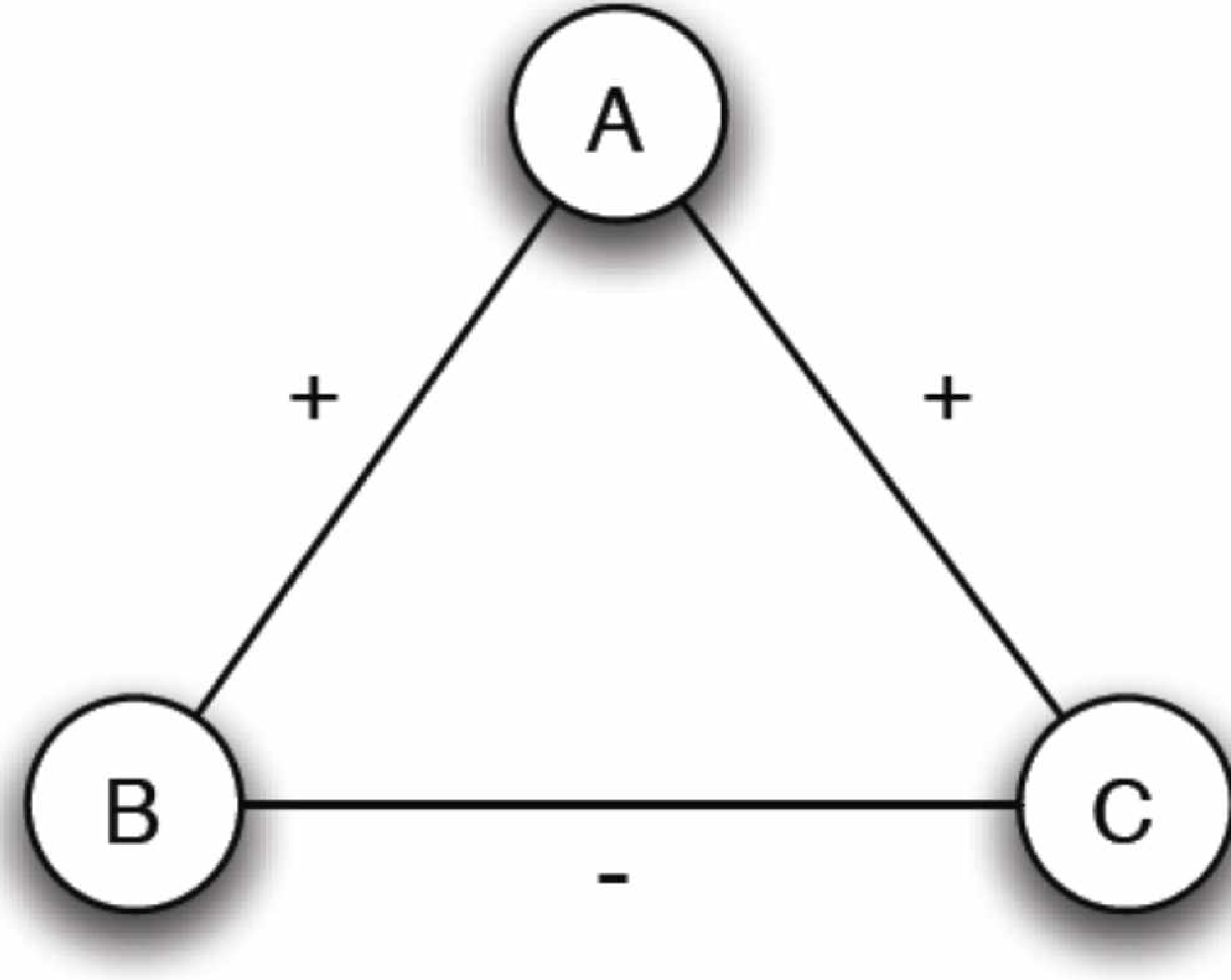
Assumptions

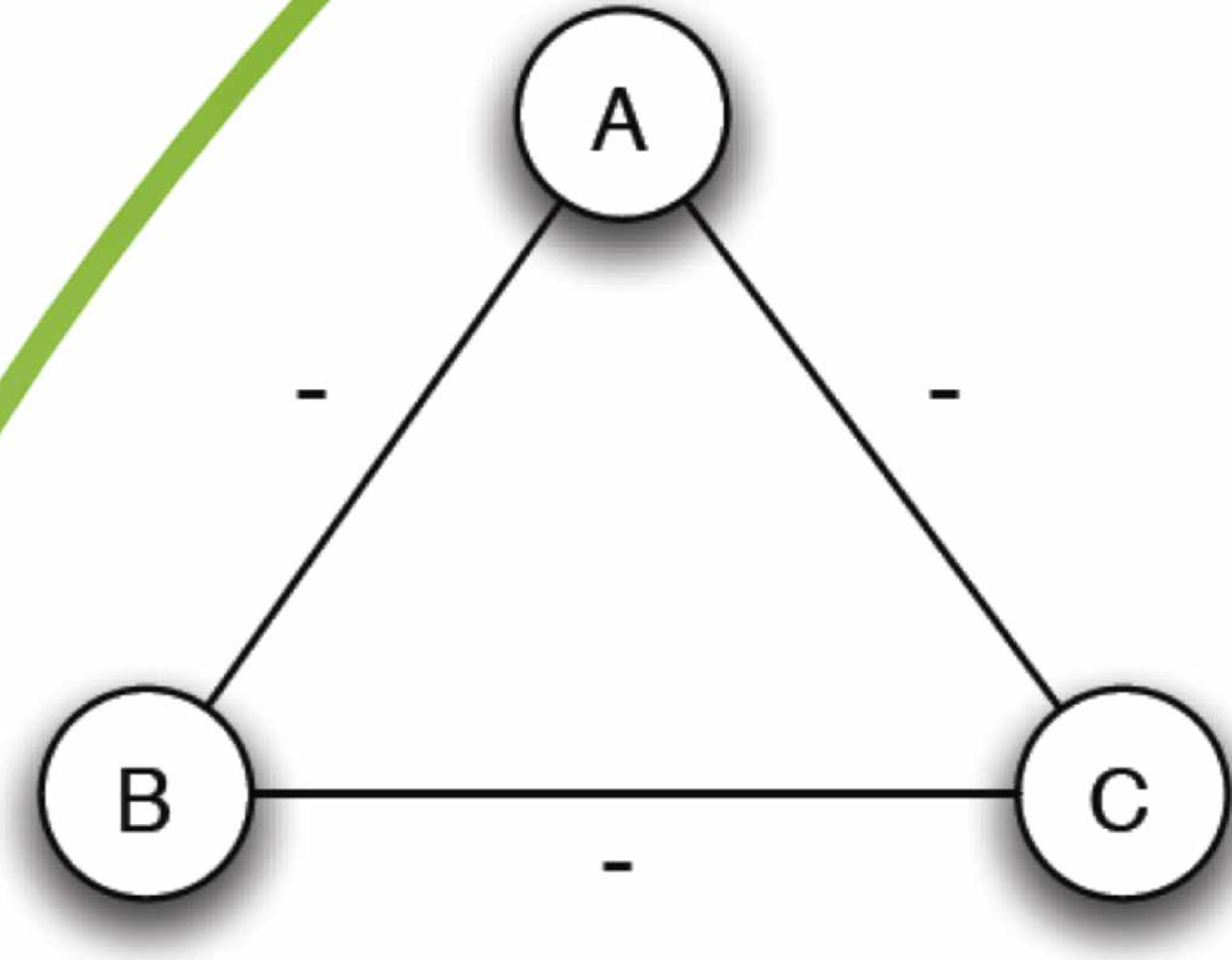
crucial idea:

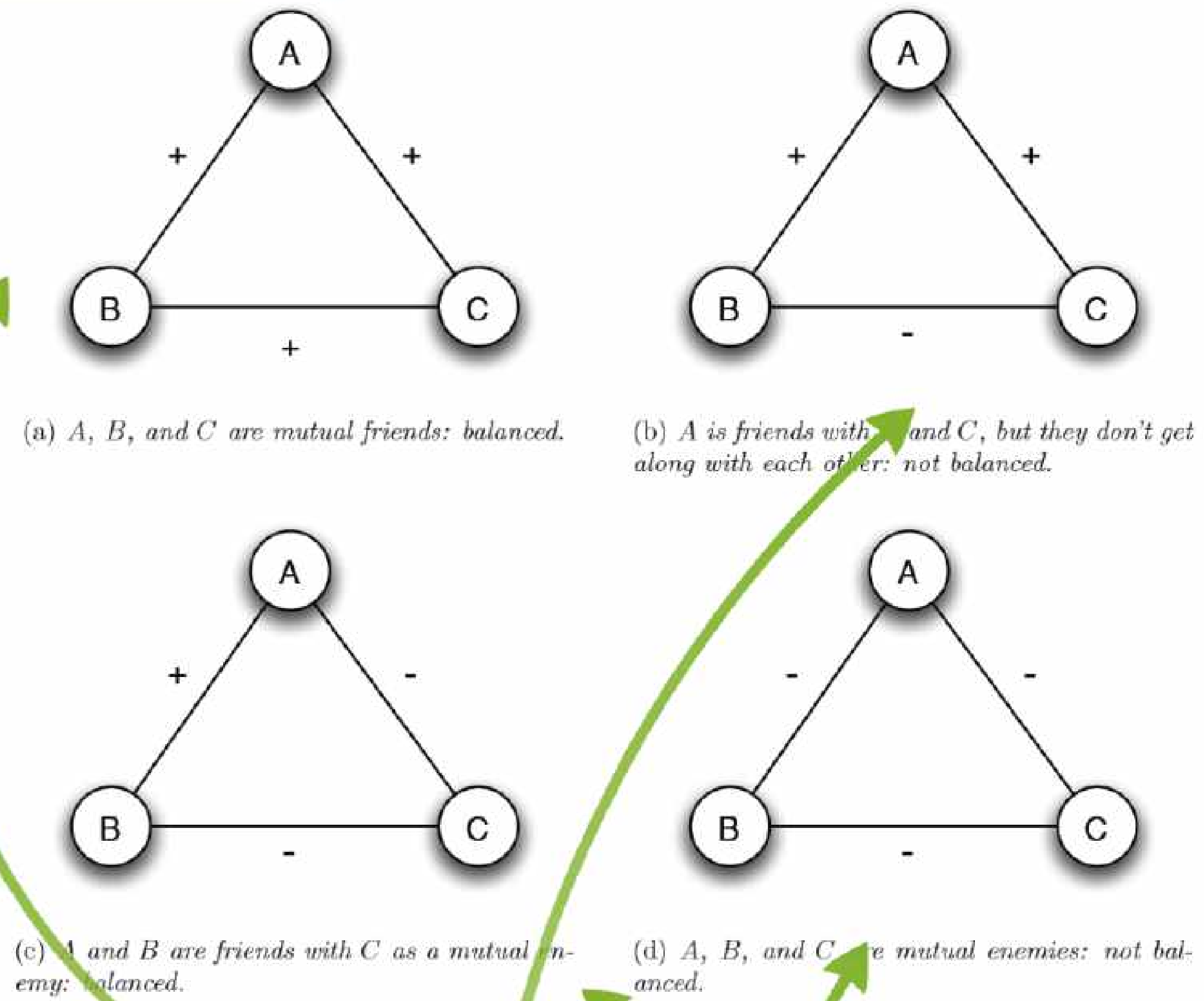
two people are either friends or enemies; but for three people, certain configurations are more plausible than others


A
B
C
+
+
+


A
+
-
B
C
-


A
+
+
B
-
C

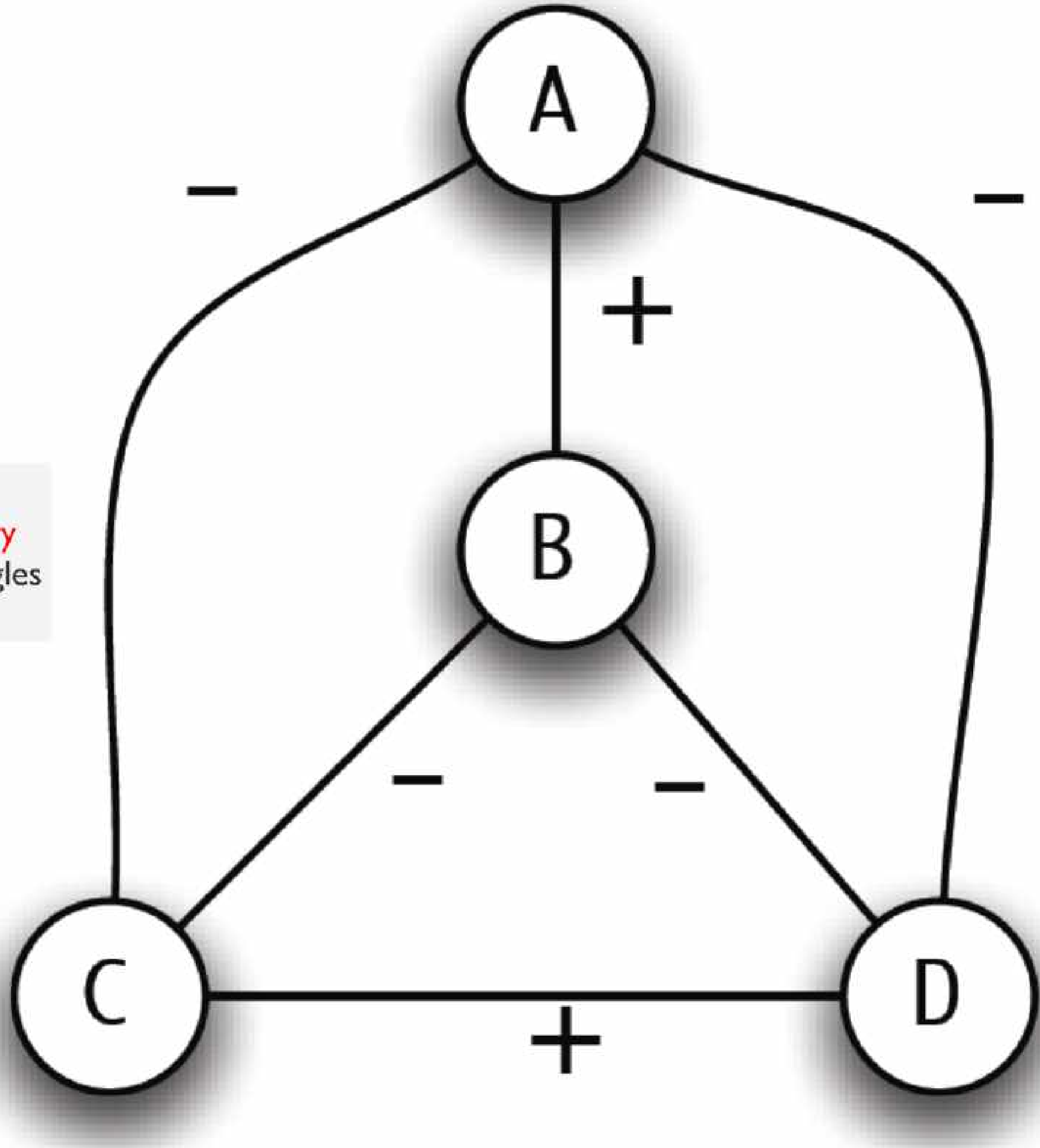

A
B
C
-
-
-



Triangles with one or three +'s are considered balanced; triangles with zero or two +'s considered unbalanced.

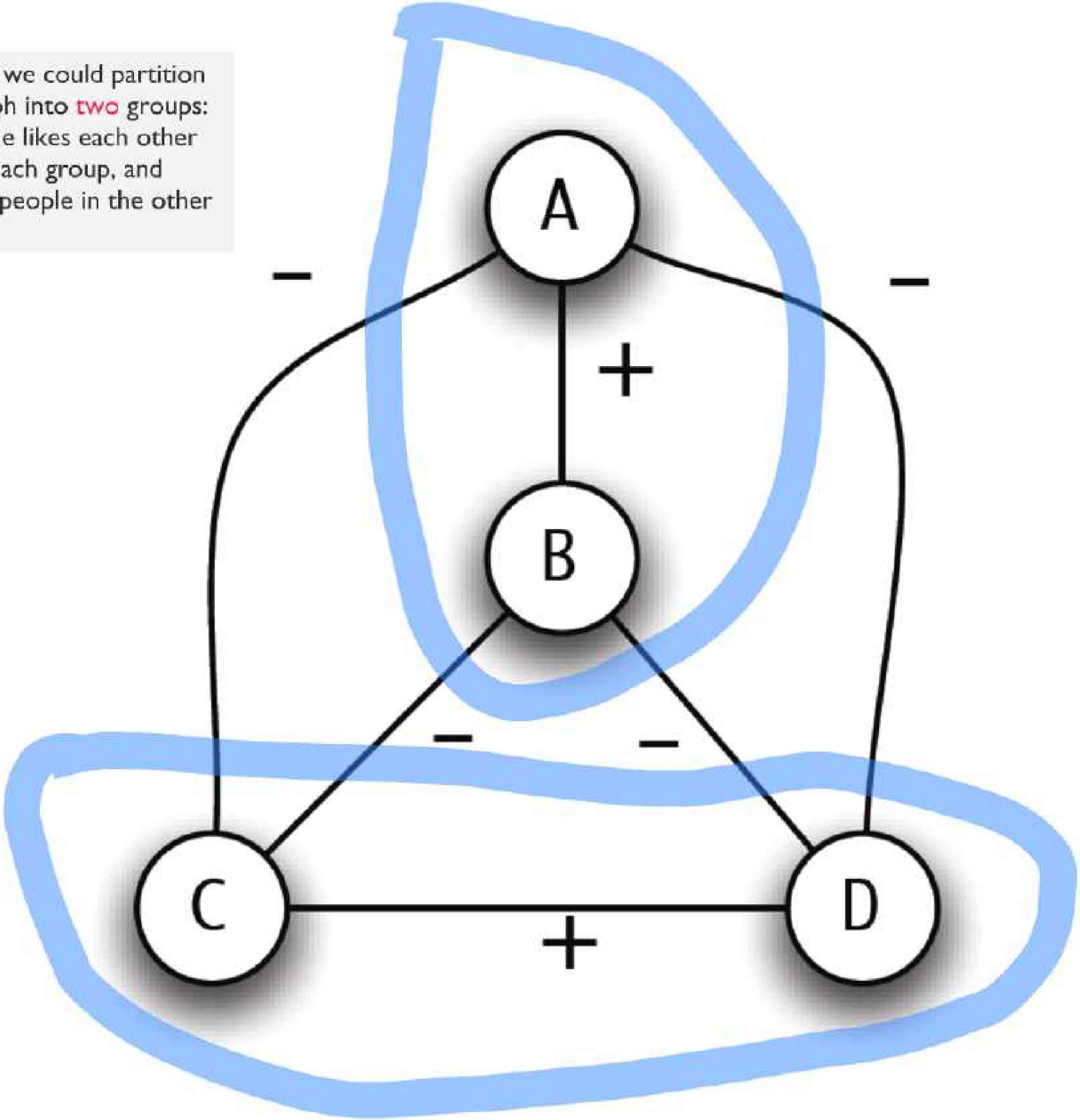
A graph is balanced if every one of its triangles is balanced

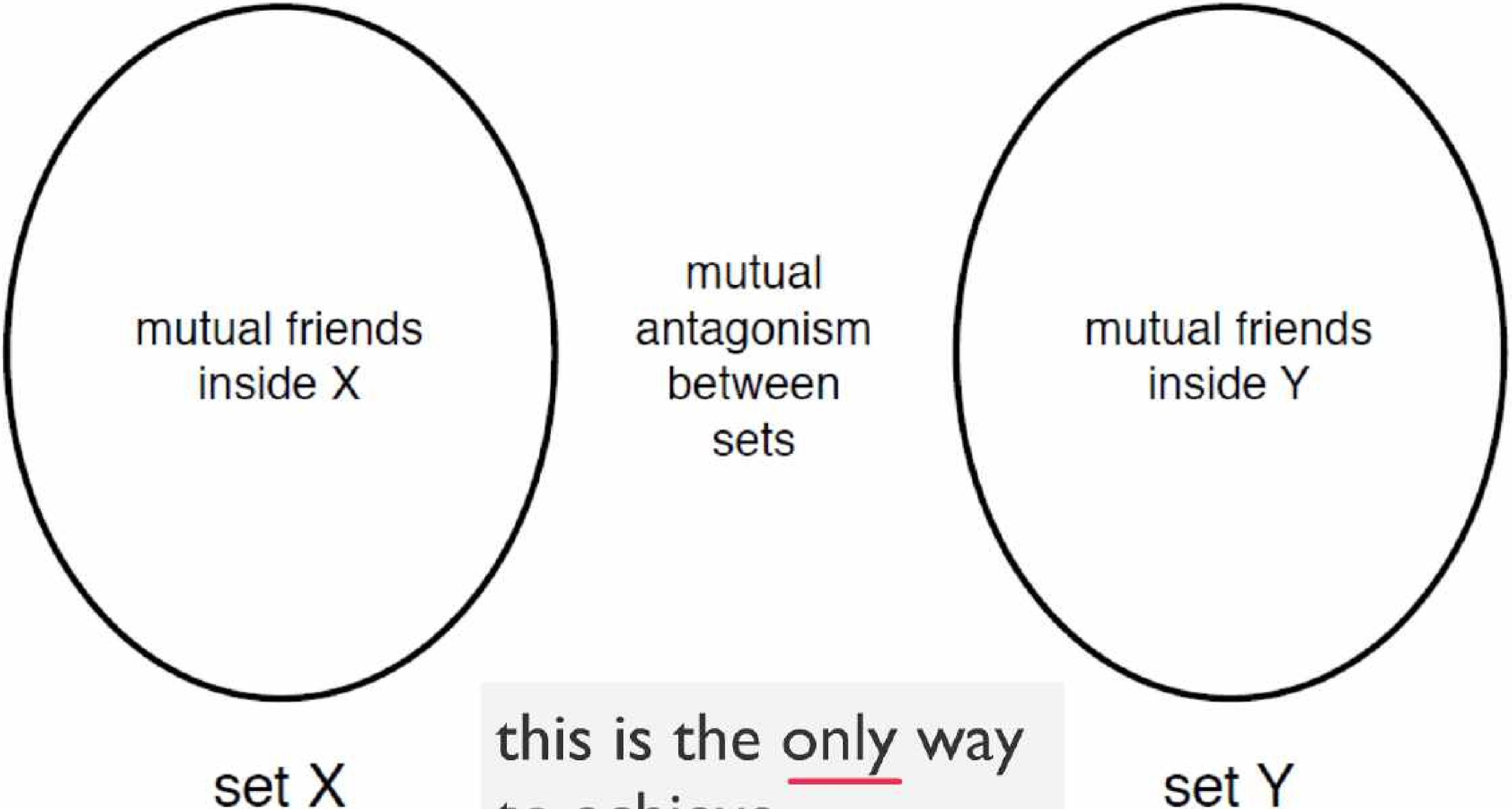
A graph is balanced if every one of its triangles is balanced



What if we could partition the graph into two groups: everyone likes each other within each group, and dislikes people in the other group

What if we could partition the graph into two groups: everyone likes each other within each group, and dislikes people in the other group





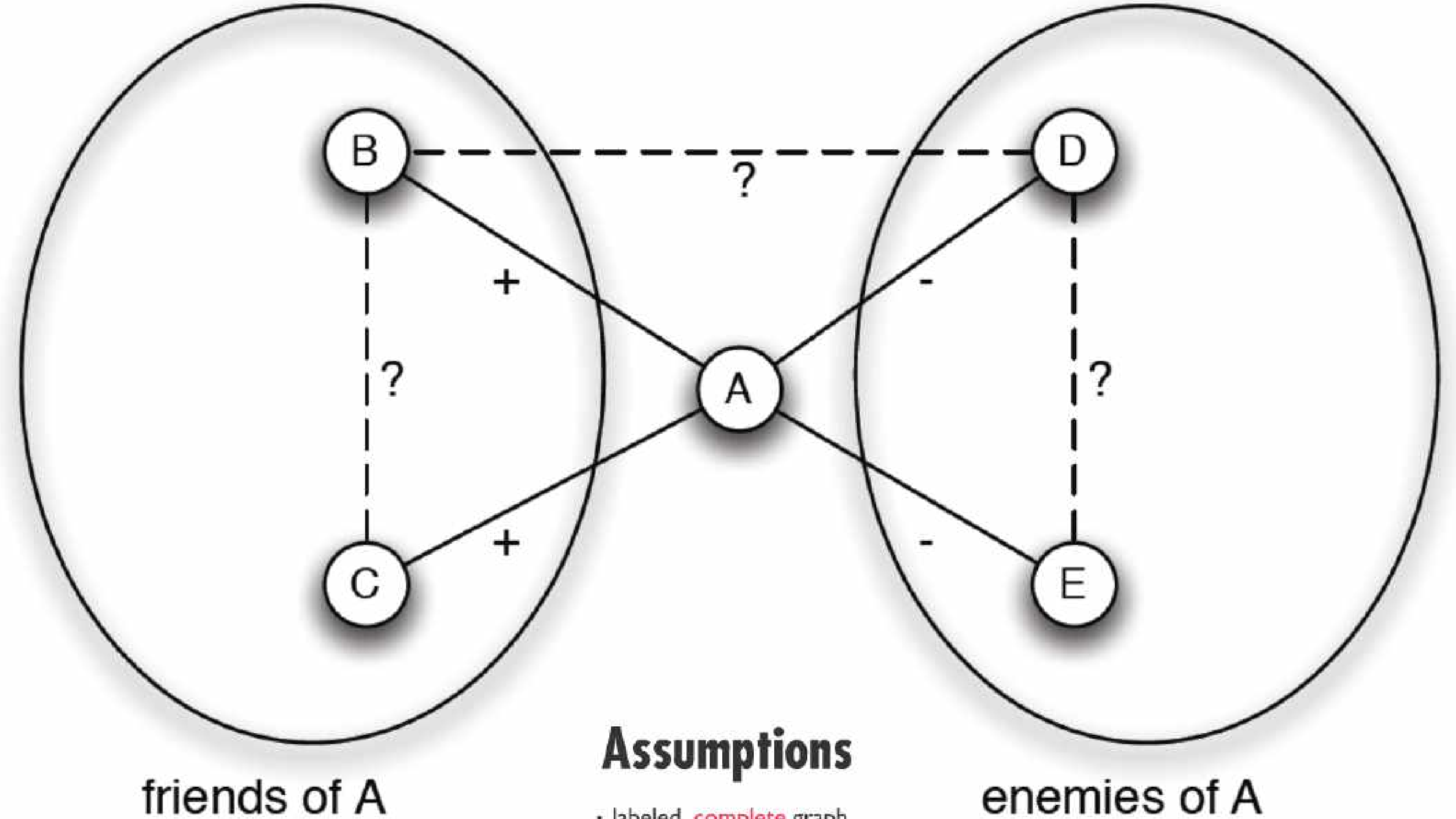
this is the only way to achieve structural balance!

If a labeled, complete graph is balanced, then either all pairs of nodes are friends, or else the nodes can be divided into two groups, X and Y, such that every pair of people in X like each other, every pair of people in Y like each other, and everyone in X is the enemy of everyone in Y.

Cartwright-Harary's Structural Balance Theorem

Assumptions

- labeled, complete graph
- we know that it is balanced



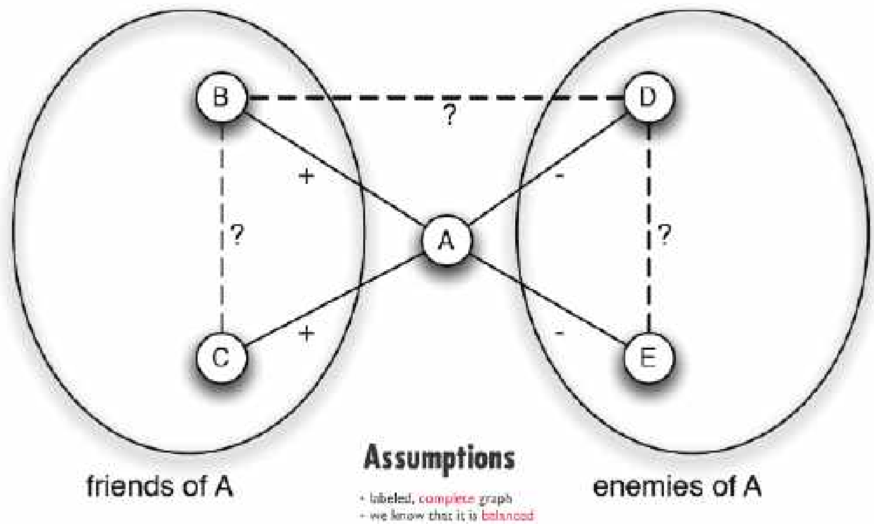
Assumptions

- labeled, complete graph
- we know that it is balanced

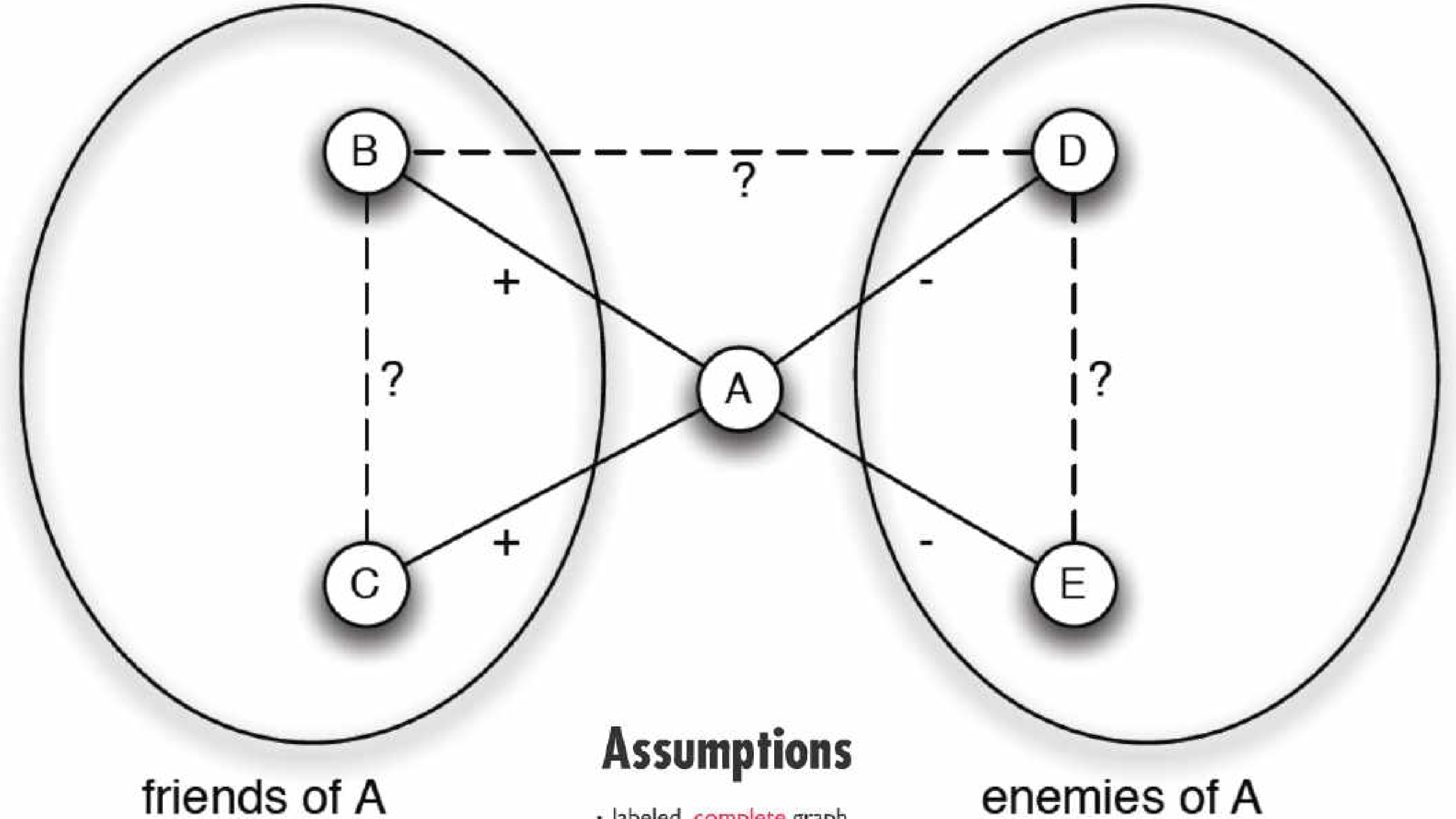
let us examine node A, that belongs to set X.

Things to Prove

- Every two nodes in X are friends.
- Every two nodes in Y are friends.
- Every node in X is an enemy of every node in Y.



let us examine node A, that belongs to set X.



Assumptions

- labeled, complete graph
- we know that it is balanced

let us examine node A, that belongs to set X.

The idea that we've discussed so far, is the analysis of stable relationships.

Marvel, Seth A., Jon Kleinberg, Robert D. Kleinberg, and Steven H.

The idea that we've discussed so far, is the analysis of stable relationships.

Marvel, Seth A., Jon Kleinberg, Robert D. Kleinberg, and Steven H. Strogatz. Continuous-time model of structural balance. Proceedings of the National Academy of Sciences 108, no. 5 (2011): 1771-1776.

"To identify and rigorously analyze a simple system that could progress to balanced graphs from generic initial configurations."

X

Entries of X are
sampled independently
from an absolutely
continuous distribution
with bounded support

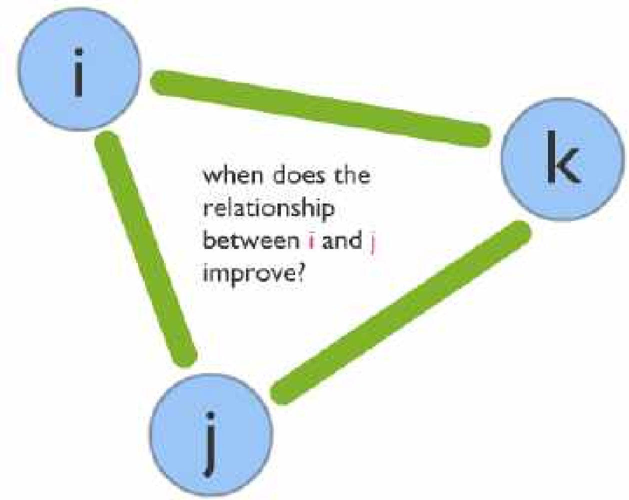
matrix of relationships

Entries of X are sampled independently from an absolutely continuous distribution with bounded support

matrix of relationships

$$\frac{\mathrm{d}X}{\mathrm{d}t} = X^2$$

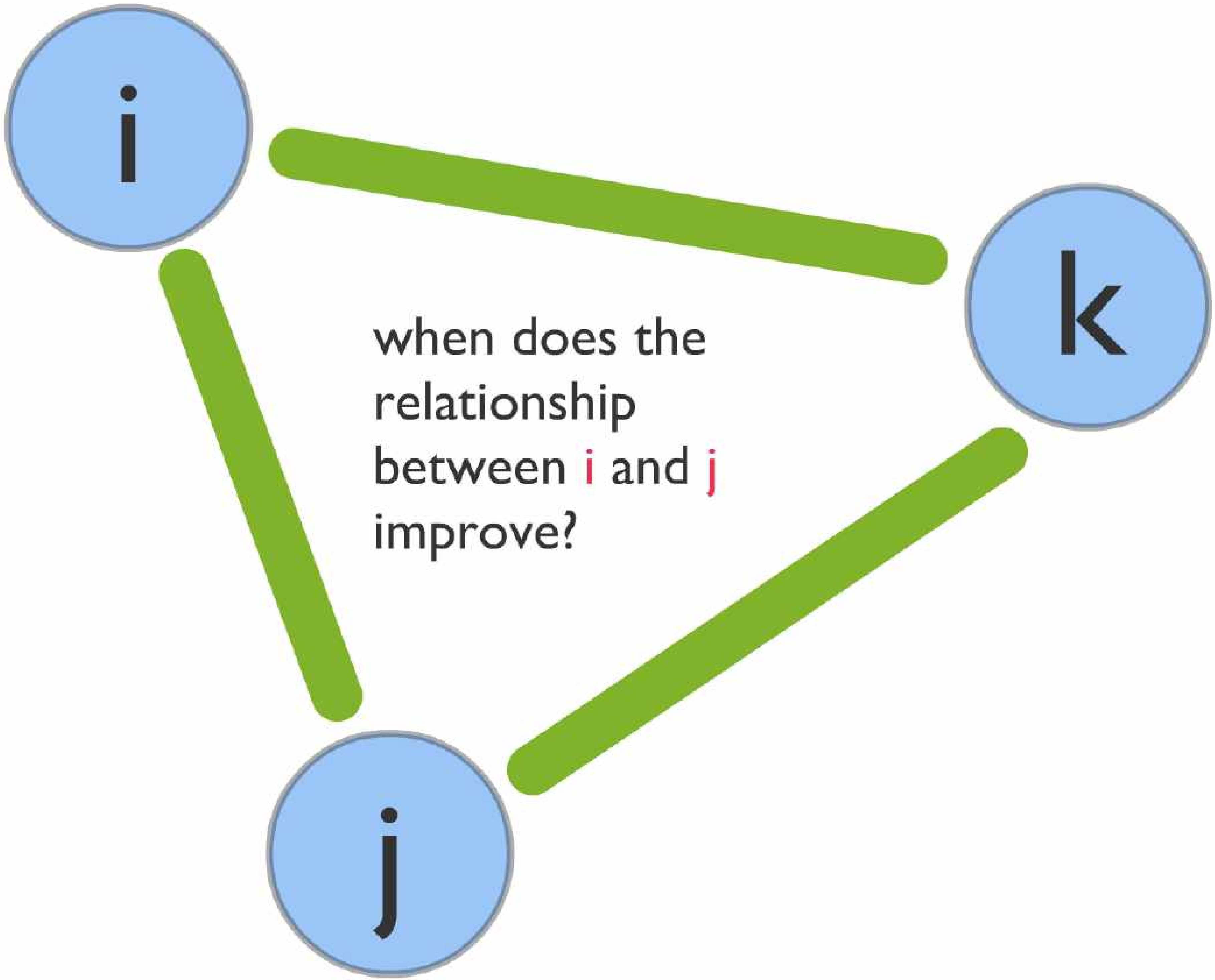
the assumption about relationship dynamics

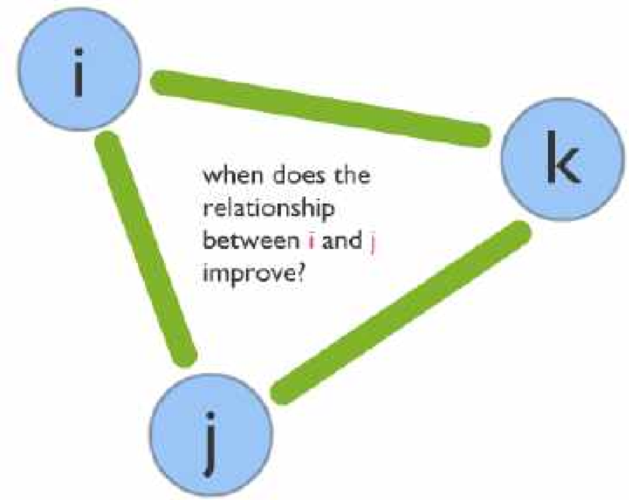


$$\frac{\mathrm{d}x_{ij}}{\mathrm{d}t} = \sum_{k} x_{ik} x_{kj}$$

all mutual friends

discrete case


i
k
j
when does the relationship between i and j improve?



$$\frac{\mathrm{d}x_{ij}}{\mathrm{d}t} = \sum_{k} x_{ik} x_{kj}$$

all mutual friends

discrete case

two factions

- Leading Eigenvalue is positive
- Leading Eigenvalue is greater than the second Eigenvalue
- Leading Eigenvector is not 0

split into positive, negative sets

$$X^* = \omega_1 \omega_1^T$$

leading eigenvector

split into positive, negative sets

$$\omega_1$$

leading eigenvector

$$\mu$$

$$\mu \leq 0$$

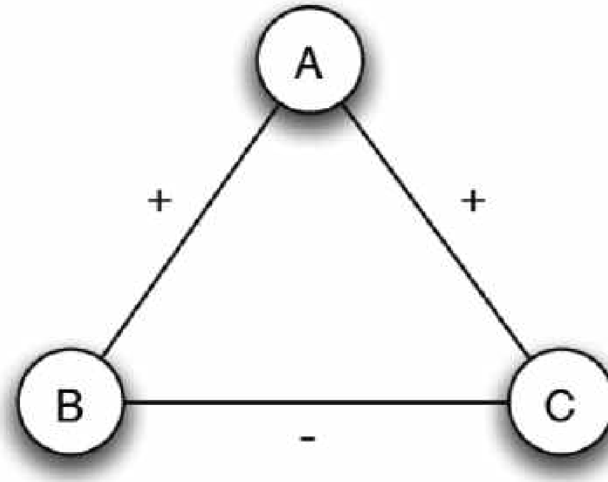
two factions

$$\mu > 0$$

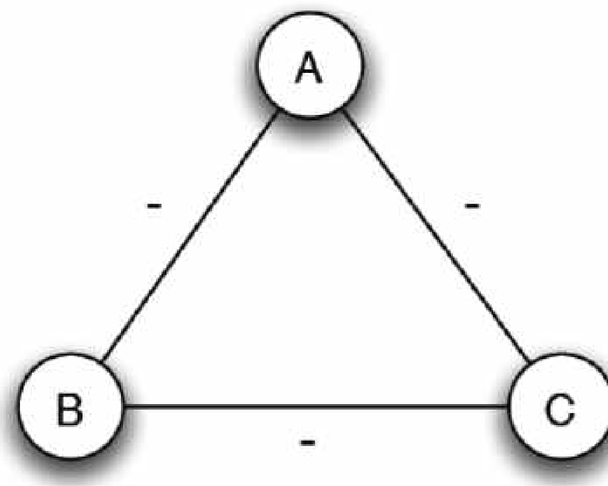
everyone's friendly

the average "friendship" strength

Can we distinguish between the two kinds of unbalanced triads?

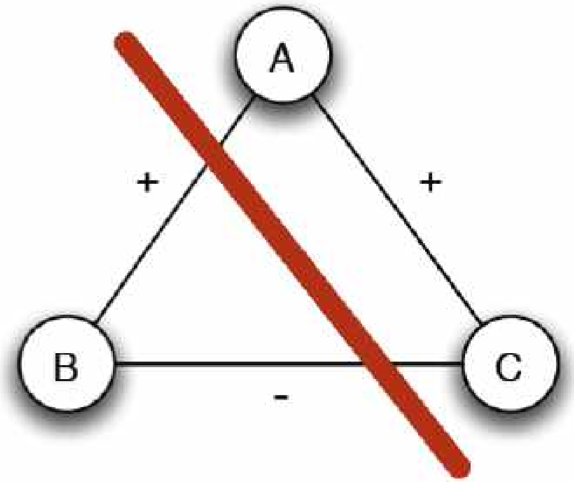


(b) *A is friends with B and C, but they don't get along with each other: not balanced.*



(d) *A, B, and C are mutual enemies: not balanced.*

What structural properties arise when we rule out only triangles with exactly two positive edges, while allowing triangles with three negative edges to be present in the network?



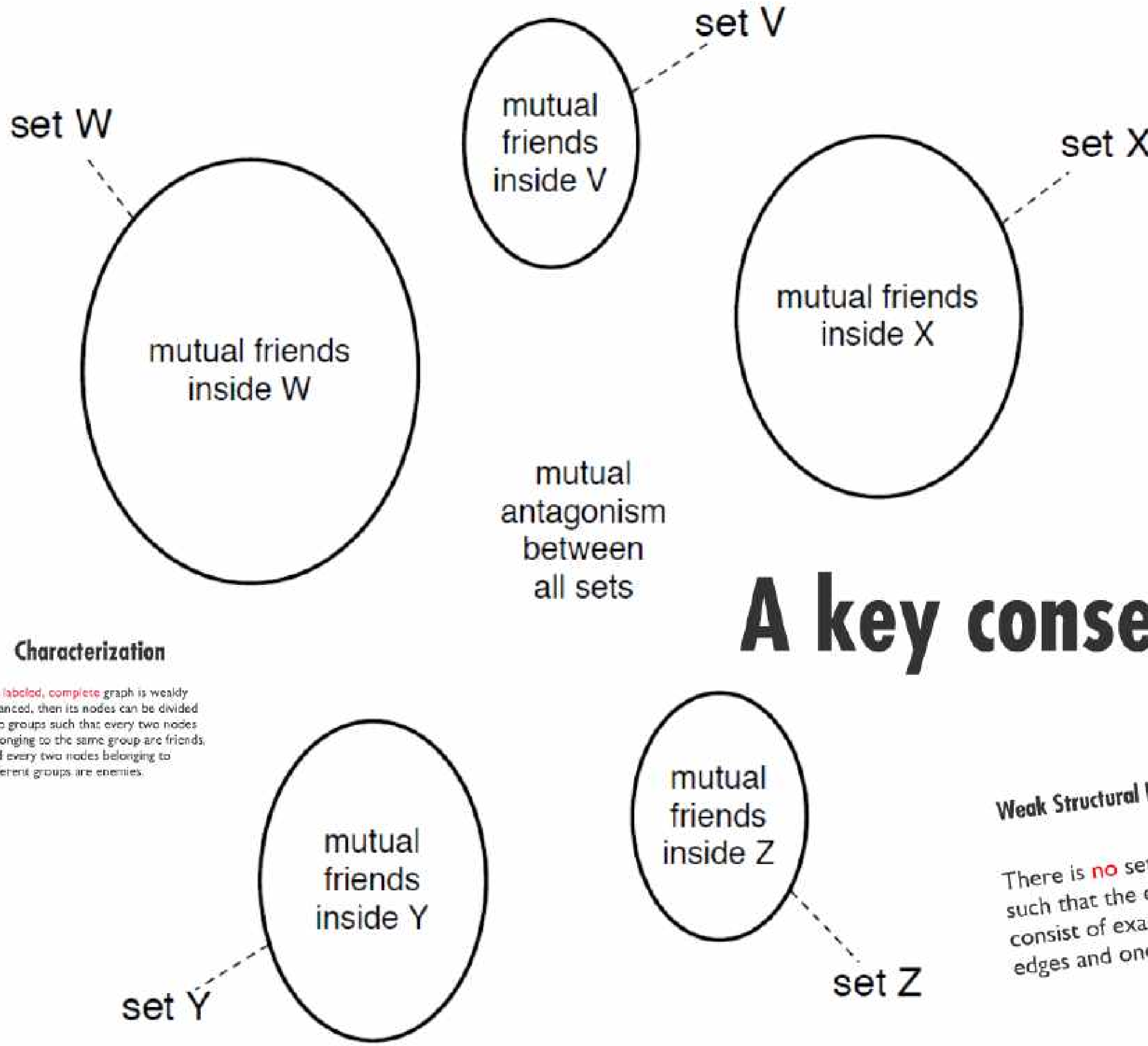
Weak Structural Balance Property

There is no set of three nodes such that the edges among them consist of exactly two positive edges and one negative edge.

Characterization

If a labeled, complete graph is weakly balanced, then its nodes can be divided into groups such that every two nodes belonging to the same group are friends, and every two nodes belonging to different groups are enemies.

A key consequence



Characterization

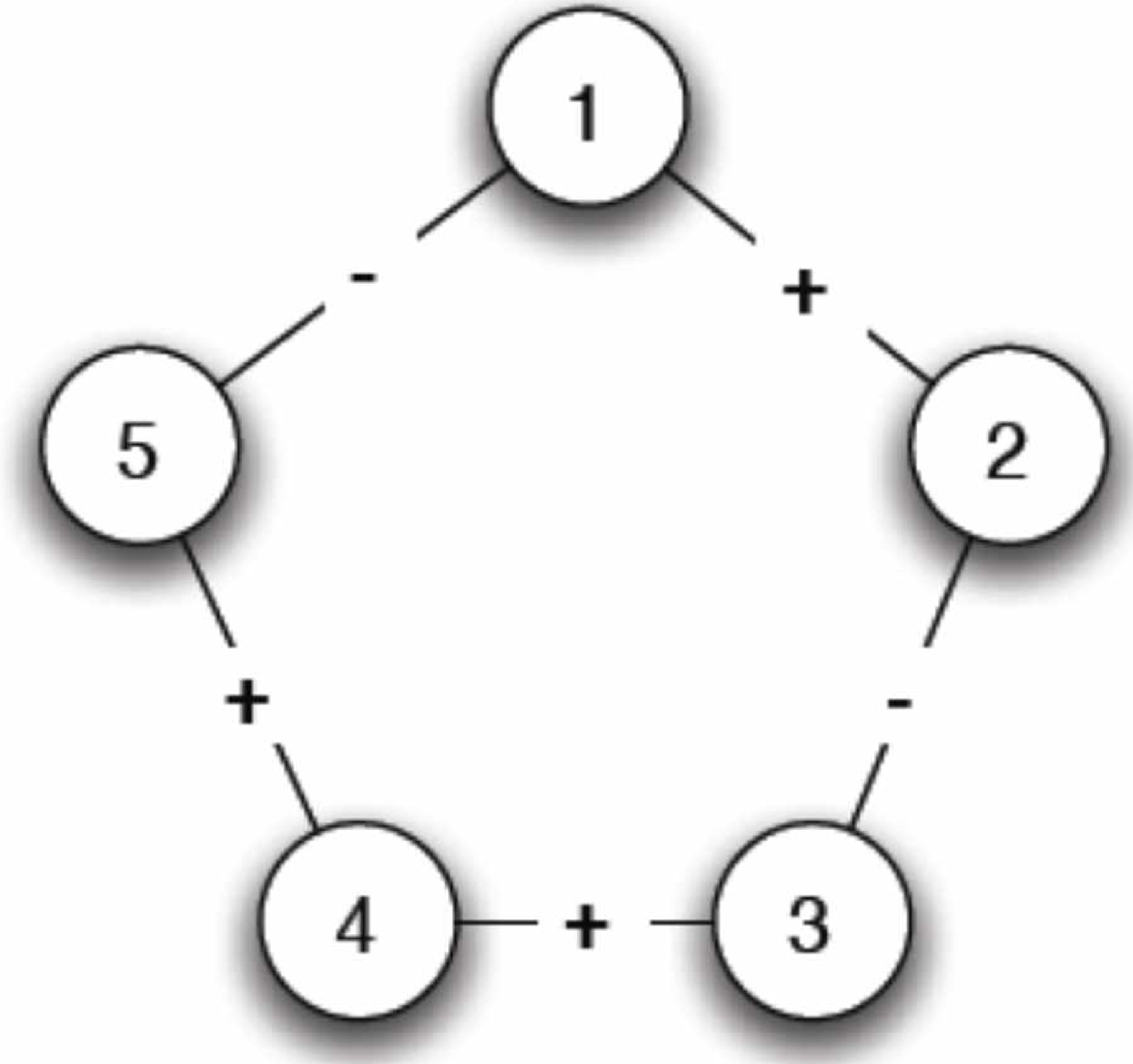
If a labeled, complete graph is weakly balanced, then its nodes can be divided into groups such that every two nodes belonging to the same group are friends, and every two nodes belonging to different groups are enemies.

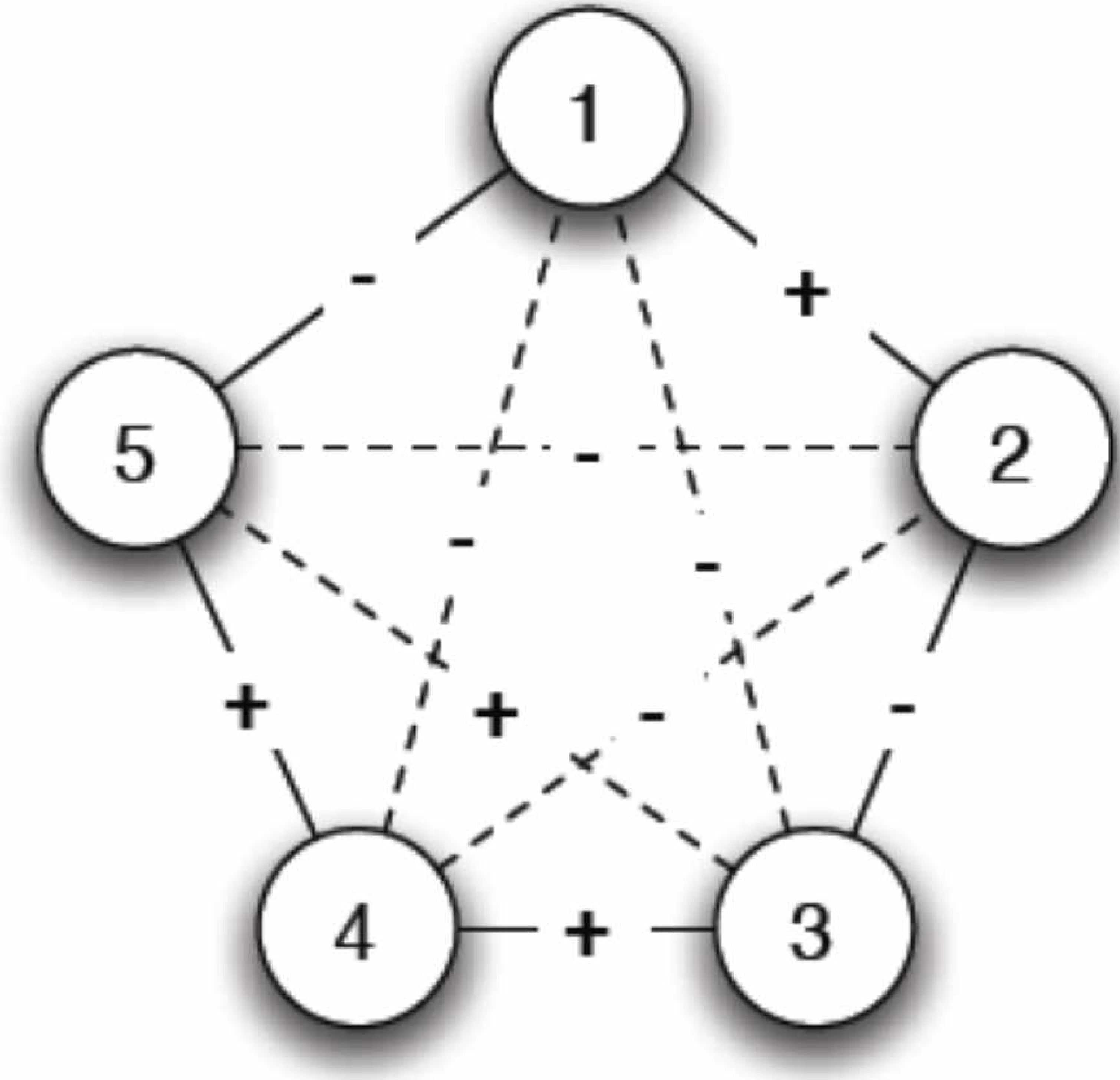
Weak Structural Balance Property

There is no set of three nodes such that the edges among them consist of exactly two positive edges and one negative edge.

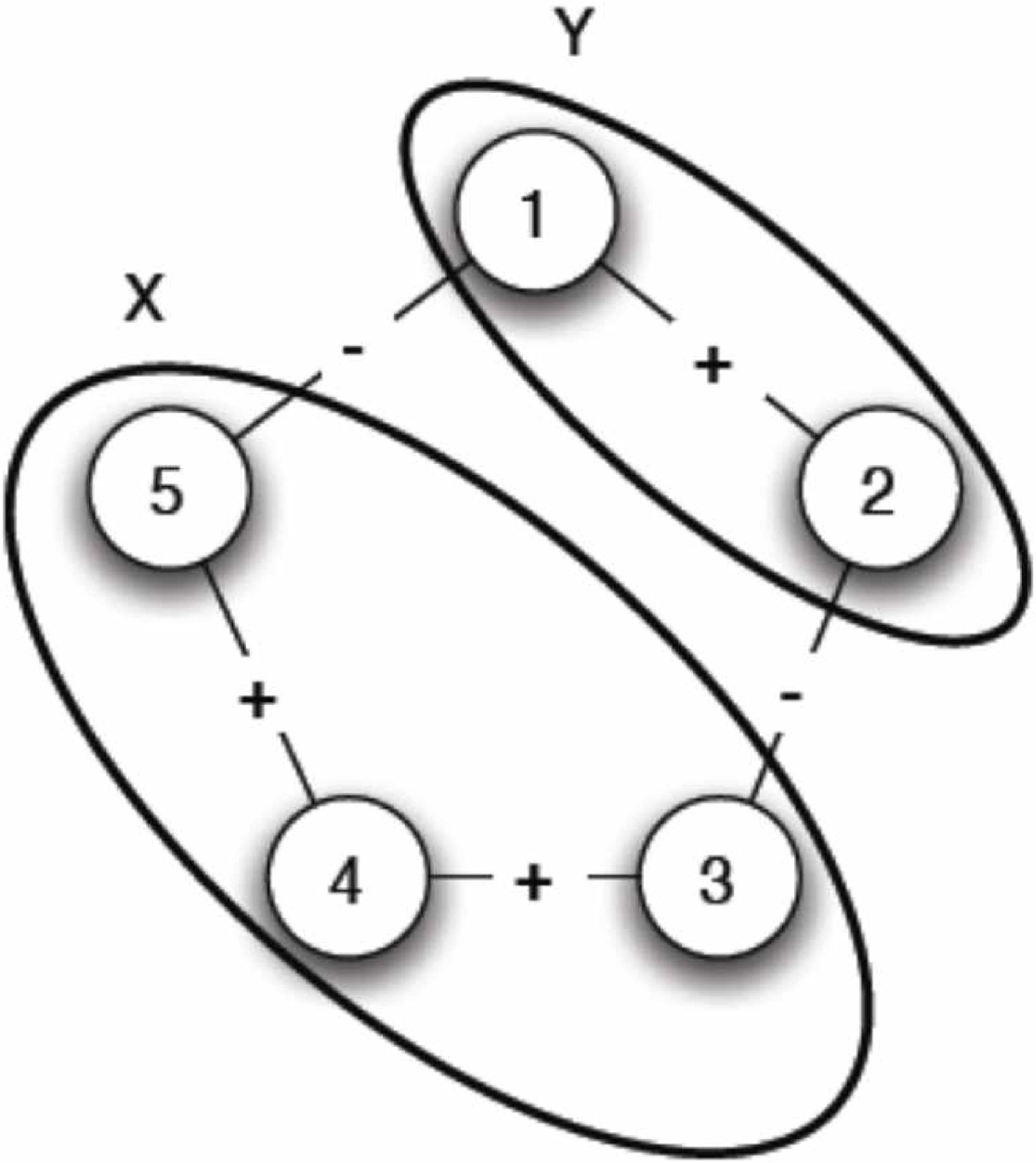
What happens when we remove the notion that the graph is complete?

The balance theorem suggests that we can think about structural balance both as a local phenomena and as a global view.


1
-
+
5
2
+
-
4
+
3

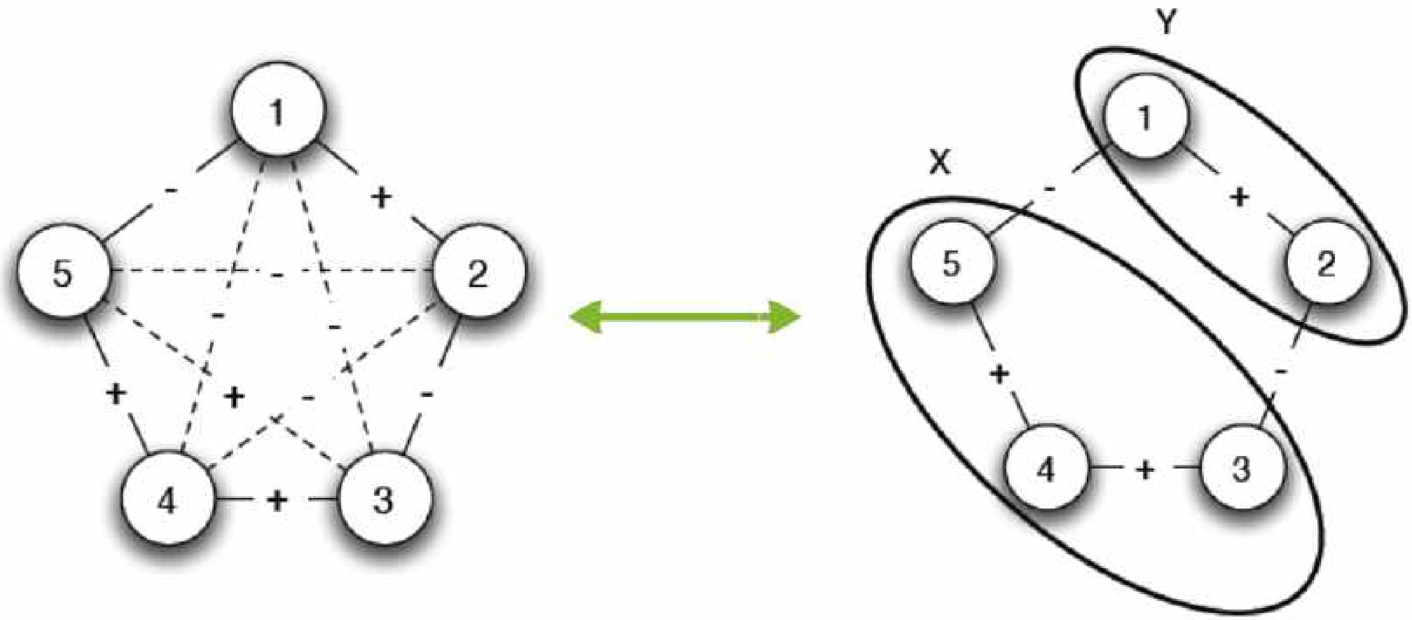


1
2
3
4
5
+
-
+
-
+
+
-
-
-
-

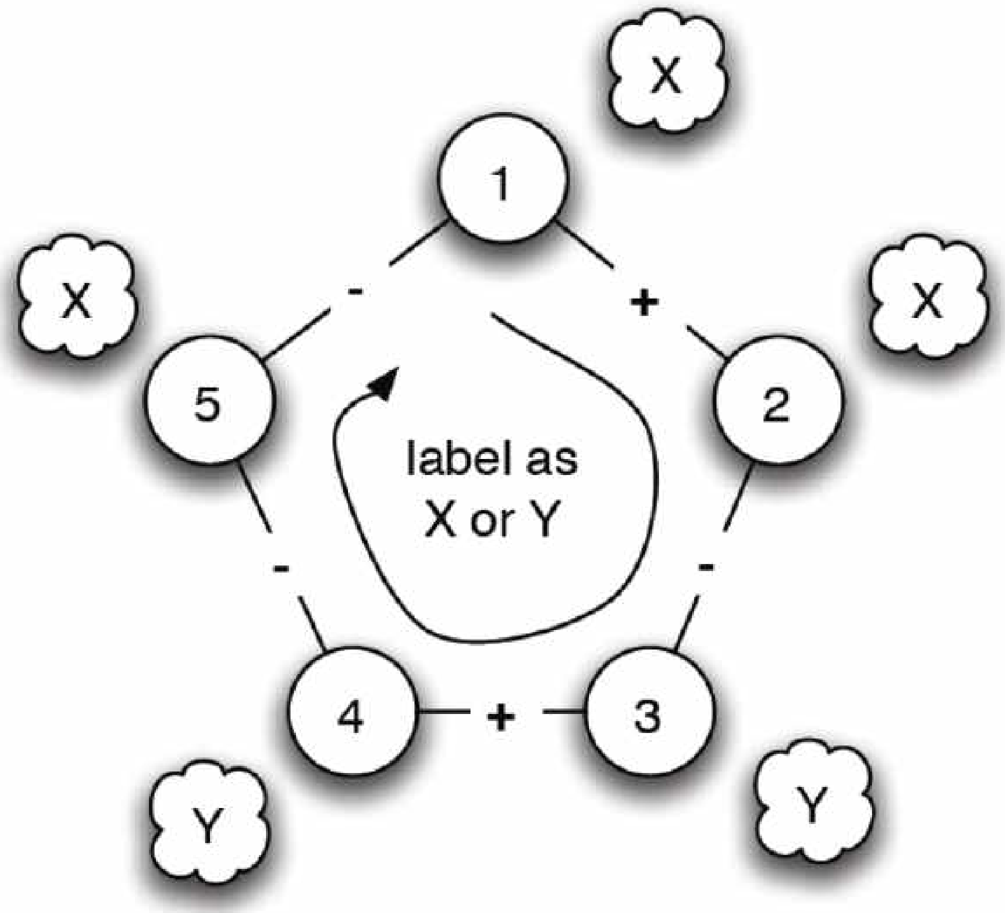


Y
X
1
2
3
4
5
+
-
+
-
+
+

these are equivalent views!




Can we examine a
graph and
determine if it is
weakly balanced?

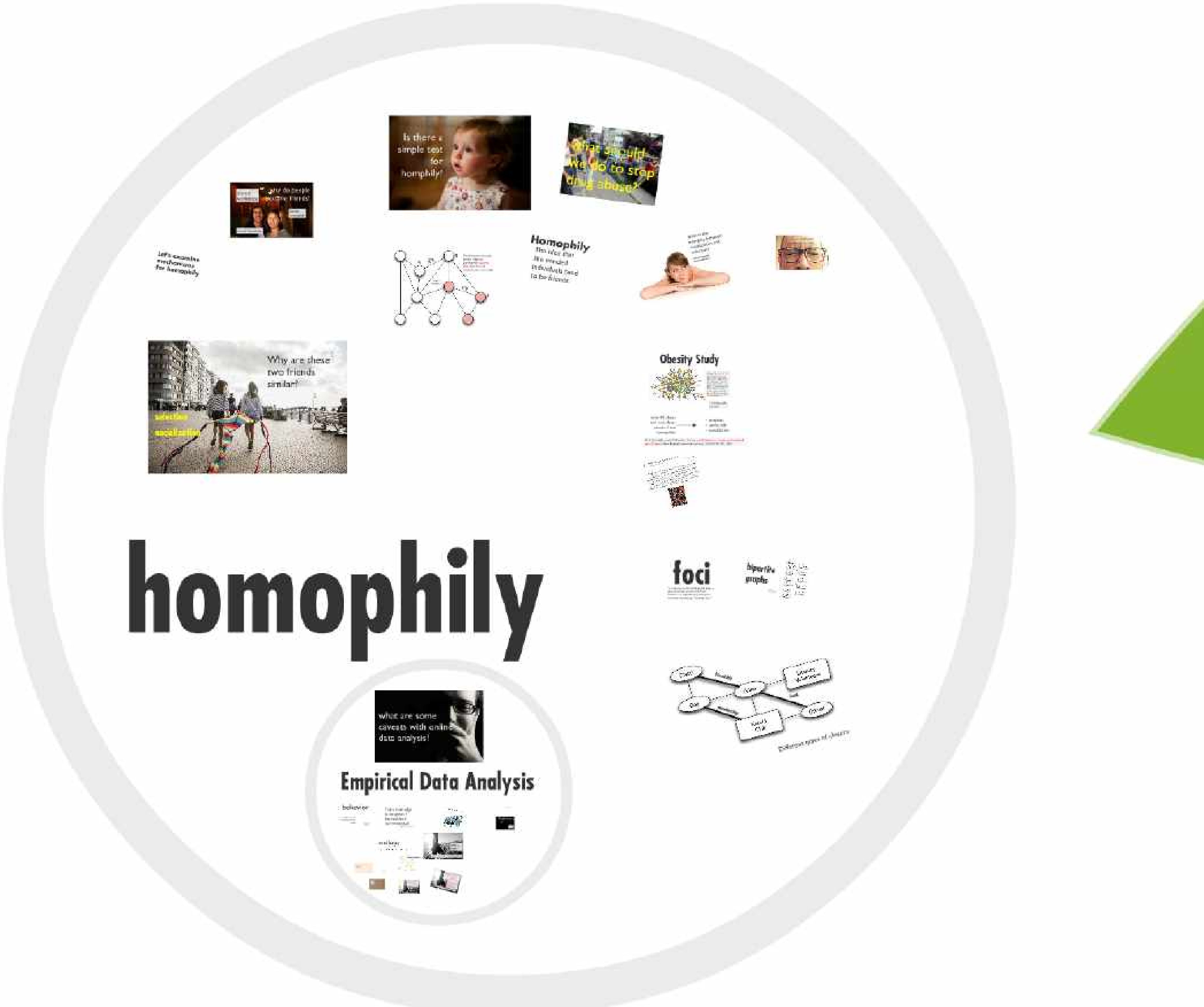


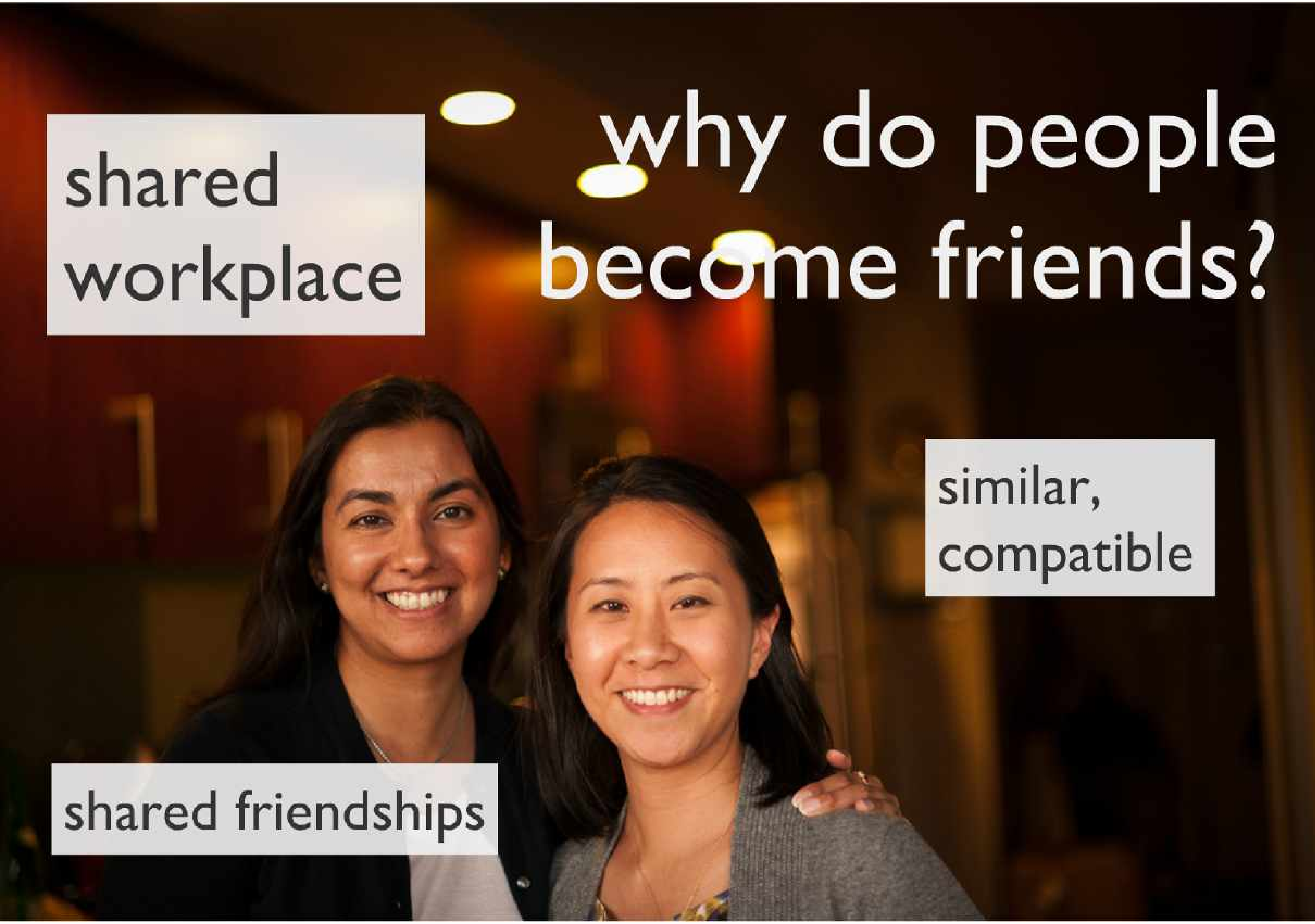
we have a cycle with an odd number of negative edges

If the graph contains a cycle with an odd number of negative edges, then this implies the graph is not balanced.

Claim

A signed graph is weakly balanced if and only if it contains no cycle with an odd number of negative edges.

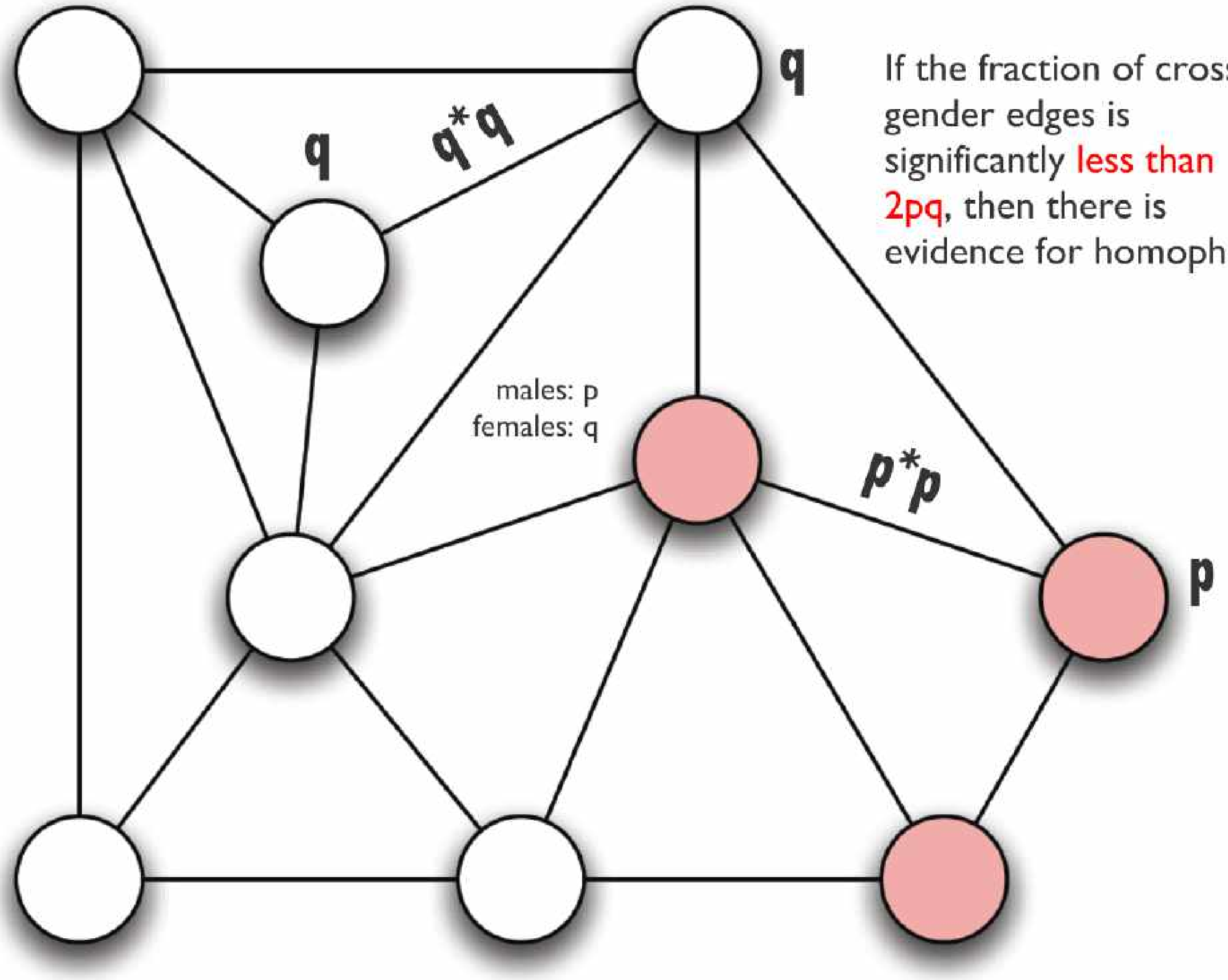

homophily
Is there a simple test for homophily?
What should we do to stop drug abuse?
Homophily
Why are these two friends similar?
selection
socialization
Obesity Study
foci
Empirical Data Analysis
what are some caveats with online data analysis?


why do people become friends?
shared workplace
similar, compatible
shared friendships

Homophily

The idea that
like minded
individuals tend
to be friends


Is there a
simple test
for
homphily?



If the fraction of cross-gender edges is significantly less than 2pq, then there is evidence for homophily

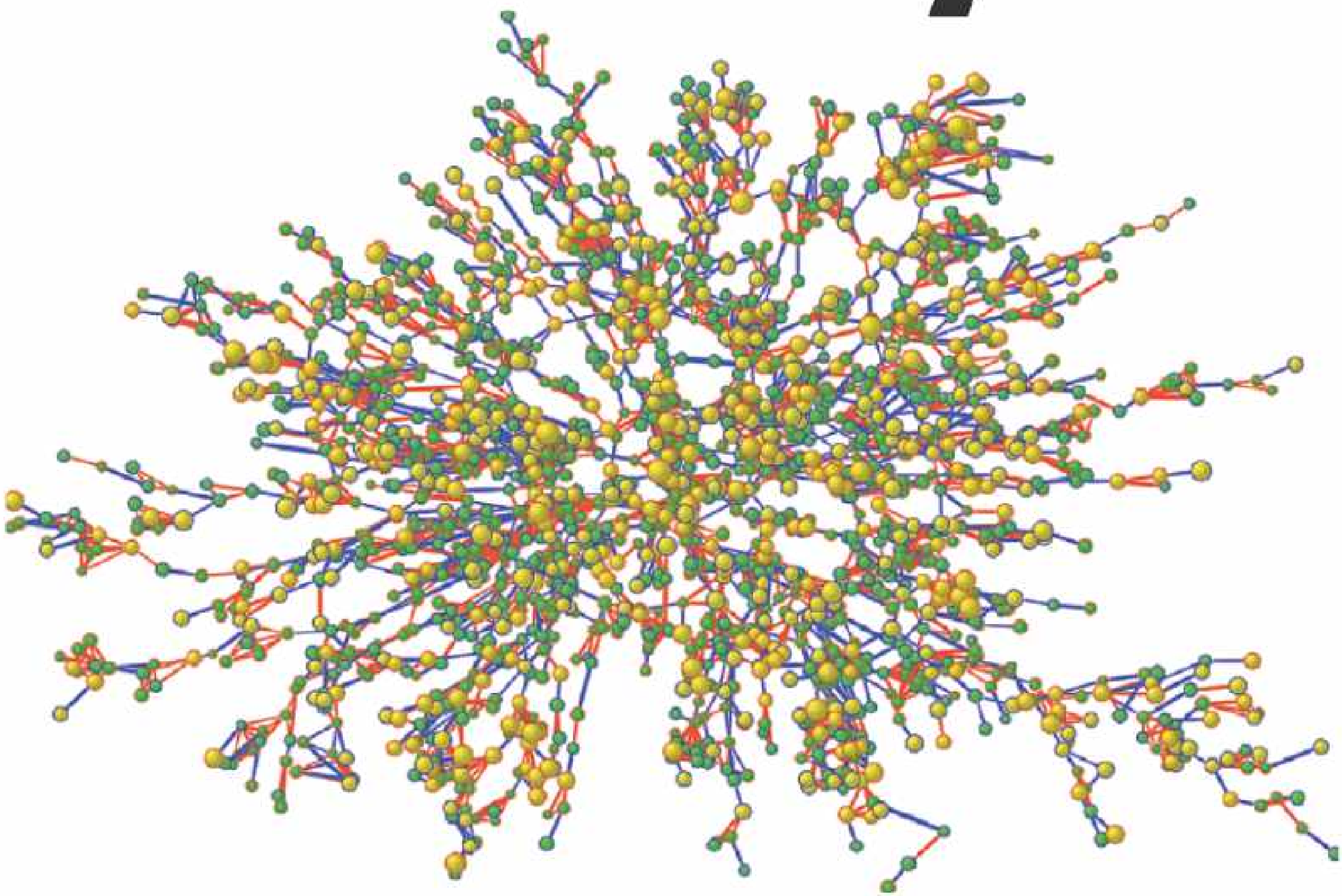

Why are these two friends similar?
selection
socialization

what is the interplay between socialization and selection?

can we distinguish between the two?


what should
we do to stop
drug abuse?

Obesity Study



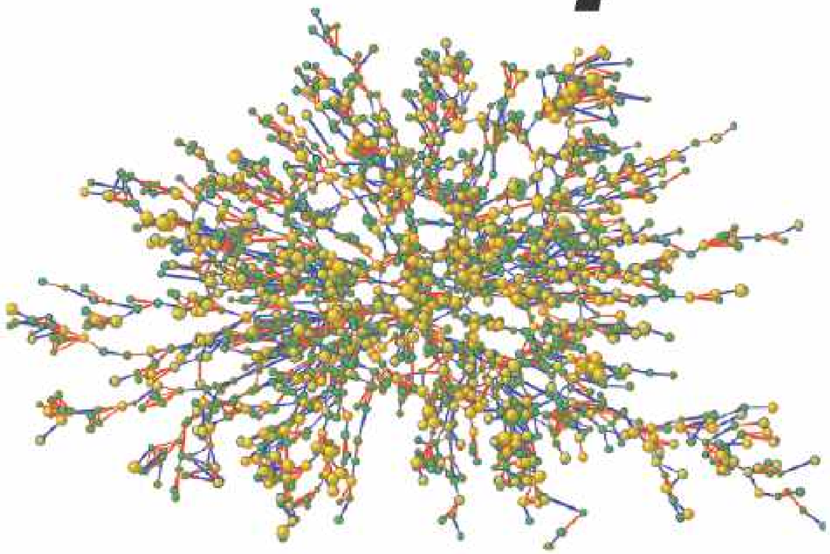
Each circle (node) represents one person in the data set. There are 2200 persons in this sub-component of the social network. Circles with red borders denote women, and circles with blue borders denote men. The size of each circle is proportional to the person's body-mass index. The interior color of the circles indicates the person's obesity status: yellow denotes an obese person (body-mass index, ≥30) and green denotes a non-obese person. The colors of the ties between the nodes indicate the relationship between them: purple denotes a friendship or marital tie and orange denotes a familial tie.

12,000 people, 32 years

why did obese and non-obese people show homophily?

- selection
- confounds
- socialization

Obesity Study



Each circle (node) represents one person in the data set. There are 2200 persons in this sub-component of the social network. Circles with red borders denote women, and circles with blue borders denote men. The size of each circle is proportional to the person's body-mass index. The interior color of the circles indicates the person's obesity status: yellow denotes an obese person (body-mass index, ≥30) and green denotes a non-obese person. The colors of the ties between the nodes indicate the relationship between them: purple denotes a friendship or marital tie and orange denotes a familial tie.

12,000 people,
32 years

why did obese and non-obese people show homophily? →

- selection
- confounds
- socialization

N. A. Christakis and J. H. Fowler. The spread of obesity in a large social network over 32 years. New England journal of medicine, 357(4):370–379, 2007.

Are Your Friends Making You Fat?

by CLIVE THOMPSON, nytimes.com

September 10th 2009

EILEEN BELLOLI KEEPS very good track of her friends. Belloli, who is 74, was born in Framingham, Mass., which is where she met her future husband, Joseph, when they were both toddlers. ("I tripped her and made her cry," recalls Joseph, a laconic and beanpole-tall 76-year-old.) The Bellolis never left Framingham, a comfortable, middle-class town 25 miles west of Boston — he became a carpenter and, later, a state industrial-safety official; and after raising four children, she taught biology at a middle school. Many of her friends from grade school never left Framingham, either, so after 60 years, she still sees a half dozen of them every six weeks.

examining the interplay between selection and socialization

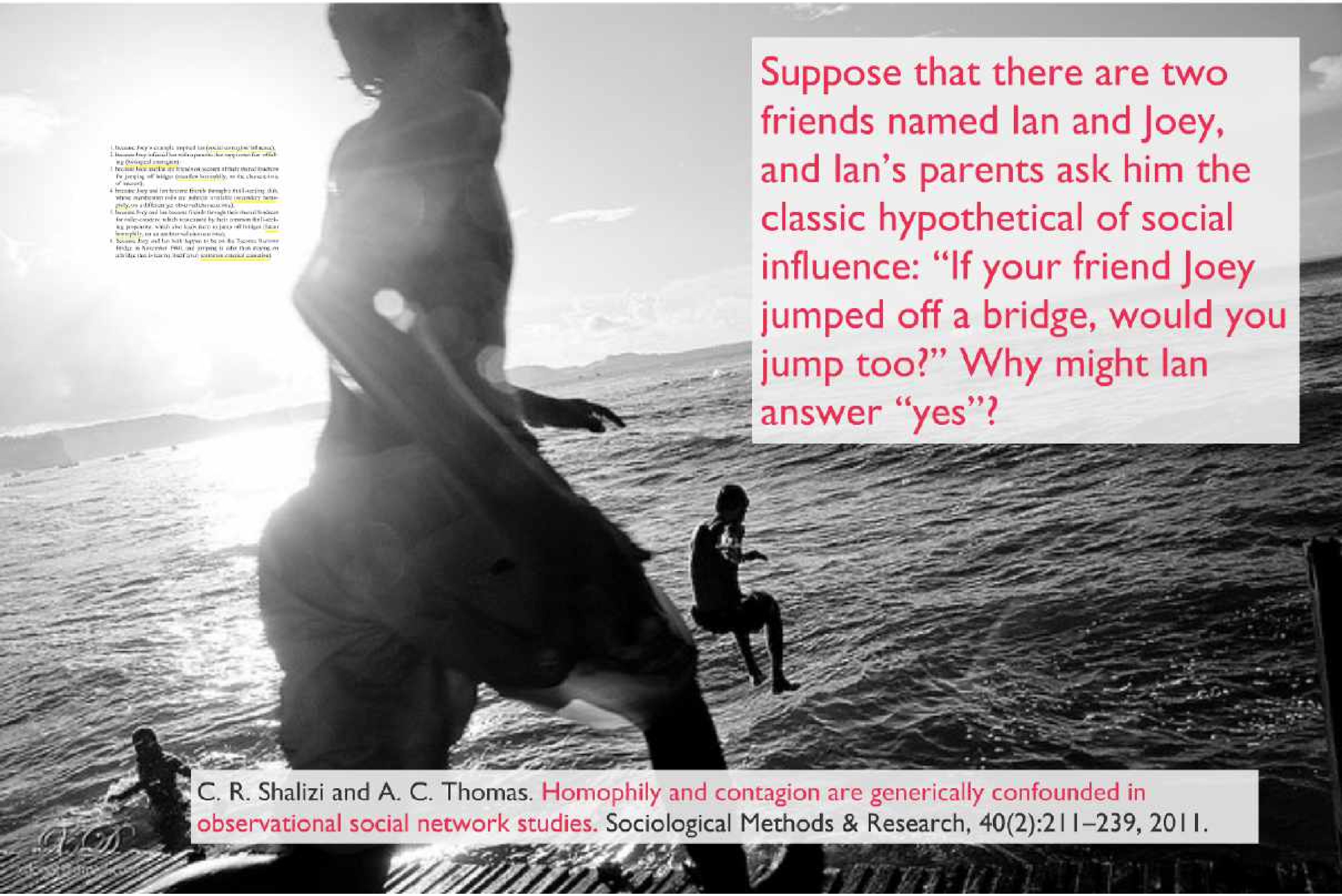
D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri. Feedback effects between similarity and social influence in online communities. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 160–168. ACM, 2008.

Catching Obesity From Friends May Not Be So Easy

by GINA KOLATA, nytimes.com

August 8th 2011

Does obesity spread like a virus through networks of friends and friends of friends? Do smoking, loneliness, happiness, depression and illegal drug use also proliferate through social networks?

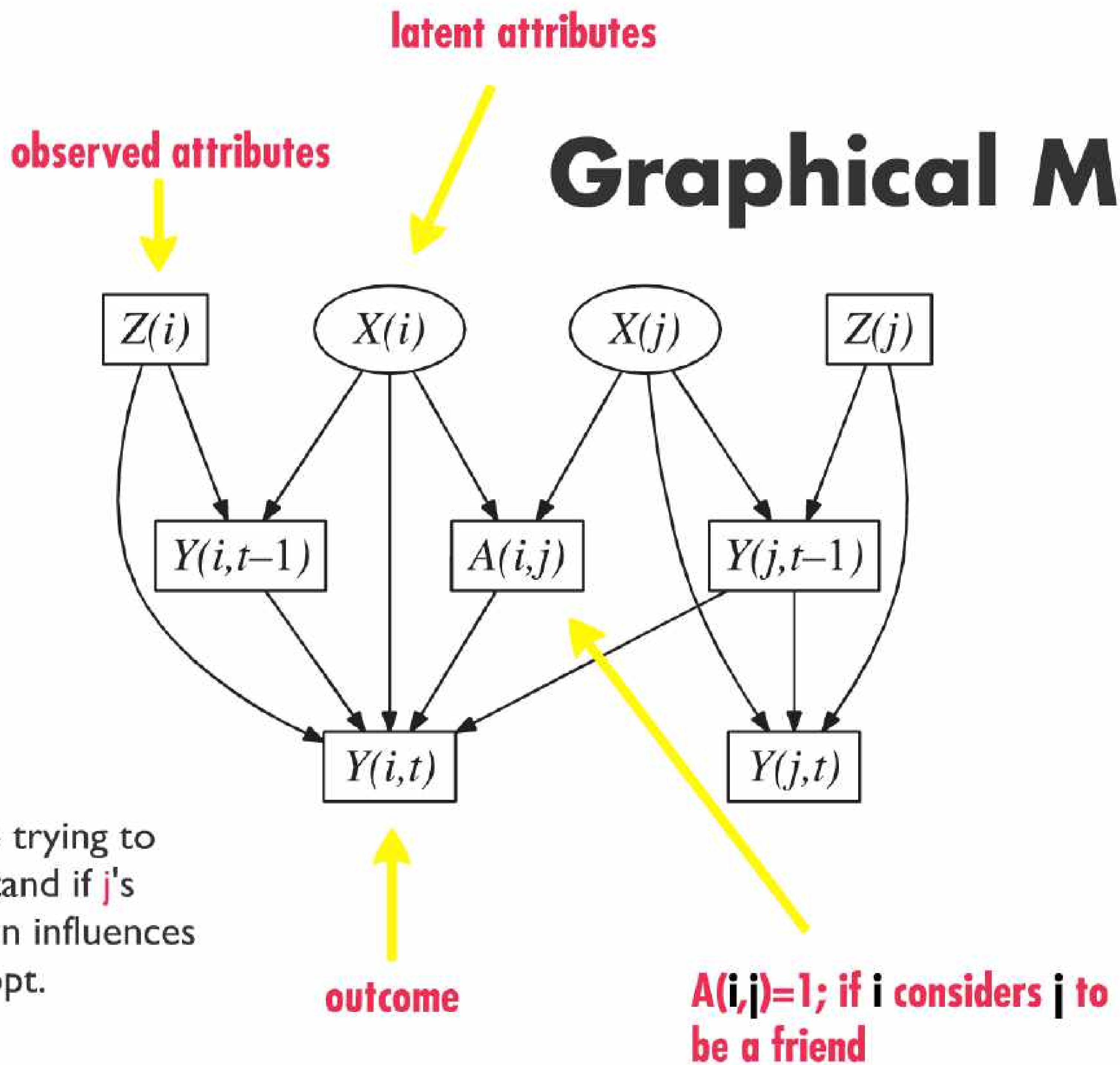


Suppose that there are two friends named Ian and Joey, and Ian's parents ask him the classic hypothetical of social influence: "If your friend Joey jumped off a bridge, would you jump too?" Why might Ian answer "yes"?

C. R. Shalizi and A. C. Thomas. Homophily and contagion are generically confounded in observational social network studies. Sociological Methods & Research, 40(2):211–239, 2011.

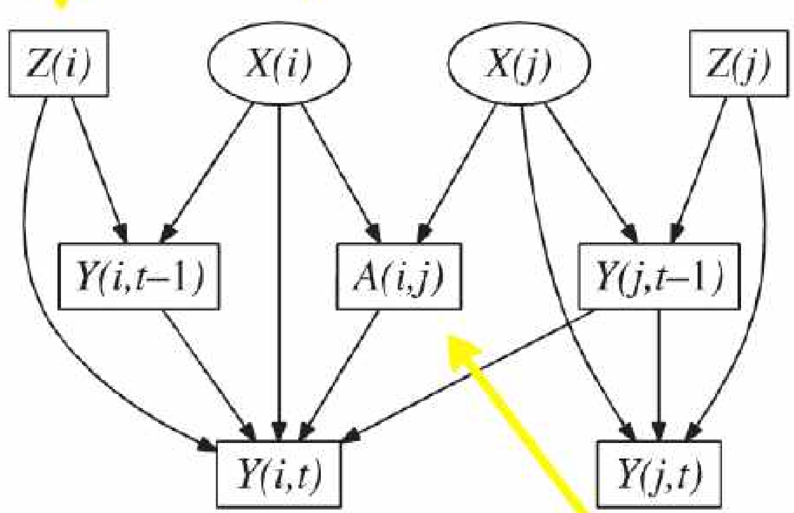
1. because Joey's example inspired Ian (social contagion/influence);
2. because Joey infected Ian with a parasite that suppresses fear of falling (biological contagion);
3. because Joey and Ian are friends on account of their shared fondness for jumping off bridges (manifest homophily, on the characteristic of interest);
4. because Joey and Ian became friends through a thrill-seeking club, whose membership rolls are publicly available (secondary homophily, on a different yet observed characteristic);
5. because Joey and Ian became friends through their shared fondness for roller-coasters, which was caused by their common thrill-seeking propensity, which also leads them to jump off bridges (latent homophily, on an unobserved characteristic);
6. because Joey and Ian both happen to be on the Tacoma Narrows Bridge in November 1940, and jumping is safer than staying on a bridge that is tearing itself apart (common external causation).

Graphical Model

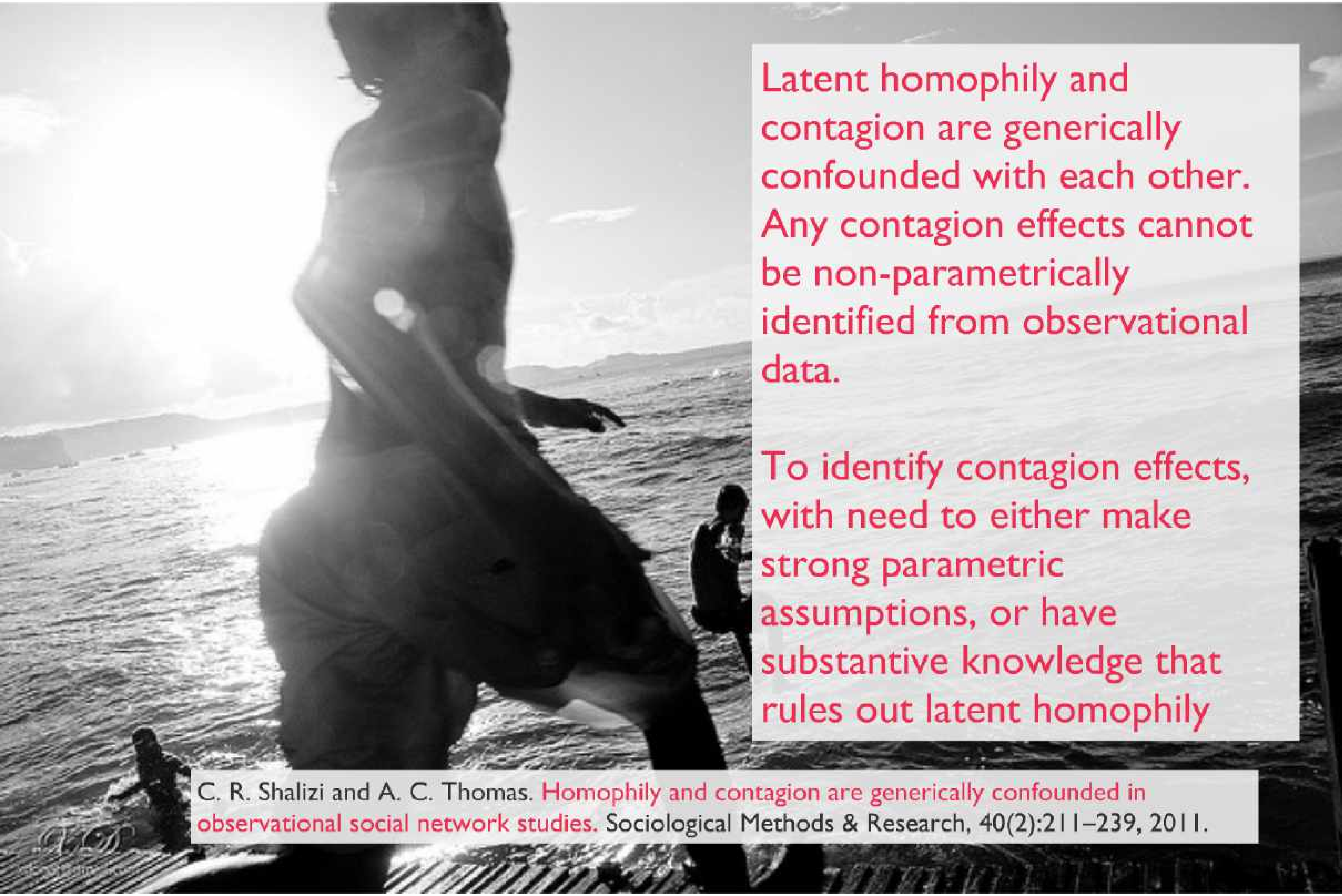


We are trying to understand if j's adoption influences i to adopt.

Graphical Model



We are trying to understand if j's adoption influences i to adopt.



Latent homophily and contagion are generically confounded with each other. Any contagion effects cannot be non-parametrically identified from observational data.

To identify contagion effects, with need to either make strong parametric assumptions, or have substantive knowledge that rules out latent homophily

C. R. Shalizi and A. C. Thomas. Homophily and contagion are generically confounded in observational social network studies. Sociological Methods & Research, 40(2):211–239, 2011.