

Social and Information Networks

University of Toronto CSC303
Winter/Spring 2020

Week 4: (2020)

Announcements

The first assignment now has six questions. There might be one more question. The first assignment is due at 2:59PM, February 14 on Markus. I moved up the time from 11:59PM to 2:59 PM so that the TA can go over any questions about solutions to the first assignment.

On piazza there was a question about 5(e). Using different names, I mean that if R has friends S and T using the method then R will start using the method. Questions?

During the week of February 10, I will lecture on Monday and Wednesday (in GB248). During the week of February 24, I will lecture on Wednesday and Friday.

We will have a common tutorial on Friday, February 14 in GB248. There will not be a tutorial during the week of February 24. After this week, we return to the standard schedule of lectures on M,F and tutorials on Wednesdays in GB248 and SS1070.

This weeks agenda

This week we continue to discuss Chapter 4 of the text. In particular, we will discuss the following:

- The probability of triadic closure (resp. focal closure, membership closure) as a function of the number of common friends (resp. the number of common interests (foci), the number of friends in a given focus)
- Schelling's segregation model.
- Chapter 5 and structural balance
- Strong structural balance and weak structural balance.

Note: Please let me know about typos (and especially errors) in the lecture slides.

LiveJournal

Note: Last lecture I said that I wasn't sure LiveJournal still existed but indeed it does exist. From the web page:

“LiveJournal is a unique place where people share their life stories, give advice and exchange ideas ”

From Wikikpedia:

LiveJournal is a Russian social networking service where users can keep a blog, journal or diary.

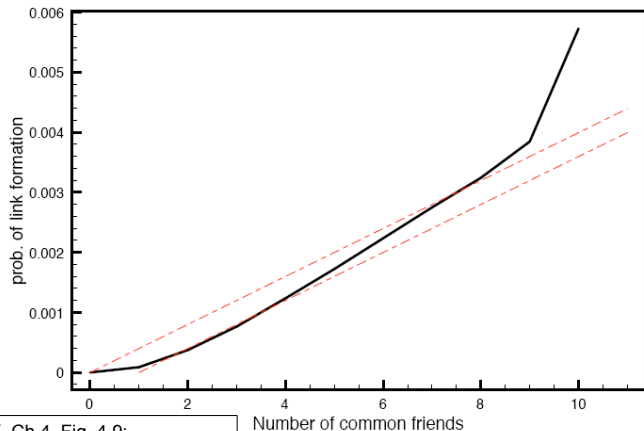
American programmer Brad Fitzpatrick started LiveJournal on April 15, 1999, as a way of keeping his high school friends updated on his activities.[6] In January 2005, American blogging software company Six Apart purchased Danga Interactive, the company that operated LiveJournal, from Fitzpatrick.

Six Apart sold LiveJournal to Russian media company SUP Media in 2007

Triadic closure: dependence on number mutual friends

- Email exchanges (over 60 days) by 22,000 students in large US university [Kossinets, Watts 2006]
- “Friends” defined as two-way email communication (prev. 60 days)
- Measure probability $T(k)$ of a new friendship emerging between a pair of students as a function of the number k of mutual friends
- That is, the probability of it happening in any given day (averaging over many such pairs)
- Compare data (black) with baseline theoretical model (red) baseline: **assume** any single mutual friend will generate a new friendship with probability p and that this will happen *independently* for each common friend. Thus $T(k) = 1 - (1 - p)^k$ **Why?**
- For **small** p , $(1 - p)^k \approx 1 - pk$ so that $T(k) \approx pk$.

Probability (per-day) of triadic closure as a function of the number of common friends



[E&K, Ch.4, Fig. 4.9;
from Kossinets and Watts, 2006]

Figure: [E&K, Fig 4.9]

Observations

- Data does not show much more propensity for friendship when going from zero to one mutual friend.
 - ▶ The second dashed red line shifts the curve over by one friend so as to better compare the actual data and baseline model.
 - ▶ Why no major impact with one common friend?
- Increasing from 1 to 9 friends shows linear curve (greater slope than baseline)
- A sharp difference going beyond 9 friends
 - ▶ The theoretical model (and its assumption of independence) no longer supported.
 - ▶ Is there some threshold of mutual friends which escalates the pressure for triadic closure?

Exercise: translate per-day probability into per-month or per-year probability

Probability of focal closure as a function of the number of common classes

Kossinetts and Watts also studied focal closure where a focus means a class in which a student is enrolled.

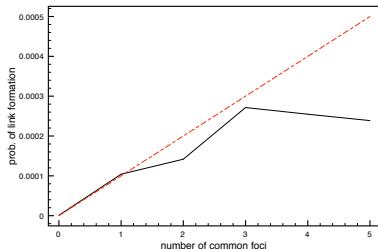


Figure: [E&K, Fig 4.10]

Clearly the theory and the actual data do not correspond especially when considering students going from 3 to 4 common classes. **Can you speculate on a reason?**

Probability of focal closure as a function of the number of common classes

Kossinets and Watts also studied focal closure where a focus means a class in which a student is enrolled.

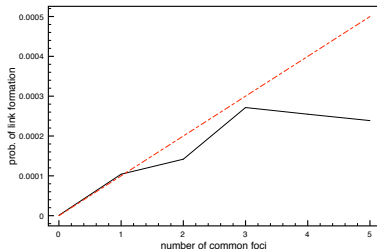


Figure: [E&K, Fig 4.10]

Clearly the theory and the actual data do not correspond especially when considering students going from 3 to 4 common classes. **Can you speculate on a reason?** If you haven't formed a friendship having attend 3 classes together, then perhaps there is a reason?

Probability of membership closure as a function of the number of common friends

The text presents two studies of membership closure where there is data concerning both person-to-person interactions and person-foci affiliations. The first study shows the probability of joining the blogging site LiveJournal (<https://www.livejournal.com>) where “friendship” is self-identified within a user’s profile.

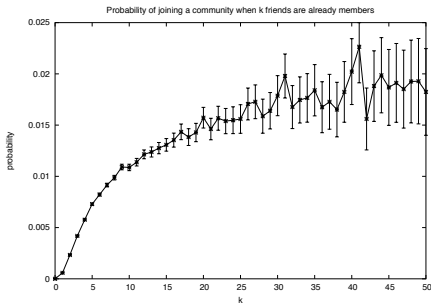


Figure: [E&K, Fig 4.11]

Second study of membership closure as a function of the number of common friends

The second study concerns Wikipedia editors and foci are specific Wikipedia pages. Here “friendship” is defined as having communicated together on a user-talk page and membership in a foci corresponds to having edited a Wikipedia page.

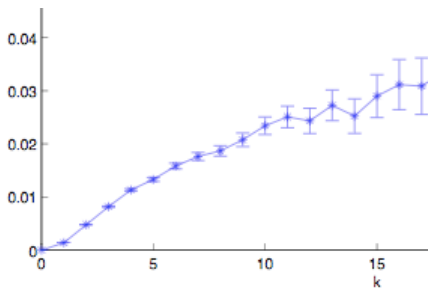


Figure: [E&K, Fig 4.12]

The interplay between selection and influence

Using the same Wikipedia data as in the previous focal closure example, The text presents one study that speaks to the manner in which selection and influence combine to result in observed homophily. Once again, the nodes are Wikipedia editors, the foci are articles, and edges correspond to communication via a user-talk page.

In addition, the study defines a numerical similarity measure between two users A and B as a small variation on the following ratio which is analogous to the way neighbourhood overlap was defined:

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

Fortunately, every action on Wikipedia is recorded and time-stamped so it is possible to conduct a meaningful longitudinal study by looking at each “time step” defined by an “action” of an editor where an action is either an article edit, or a communication.

Average level of similarity before and after the first Wikipedia communication

The figure below plots the level of similarity as a function of the number of edits before and after the first communication. Time 0 is defined to be the time of the first interaction between a pair (A, B) of editors. This is then averaged over all the (A, B) plots.

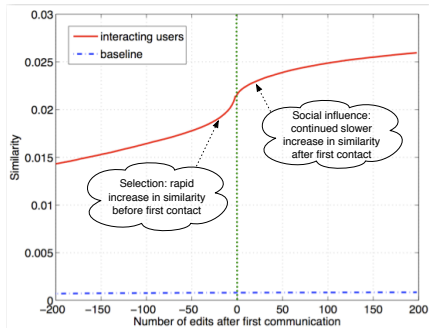


Figure: [E&K, Fig 4.13]

Observations on similarity vs. interactions (Figure 4.13)

There are a number of interesting observations and caveats regarding Figure 4.13. First some notable observations.

- The level of similarity is increasing over “time” before and after the first interaction.
- The steepest increase in similarity occurs just before the first interaction suggesting that selection is playing a pronounced role in forming this “friendship link” in the networks that are being dynamically created.
- The bottom dashed line indicates the level of similarity for those who never communicate. Clearly those who eventually interact evidence more similarity suggesting some significant similarity factors outside of what is being studied.

Some caveats

- Like any averaging of individual data, we cannot say why any particular pair of editors have decided to communicate.
- Because the defined time 0 corresponds to different moments in “real time” for each pair, we cannot understand to what extent real time events may also be a factor leading communication.
- In this study, links are never eliminated. Other “fully dynamic” network settings would have node and/or links that are not permanent.
- The biggest question about such a study is the extent to which any observations may or may not extend to different settings. In what settings do we have the same kind of detailed time stamping of events?

Is some degree of segregation “natural”?

A dramatic example of homophily (i.e., we tend to live “adjacent” to people similar to ourselves) can be found at the end of Chapter 4. The chapter ends with a discussion of Schelling’s segregation model that provides an explanation as to how racial neighbourhood segregation can evolve when driven by individuals wanting to be near “people similar to themselves”. Schelling’s model and his simulations led him to a fundamental observation:

Segregation can and will happen even if there is no explicit individual desire to avoid (say) people of a different race. All that is needed is some desire to be near enough similar people.

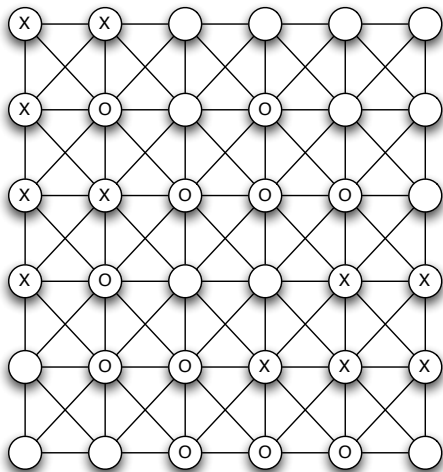
This observation isn’t restricted to racial segregation but as can also witness neighbourhoods that are largely or significantly based on ethnicity.

In addition to the importance of this fundamental observation, the model provides an interesting study of network dynamics, homophily driven by selection, and how local decisions lead to global structure in a network.

Of course, Schelling’s model does not preclude the presence of other economic and political factors, nor does it preclude explicit racism.

The Schelling model

The model itself is quite simple but still hard to analyze analytically. In this model, we view two classes of individuals (X and O) living in a grid. More specifically, individuals occupy some subset of the nodes as depicted in figure 4.15 of the text.



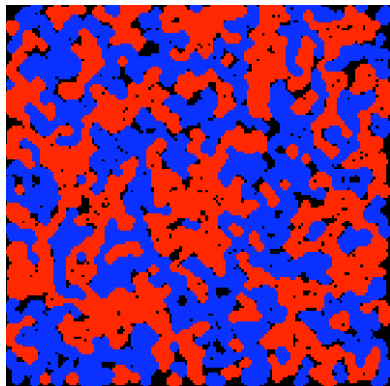
The dynamics of the Schelling model

Schelling then hypothesizes that every individual wants to have at least some threshold t of his/her neighbours to be immediate neighbours. When an individual's threshold is not met, they move. There are different versions of the model depending on the order in which individuals move and where they randomly move to in order to satisfy the desire for similarity. The claim is that the results do not change qualitatively.

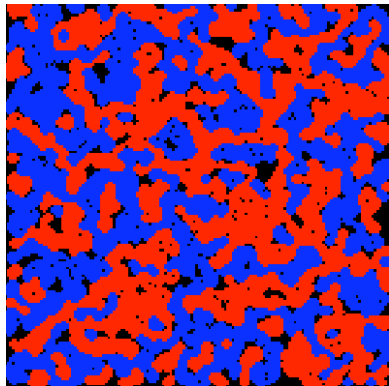
To observe the dynamics, simulations of the network are conducted for different threshold values. What is very apparent is the segregated structure of the network as it evolves.

The specific grid is a 150 by 150 grid (i.e., 12,500 cells, with 10,000 cells occupied) with both groups equally represented. The following figures show the results for thresholds $t = 3$ (i.e. an individual desires less than a majority of his/her neighbours to be similar) and $t = 4$.

Simulations for $t = 3$



(a) *A simulation with threshold 3.*



(b) *Another simulation with threshold 3.*

Figure 4.17: Two runs of a simulation of the Schelling model with a threshold t of 3, on a 150-by-150 grid with 10,000 agents of each type. Each cell of the grid is colored red if it is occupied by an agent of the first type, blue if it is occupied by an agent of the second type, and black if it is empty (not occupied by any agent).

Simulation for $t = 4$

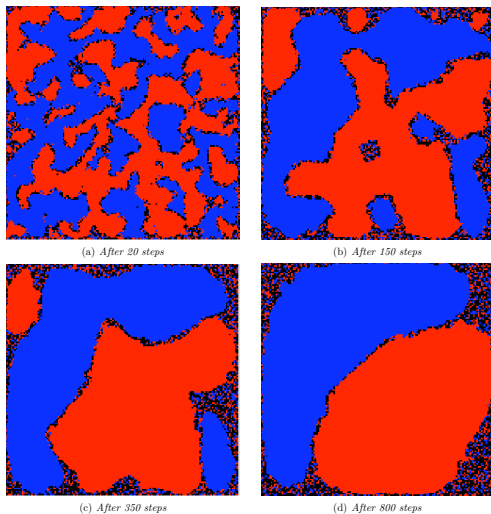
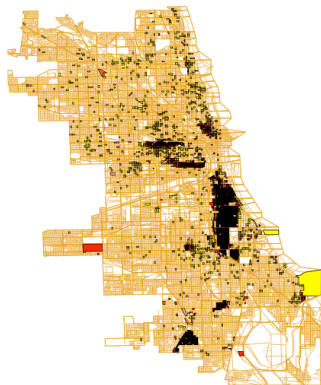


Figure 4.19: Four intermediate points in a simulation of the Schelling model with a threshold t of 4, on a 150-by-150 grid with 10,000 agents of each type. As the rounds of movement progress, large homogeneous regions on the grid grow at the expense of smaller, narrower regions.

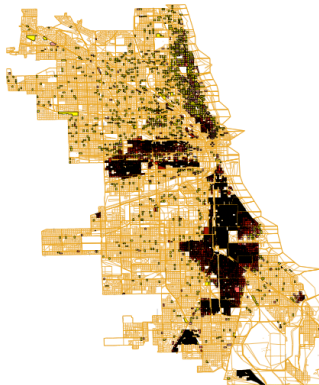
Some concluding comments on the Schelling study

- The model is not constructed so as to build in segregation. More specifically, the model allows for stable configurations that are well integrated.
- However, given a random starting configuration, the simulations show that people will gravitate to a segregated structure.
- There is a compounding effect of the model dynamics. Namely, when one person leaves, it can result in other neighbours falling below their threshold and hence a new desire to leave the current location has been created.
- The word segregation is a term with a very negative connotation due to the use of the term with respect to racial (e.g., Jim Crow legislation in the US) and religious segregation (e.g., ghettos in Europe) which was forced by governments.
- What if we used the word "clustered" instead of segregated? Do we think that neighbourhoods that are concentrated along say ethnic lines is a bad thing?

The reality of neighbourhood segregation in Chicago (1970s)



(a) *Chicago, 1940*



(b) *Chicago, 1960*

And how integrated or segregated is Toronto?

At some level (i.e., Metro Toronto), Toronto may be the most ethnically diverse city as is claimed. But at a more detailed level, many neighbourhoods are far from being “integrated”.

It seems much easier to talk about Chicago (as we all know about racial segregation in the US) but perhaps more difficult to talk about Canada and Toronto. I am posting a September 2018 newspaper article and a February 2019 talk by David Hulchanski (UT Faculty of Social Work) that describes the changes in income levels and neighbourhoods in Toronto. The title of the article more or less summarizes his major conclusion: “Toronto is segregated by race and income. And the numbers are ugly”.

I call attention to this (and other similar studies) to indicate that while homophily is a factor (especially with regard to ethnicity), there are clearly many other factors that are prevalent and arguably dominant.

Structural balance: positive and negative links

Now for a relatively less sensitive (political) topic, who we like and dislike. As previously mentioned so far we have restricted attention to social networks where all edges reflect some positive degree of friendship, collaboration, communication, etc.

Chapter 5 now explores some interesting aspects of networks where edges can be both positive and negative. This is, of course, quite natural in that people (countries) often have enemies as well as friends (allies). We also have companies that can be aligned in some way or can be competitors. Here the meaning of a edge can seemingly change any general observations more than when there are only positive edges.

Following the development stemming from the distinction between strong and weak ties, we would like to see what we can infer about a network given that some edges are positive and some are negative. More specifically, what can be assumed from certain types of triadic closures? How can local properties (e.g., how edges of a triangle are labelled) can have global implications (i.e., provable results about network structure)?

Some initial assumptions

We start with a strong assumption:

Assume the network is a complete (undirected) graph. That is, as individuals we either like or dislike someone. Furthermore, this is not nuanced in the sense that there is no differentiation as to the extent of attraction/repulsion).

Later in the chapter, the text considers the issue of networks that are not complete networks. The text also reflects a little on the nature of directed networks (when discussing the *weak balance property*) but essentially this chapter is about undirected networks.

Note: We can assume the graph is connected since otherwise we can consider each connected component separately.

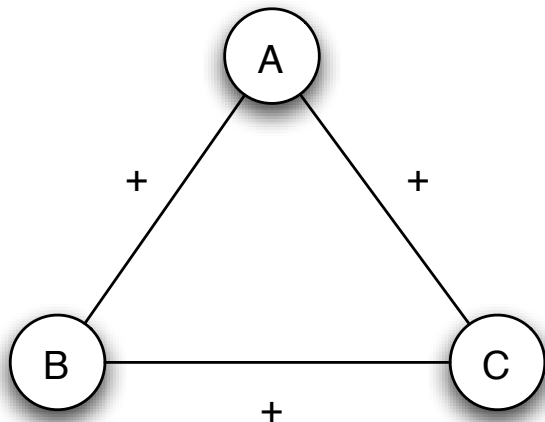
Types of instability

Thinking of networks as people with likes and dislikes of other people (rather than some other possible interpretations), we can consider the 4 different types of labelled triangles in the graph, depending on the number of positive (+) and negative (-) edges. That is, any completely labelled triangle can have 0,1,2, or 3 positive edges and due to the symmetry of a triangle that is all the information we have about any particular triangle.

Using a central idea from social psychology, two of the four triangle labellings are considered relatively stable (called *balanced*) and the other two relatively unstable (*not balanced*).

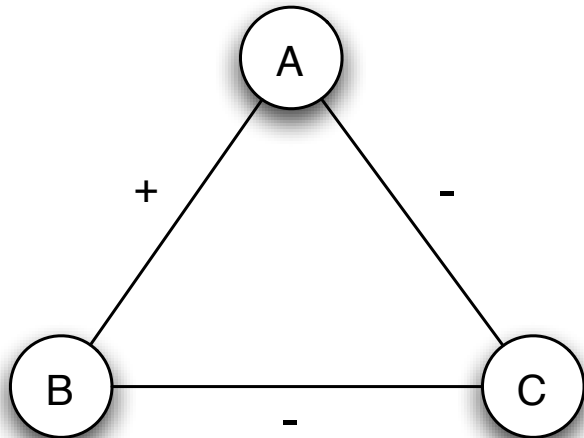
Here follows the four types of triangles as depicted in Figure 5.1 of the text:

A natural stable configuration



In this case, A, B, C are mutual friends and that naturally indicates that they would likely remain so.

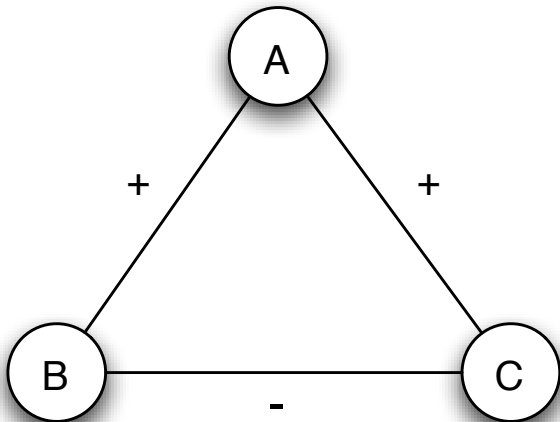
The second stable configuration



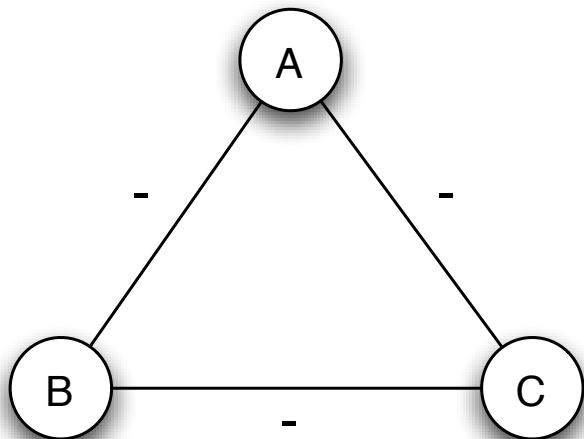
This may be a slightly less obvious stable situation where *A* and *B* are friends and if anything that friendship is reinforced by a mutual dislike for *C*.

A natural unstable configuration

In this case, A has two friends B and C who unfortunately do not like each other. The claim here is that the stress of this situation will encourage A to either try to have B and C become friends or else for A to take sides with B or C and thus eliminate a friendship so as to move toward the previous stable configuration.

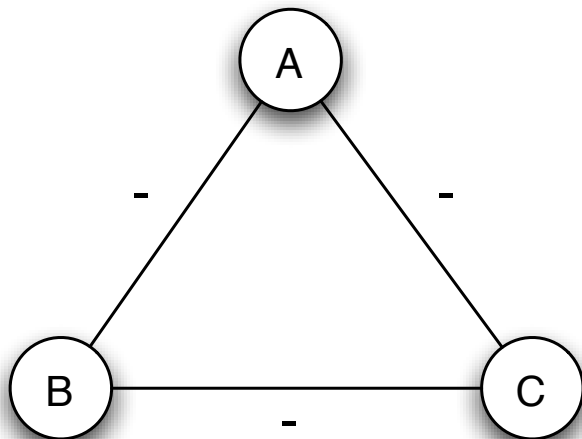


A somewhat less obvious unstable configuration



Why is this called unstable?

A somewhat less obvious unstable configuration



Why is this called unstable? The instability here is sometimes explained by the phenomena that “the enemy of my enemy becomes my friend” as we sometimes see in international relations. This is less convincing than the other type of instability and we will return to this situation soon.

The strong structural balance property

The underlying behavioural theory is that these unstable triangles cause stress and hence the claim that such unbalanced triangles are not common.

In order to try to understand if this theory tells us anything about the global structure of the network, we can make the following strong balance assumption (much as we made the strong triadic closure assumption).

Strong structural balance property: Every triangle in the network is balanced.

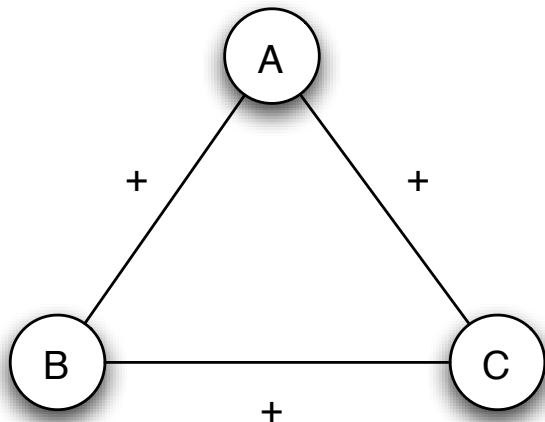
Recall that we started off with the assumption that the network is a complete graph with every edge labelled so we are assuming a property for all n choose 3 triangles. Of course, we cannot expect this property to hold but just as the strong triadic closure property was an extreme assumption, we can hope that this strong assumption will also suggest or predict useful information about the network.

End of Monday, January 27 lecture and today's agenda

We ended here but noted that I skipped over the Schelling segregation model which we will do at the end of today's lecture.

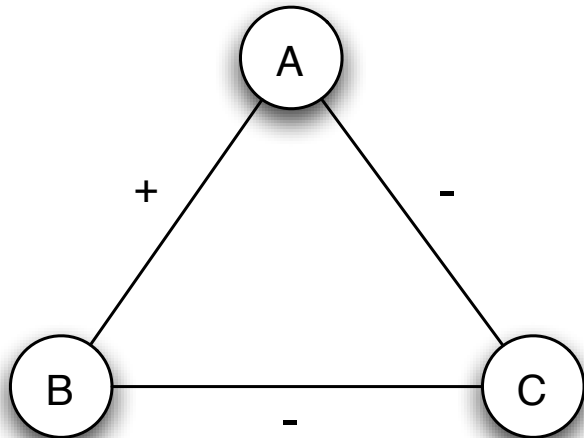
- Quick review of balanced vs unbalanced triangles. Balance as a type of equilibrium.
- Consequence of the strong structural balance property: A characterization of the networks that satisfy the property.
- Harary's Balance Theorem
- Determining when unlabelled edges can and cannot be labelled to create a strongly balanced network
- Weak structural balance
- Characterization of weak structural balance.
- The Schelling segregation model.

A natural stable configuration



In this case, A, B, C are mutual friends and that naturally indicates that they would likely remain so.

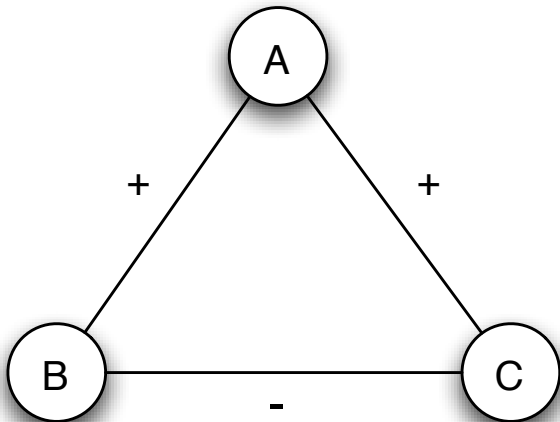
The second stable configuration



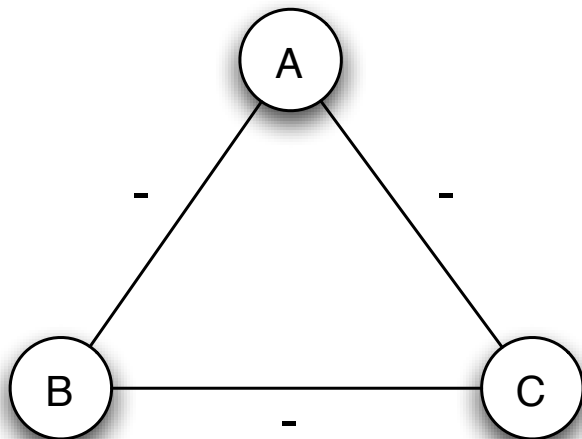
This may be a slightly less obvious stable situation where *A* and *B* are friends and if anything that friendship is reinforced by a mutual dislike for *C*.

A natural unstable configuration

In this case, A has two friends B and C who unfortunately do not like each other. The claim here is that the stress of this situation will encourage A to either try to have B and C become friends or else for A to take sides with B or C and thus eliminate a friendship so as to move toward the previous stable configuration.

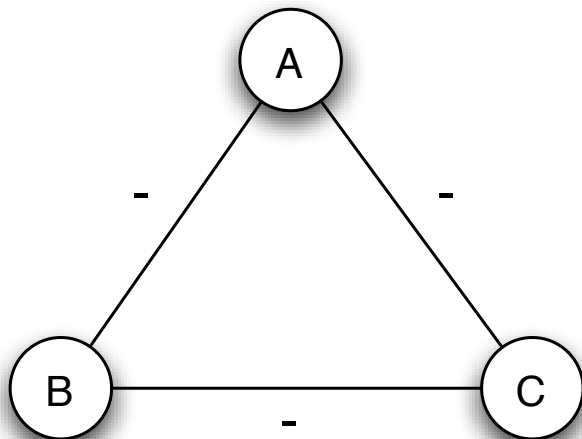


A somewhat less obvious unstable configuration



Why is this called unstable?

A somewhat less obvious unstable configuration



Why is this called unstable? The instability here is sometimes explained by the phenomena that “the enemy of my enemy becomes my friend” as we sometimes see in international relations. This is less convincing than the other type of instability and we will return to this situation soon.

Balance as a form of equilibrium

In our last lecture, while introducing these labelled triangles, it was suggested that one way to further justify the distinction between balanced and unbalanced triangles is to view balance (resp non balance) as a desirable (undesirable) situation.

In a balanced (resp. unbalanced) configuration, any single change in a relation (i.e. edge label) will lead to an unbalanced (resp. balanced) configuration.

That is, balance is a form equilibrium.

Later in the term, we will discuss stable matchings. (How many have seen this in CSC304 or elsewhere?) We view stable matchings as an equilibrium. In stable matchings (as in balanced triangles), it is a pair of “agents” that we consider in a single change. We discuss stable matchings later in this course.

Consequence of the strong structural balance property: A provable characterization of networks that satisfy the property

One simple (idealistic) way to construct a network satisfying the property is to assume that there are no enemies; everyone is a friend. Is this the only way?

Consequence of the strong structural balance property: A provable characterization of networks that satisfy the property

One simple (idealistic) way to construct a network satisfying the property is to assume that there are no enemies; everyone is a friend. Is this the only way?

Suppose that we had two communities of active political people (e.g. X = the “base” for candidate or political party R , and Y = the “base” for candidate or political party U). In the world of highly politicized politics, it isn't too far of a stretch to think that everyone within a community are friends and everyone dislikes people in the other community. This kind of network would also clearly satisfy the property.

So far then, we have two possibilities, the network is a clique with all positive edges, or the network is composed of two positive cliques with a complete bipartite graph of negative edges between the communities. Are there other possible ways to have the strong balance property?

Harary's Balance Theorem

Are there other possible ways to have the strong balance property?

Harary's Balance Theorem

Are there other possible ways to have the strong balance property?

Perhaps surprisingly, in a complete network, these two types of networks (no enemies and two opposing communities) are the only possibilities.

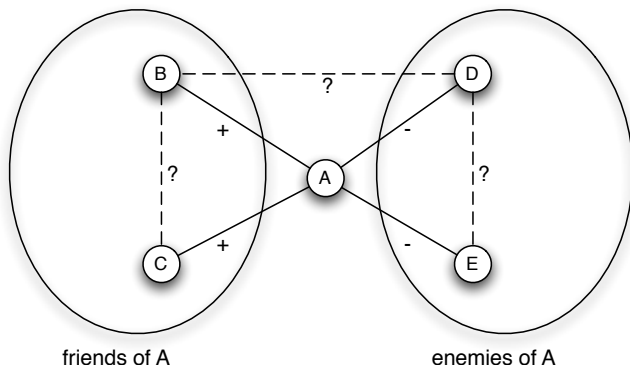
This is a theorem and the proof is not difficult as we will show using the figure 5.4 in the text.

Proof

We assume that the network satisfies the strong balance property. If there are no enemies, then we are done. So suppose there is at least one negative edge and for definiteness let's say that edge is adjacent to node A . Let X be all the friends of A and Y all of its enemies. So every node is in either X or Y since every edge is labelled.

Proof of balance theorem continued

Consider the three possible triangles as in the figure. It is easy to see that in order to maintain structural balance, B and C must be friends as must D and E , whereas B and D (also C and E) must be enemies.



Strong structural balance in networks that are not complete

We will depart from the order of topics in chapter 5, and consider the issue of networks that are not complete. Is there a meaningful sense in which a (non-complete) network is or is not structurally balanced?

One possibility is to ask whether or not there is a way to complete the graph so that it becomes structurally balanced. Of course, if there is already an unbalanced triangle then there is no way to complete the graph into one satisfying the strong balance property.

Aside: Of course, this immediately raises the question as to how many existing edge labels need to be changed so that a complete network is balanced (or an incomplete network can be made to be balanced)? And will networks tend to dynamically evolve into balanced networks. But for now we will assume that all existing labels are permanent.

How to label missing edges?

Note that when considering the strong triadic property, if all existing triangles satisfied the strong triadic property, then there was always a trivial way to assign labels to unlabelled edges by simply making each unlabelled edge a weak link.

Question: If all existing triangles are balanced, is there always a way to complete a network so as to form a strongly balanced network?

How to label missing edges?

Note that when considering the strong triadic property, if all existing triangles satisfied the strong triadic property, then there was always a trivial way to assign labels to unlabelled edges by simply making each unlabelled edge a weak link.

Question: If all existing triangles are balanced, is there always a way to complete a network so as to form a strongly balanced network?

It is easy to see that this is not always possible. For example, consider a network which is a 4 node cycle having 3 positive edges and one negative edge. Any way to label a “diagonal edge” will lead to an imbalance.

We are then led to the following

Question: Can we determine when there is an efficient algorithm to complete the network so as to satisfy the strong balance property? And if there is a completion, how efficiently can one be found?

Determining when and how to complete a network to satisfy the strong balance property

Clearly, if the existing edges are all positive links then there is a trivial way to complete the graph by simply making all missing edges to be positive edges.

So the interesting case is when there are existing negative edges. In this case, the characterization of strongly balanced networks tells us that when the graph is completed, the graph structure must be that of two opposing communities, with only positive edges within each community and only negative edges for links between the communities.

The previous example of a 4 node cycle is a clue as to how to proceed. That example can be stated as follows: if a network contains a 4 node cycle with one negative edge then it cannot be completed (to be strongly balanced) . More generally, if a network contains a cycle (of any length) with one negative edge, it cannot be completed. And even more generally, if a network contains a cycle having an odd number of negative edges it cannot be completed. **Why?**

Consequence of an odd cycle

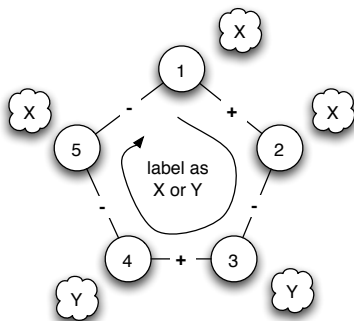


Figure 5.10: If a signed graph contains a cycle with an odd number of negative edges, then it is not balanced. Indeed, if we pick one of the nodes and try to place it in X , then following the set of friend/enemy relations around the cycle will produce a conflict by the time we get to the starting node.

The algorithm for determining if a partially labelled network can be completed to the strongly balanced

Lets call a cycle with an odd number of edges an odd cycle. The desired algorithm will either find an odd cycle (certifying that the network cannot be completed) or it will return a bipartiton of the nodes satisfying the Balance Theorem. This then also determines if a complete network is balanced.

We proceed as follows:

- Suppose $G = (V, E)$ is the given connected network and let $G^+ = (V, E^+)$ where $E^+ = \{e \in E \text{ such that } e \text{ is a positive link.}\}$
- We consider the connected components $\mathcal{C} = C_1, \dots, C_r$ of G^+ and let T_1, \dots, T_r be spanning trees for these componenets.
- Note that all edges between any C_i, C_j must be labelled as negative edges (or else they would have been merged into a larger connected component in $G^+..$

Completing the algorithm

- Otherwise, consider the graph $G^- = \{C, E^-\}$ whose nodes are the components of G^+ and whose edges are negative edges in G .
- Since G is connected, G^- is connected.
- if G^- has a cycle with an odd number of negative edges, then by following positive edges in each C_i we have such a cycle in G . We then again have a witness that G cannot be completed.
- Otherwise we are showing that G^- is bipartite and this gives us the bipartition we need for the balance theorem.
- A graph has an odd cycle iff the graph is not bipartite. Breadth first search can be used to determine whether or not a graph is bipartite (equivalently has a 2-colouring). Hence this development is efficient.

For the remainder of this lecture, we return to the assumption that our networks are undirected complete graphs.

Friends-enemies vs trust-distrust

There is always an ambiguity in social networks as to how to interpret links. Is a friend as we might traditionally mean a “good friend”, or is it a friend as in Facebook friends which in many cases are just people you know? And as we have seen we also use social network links to mean collaboration or communication rather than friendship.

This is both the power of network modeling (i.e., that results can carry over to different settings) and also the danger of misinterpreting results for one type of setting to apply to another.

In chapter 5, we see another instance of the ambiguity where instead of the friend-enemy relation, one can interpret an edge label as a trust-distrust relation.

To what extent should we expect intuition for friendship to carry over to trust? As discussed in the text, one distinction between these settings is that trust may be more of a directed edge concept relative to friendship. (Of course, even for friendship the relation may not be symmetric which is why maybe we should reserve the term of “friend” for a good friend.)

The ambiguity in the trust-distrust relation

Ignoring the fact that trust might not be at all symmetric, there is an additional ambiguity in the trust-distrust terminology. Namely, the text considers two possible interpretations that are meaningful even in the context of a single setting as in the online product rating site Epinions.

- 1 If trust is aligned with agreement on political issues as in the ratings of political commentary, then the four cases of balanced and unbalanced triangles still seem to apply. In particular, if A distrusts B and B distrusts C , it is reasonable to assume that A trusts C and hence a triangle having three negative labels is not stable.
- 2 However, if A distrusts B is aligned with A believing that he/she is more knowledgeable than B about a certain product, then a triangle having three negative labels is stable.

This suggests that it is reasonable to study a weaker form of structural balance.

A weaker form of structural balance

It is then interesting to consider a weaker form of structural balance where the only unstable triangles are those having two positive labels.

This leads to the following definition (analogous to the strong structural balance property):

A network satisfies the weak structural balance property if it does not contain any triangles with exactly two positive edges. This in turn leads to the following

Question: Is there a characterization of which (complete) networks satisfy the weak structural balance property?

A weaker form of structural balance

It is then interesting to consider a weaker form of structural balance where the only unstable triangles are those having two positive labels.

This leads to the following definition (analogous to the strong structural balance property):

A network satisfies the weak structural balance property if it does not contain any triangles with exactly two positive edges. This in turn leads to the following

Question: Is there a characterization of which (complete) networks satisfy the weak structural balance property?

Since every network that satisfies the strong balance property must also satisfy the weak balance property, the characterization of strongly balanced networks must be a special case of weakly balanced networks. Indeed we have the following characterization:

Theorem: A network $G = (V, E)$ satisfies the weak structural balance property iff $V = V_1 \cup V_2 \dots V_r$ such that all edges within any V_i are positive edges and all edges between V_i and V_j ($i \neq j$) are negative edges.

Proof of the characterization of weak structural balance

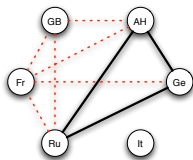
Clearly if the network $G = (V, E)$ has the network structure specified in the Theorem, then the network satisfies the weak balance property. The converse (that the weak balance property implies the network structure) is a reasonably simple inductive argument (say with respect to the number of edges).

Consider any node A and let X be all the friends of A . The following two claims are easy to verify:

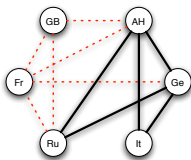
- Any $B, C \in X$ are friends
- If $B \in X$ and $D \notin X$, then B and D are enemies.

Upon removing the nodes in X , the induced network G' of the remaining nodes still must satisfy the weak structure balance property and hence by the induction hypothesis must have the stated network structure.

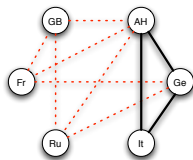
The evolution of European alliances preceding WW I



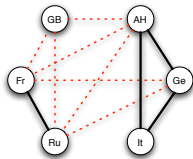
(a) *Three Emperors' League 1872-81*



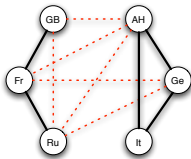
(b) *Triple Alliance 1882*



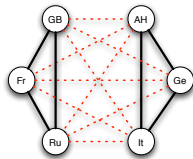
(c) *German-Russian Lapse 1890*



(d) *French-Russian Alliance 1891-94*



(e) *Entente Cordiale 1904*



(f) *British Russian Alliance 1907*

Figure 5.5: The evolution of alliances in Europe, 1872-1907 (the nations GB, Fr, Ru, It, Ge, and AH are Great Britain, France, Russia, Italy, Germany, and Austria-Hungary respectively). Solid dark edges indicate friendship while dotted red edges indicate enmity. Note how the network slides into a balanced labeling — and into World War I. This figure and example are from Antal, Krapivsky, and Redner [20].