Networks in Their Surrounding Contexts

Ana Paula Couto
Computer Science Department
Universidade Federal de Minas Gerais

Based on http://www.cs.uoi.gr/~tsap/teaching/2013-cs-l14/references.html
and https://www.ismll.uni-hildesheim.de/lehre/cmie-11w/script/lecture6.pdf

So far…

- Factors that affect the formation of links, based on similarities
  - Characteristics of the nodes
  - Behaviors and activities nodes engage in
- Contexts outside the network

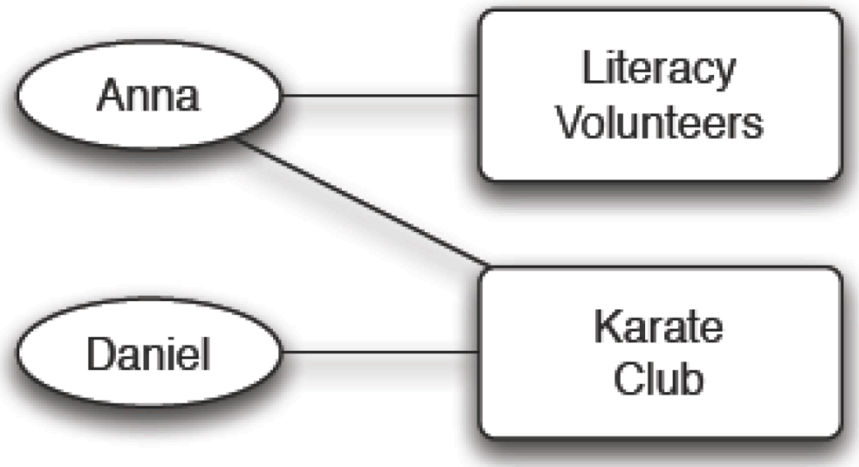
- It is possible to put these contexts into the network itself
  - Larger network: people and context as nodes

Affiliation

- Represent the set of activities a person takes part in (a general view of "activity")
  - Ex: part of a particular company, organization, frequenting a particular place, hobby
- Refer to activities as foci: "focal points" of social interaction

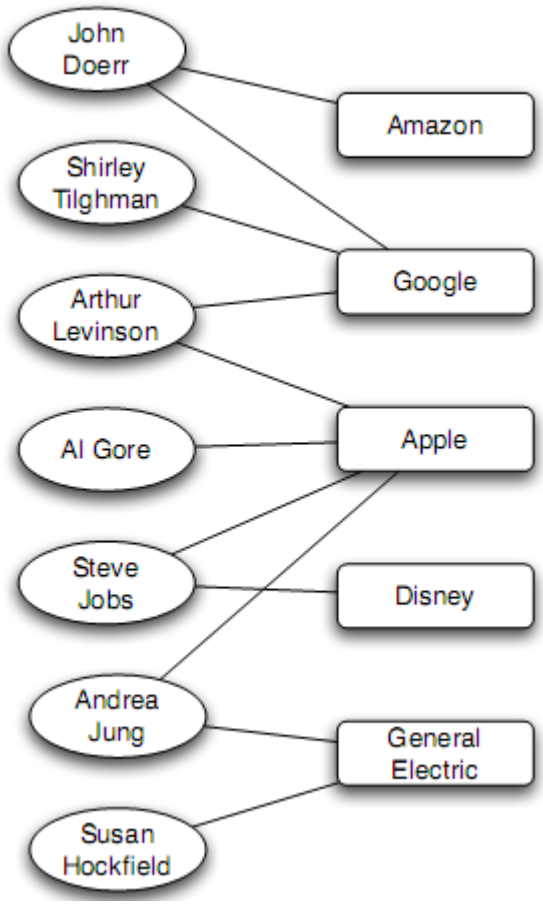
Affiliation

- Affiliation network:
  - bipartite graph
    - nodes divided into 2 sets
    - no edges joining a pair of nodes that belong to the same set
  - people affiliated with foci
- Ex:
  - Anna participates in both of the social foci on the right
  - Daniel participates in only one



Affiliation

Example:
Board of directors



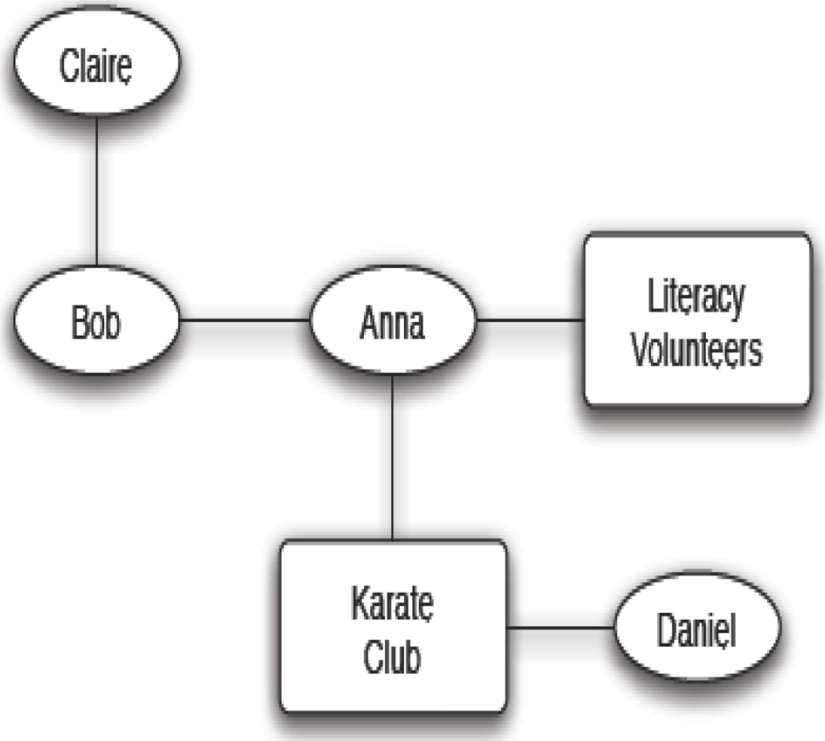
1) Companies implicitly links by having the same person sit on both their boards
2) People implicitly linked by serving together on a aboard

Co-Evolution of Social and Affiliation Networks

- Social networks change over time
  - new friendship links are formed
- Affiliation networks change over time
  - people become associated with new foci
- Co-evolution reflects interplay between selection and social influence
  - 2 people participate in a shared focus can become friends
  - if 2 people are friends, they can share their foci
- How to represent co-evolution with a single network?

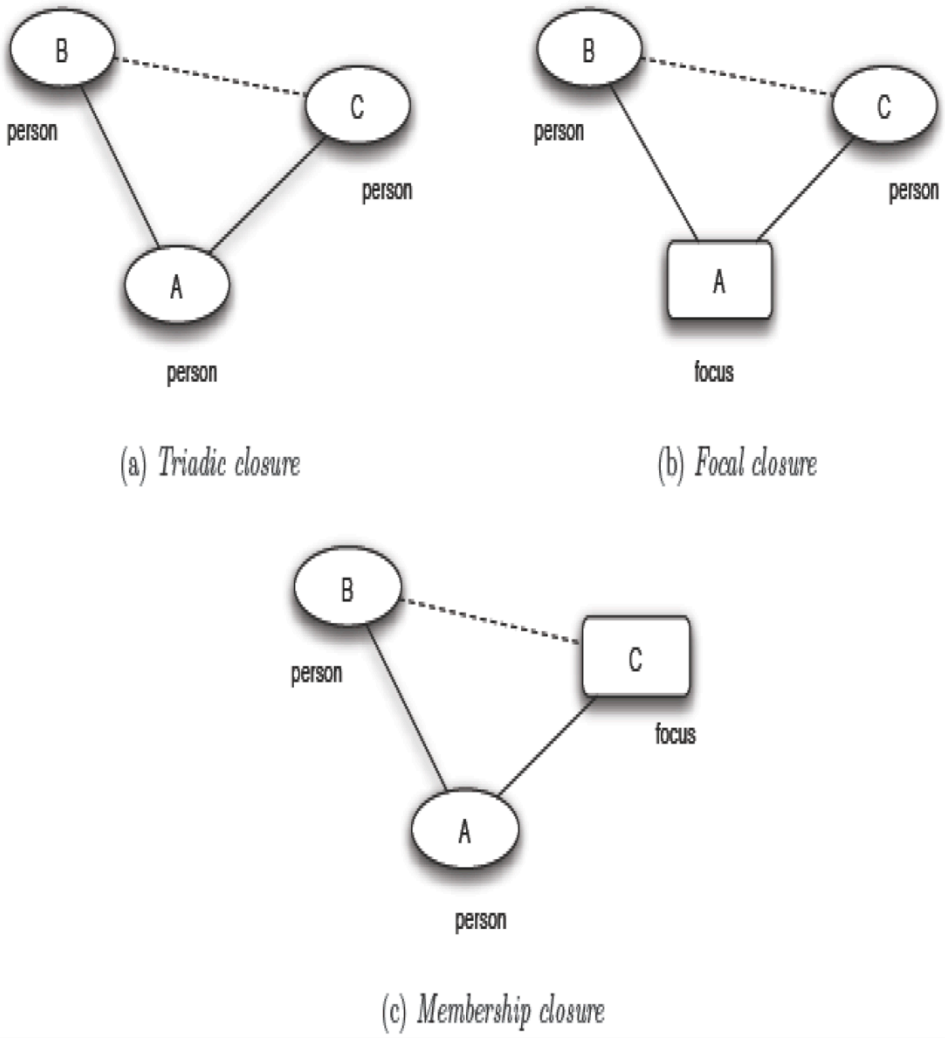
Social-affiliation Networks

- Social-affiliation network contains:
  - a social network on the people and
  - an affiliation network on the people and foci



Social-affiliation Networks

- In social-affiliation networks link formation as a closure process
- Several options for "closing" B-C
  - triadic closure: A, B, and C represent a person (already examined)
  - focal closure: B and C people, A focus
    - **selection**: B links to similar C (common focus)
  - membership closure: A and B people, C focus
    - **social influence**: B links to C influenced by A



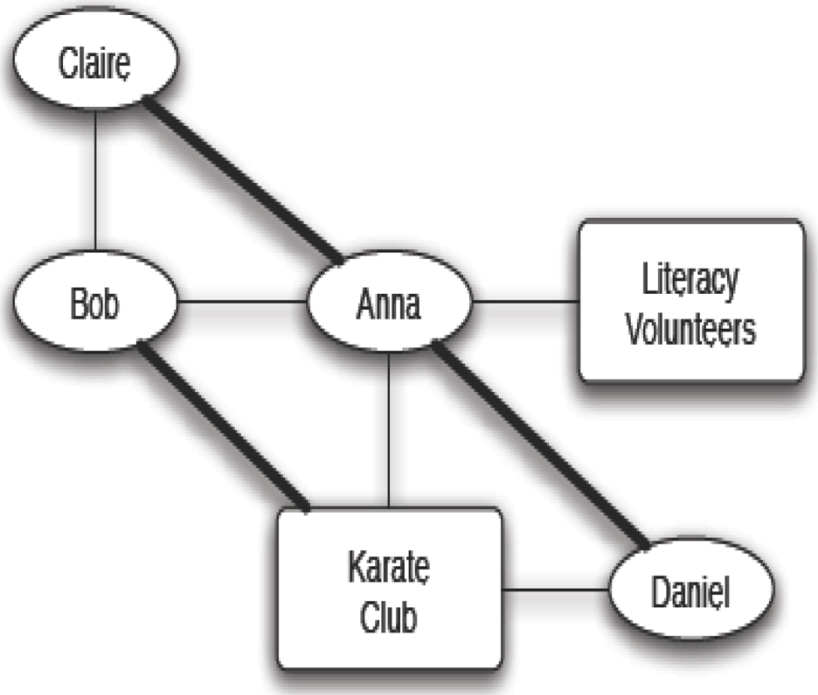
(a) *Triadic closure*

(b) *Focal closure*

(c) *Membership closure*

Co-evaluation of Social and Affiliation Networks

- Bob introduces Anna to Claire
- Karate "introduces" Anna to Daniel
- Anna introduces Bob to Karate



Edges with **bold** are the newly formed

Tracking Link Formation in On-Line Data

- Story so far: a set of mechanisms that lead to the formation of links
  - triadic closure
  - focal closure
  - membership closure
- Tracking these mechanisms in large populations
  - their accumulation observable in the aggregate

Tracking Link Formation in On-Line Data

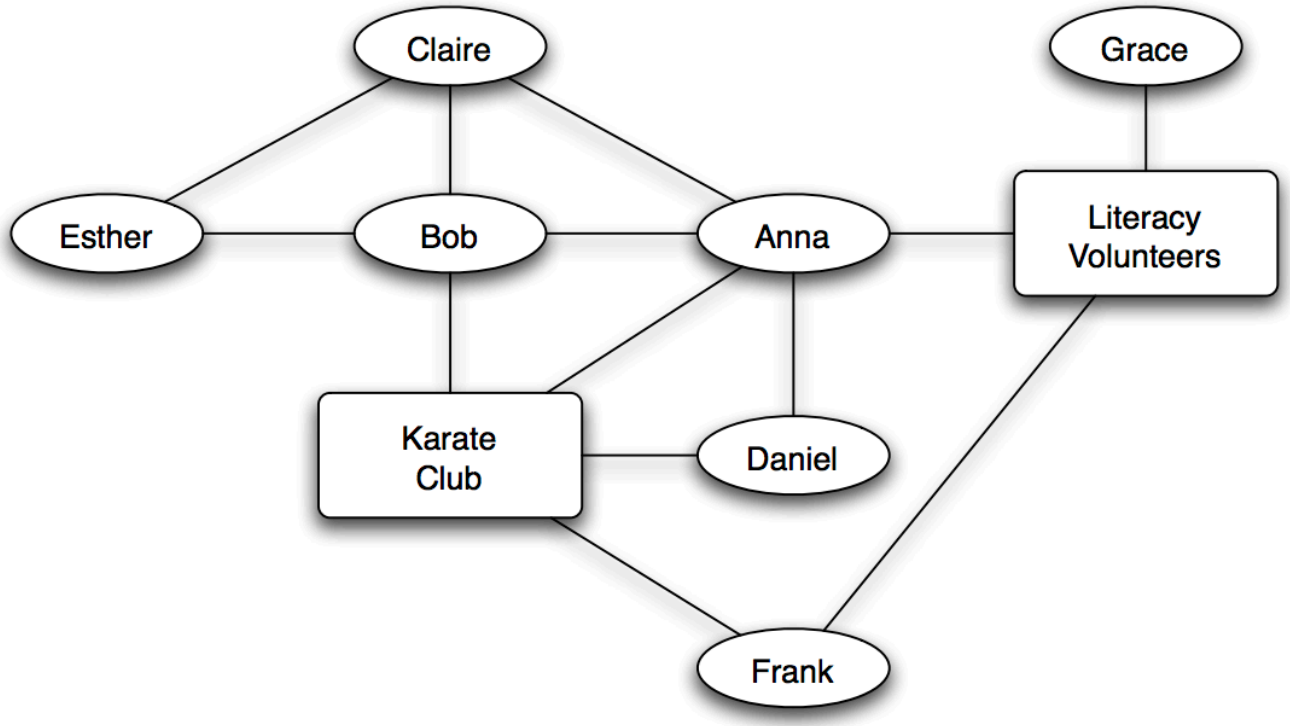


Figure 4.8: A larger network that contains the example from Figure 4.7. Pairs of people can have more than one friend (or more than one focus) in common; how does this increase the likelihood that an edge will form between them?

Tracking Triadic Closure

- Likelihood of link as a function of common friends?

1. Two snapshots of the network
2. For each k, find all pairs of nodes with k common friends in the first snapshot, but not directly connected
3. T(k): fraction of these pairs connected in the second snapshot
   - empirical estimate of probability that a link will form between two people with k common friends
4. Plot T(k) as a function of k
   - T(0) is the rate of link formation when it does not close a triangle

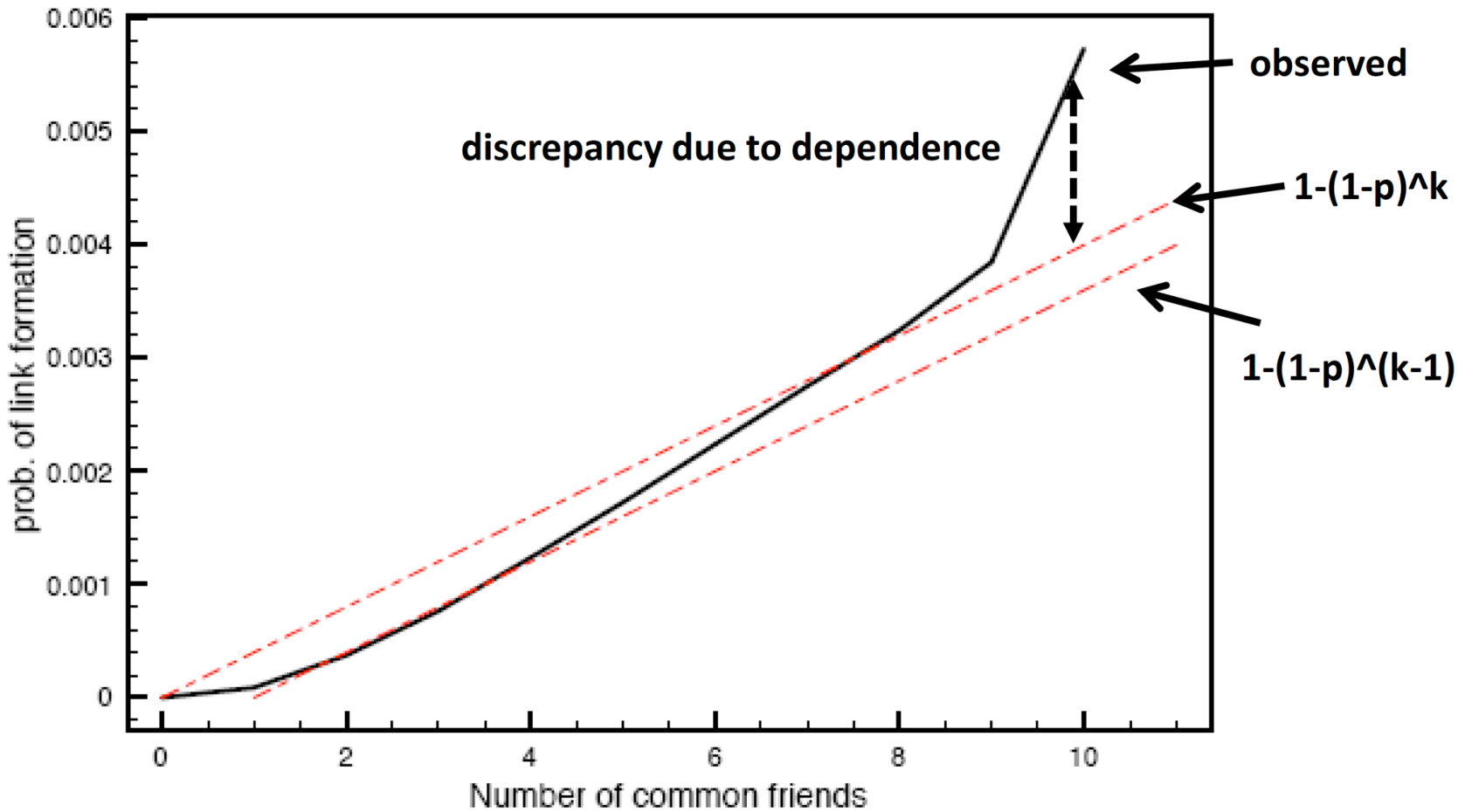
Tracking Triadic Closure

- Kossinets and Watts computed T(k)
  - full history of e-mail communication ("who-talks-to-whom")
  - a one-year period
  - 22,000 students at a large U.S. university
  - observations in each snapshot were one day apart (average over multiple snapshots)

Tracking Triadic Closure

- Interpret the result compared to a baseline
- Assume that each common friend that 2 people have, gives them an independent probability **p** of forming a link
  - 2 people have k friends in common => the probability they fail to form a link is **(1-p)^k**
  - 2 people have k friends in common => probability that they form a link is **1-(1-p)^k**

Tracking Triadic Closure



Having two common friends produces significantly more than twice the effect compared to a single common friend

Microscopic evolution of social networks, Leskovec et al 2008

Tracking Focal and Membership Closure

- For a focal closure, what is the probability that two people form a link as a function of the number of foci they are jointly affiliated with?

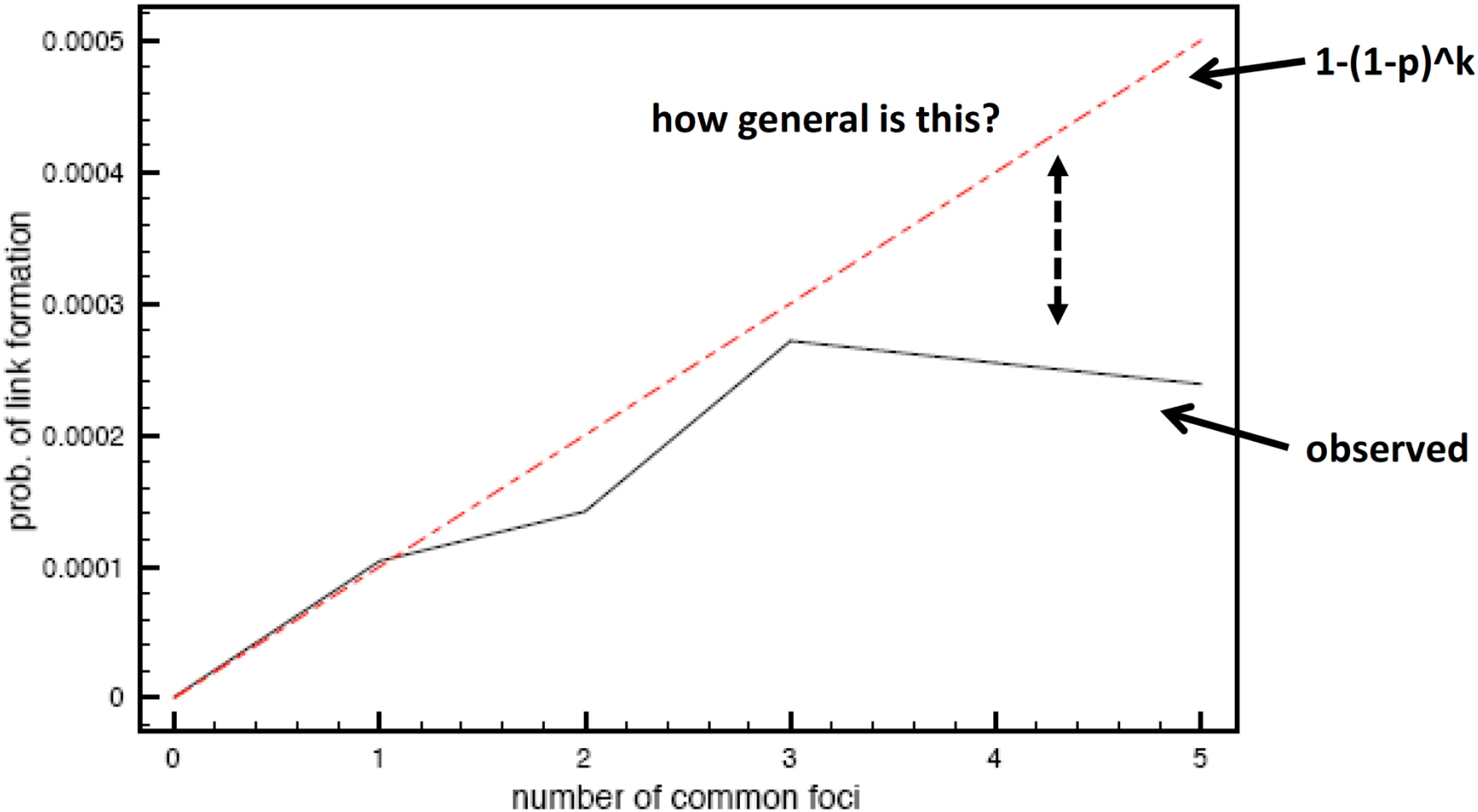
- For membership closure, what is the probability that a person becomes involved with a particular focus as a function of the number of friends who are already involved in it?

Tracking Focal Closure

- Likelihood of link formation as a function of the number of common foci?

- Kossinets and Watts supplemented their university e-mail dataset with information about the class schedules
  - each class became a focus
  - students shared a focus if they had taken a class together

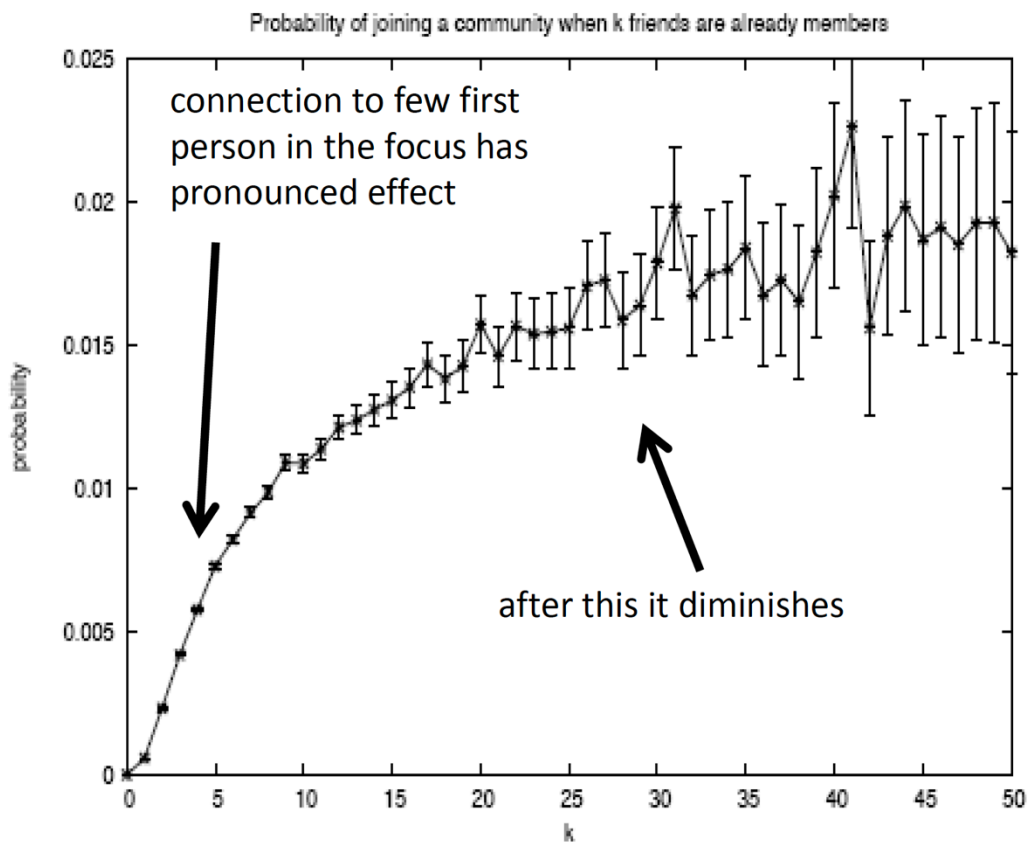
Tracking Focal Closure



A single shared class same effect as a single shared friend, then different subsequent shared classes after the first produce a diminishing returns effect

Tracking Membership Closure

- Blogging site LiveJournal
  - social network (friendship links)
  - foci correspond to membership in user-defined communities

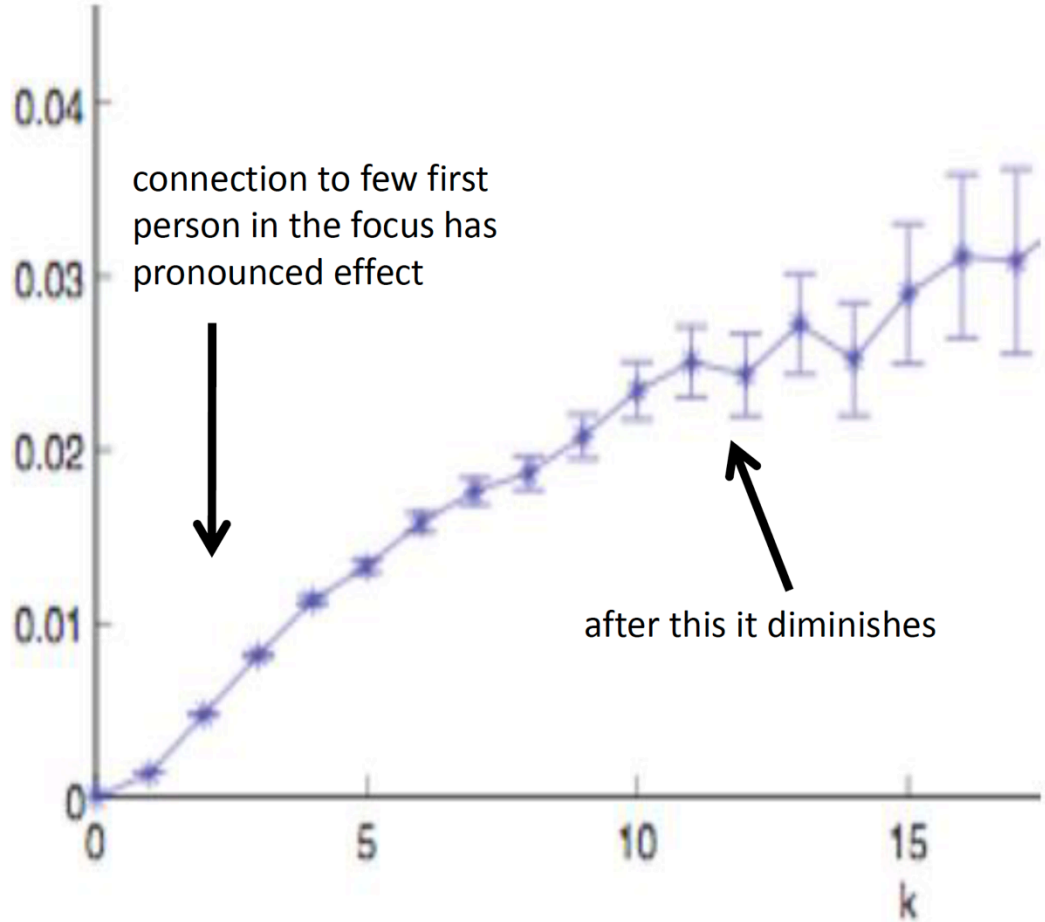


probability of joining a LiveJournal community as a function of the number of friends who are already members

Tracking Membership Closure

- Wikipedia editors
  - link editors when they communicated (user talk page)
  - each Wikipedia article defines a focus (editor associated with the articles he/she edited)

Again, an initial increasing effect: the probability of editing a Wikipedia article is more than twice as large when you have two connections into the focus than one



probability of editing a Wikipedia articles as a function of the number of friends who have already done so

Quantifying the Interplay between Selection and Social Influence

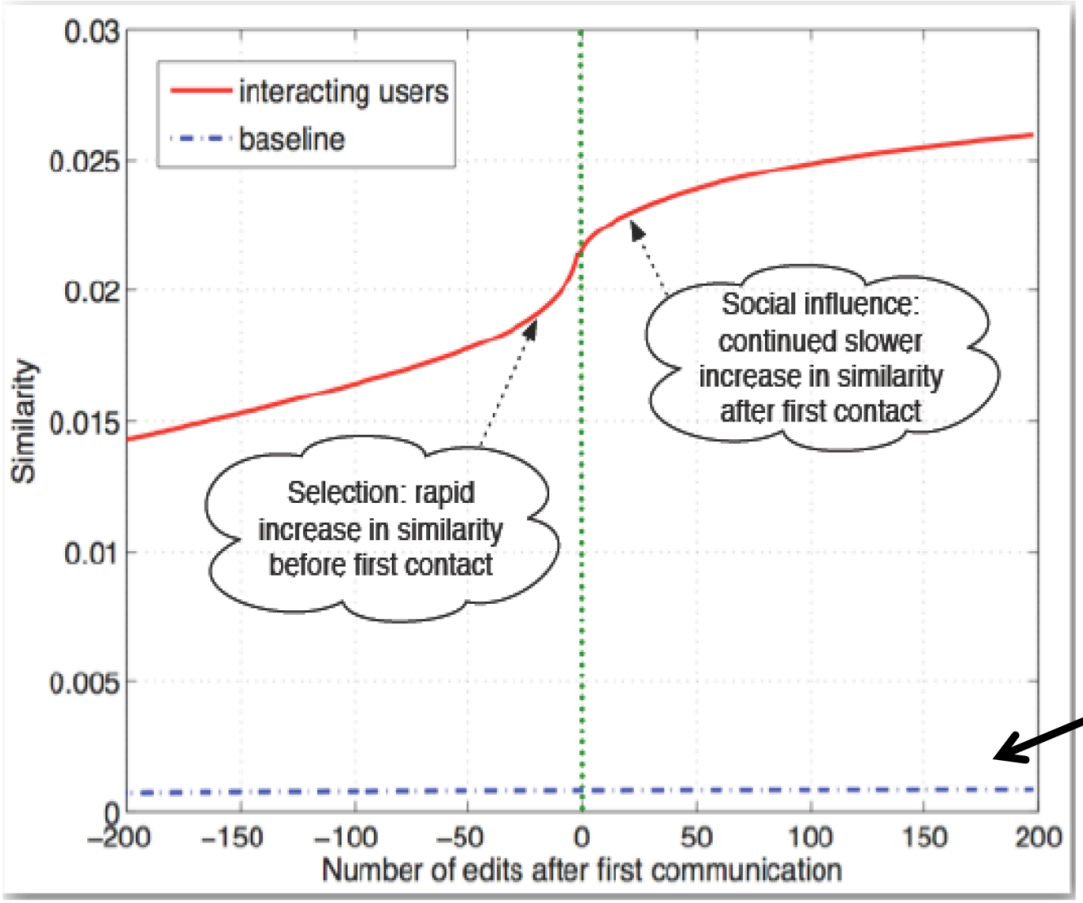
- How selection and social influence work together to produce homophily?
  - How do similarities in behavior between two Wikipedia editors relate to their pattern of social interaction over time?
  - Similarity between 2 Wikipedia editors A, B:

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

- Is homophily (similarity) due to editors connected (talk) with those edited the same articles (selection), or because editors are led to edit articles by those they talk to (social influence)?

Quantifying the Interplay between Selection and Social Influence

Record similarity over time for each pair of editors A and B who have ever talked

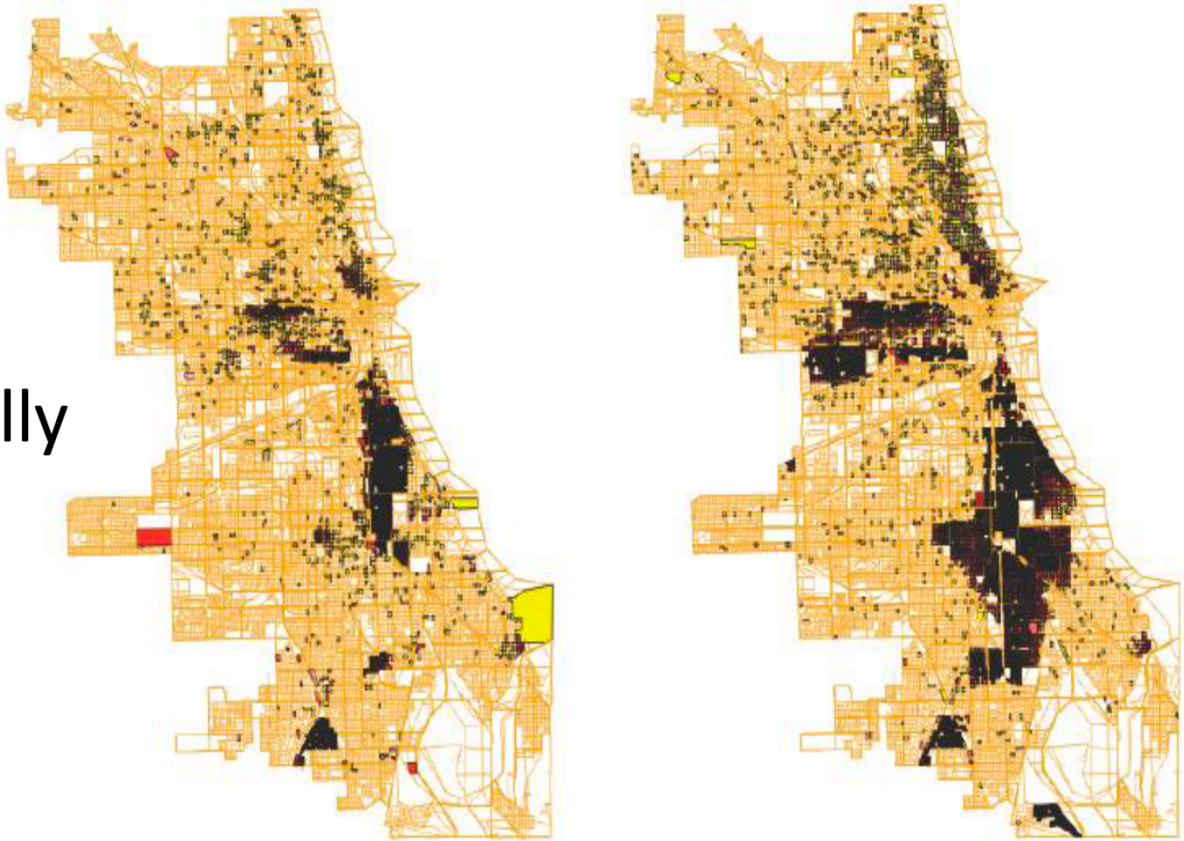


similarity of non-interacting pairs

"tick" in time whenever either A or B performs an action (editing or talking). Time 0 is the point at which they first talked

A Spatial Model of Segregation

- One of the most strong effects of homophily is in the formation of ethnically and racially homogeneous neighborhoods in cities
  - a process with a dynamic aspect
  - what mechanisms?



(a) *Chicago, 1940* (b) *Chicago, 1960*

In blocks colored yellow and orange the percentage of African-Americans is below 25, while in blocks colored brown and black the percentage is above 75

The Shelling Model

- How global patterns of spatial segregation can arise from the effect of homophily operating at a local level (Thomas Schelling)
  - an intentionally simplified mechanism
  - works even when no one individual explicitly wants a segregated outcome

The Shelling Model

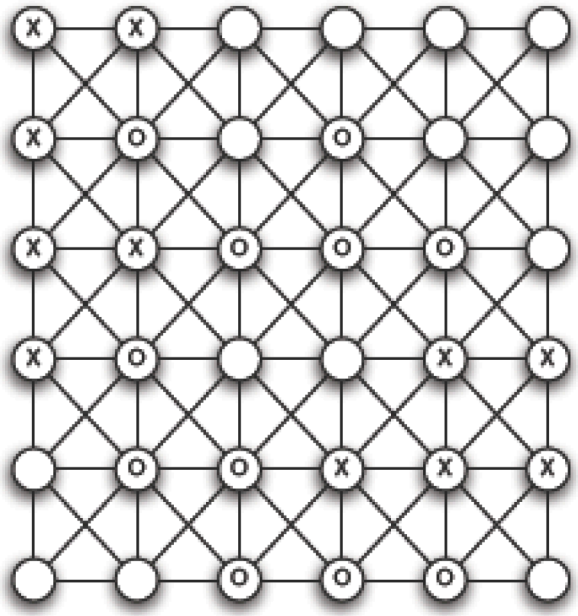
- Model assumptions:
  - Population of individuals called agents
  - Each agent of type X or type O
  - The two types represent some characteristic as basis for homophily (race, ethnicity, country of origin, or native language)
  - Agents reside in cells of a grid (simple model of a 2-D city map)
  - Some cells contain agents while others are unpopulated
  - Cell's neighbors: cells that touch it (including diagonal contact)

X	X				
X	O		O		
X	X	O	O	O	
X	O			X	X
	O	O	X	X	X
		O	O	O	

The Shelling Model

x	x				
x	o		o		
x	x	o	o	o	
x	o			x	x
	o	o	x	x	x
		o	o	o	

(a) *Agents occupying cells on a grid.*



(b) *Neighbor relations as a graph.*

Cells are the nodes and edges connect neighboring cells.
We will continue with the geometric grid rather than the graph.

The Dynamics of Movement

- Unsatisfied agents move in rounds
  - consider unsatisfied agents in some order
    - random or row-sweep
  - unsatisfied agents move to an unoccupied cell where will be satisfied
    - random or to nearest cell that satisfies them
  - may cause other agents to be unsatisfied
  - deadlocks may appear (no cell that satisfies)
    - stay or move randomly
- All variations have similar results
- Ex (figure):
  - t=3, one round, row-sweep, move to nearest cell, stay when deadlocks

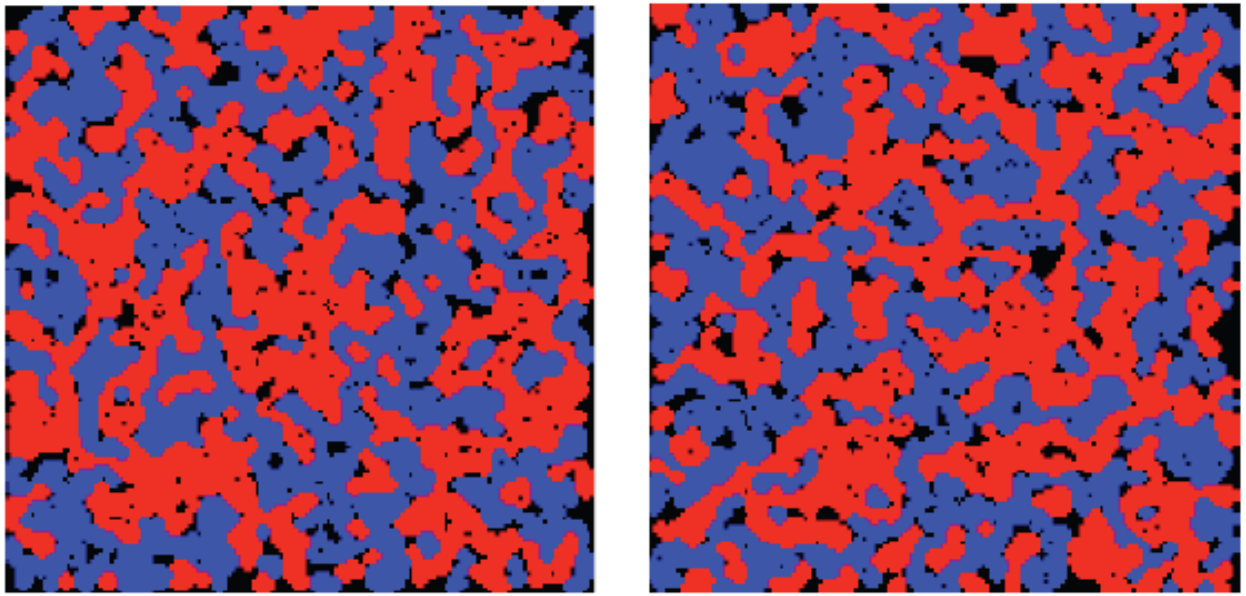
X1*	X2*				
X3	O1*		O2		
X4	X5	O3	O4	O5*	
X6*	O6			X7	X8
	O7	O8	X9*	X10	X11
		O9	O10	O11*	

(a) *An initial configuration.*

X3	X6	O1	O2		
X4	X5	O3	O4		
	O6	X2	X1	X7	X8
O11	O7	O8	X9	X10	X11
	O5	O9	O10*		

(b) *After one round of movement.*

Larger Examples



(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, unsatisfied agents move to a random location
- ~50 rounds, all satisfied
- Different random starts
- Large homogeneous regions

Interpretations of the Model

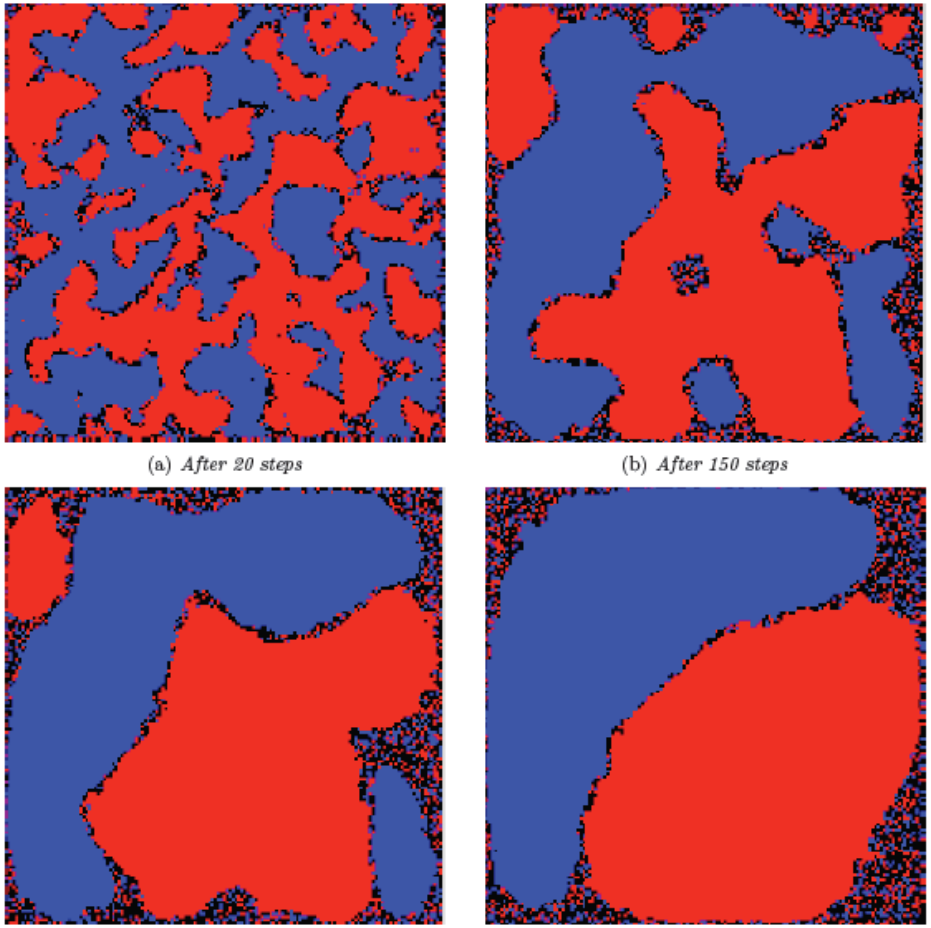
- Spatial segregation is taking place even though no individual agent is seeking it
  - agents just want to be near **t** others like them
  - when t=3, agents are satisfied being minority among its neighbors (5 neighbors of the opposite type)
- Ex (figure):
  - a checkerboard 4x4 pattern can make all agent satisfied (even for large grids)
  - we don't see this result in simulations

X	X	0	0	X	X
X	X	0	0	X	X
0	0	X	X	0	0
0	0	X	X	0	0
X	X	0	0	X	X
X	X	0	0	X	X

Interpretations of the Model

- More typically, agents form larger clusters
  - agents become unsatisfied and attach to larger clusters (where higher probability to be satisfied)
- The overall effect:
  - local preferences of individual agents have produced a global pattern that none of them necessarily intended

Larger Examples



(a) *After 20 steps*

(b) *After 150 steps*

(c) *After 350 steps*

(d) *After 800 steps*

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.

Schelling Model and Homophily

- The Schelling model is an example that, as homophily draws people together along immutable characteristics (race or ethnicity), it creates a natural tendency for mutable characteristics (decision about where to live) to change in accordance with the network structure