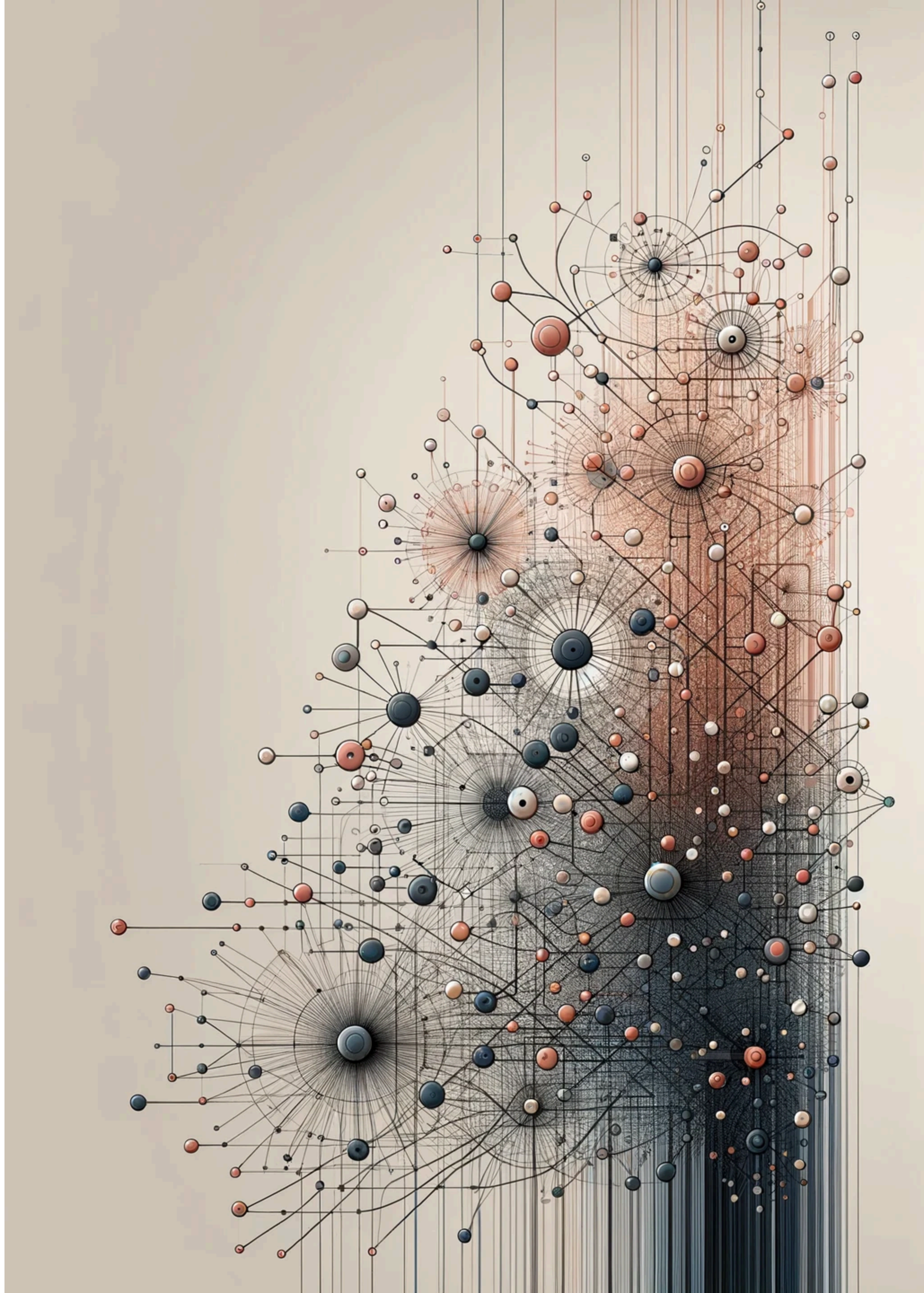
UNIVERSITÀ DI TORINO

Analisi e Visualizzazione delle Reti Complesse

NS16 - Cascading behaviors in networks

Prof. Rossano Schifanella



Agenda

- Diffusion in Networks
- Modeling Diffusion through a Network
- Cascades and Clusters
- Diffusion, Thresholds, and the Role of Weak Ties
- Extensions of the Basic Cascade Model
- Knowledge, Thresholds, and Collective Action
- The Cascade Capacity

The diffusion of innovations

- Let's consider how new behaviors, practices, opinions, conventions, and technologies spread from person to person through a social network, as people influence their friends to adopt new ideas
- diffusion of innovation: long list of classic studies in the middle of the 20th century
    - focus on informational and network effects
- **direct benefit effects** to understand how technologies spread: incentives to adopt telephone, fax, emails based on friends that already adopted those technologies
- we need to take into account **network effects** on a **local network level**
- Objective of this lecture: **formulating a model for the spread of an innovation through a social network**

UNIVERSITÀ DI TORINO

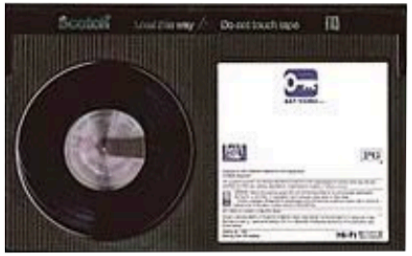
A diffusion of a new behavior

- Assumption: individuals make decisions based on the choices of their neighbors
  - focus on links
- In this lecture, let's focus on **direct-benefit effects** instead of informational effects
- Natural model introduced by Stephen Morris in 2000
- Reading material:
  - Stephen Morris. Contagion. Review of Economic Studies, 67:57–78, 2000.

Examples

VHS

vs



Betamax

Blu-ray Disc

vs

HD DVD

A networked coordination game

- It is natural to use a **coordination game**
  - each node has a choice between two possible behaviors, A and B
  - players have an incentive to adopt the same behavior

$$
\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{w} \\
 & & \multicolumn{1}{c}{A} & \multicolumn{1}{c}{B} \\
\cline{3-4}
\multirow{2}{*}{v} & A & a, a & 0, 0 \\
\cline{3-4}
 & B & 0, 0 & b, b \\
\cline{3-4}
\end{array}
$$

UNIVERSITÀ DI TORINO

A networked coordination game

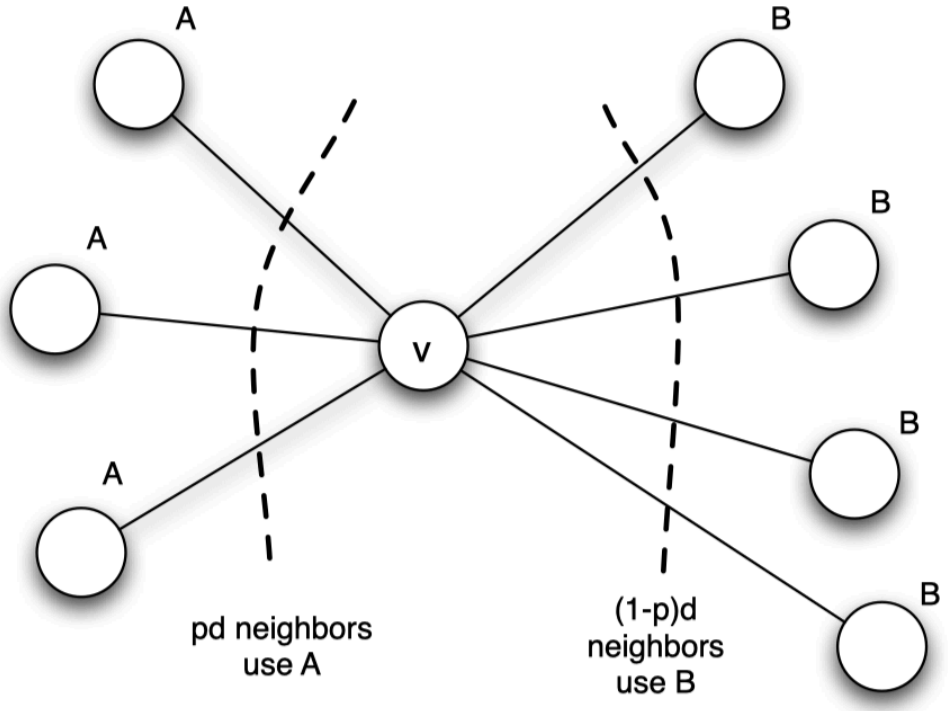
p `fraction of neighbors adopting` A

$1 - p$ `fraction of neighbors adopting` B

d `is the number of neighbors`

Node v chooses A if

$$pda \geq (1-p)db$$
$$pa \geq b - pb$$
$$(a+b)p \geq b$$
$$p \geq \frac{b}{a+b} = q$$



UNIVERSITÀ DI TORINO

Threshold rule

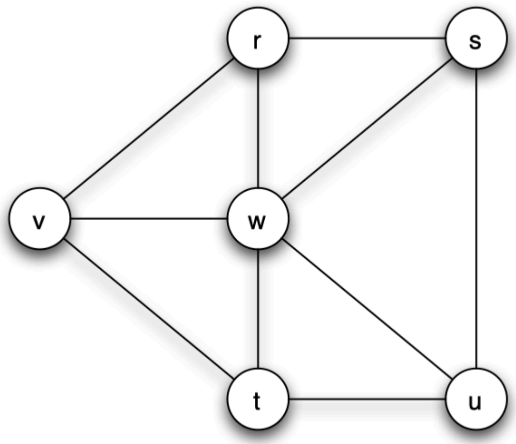
- In order to adopt A or B, we just need to check if $p \geq q$.
- Very simple - and myopic - model of individual decision making
- It is a research question to think about richer models that incorporate more long-range considerations

Cascading behavior

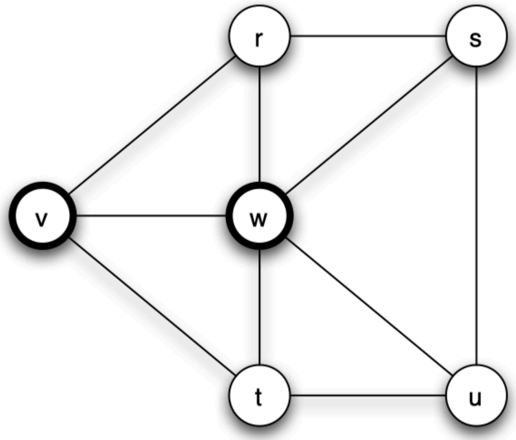
- Two obvious equilibria in the network:
  - one in which everyone adopts A
  - another in which everyone adopts B
- We want to understand how easy is to get one of these equilibria
- Also, we want to understand if other intermediate equilibria exist and how they look like
- **Assumptions**:
  - everyone is using B at the beginning
  - S: small set of **initial adopters** of A
- Will the spread of A make everyone to switch to the new technology, or will the spread stop?
- **Answer:** it depends on the network structure, the choice of nodes in S, and the value of q.

Example

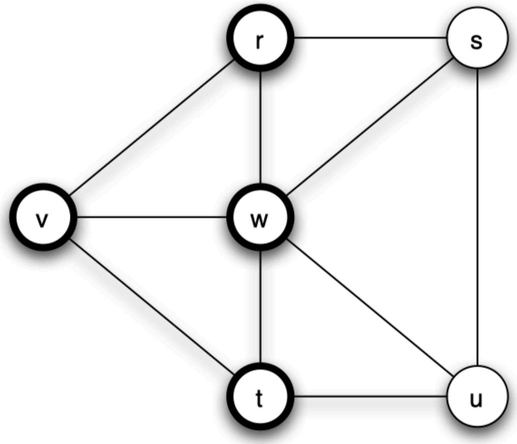
- $a = 3, b = 2 \Rightarrow q = \frac{2}{5}$
- $S = \{u, v\}$
- **complete cascade**



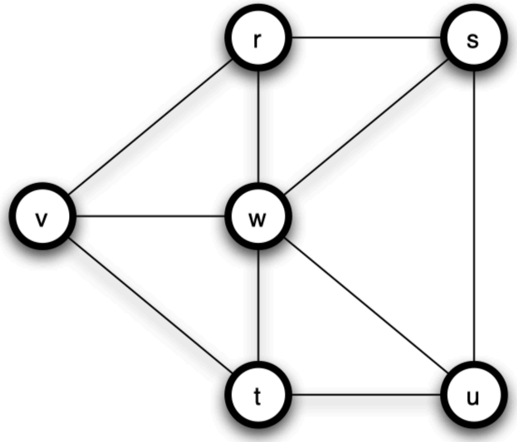
(a) *The underlying network*



(b) *Two nodes are the initial adopters*



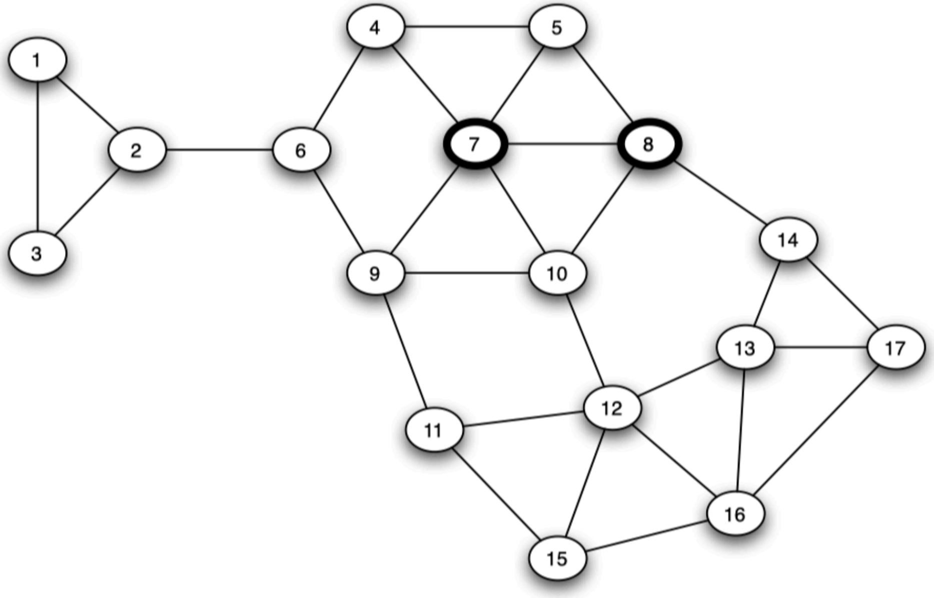
(c) *After one step, two more nodes have adopted*



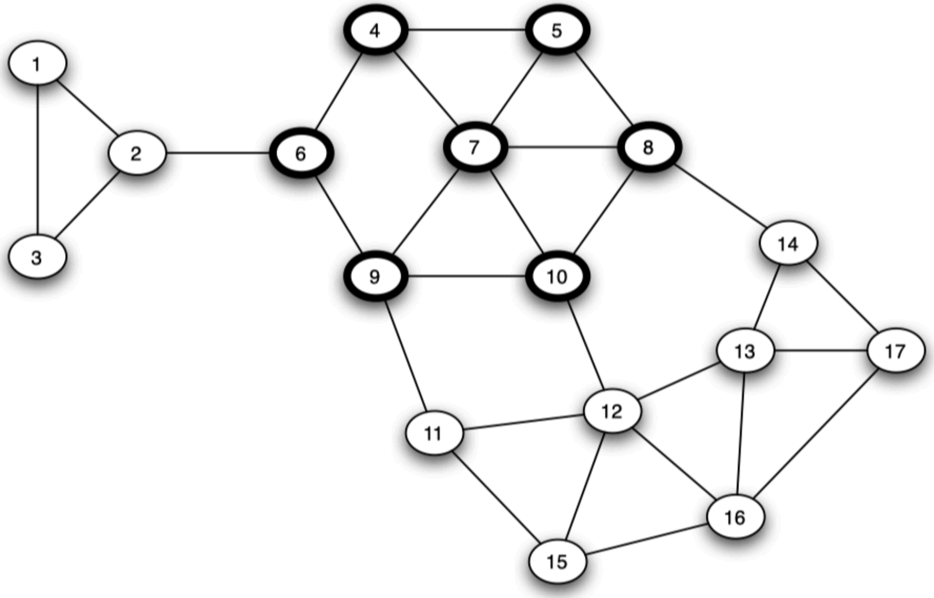
(d) *After a second step, everyone has adopted*

Example

- $a = 3, b = 2 \Rightarrow q = \frac{2}{5}$
- $S = \{u, v\}$
- The diffusion of A stops here
  - **partial cascade**



(a) *Two nodes are the initial adopters*



(b) *The process ends after three steps*

Recap

- Consider a set of initial adopters who start with a new behavior A, while every other node starts with behavior B.
- Nodes then repeatedly evaluate the decision to switch from B to A using a threshold of q.
- If the resulting cascade of adoptions of A eventually causes every node to switch from B to A, then we say that the set of initial adopters causes a **complete cascade at threshold** q

Definition

- We say that a **cluster of density** p is **a set of nodes** such that each node in the set has **at least a** p **fraction of its network neighbors in the set**.

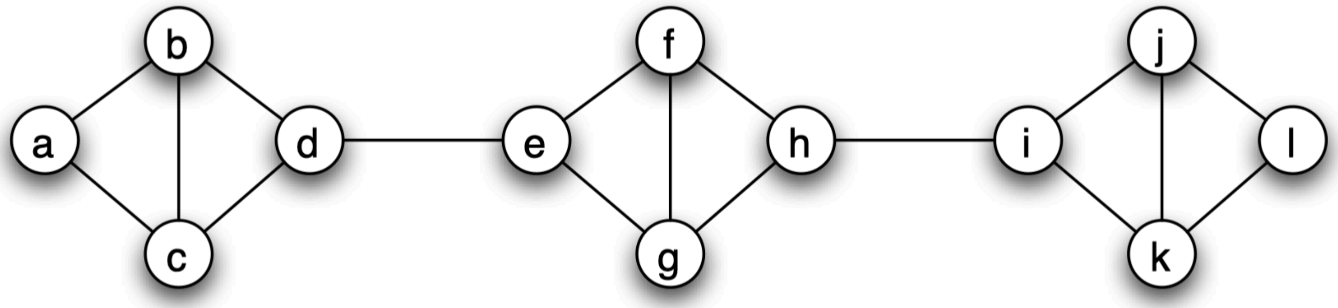
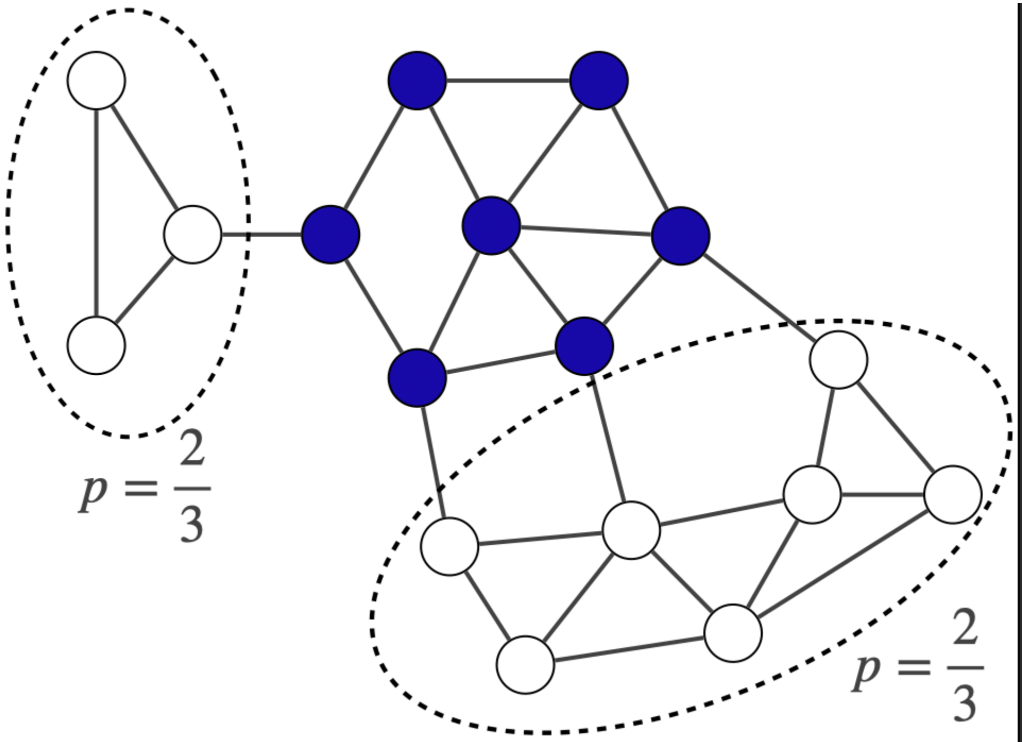


Figure 19.6: A collection of four-node clusters, each of density 2/3.

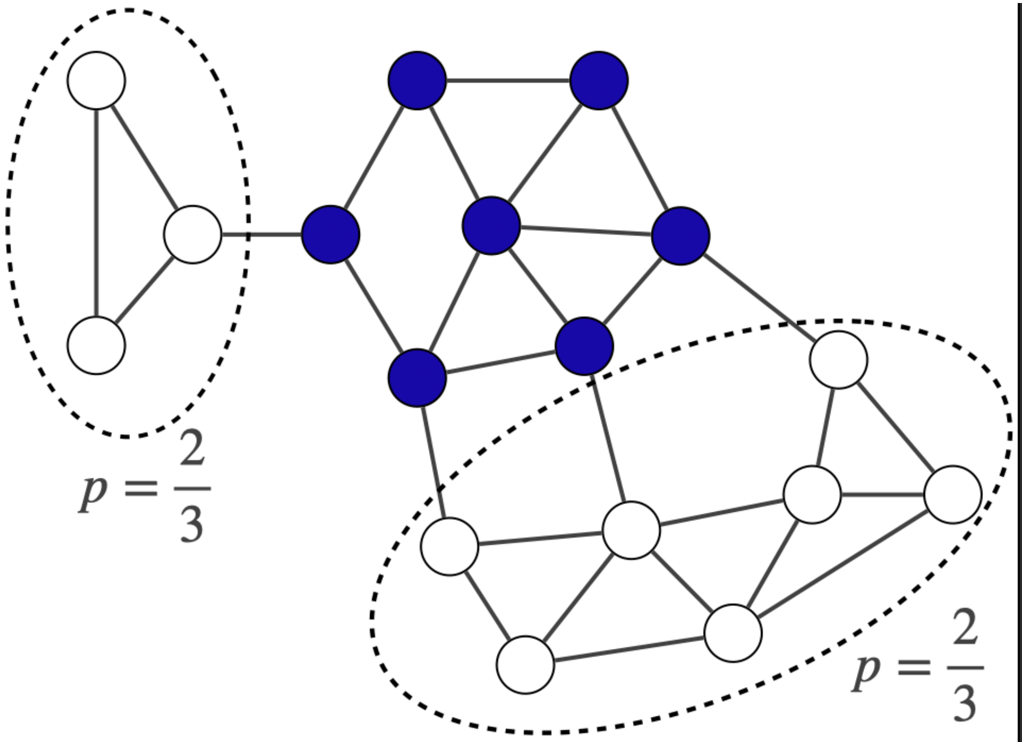
Stopping cascades

- What prevents cascades from spreading?
  - **Homophily** can serve as a barrier to diffusion: it is hard for innovation to arrive from outside densely connected communities
- Let's try to quantify this intuition:
  - **Cluster of density** p



Internal cohesion

- Each node in a cluster has a prescribed fraction of its friends residing in the cluster: **cohesion**
- Nodes in the same clusters do not necessarily have much in common
  - Any network is a cluster of density $p = 1$.
  - The union of two clusters of density p is still a cluster of density p.
- In fact, clusters in networks can exist simultaneously at different scales.

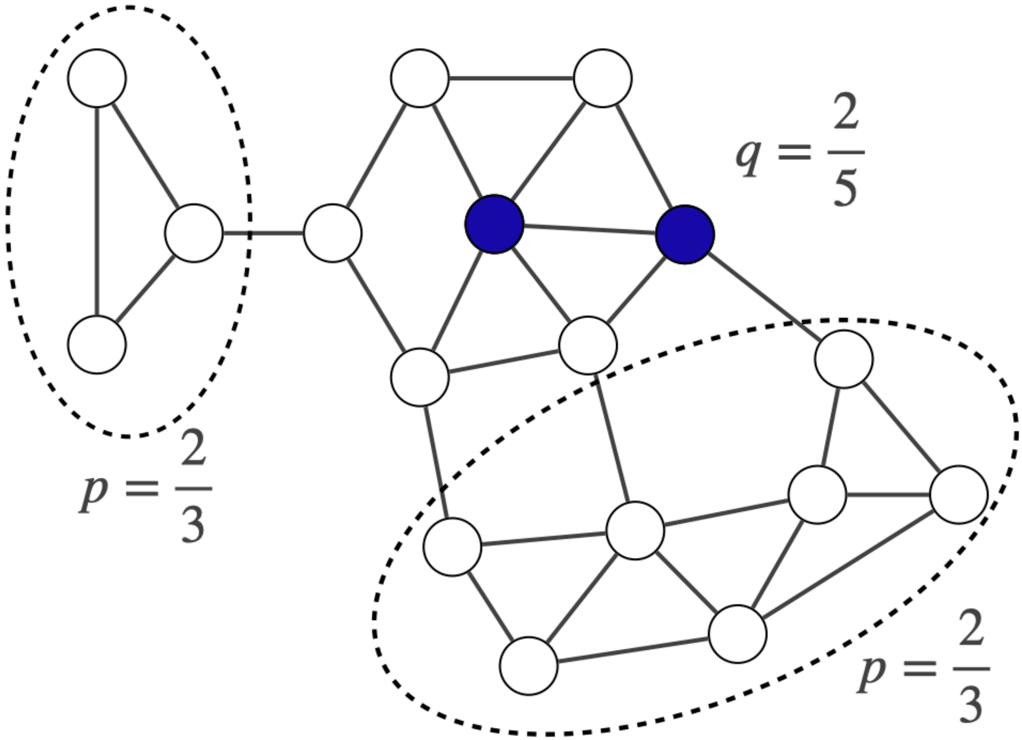


Relationship between clusters and cascades

- A cascade stops **iff** it runs into a dense cluster
  - i.e., **clusters are natural obstacles to cascades**
- **Claim**: Let S be the set of **initial adopters** of A, with a threshold of q

  -(i) if the remaining network contains a cluster of density $p > (1 - q)$ then **S cannot cause a complete cascade**

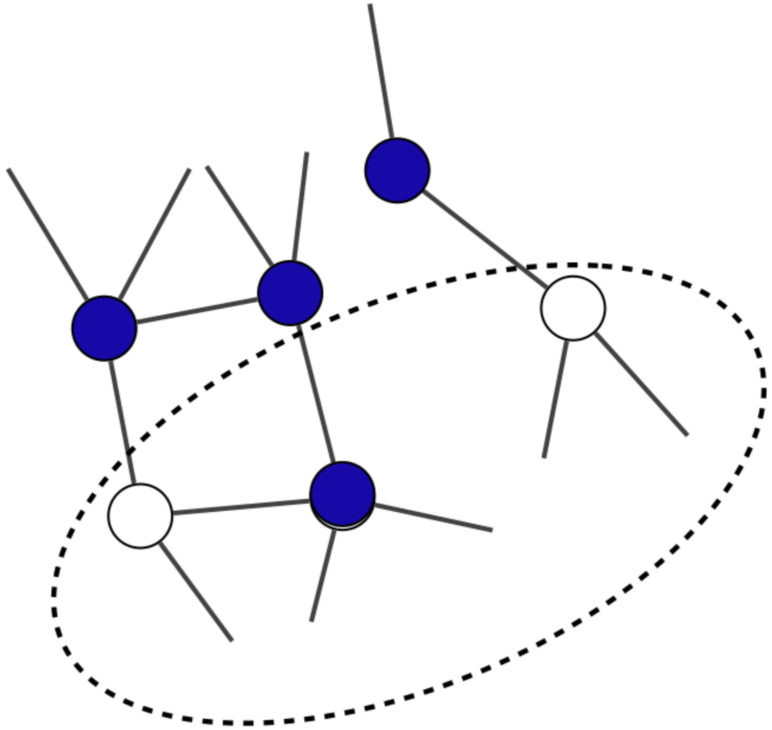
  (ii) if S fails to cause a cascade, then **there is a cluster of density $p > (1 - q)$ in the remaining network**



Clusters implies no cascade

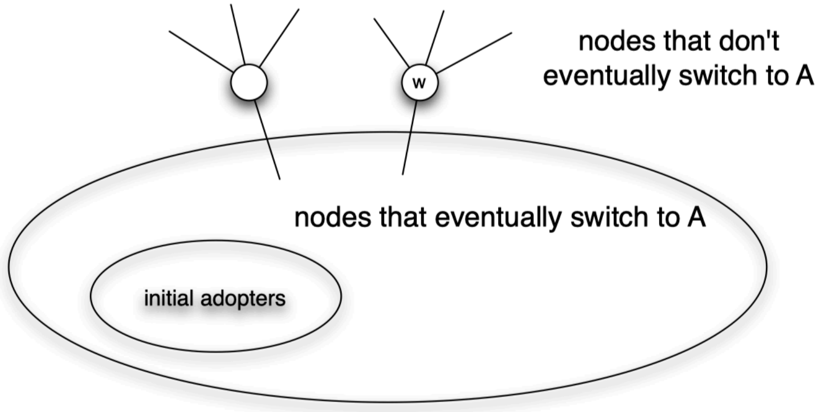
Proof by contradiction:

- some node inside the cluster C of density $p > (1 - q)$ will adopt A at time t (the first to adopt in its cluster);
- this means that the decision was made at time $t - 1$, when no other node in the cluster C adopted A;
- since the cluster C has density $p > (1 - q)$, less than a q fraction of nodes are outside the cluster;
- these nodes are the only that could have adopted A and they were outside the cluster at time $(t - 1)$
- so at time $t - 1$ it is impossible that the threshold rule fired for v since $p < q$: **contradiction**



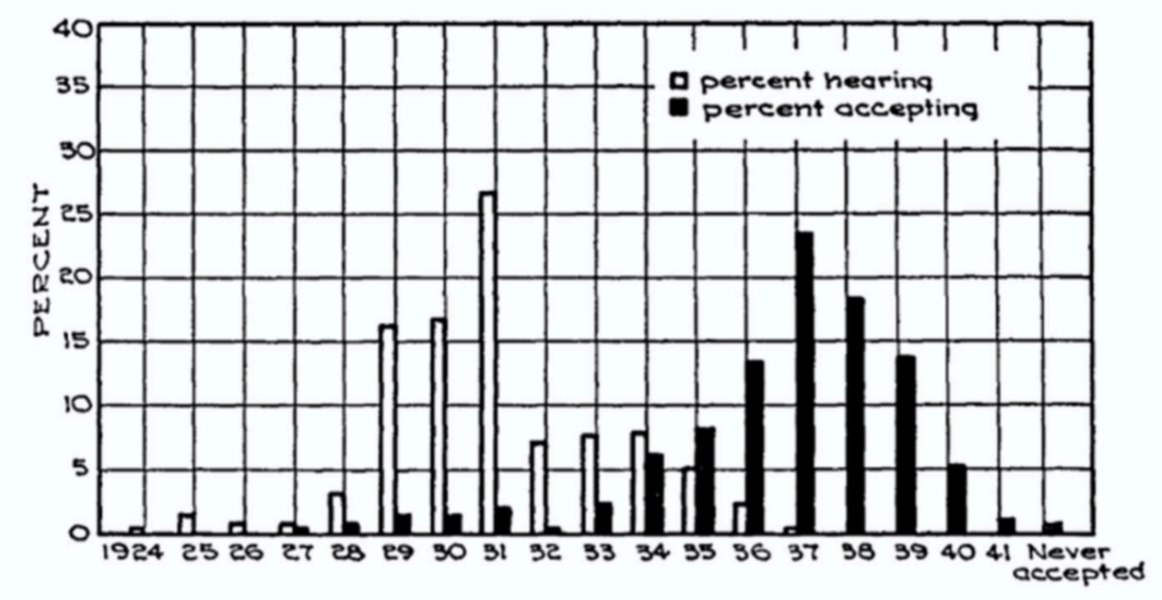
Clusters implies no cascade

- Let S be the set of initial adopters of A
- The spreading process stops: let S' be the maximum set of nodes that switched to A
- Let R be the set of nodes still using B at the end of the process
- Pick one node in R: it does not want to switch to A
    - the fraction of its friends using A is $< q$
    - the fraction of its friends belonging to R is $p > (1-q)$
    - this holds for every node in R
    - R is a cluster of density $p > (1-q)$



Viral marketing

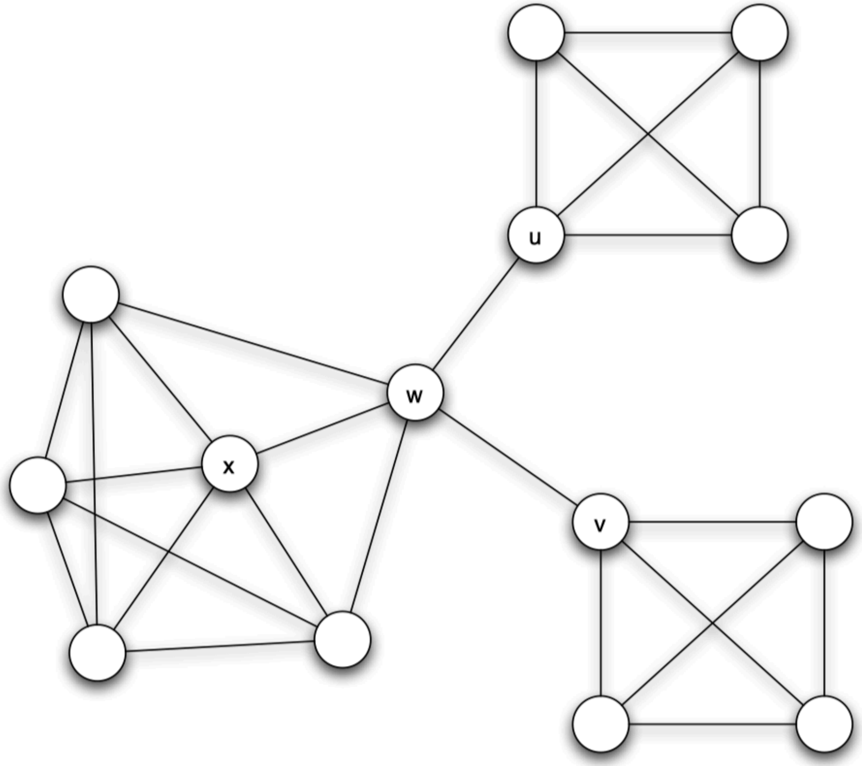
- Lesson learned: tightly-knit communities in the network can work to stop the spread of an innovation
  - We have coexistence of both behaviors (or also competing products)
- In case of coexistence, if a firm can rise the quality of product A, than the spread of the new technology can produce a complete cascade
- Alternative strategy (when the firm cannot rise the quality of the product): trying to **convince a small number of key people** in the part of the network still using B
  - how to choose key people to trigger a viral marketing strategy?
  - their position in the underlying network can be very relevant - let's explore this now



- In the Ryan-Gross study (1943) on the adoption of hybrid seed corn, it is possible to spot the difference between learning about a new idea and actually deciding to adopt it
- It is like nodes at the border in a cluster: they have been exposed to the idea, but still they decide not to adopt it.
- Reading material: Bryce Ryan and Neal C. Gross. The diffusion of hybrid seed corn in two Iowa commu- nities. Rural Sociology, 8:15–24, 1943.

The role of weak ties

- Threshold models highlight some important implications of the **strength of weak ties** theory
- They receive very fresh ideas from other communities; not enough for adoption and spread (try with $q = \frac{1}{2}$)
- Bridges and weak ties are great for spreading rumors or jokes across the network, but not for diffusion of innovation or social mobilization
- Strong ties can have more significant role for others in the community to take actions
- Reading material: Damon Centola and Michael Macy. Complex contagions and the weakness of long ties. American Journal of Sociology, 113:702–734, 2007.



UNIVERSITÀ DI TORINO

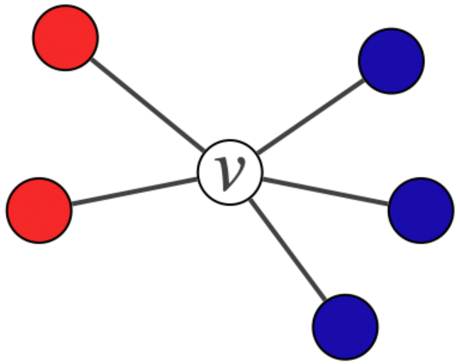
Extensions of the Basic Cascade Model

Heterogeneous thresholds

- Let's suppose each person gives values to A and B subjectively

$$
\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{w} \\
 & & A & B \\
\hline
\multirow{2}{*}{v} & A & a_v, a_w & 0, 0 \\
\hline
 & B & 0, 0 & b_v, b_w \\
\hline
\end{array}
$$

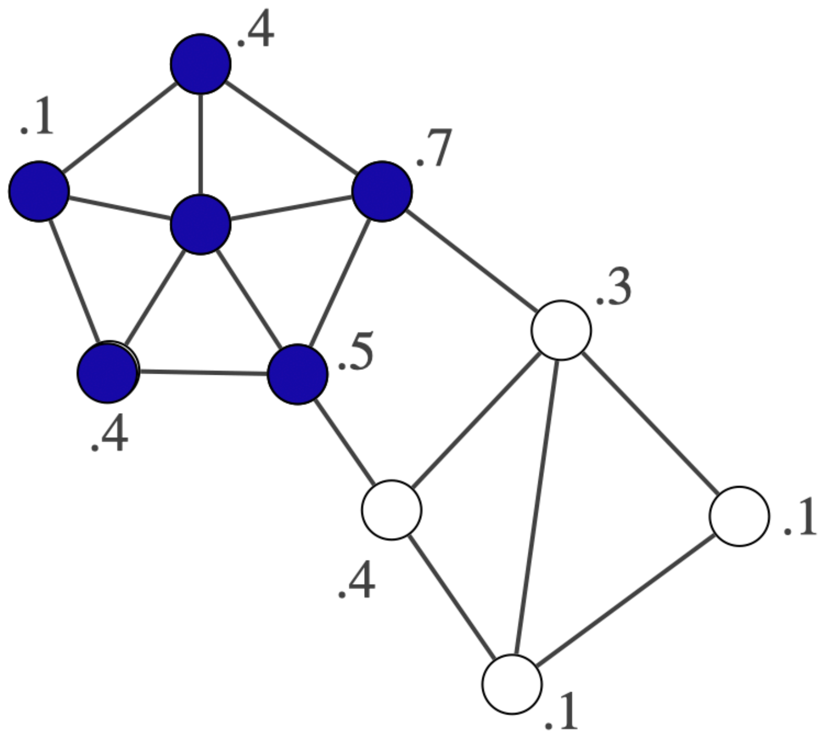
Heterogeneous thresholds



- p fraction of neighbors adopting A
- $1 - p$ fraction of neighbors adopting B
- d is the number of neighbors
- The node chooses A if

$$pda_v \geq (1-p)db_v \Rightarrow p \geq \frac{b_v}{a_v + b_v} = q_v$$

- Watts and Dodds: we need to take into account not just the power of influential nodes, but also the extent to which these influential nodes have access to easily **influenceable** people.
- Reformulating the notion of **blocking clusters**: set of nodes for which each node v has a fraction $p > (1 - q_v)$ of its friends inside the set.
- The notion of density becomes **heterogeneous** as well: each node has a different requirement for the fraction of friends it needs to have in the cluster.
- Reading material: Duncan J. Watts and Peter S. Dodds. Networks, influence, and public opinion formation. Journal of Consumer Research, 34(4):441–458, 2007.



Knowledge, Thresholds, and Collective Action

Integrating network effects

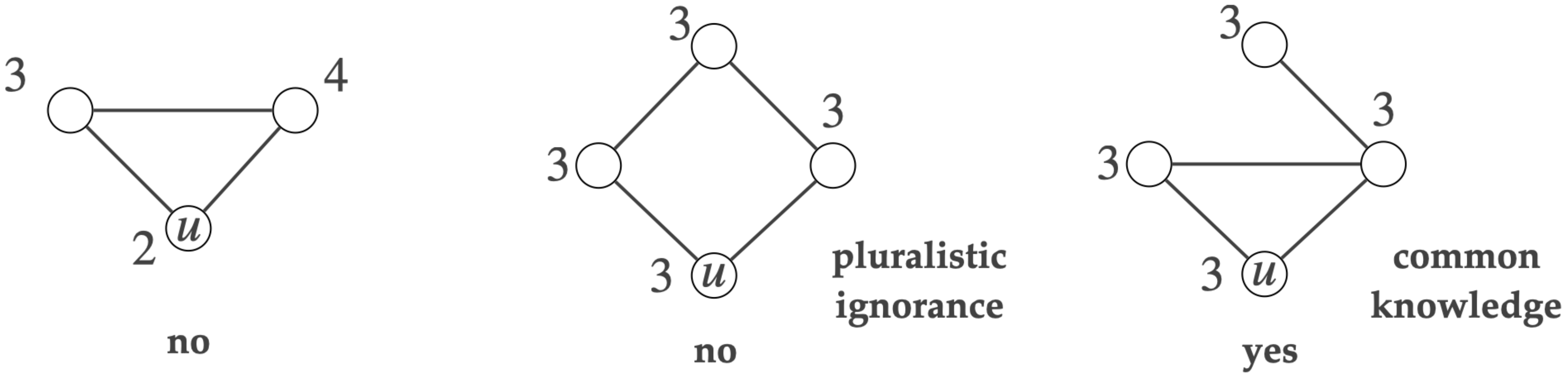
- In many domains you need to integrate network effects at both the **population level** and the **local network level**
  - collective actions and pluralistic ignorance
  - a model for the effect of knowledge on collective actions
  - common knoledge and social institutions

Collective actions and pluralistic ignorance

- **Objective:** you want to organize a revolt under a repressive regime. Which is the probability to succeed?
- **Collective actions:** benefits only if enough people participate - but you can talk to a very limited number of persons you trust
- **Pluralistic ignorance:** when people have wildly erroneous estimates about the prevalence of certain opinions in the population at large

A model for the effect of knowledge on collective actions

- Each individual has a personal threshold of k: "I will show up for the protest if I am sure that at least k people in total will show up"
- Assumption: every node knows the thresholds of all its neighbors in the network.
- Is it safe for u to join the protest in the following configurations?



Common knowledge and social institutions

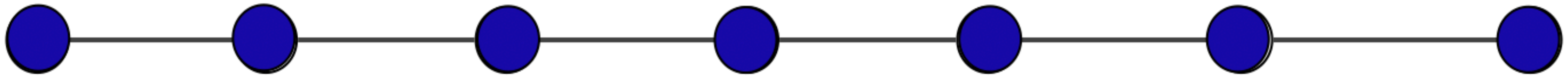
- Some social institutions have the purpose of helping people to achieve common knowledge
- Widely-publicized speech, article in a high circulation newspaper, and so on: people receive the message AND that a lot of other people are receiving the message
- Freedom of the press and freedom of assembly are important social institutions!
- Marketing: commercials during a very popular event (as the famous Apple Macintosh commercial during the 1984 Super Bowl)

The Cascade Capacity

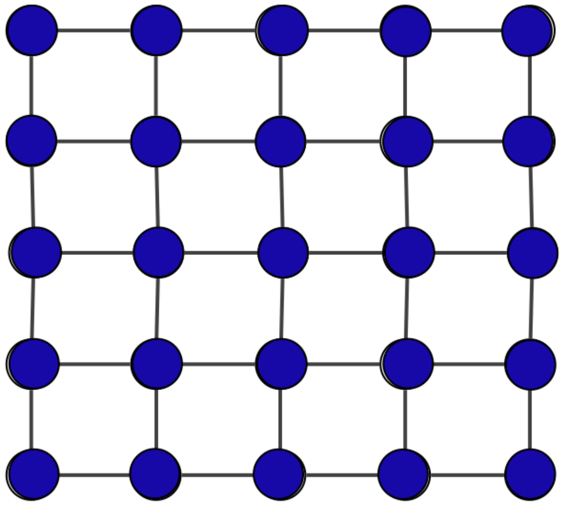
Infinite Networks

- Understanding how different network structures are more or less prone to cascades
- **Cascade capacity:** the **maximum threshold** at which a small set of initial adopters can cause a complete cascade
- Let's consider **infinite graphs**
  - nodes have finite number of neighbors
- Let's say that a finite set S of initial adopters of A causes a cascade in G with threshold q **if eventually every node in** G **adopts** A
- Example: Infinite path

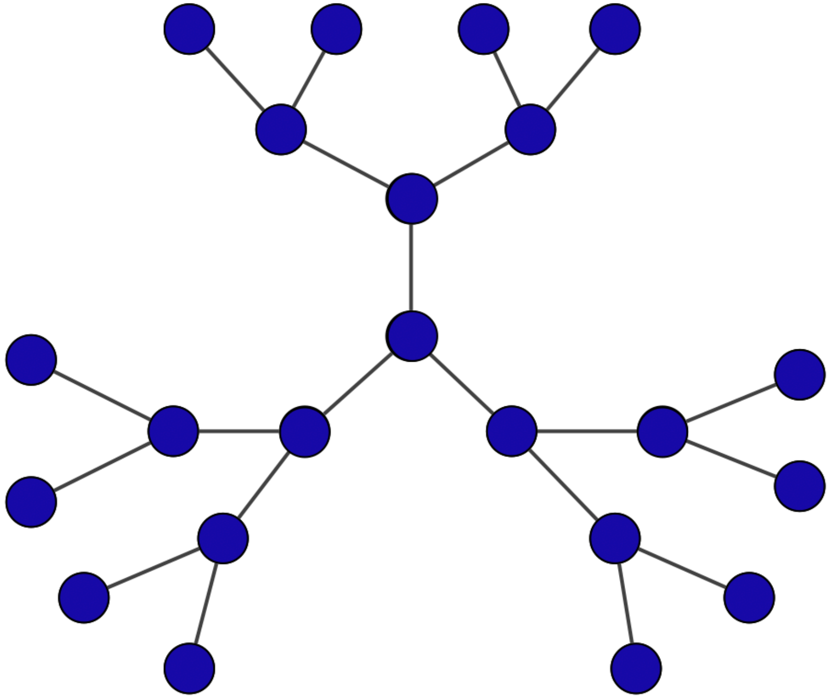
$$q \leq \frac{1}{2}$$



Infinite grid $q \leq \frac{1}{4}$



Infinite tree $q \leq \frac{1}{3}$



How large can the cascade capacity be?

Claim: there is no network which the cascade capacity exceeds $\frac{1}{2}$

Proof by contradiction:

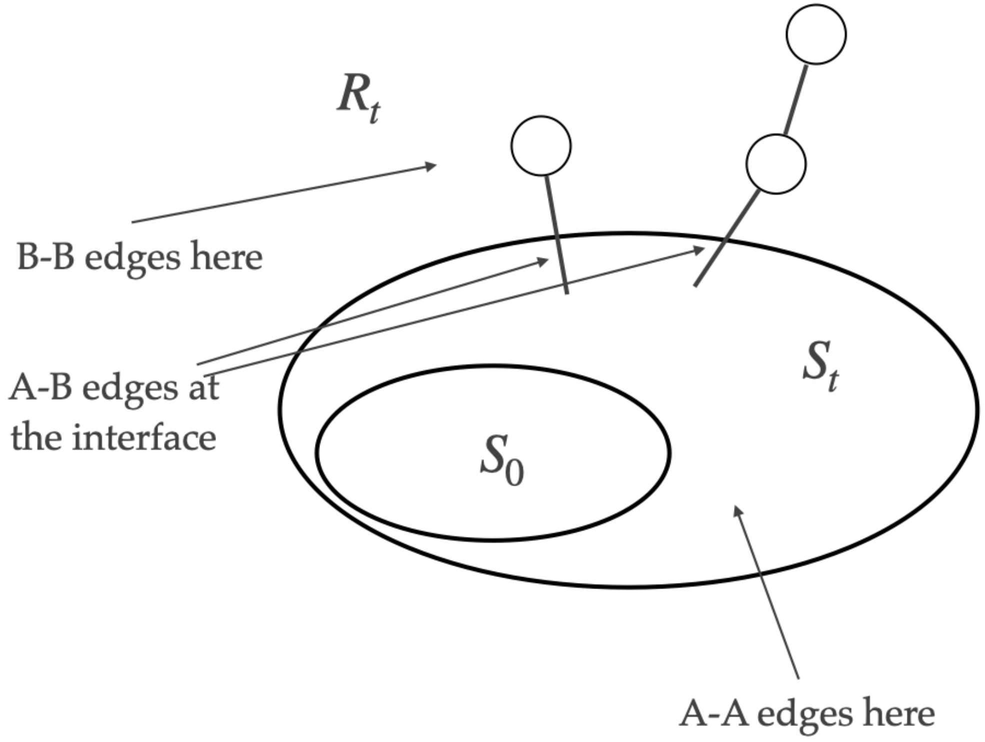
Let's suppose such G exists with $q > \frac{1}{2}$

We want to find a contradiction, arguing that nodes stop switching from $B \rightarrow A$ after a finite number of steps.

UNIVERSITÀ DI TORINO

Focus on the interface

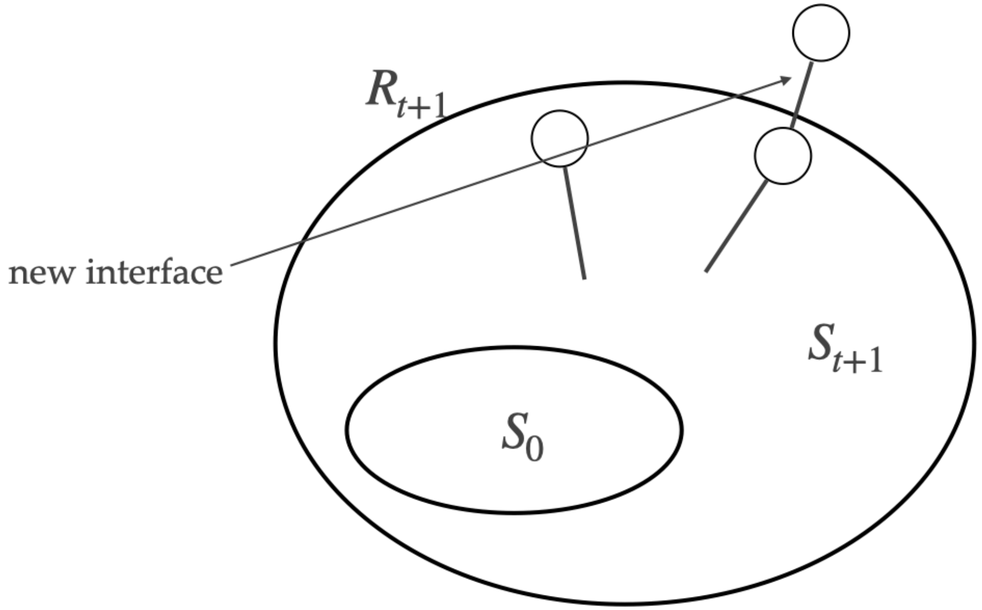
- S_0: finite set of initial adopters of A
- S_t: set of adopters of A, potentially larger than S_{t-1}
  - size of the **interface** at time t: I_t
- We show that $I_t < I_{t+1}$
- Hence, if the size of the interface strictly decreases, the diffusion process will terminate after I_0 steps



UNIVERSITÀ DI TORINO

Size of the interface

- at time $t+1$ some nodes in R switches from $B \to A$
- Focus on a node that switches from $B \to A$
    - it has edges to B nodes
    - and edges with A nodes
    - recall that $q > \frac{1}{2}$, so this node has more edges to nodes in S than to nodes in R
    - more edges leaving the interface than edges joining it
    - $\Rightarrow I_{t+1} < I_t$



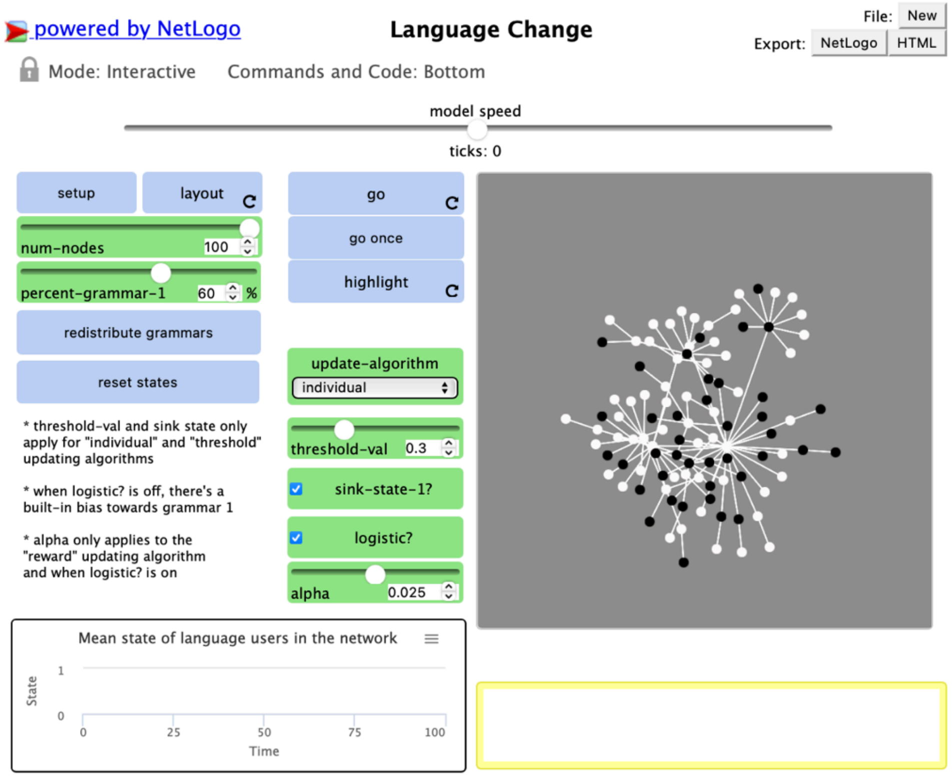
Observations

- When $q > \frac{1}{2}$ no finite set of nodes can cause a complete cascade in any network
- This threshold corresponds intuitively to a "lower quality" technology: a person will switch to it only if more than 50% of their friends already have.

Another extension of the model

- Extending the model: allow people to **adopt both A and B** (compatibility)
- It is referred as the **bilingual option**
- Although the model is amazingly simple, we have very surprising and complex findings
  - compatibility can be a strategy of a firm to enter in a market and progressively cut out the competitor's product

Language Change (netlogo model)



- [Play with the model]

Reading material

[ns2] **Chapter 19 (19.1-19.7)** Cascading behavior in Networks

UNIVERSITÀ DI TORINO

Q&A