Network Effects and Cascading Behavior

How the Class Fits Together

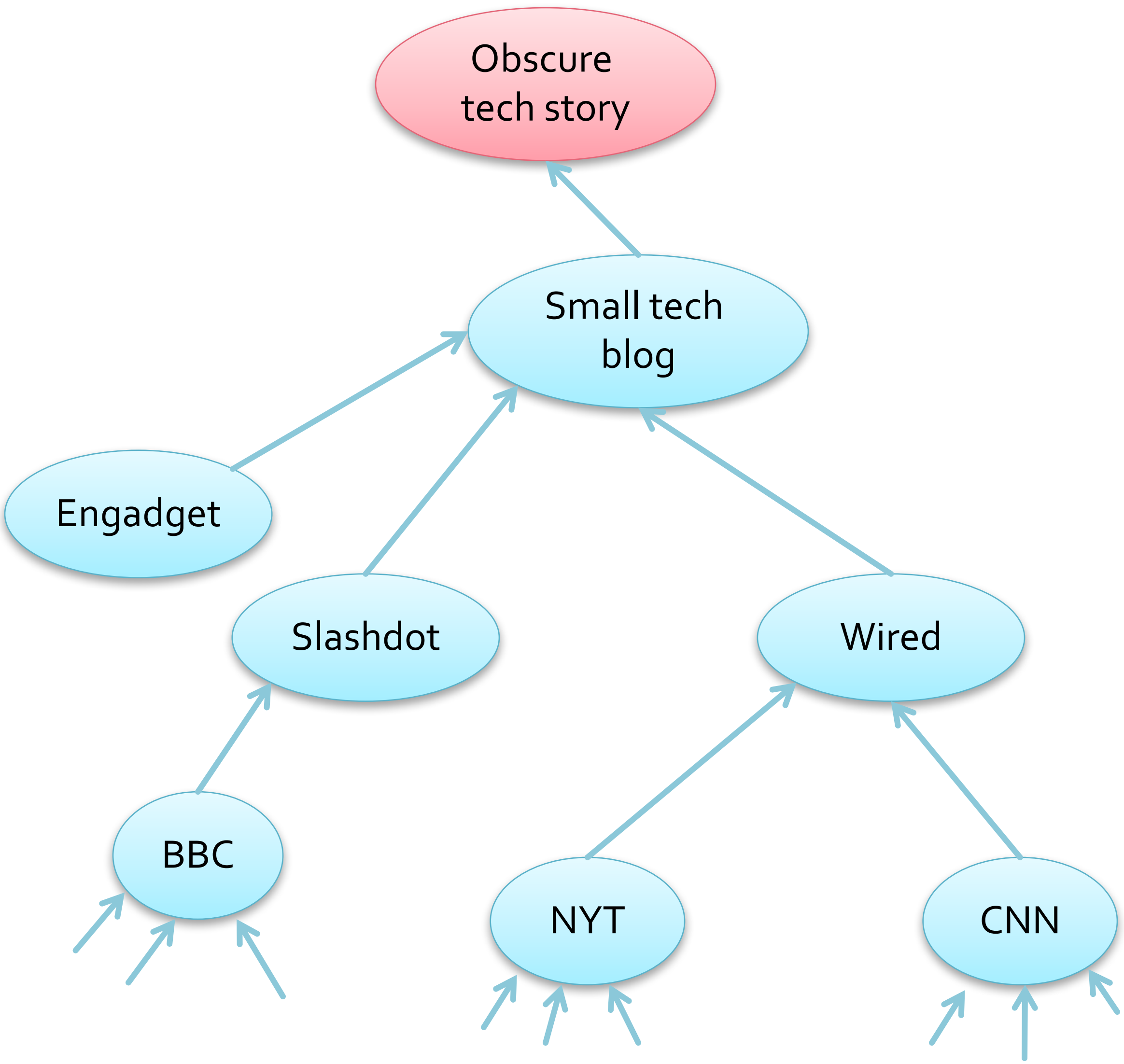
Observations	Models	Algorithms
Small diameter, Edge clustering	Erdös-Renyi model, Small-world model	Decentralized search
Patterns of signed edge creation	Structural balance, Theory of status	Models for predicting edge signs
Viral Marketing, Blogosphere, Memetracking	Independent cascade model, Game theoretic model	Influence maximization, Outbreak detection, LIM
Scale-Free	Preferential attachment, Copying model	PageRank, Hubs and authorities
Densification power law, Shrinking diameters	Microscopic model of evolving networks	Link prediction, Supervised random walks
Strength of weak ties, Core-periphery	Kronecker Graphs	Community detection: Girvan-Newman, Modularity

Spreading Through Networks

- **Spreading through networks:**
  - Cascading behavior
  - Diffusion of innovations
  - Network effects
  - Epidemics
- **Behaviors that cascade from node to node like an epidemic**

- **Examples:**
  - **Biological:**
    - Diseases via contagion
  - **Technological:**
    - Cascading failures
    - Spread of information
  - **Social:**
    - Rumors, news, new technology
    - Viral marketing

Information Diffusion: Media



Twitter & Facebook post sharing

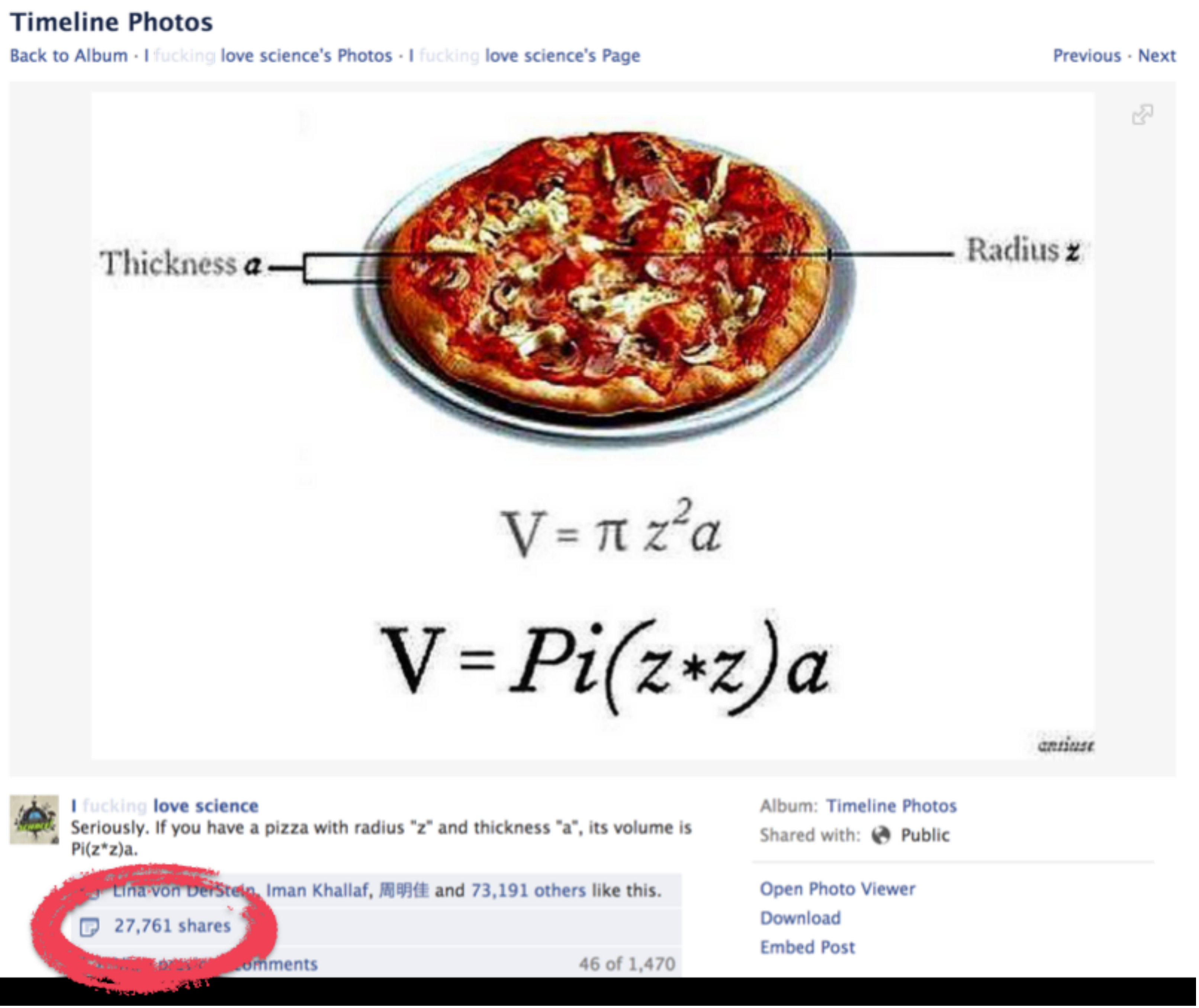
Lada Adamic shared a link via Erik Johnston.
January 16, 2013

When life gives you an almost empty jar of nutella, add some ice cream... (and other useful tips)

50 Life Hacks to Simplify your World
twistedsifter.com

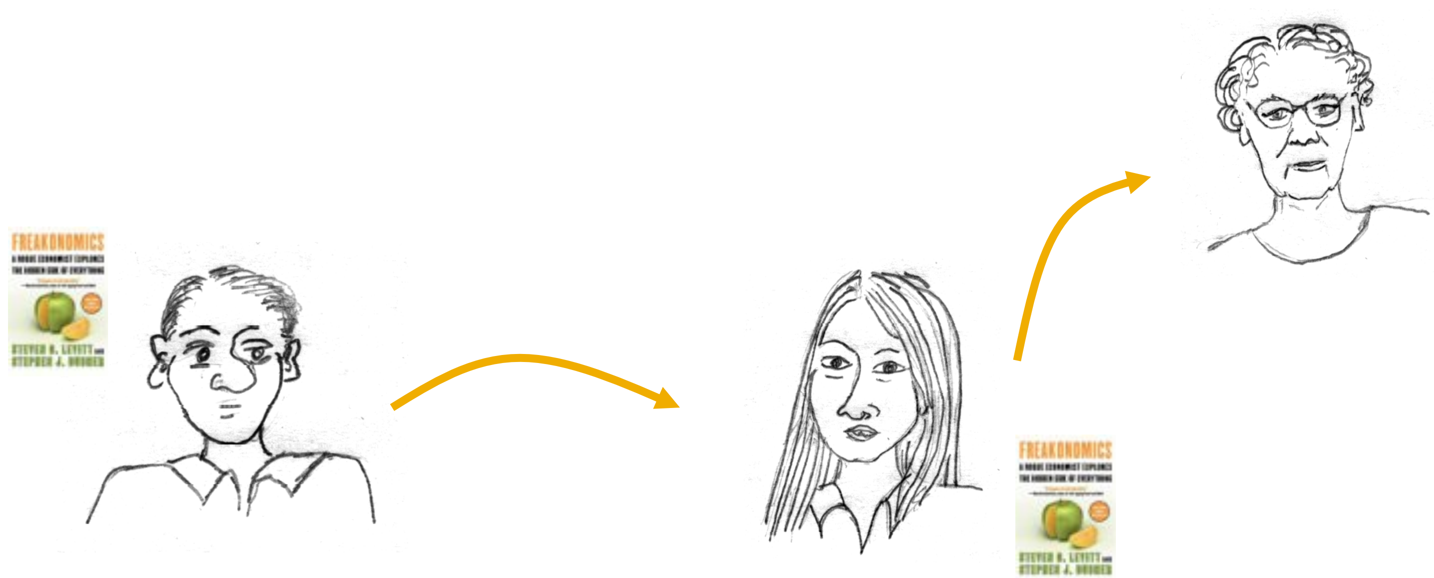
Life hacks are little ways to make our lives easier. These low-budget tips and trick can help you organize and de-clutter space; prolong and preserve your products; or teach you...

Like · Comment · Share 40 3 25

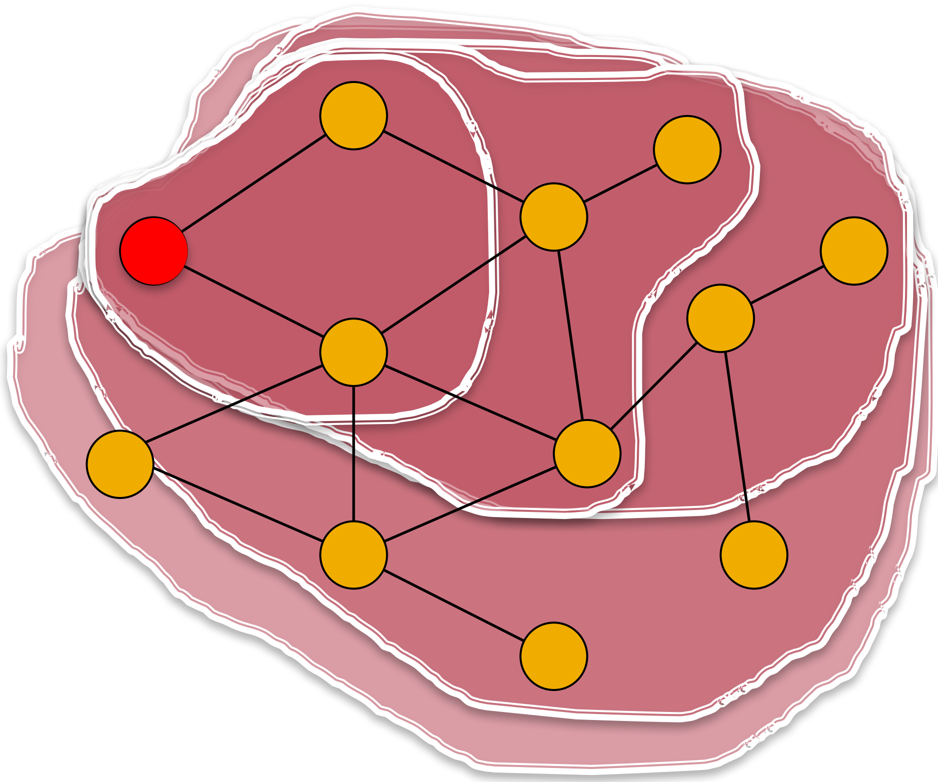

Timeline Photos
Back to Album · I fucking love science's Photos · I fucking love science's Page
Previous · Next
Thickness a
Radius z
V = π z²a
V = Pi(z*z)a
I fucking love science
Seriously. If you have a pizza with radius "z" and thickness "a", its volume is Pi(z*z)a.
Iman Khallaf, 周明佳 and 73,191 others like this.
27,761 shares
46 of 1,470
Album: Timeline Photos
Shared with: Public
Open Photo Viewer
Download
Embed Post

Diffusion in Viral Marketing

- **Product adoption:**
  - **Senders and followers of recommendations**

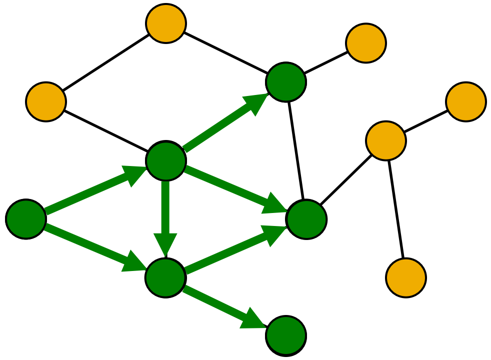


Spread of Diseases (e.g., Ebola)

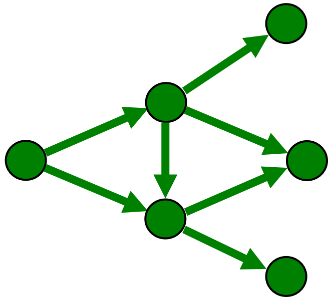


Network Cascades

- **Contagion that spreads over the edges of the network**
- It creates a propagation tree, i.e., **cascade**



Network

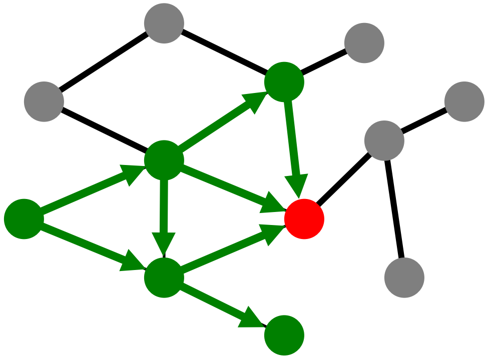


Cascade
(propagation graph)

Terminology:
- Stuff that spreads: Contagion
- "Infection" event: Adoption, infection, activation
- We have: Infected/active nodes, adoptors

How Do We Model Diffusion?

- **Decision based models (first!):**
  - Models of product adoption, decision making
    - A node observes decisions of its neighbors and makes its own decision
  - **Example:**
    - You join demonstrations if *k* of your friends do so too
- **Probabilistic models (next):**
  - **Models of influence or disease spreading**
    - An infected node tries to "push" the contagion to an uninfected node
  - **Example:**
    - You "catch" a disease with some prob. from each active neighbor in the network



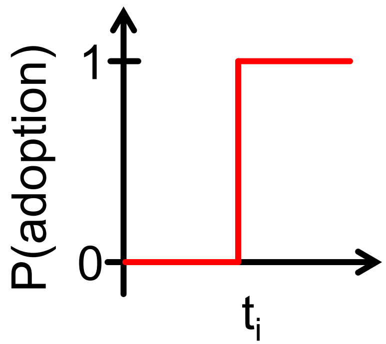
Granovetter's Model of Collective Action

Decision Based Models

- Collective Action [Granovetter, '78]
  - **Model where everyone sees everyone else's behavior (that is, we assume a complete graph)**
  - **Examples:**
    - Clapping or getting up and leaving in a theater
    - Keeping your money or not in a stock market
    - Neighborhoods in cities changing ethnic composition
    - Riots, protests, strikes
- **How does the number of people participating in a given activity grow or shrink over time?**

Collective Action: The Model

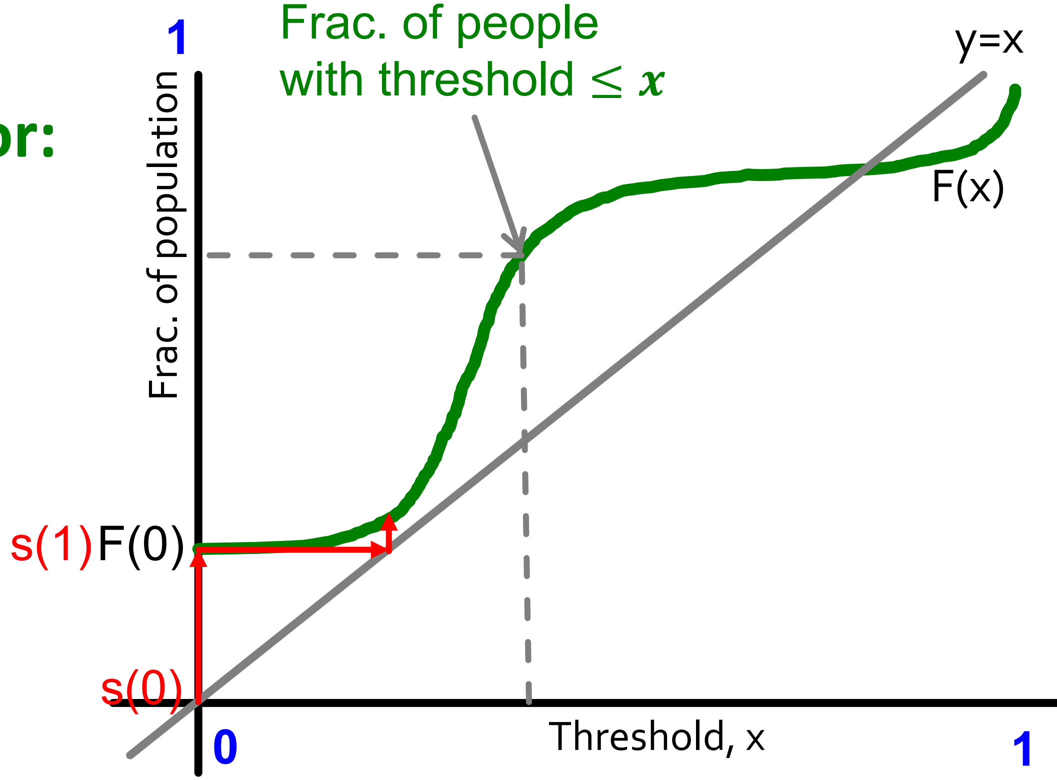
- ***n* people – everyone observes all actions**
- Each person ***i*** has a threshold $\boldsymbol{t_i}$ $(0 \leq t_i \leq 1)$
  - Node ***i*** will adopt the behavior iff at least $\boldsymbol{t_i}$ fraction of people have already adopted:
    - **Small $\boldsymbol{t_i}$**: early adopter
    - **Large $\boldsymbol{t_i}$**: late adopter
  - Time moves in discrete steps



- **The population is described by $\{t_1,\ldots,t_n\}$**
  - **F(x)** … fraction of people with threshold $\boldsymbol{t_i} \leq \boldsymbol{x}$
    - F(x) is given to us. F(x) is a property of the contagion.

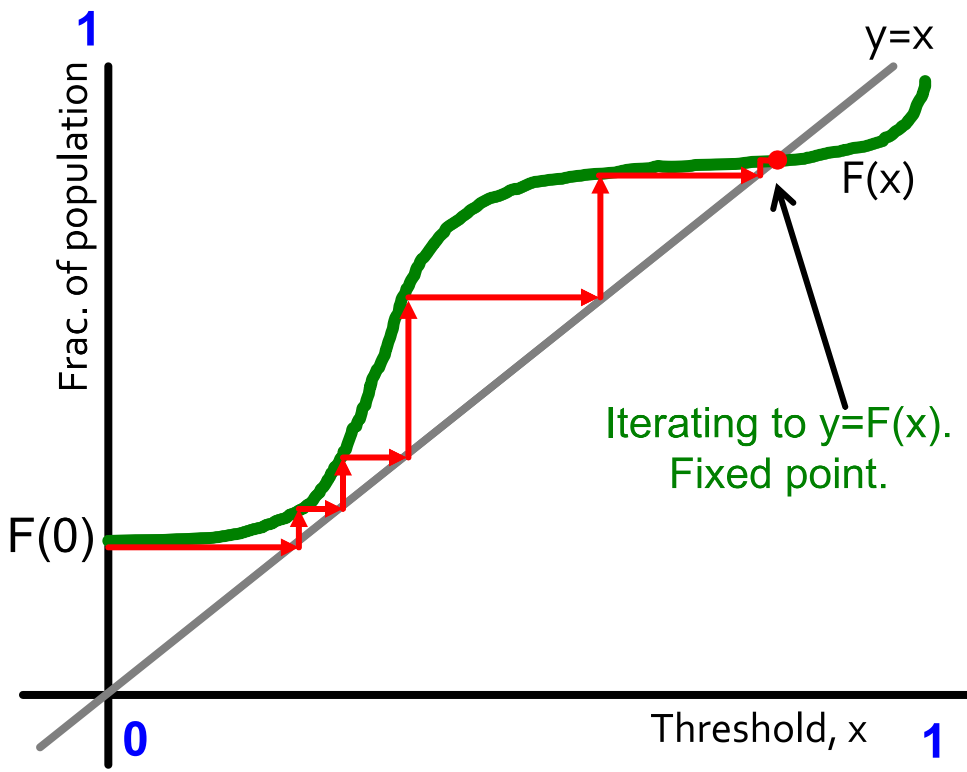
Collective Action: Dynamics

- **F(x)** … fraction of people with threshold $t_i \leq x$
  - **F(x) is non-decreasing:** $F(x+\varepsilon) \geq F(x)$
- **The model is dynamic:**
  - **Step-by-step change in number of people adopting the behavior:**
    - **F(x)** … frac. of people with threshold ≤ **x**
    - **s(t)** … frac. of people participating at time **t**
  - **Simulate:**
    - s(0) = 0
    - s(1) = F(0)
    - s(2) = F(s(1)) = F(F(0))



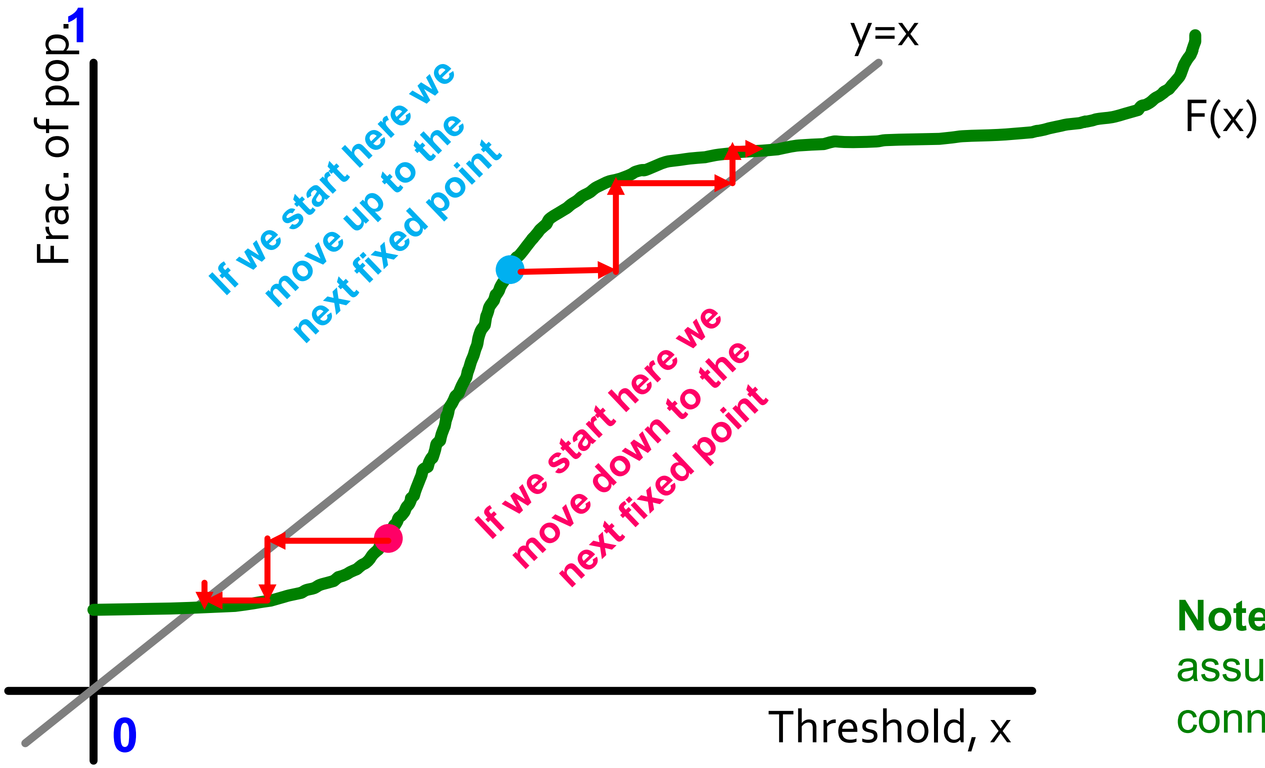
Collective Action: Dynamics

- **Step-by-step change in number of people :**
  - **F(x)** … fraction of people with threshold ≤ **x**
  - **s(t)** … number of participants at time **t**
- **Easy to simulate:**
  - $s(0) = 0$
  - $s(1) = F(0)$
  - $s(2) = F(s(1)) = F(F(0))$
  - $s(t+1) = F(s(t)) = F^{t+1}(0)$
- **Fixed point: F(x)=x**
  - Updates to **s(t)** to converge to a stable fixed point **x=y**
  - There could be other fixed points but starting from **0** we only reach the first one



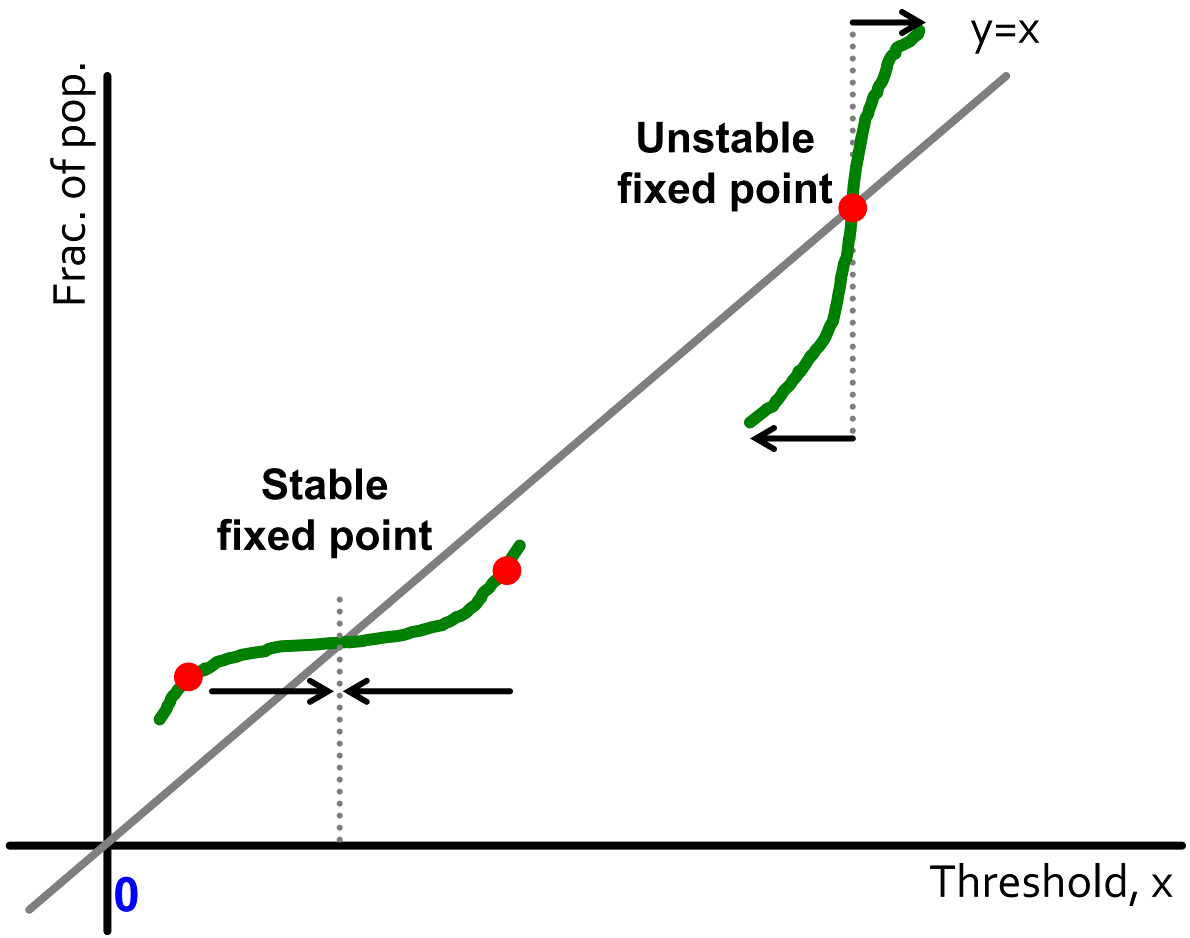
Starting Elsewhere

- **What if we start the process somewhere else?**
  - We move up/down to the next fixed point
  - **How is market going to change?**



Note: we are assuming a fully connected graph

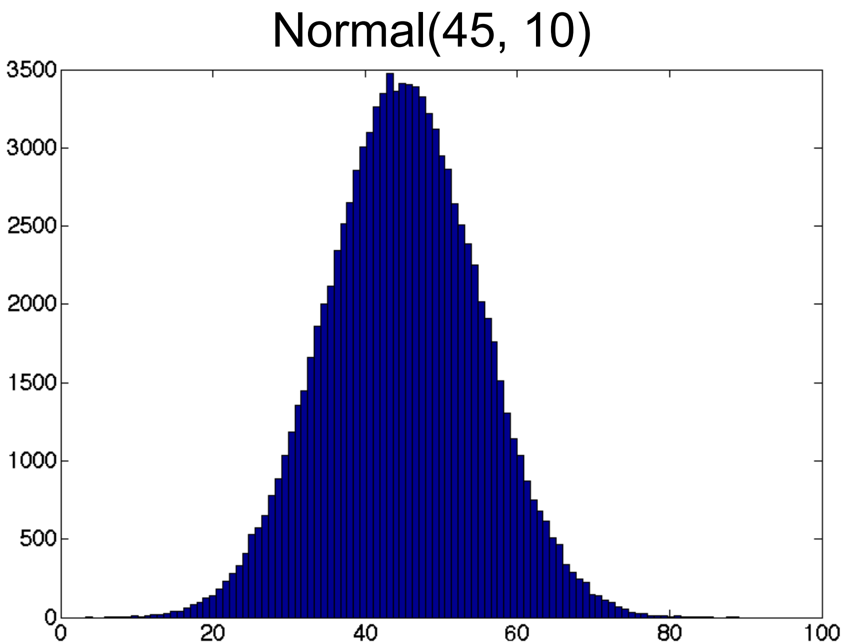
Stable vs. Unstable Fixed Point



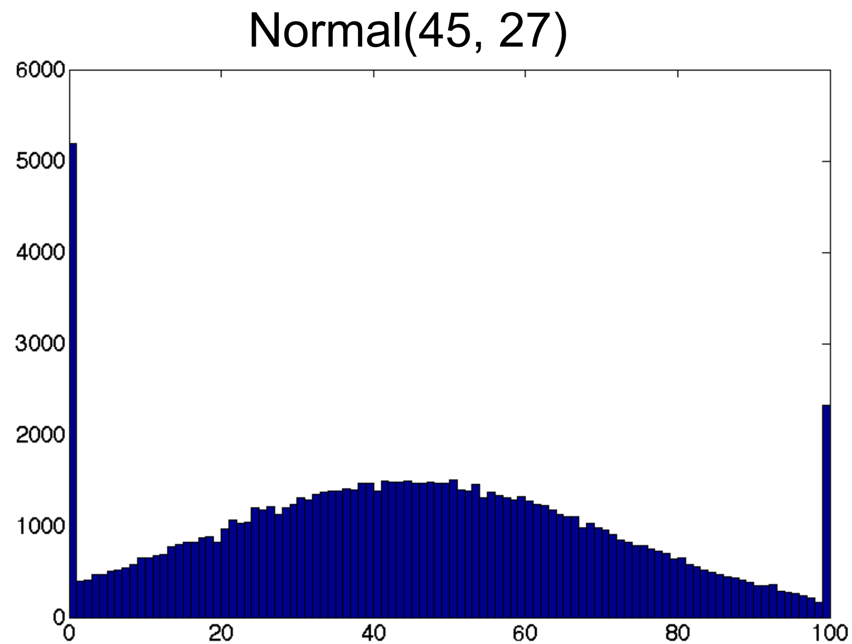
Discontinuous Transition

- Each threshold t_i is drawn independently from some distribution ***F(x) = Pr[thresh ≤ x]***
  - **Suppose:** Truncated normal with **μ=45**, variance **σ**

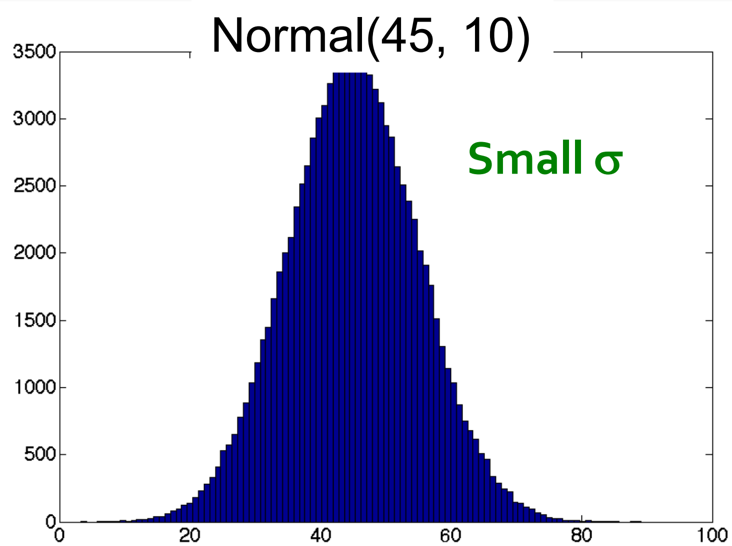
Small σ:

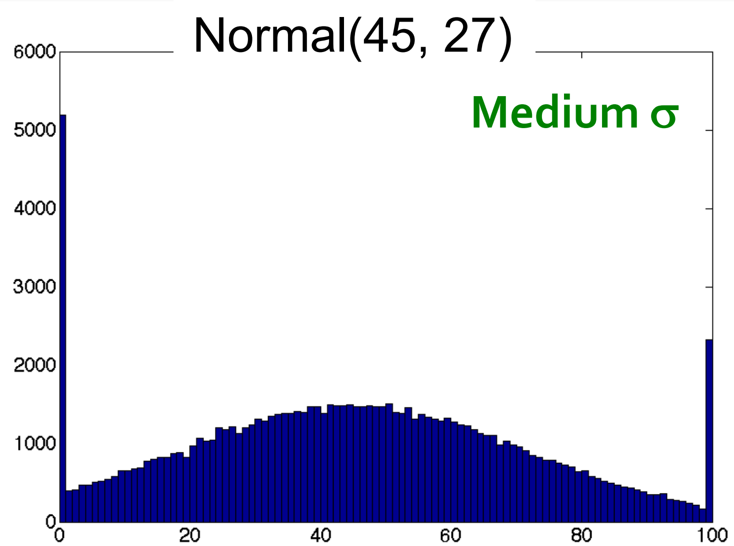


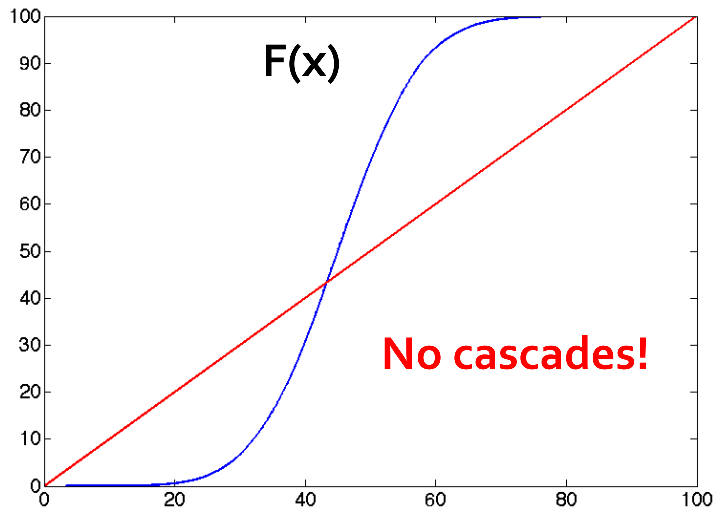
Large σ:

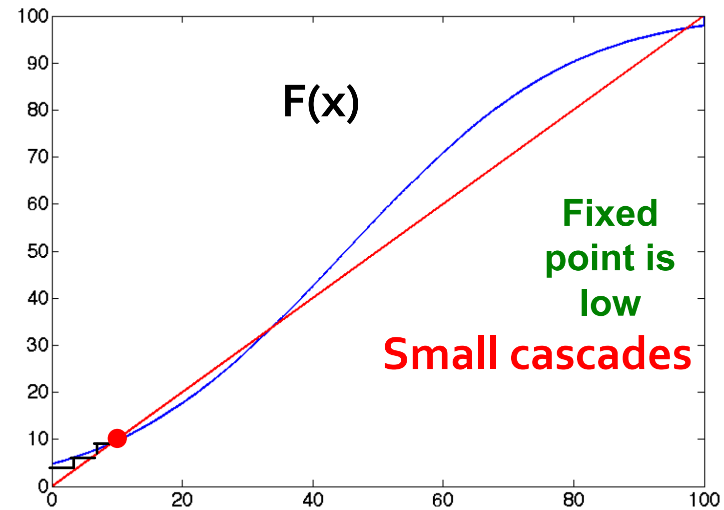


Discontinuous Transition



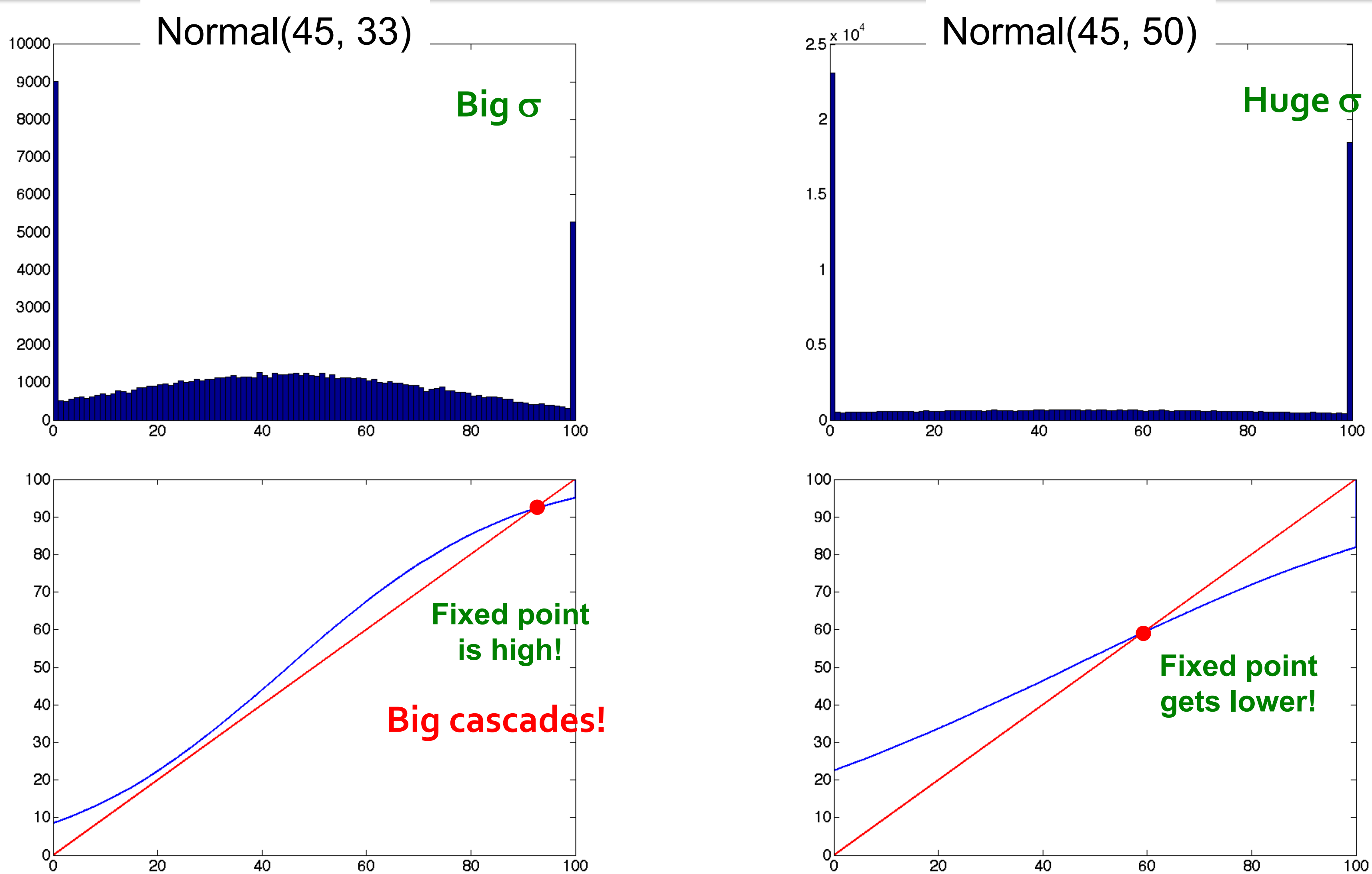






Bigger variance lets you build a bridge from early adopters to mainstream

Discontinuous Transition



But if we increase the variance the fixed point starts going down!

Weaknesses of the Model

- **No notion of social network:**
  - Some people are more influential
  - It matters who the early adopters are, not just how many
- **Models people's awareness** of size of participation **not just actual number of people participating**
  - Modeling perceptions of who is adopting the behavior vs. who you believe is adopting
  - Non-monotone behavior – dropping out if too many people adopt
  - People get "locked in" to certain choice over a period of time
- **Modeling thresholds**
  - Richer distributions
  - Deriving thresholds from more basic assumptions
    - Game theoretic models

Pluralistic Ignorance

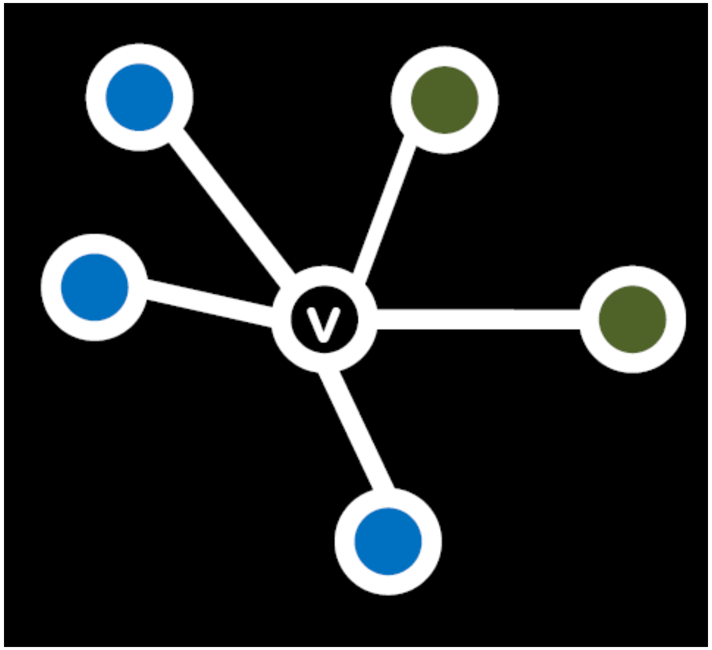
- <u>**Dictator tip:**</u> **Pluralistic ignorance** – erroneous estimates about the prevalence of certain opinions in the population
  - Survey conducted in the U.S. in 1970 showed that while a clear minority of white Americans at that point favored racial segregation, significantly more than 50% **believed** that it was favored by a majority of white Americans in their region of the country

Decision Based Model of Diffusion

[Morris 2000]

Game Theoretic Model of Cascades

- **Based on 2 player coordination game**
  - 2 players – each chooses technology A or B
  - Each person can only adopt **one** "behavior", **A** or **B**
  - You gain more payoff if your friend has adopted the **same** behavior as you



Local view of the network of node **v**

Example: VHS vs. BetaMax

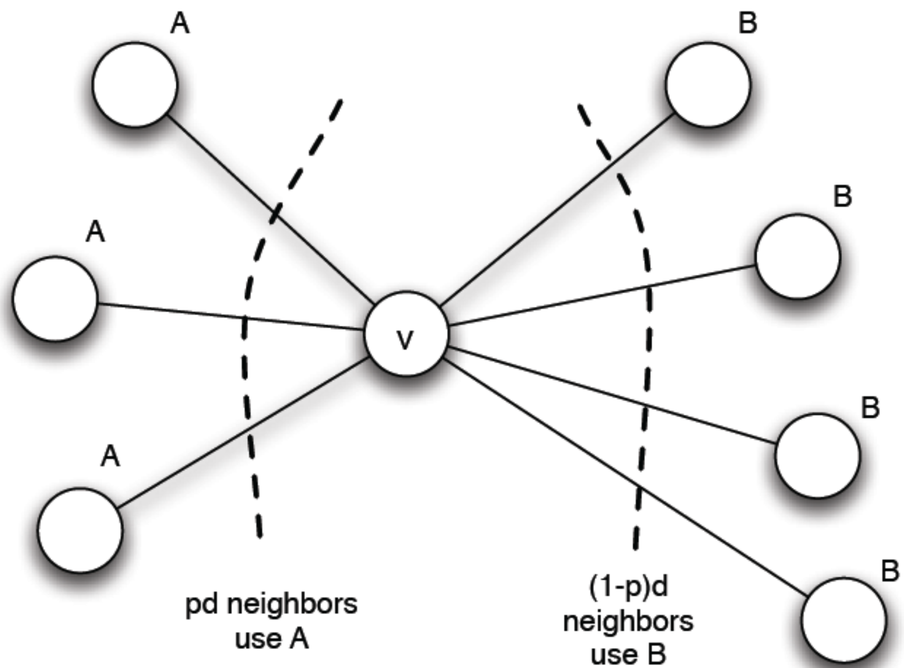
Example: BlueRay vs. HD DVD

The Model for Two Nodes

- ***Payoff matrix:***
  - If both **v** and **w** adopt behavior **A**, they each get payoff ***a > 0***
  - If ***v*** and ***w*** adopt behavior ***B***, they reach get payoff ***b > 0***
  - If ***v*** and ***w*** adopt the opposite behaviors, they each get ***0***

- **In some large network:**
  - Each node ***v*** is playing a copy of the game with each of its neighbors
  - **Payoff**: sum of node payoffs per game

Calculation of Node *v*



Threshold:
v chooses ***A*** if

$$p > \frac{b}{a+b} = q$$

p… frac. v's nbrs. with A
***q*… payoff threshold**

- Let **v** have ***d*** neighbors
- Assume fraction ***p*** of ***v***'s neighbors adopt ***A***
  - *Payoff*$_v$ $= a \cdot p \cdot d$ if *v* chooses A
    $= b \cdot (1-p) \cdot d$ if *v* chooses B
- **Thus: *v* chooses A if: *a·p·d > b·(1-p)·d***

Example Scenario

Scenario:

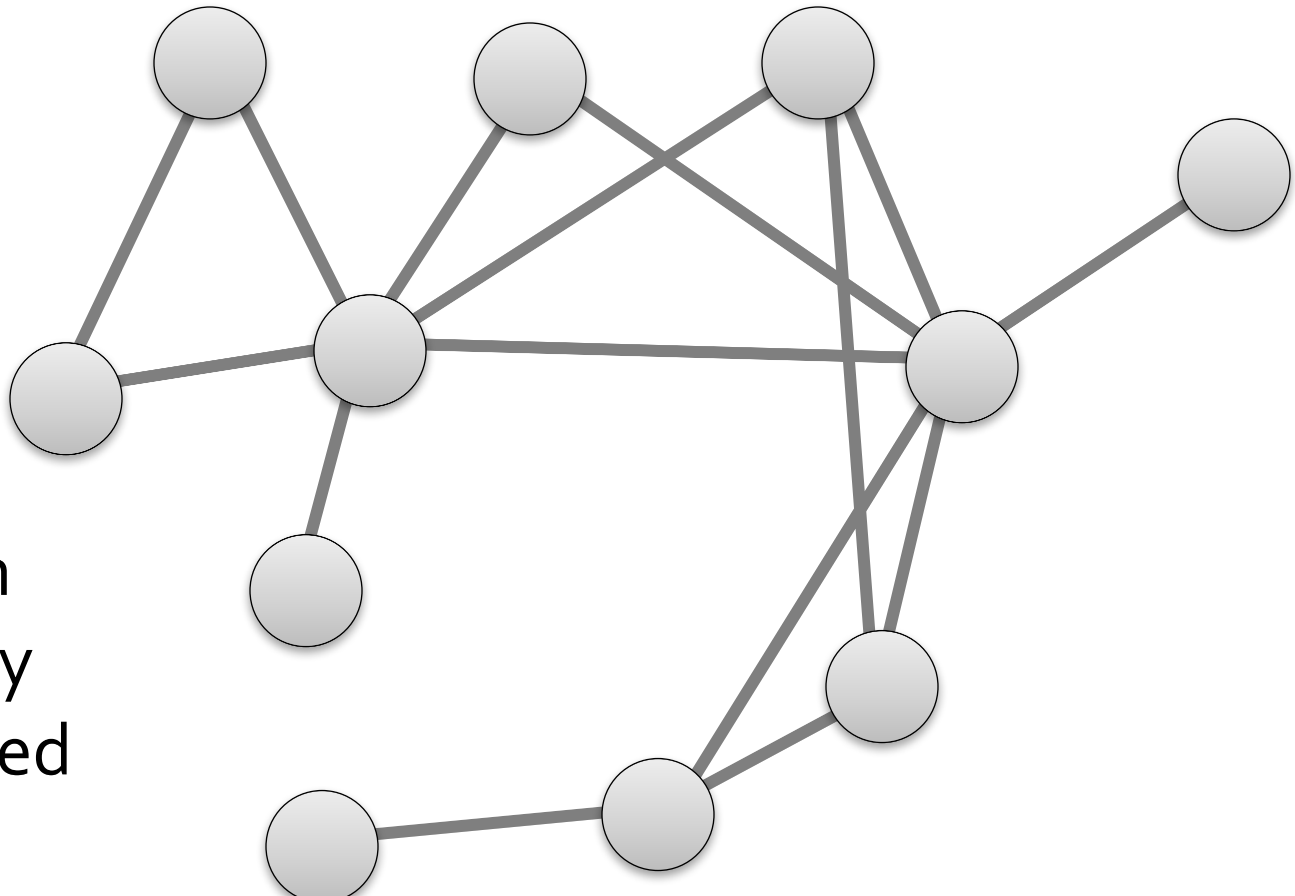
- Graph where everyone starts with all ***B***
- Small set ***S*** of early adopters of ***A***
  - Hard-wire ***S*** – they keep using ***A*** no matter what payoffs tell them to do

- **Assume payoffs are set in such a way that nodes say:**
  If more than q=50% of my friends take A I'll also take A

  This means: $\mathbf{a = b - \varepsilon}$ ($\varepsilon > 0$, small positive constant) and $\mathbf{q = 1/2}$

Example Scenario

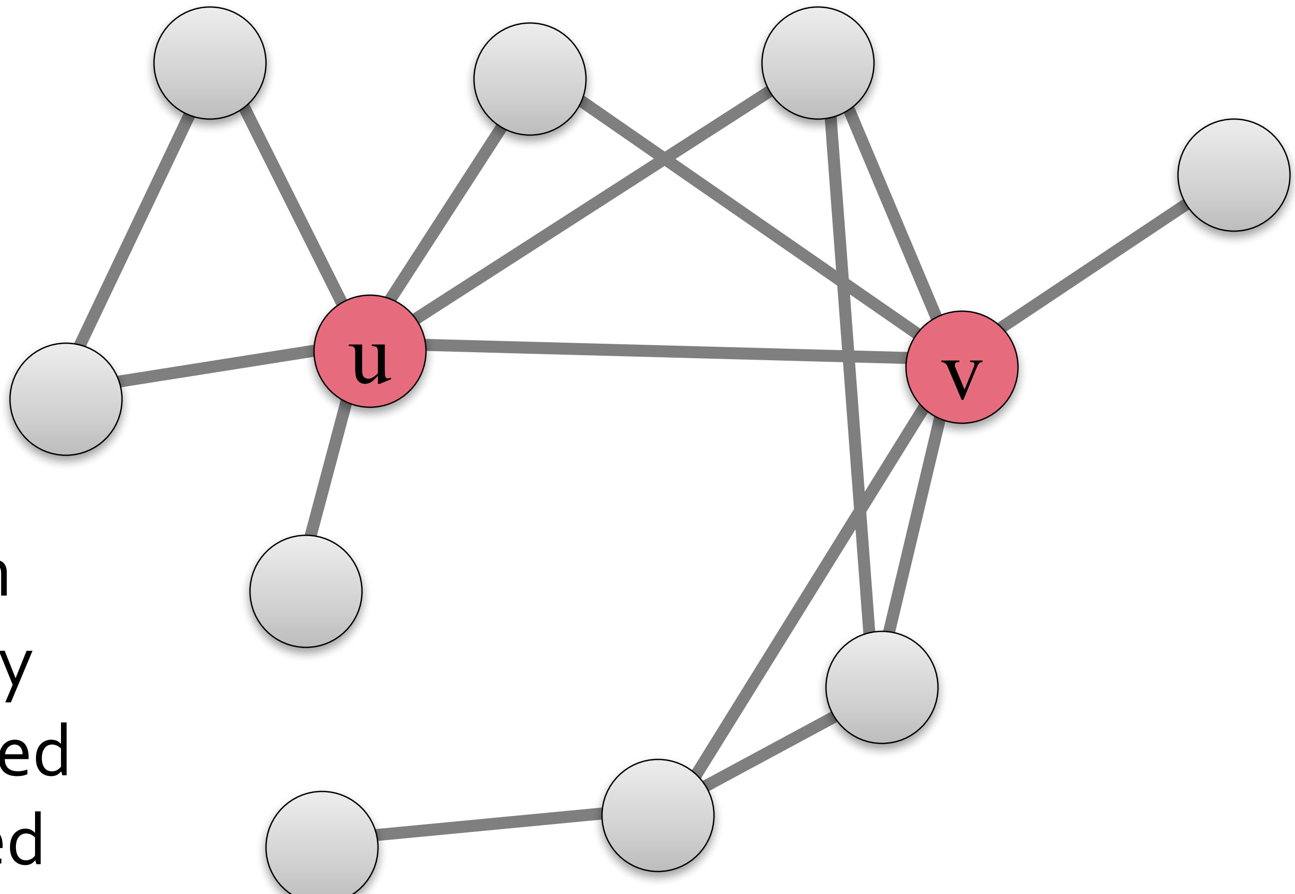
$$S = \{u, v\}$$



If **more** than **q=**50% of my friends are red I'll be red

Example Scenario

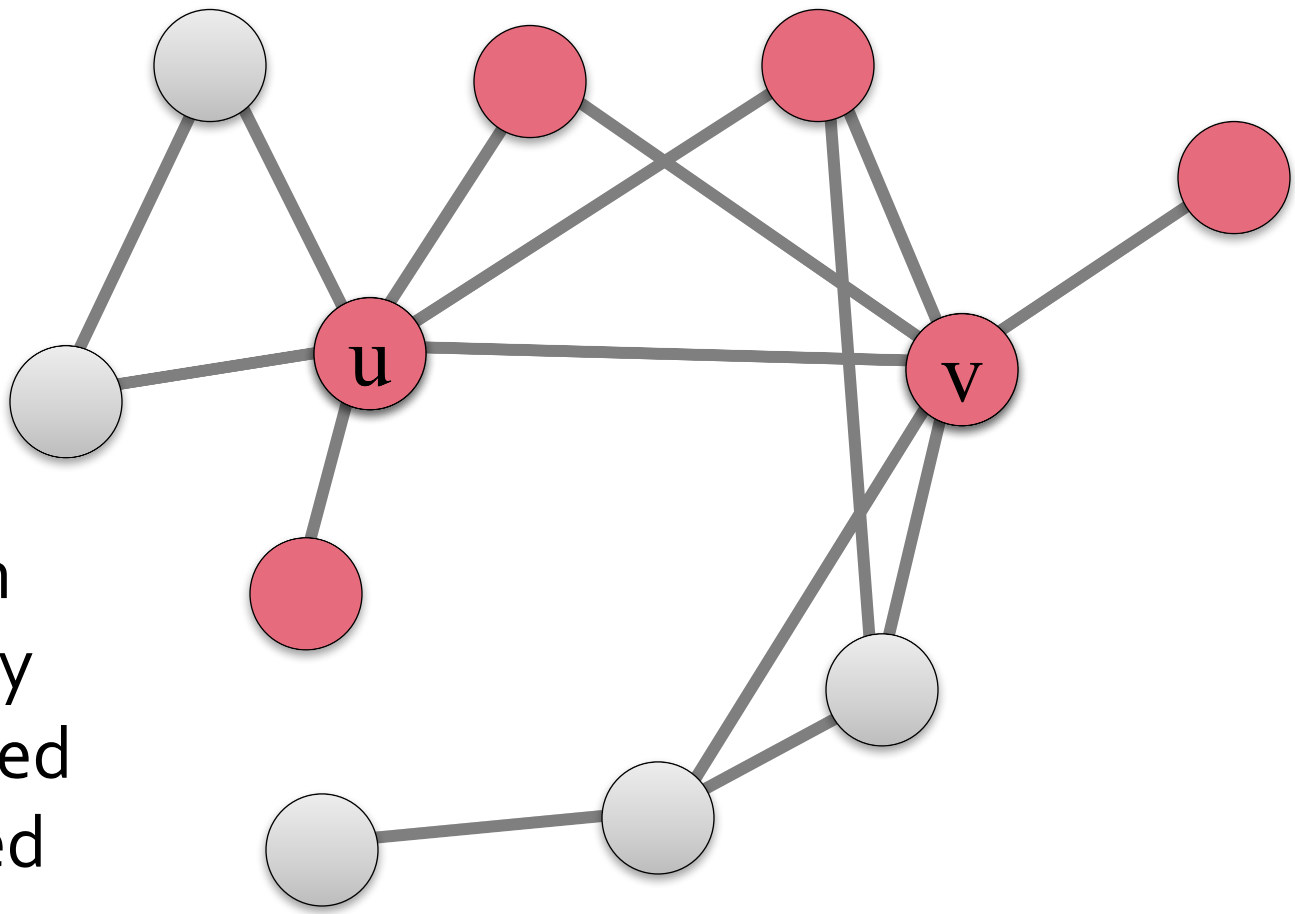
$S = \{u, v\}$



If **more** than **q=**50% of my friends are red I'll also be red

Example Scenario

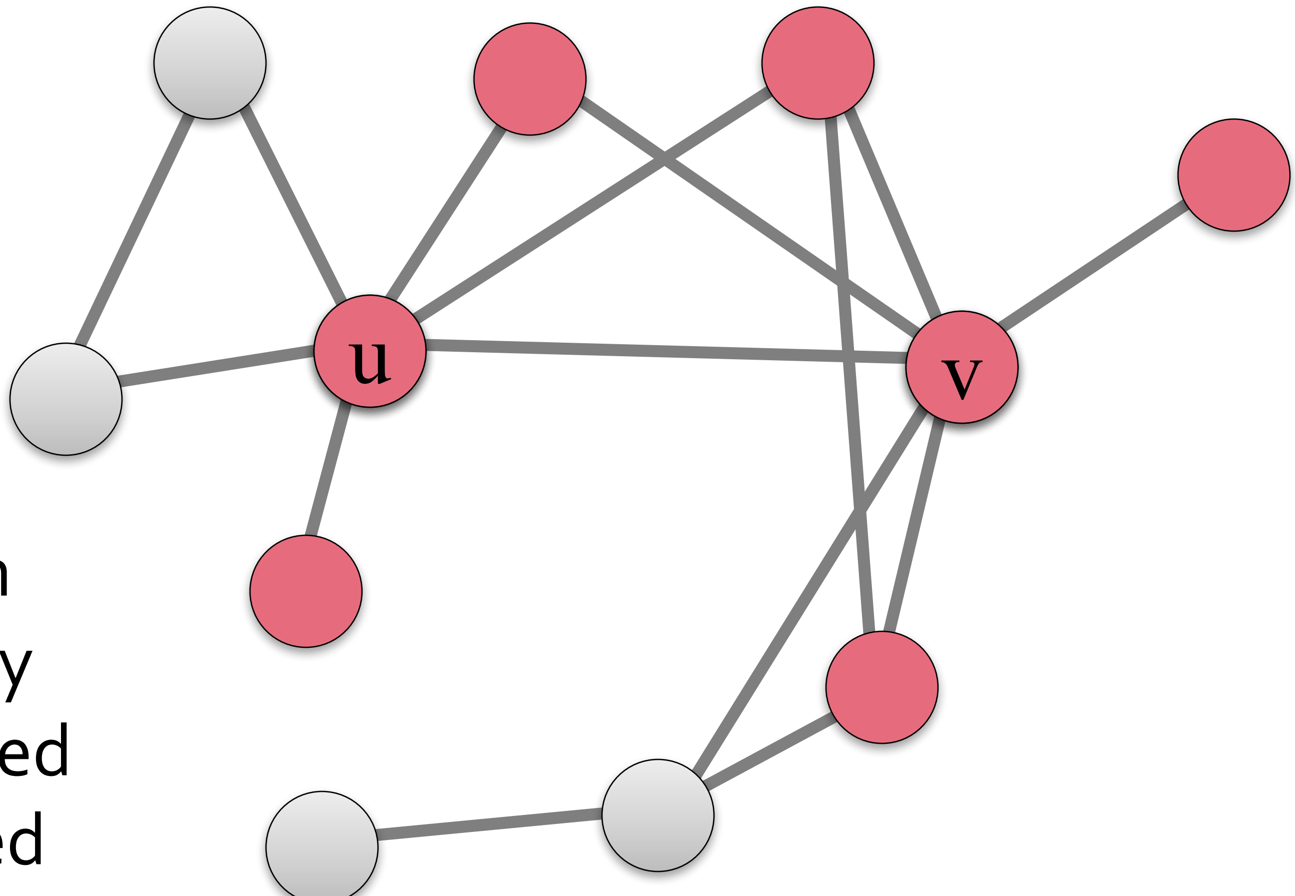
$S = \{u, v\}$



If **more** than **q=**50% of my friends are red I'll also be red

Example Scenario

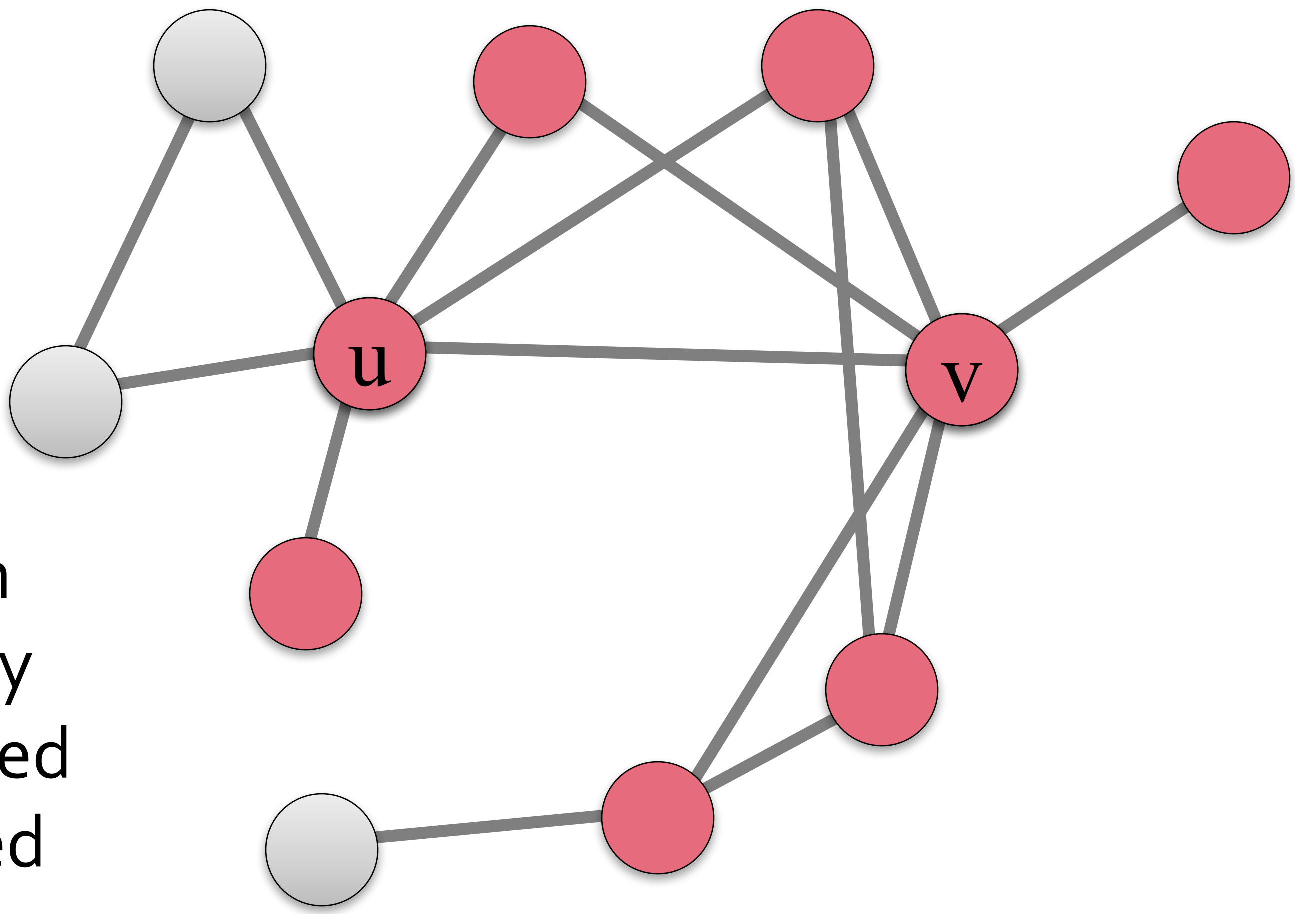
$S = \{u, v\}$



If **more** than **q=**50% of my friends are red I'll also be red

Example Scenario

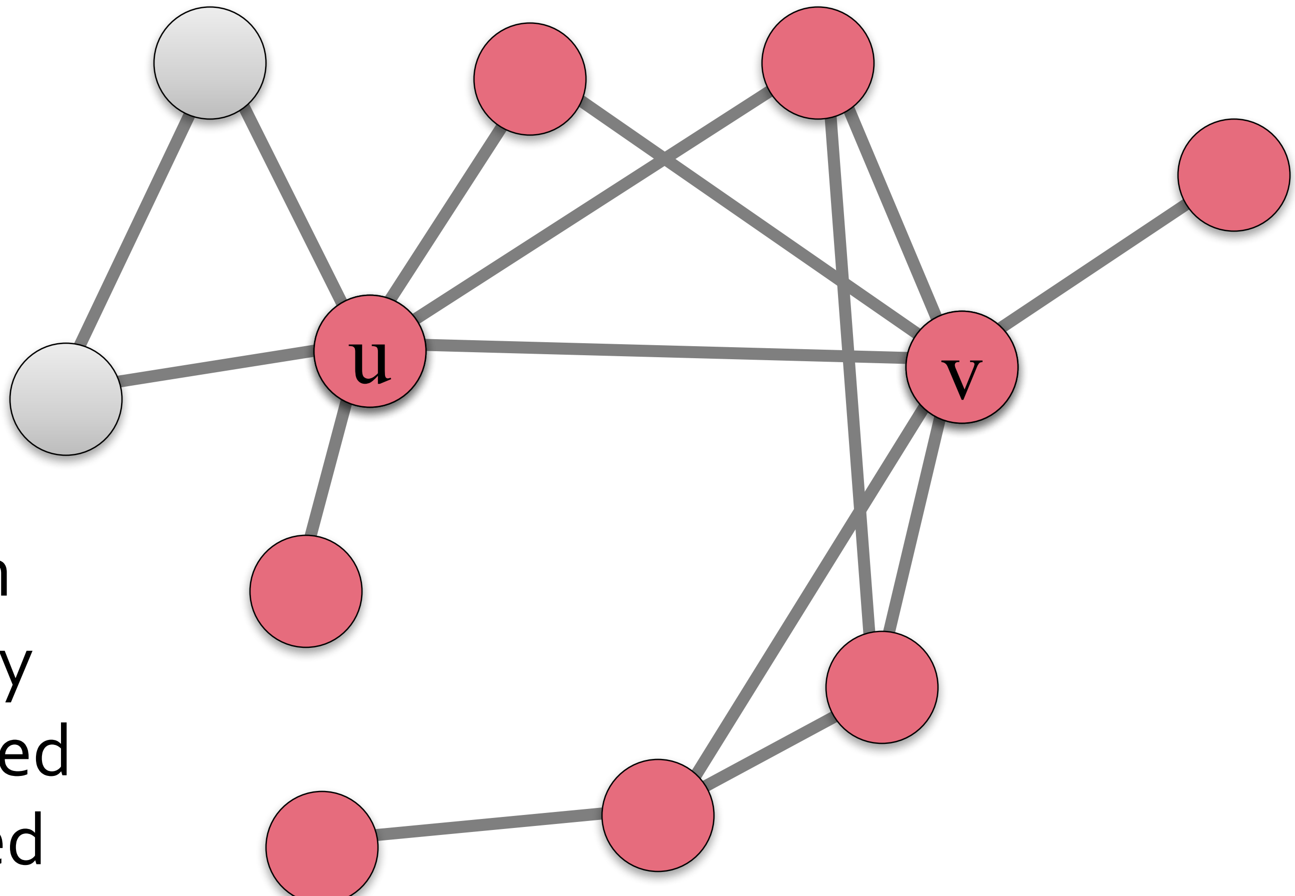
$S = \{u, v\}$



If **more** than **q=**50% of my friends are red I'll also be red

Example Scenario

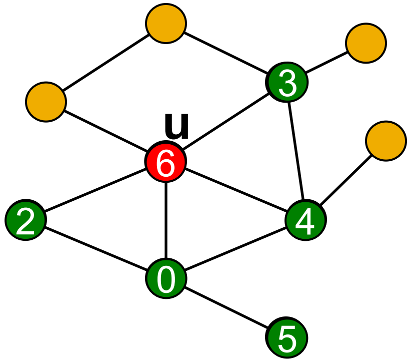
$S = \{u, v\}$



If **more** than **q=**50% of my friends are red I'll also be red

Monotonic Spreading

- **Observation: Use of A spreads monotonically**
  (Nodes only switch **B→A**, but never back to **B**)
- **Why?** Proof sketch:
  - **Nodes keep switching from B to A:** $B \to A$
  - Now, suppose some node switched back from ***A→B***, consider the **first** node ***u*** (not in **S**) to do so (say at time ***t***)
  - Earlier at some time ***t'*** (***t'***<***t***) the same node ***u*** switched ***B→A***
  - So at time ***t' u*** was above threshold for ***A***
  - But up to time ***t*** no node switched back to **B**, so node ***u*** could only have more neighbors who used ***A*** at time ***t*** compared to ***t'***.
    There was no reason for *u* to switch at the first place!



!! Contradiction !!

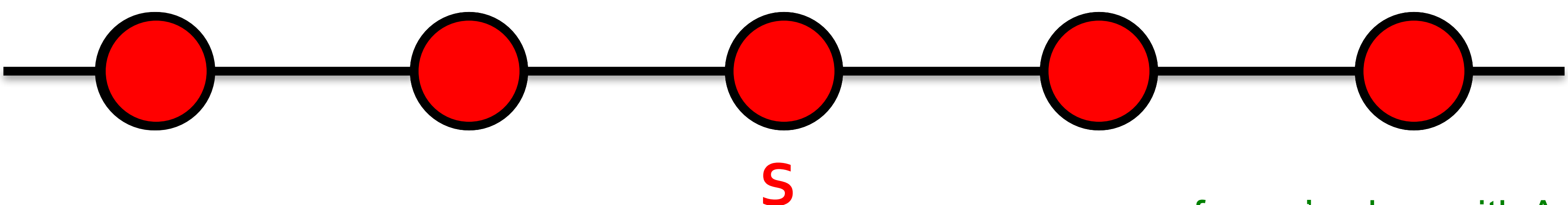
Infinite Graphs

v chooses *A* if *p>q*

$$q = \frac{b}{a+b}$$

- **Consider infinite graph *G***
  - (but each node has finite number of neighbors!)
- We say that a finite set ***S*** causes a cascade in *G* with **threshold *q*** if, when ***S*** adopts ***A***, eventually **every node in G adopts *A***
- Example: **Path**

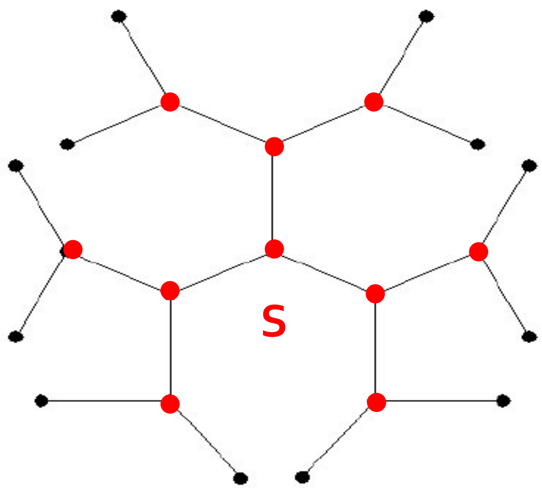
If ***q<1/2*** then cascade occurs



p… frac. v's nbrs. with A
***q*… payoff threshold**

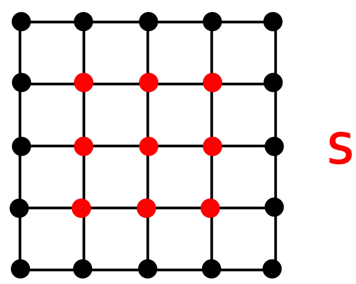
Infinite Graphs

- **Infinite Tree:**



If *q<1/3* then cascade occurs

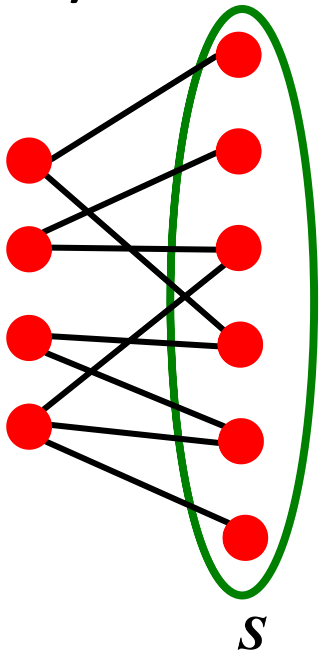
- **Infinite Grid:**



If *q<1/4* then cascade occurs

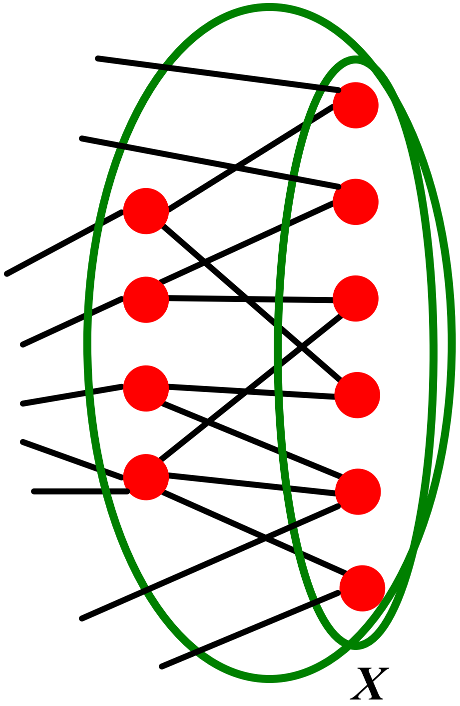
Cascade Capacity

- Def:
  - The **cascade capacity** of a graph ***G*** is the **largest *q*** for which some **finite set S** can cause a **cascade**
- Fact:
  - There is no (infinite) ***G*** where cascade capacity **> ½**
- **Proof idea:**
  - Suppose such **G** exists: **q>½**, finite **S** causes cascade
  - **Show contradiction:** Argue that nodes stop switching after a finite # of steps



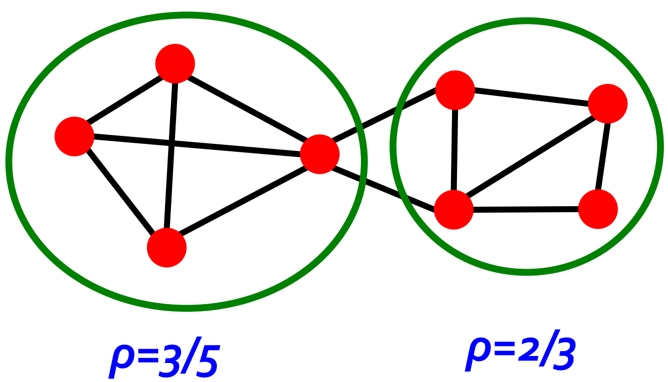
Cascade Capacity

- **<u>Fact:</u>** There is no ***G*** where cascade capacity **> ½**
- **Proof sketch:**
  - Suppose such **G** exists: **q>½**, finite **S** causes cascade
  - **Contradiction:** Switching stops after a finite # of steps
    - <u>Define</u> "potential energy"
    - Argue that it starts finite (non-negative) and strictly decreases at every step
  - "Energy": = **$|d^{out}(X)|$**
    - **$|d^{out}(X)|$** := # of outgoing edges of active set *X*
    - The only nodes that switch have a strict majority of its neighbors in ***S***
    - **$|d^{out}(X)|$** strictly decreases
    - It can do so only for a finite number of steps



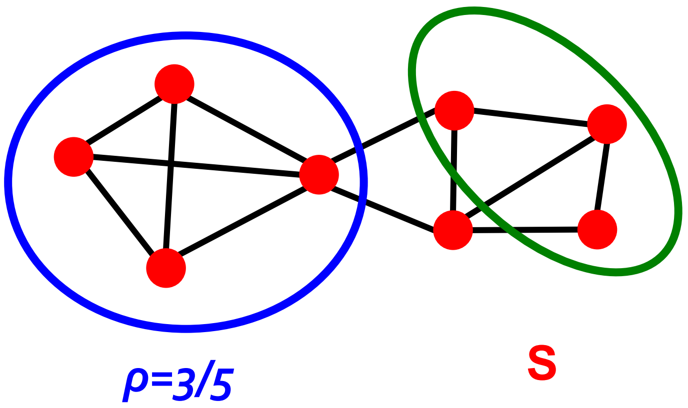
Stopping Cascades

- **What prevents cascades from spreading?**
- Def: **Cluster of density *ρ*** is a **set of nodes *C*** where **each node** in the set has at least ***ρ*** fraction of neighbors in ***C***



Stopping Cascades

- **Let S be an initial set of adopters of A**
- All nodes apply threshold ***q*** to decide whether to switch to **A**
- **Two facts:**
  - 1) If ***G\S*** contains a cluster of density ***>(1-q)*** then ***S*** cannot cause a cascade
  - 2) If **S** fails to create a cascade, then there is a cluster of density ***>(1-q)*** in **G\S**



No cascade if q>2/5

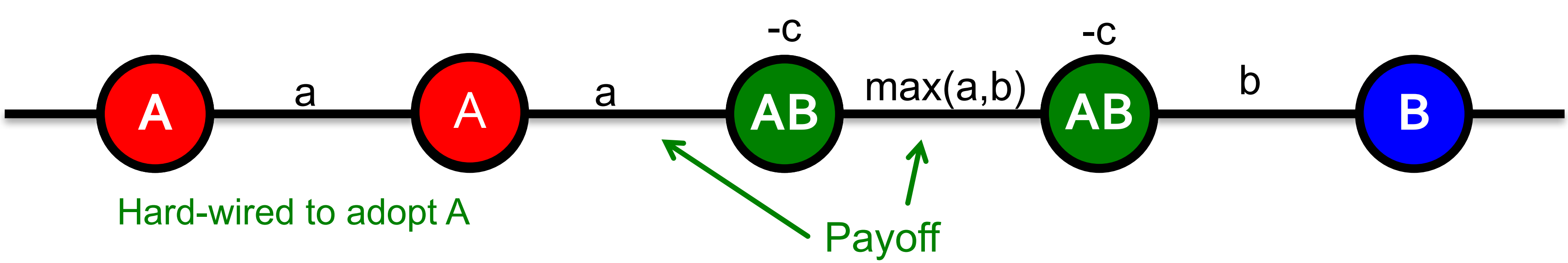
Extending the Model: Allow People to Adopt A and B

Cascades & Compatibility

- **So far:**
  - Behaviors ***A*** and ***B*** compete
  - Can only get utility from neighbors of same behavior: ***A-A*** **get a**, ***B-B*** **get *b***, ***A-B*** **get *0***
- **Let's add an extra strategy "*AB*"**
  - ***AB-A*** : gets ***a***
  - ***AB-B*** : gets ***b***
  - ***AB-AB*** : gets **max(*a*, *b*)**
  - **Also:** Some **cost *c*** for the effort of maintaining both strategies (summed over all interactions)
    - Note: a given node can receive *a* from one neighbor and *b* from another by playing AB, which is why it could be worth the cost *c*

Cascades & Compatibility: Model

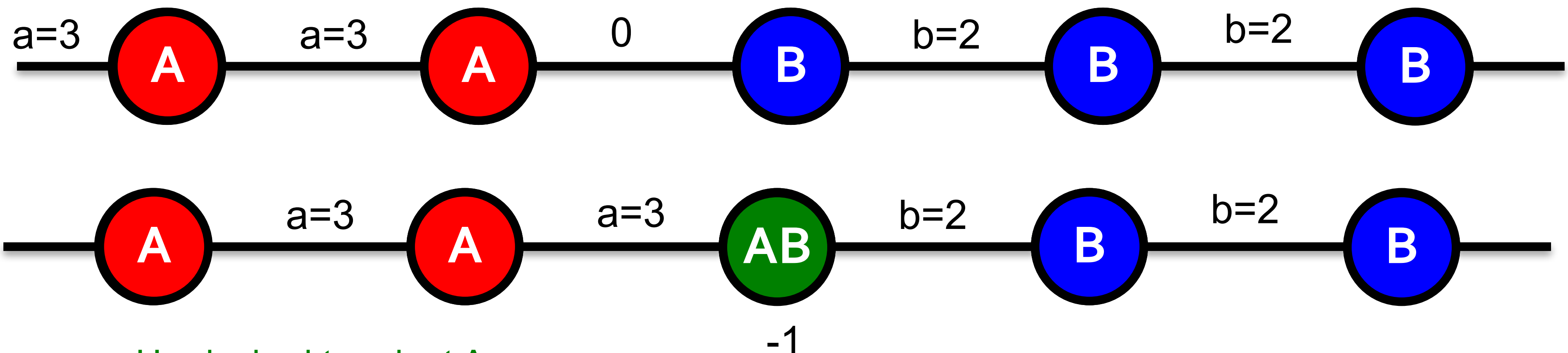
- Every node in an infinite network starts with ***B***
- Then a finite set ***S*** initially adopts ***A***
- Run the model for ***t=1,2,3,…***
  - Each node selects behavior that will optimize payoff (given what its neighbors did in at time ***t-1***)



- **How will nodes switch from *B* to *A* or *AB*?**

Example: Path Graph (1)

- **Path graph:** Start with **B**s, **a** > **b** (**A** is better)
- **One node switches to A – what happens?**
  - **With just A, B:** A spreads if ***a*** > ***b***
  - **With A, B, AB: Does A spread?**
- **Example: a=3, b=2, c=1**

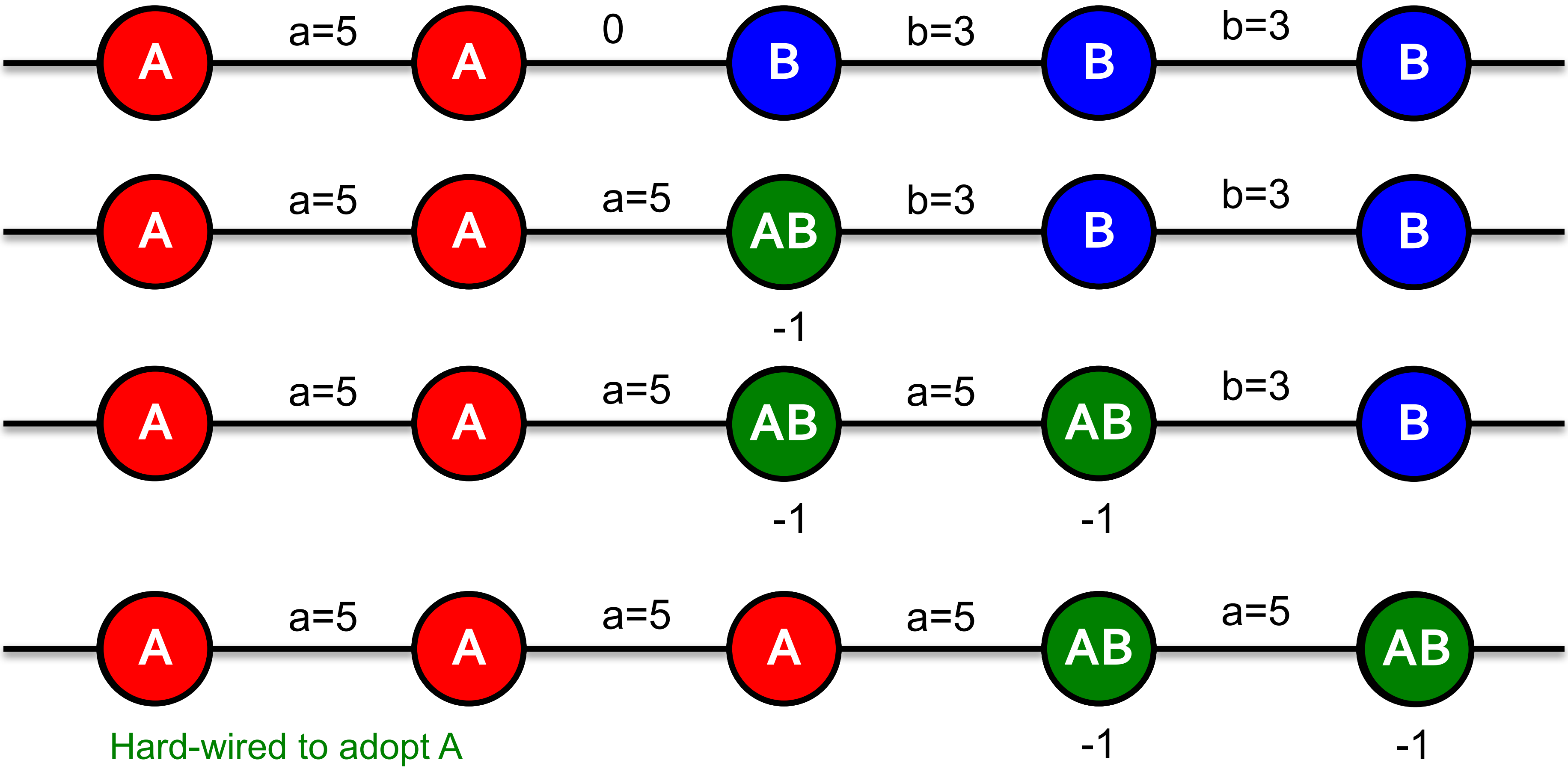


Hard-wired to adopt A

Cascade stops

Example: Path Graph (2)

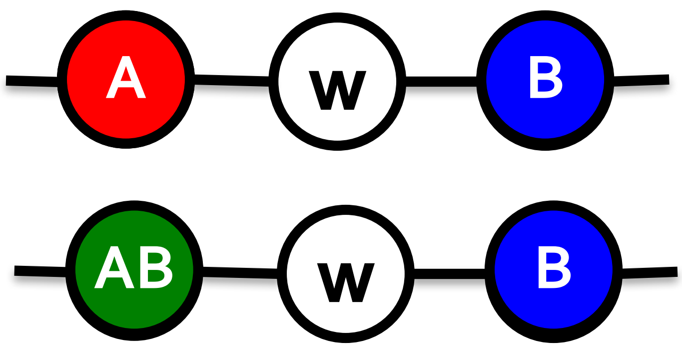
- **Example: a=5, b=3, c=1**



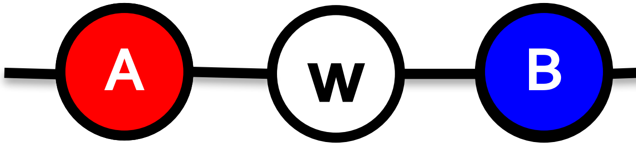
Cascade never stops!

What about in a general case?

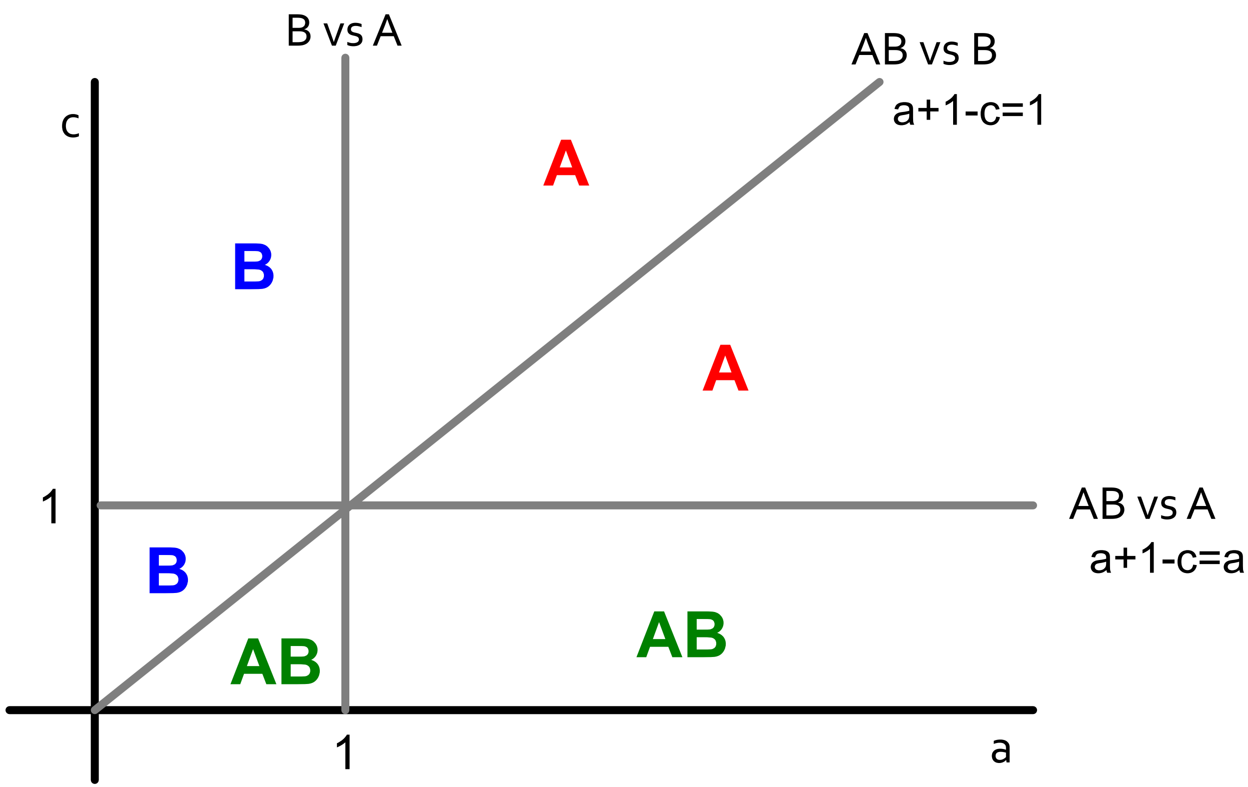
- Let's solve the model in a general case:
  - **Infinite path, start with all Bs**
  - **Payoffs for *w*: A**:a, **B**:1, **AB**:a+1-c
- **For what pairs (c,a) does A spread?**
  - **We need to analyze two cases for node *w*:** Based on the values of a and c, what would w do?



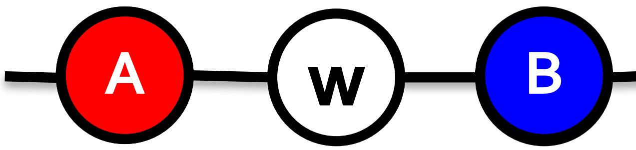
For what pairs (c,a) does A spread?



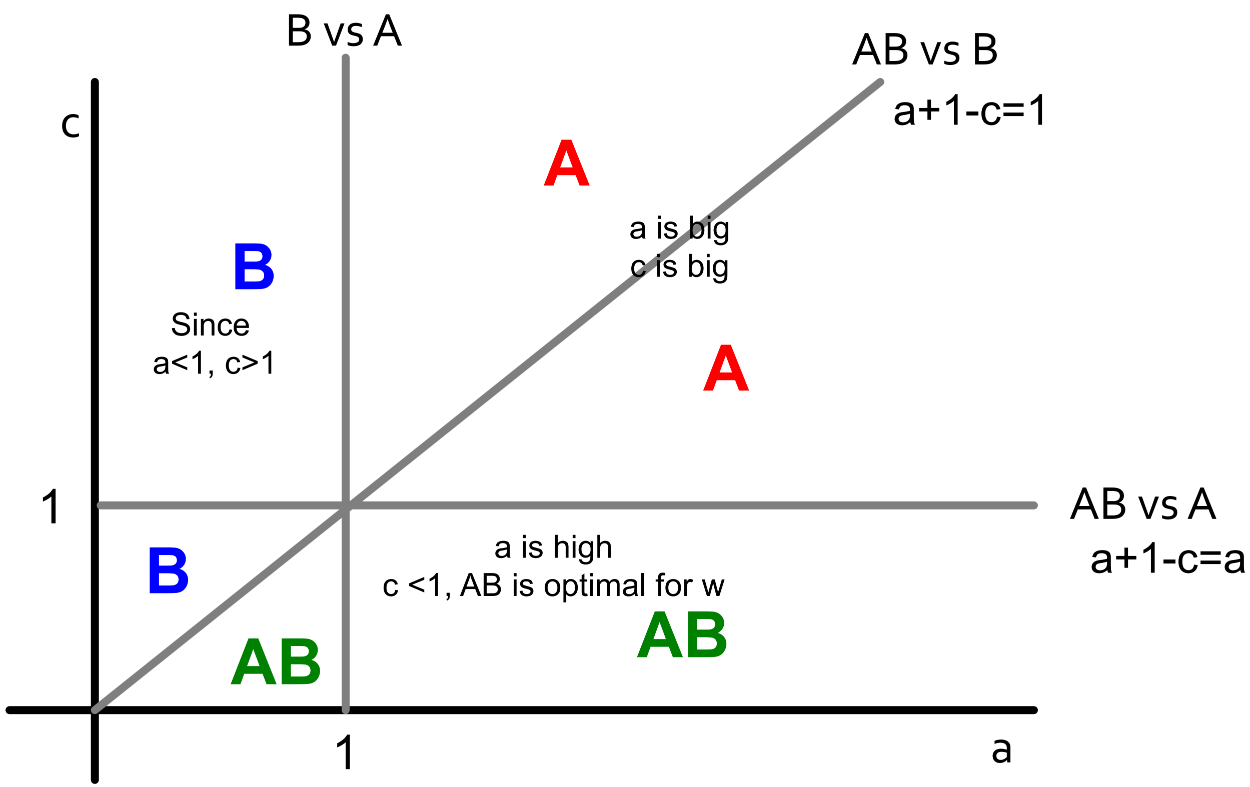
- **Infinite path, start with Bs**
- **Payoffs for *w*:** **A**:a, **B**:1, **AB**:a+1-c
- What does node ***w*** in **A-*w*-B** do?



For what pairs (c,a) does A spread?

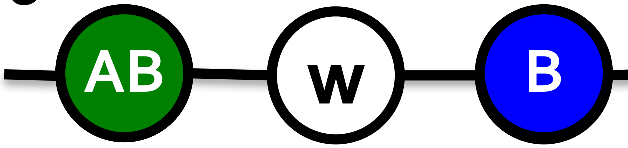


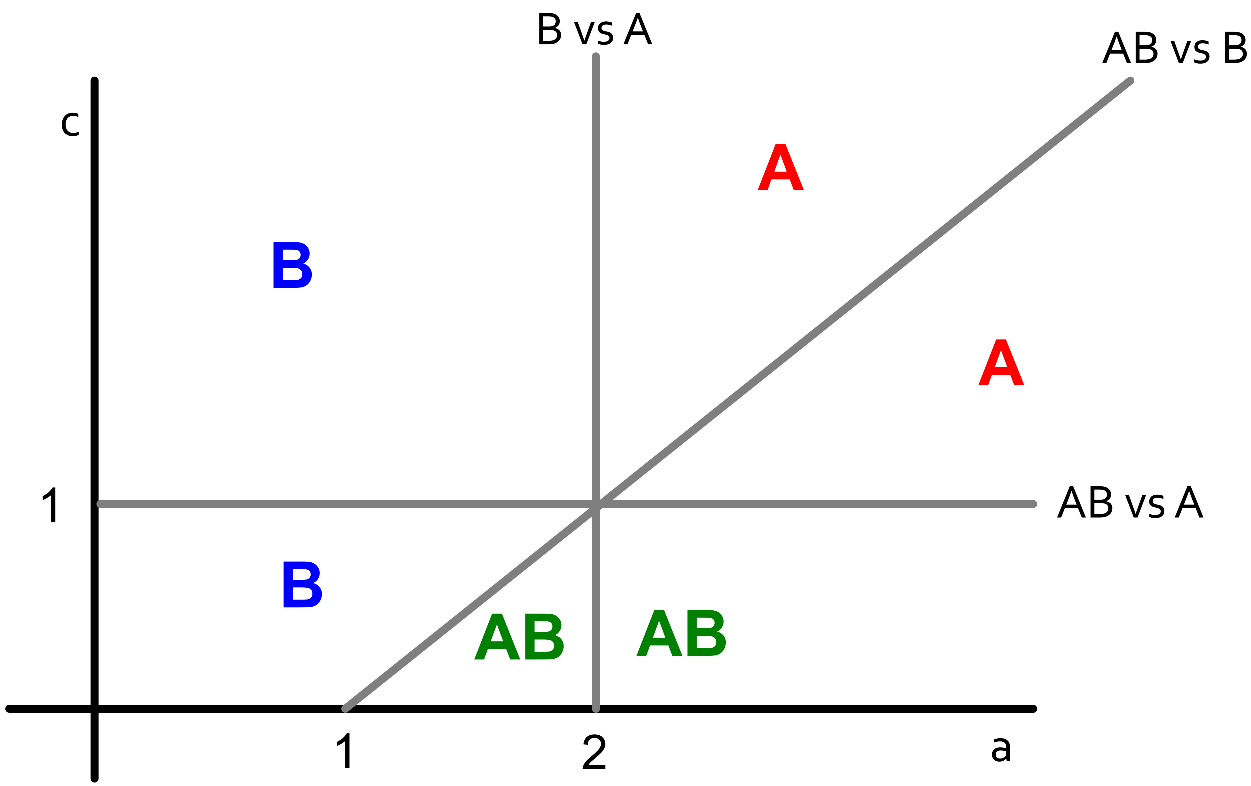
- **Infinite path, start with Bs**
- **Payoffs for *w*:** **A**:a, **B**:1, **AB**:a+1-c
- What does node ***w*** in **A-*w*-B** do?



For what pairs (c,a) does A spread?

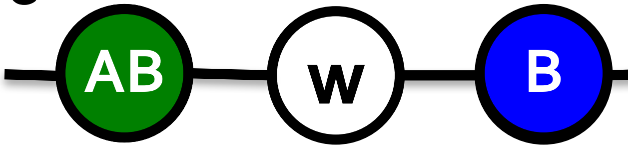
- **Same reward structure as before but now payoffs for *w* change:** **A**:a, **B**:1+1, **AB**:a+1-c
- **Notice: Now also AB spreads**
- What does node *w* in **AB-*w*-B** do?

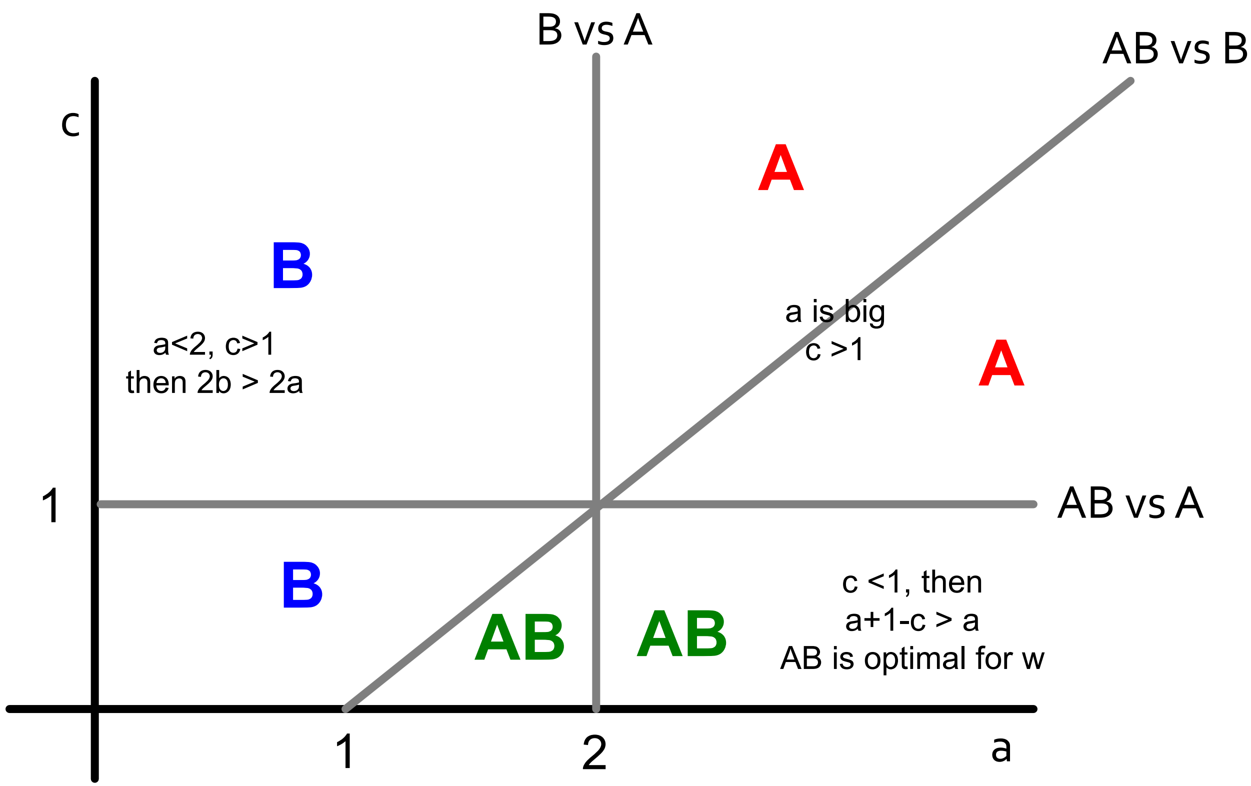




For what pairs (c,a) does A spread?

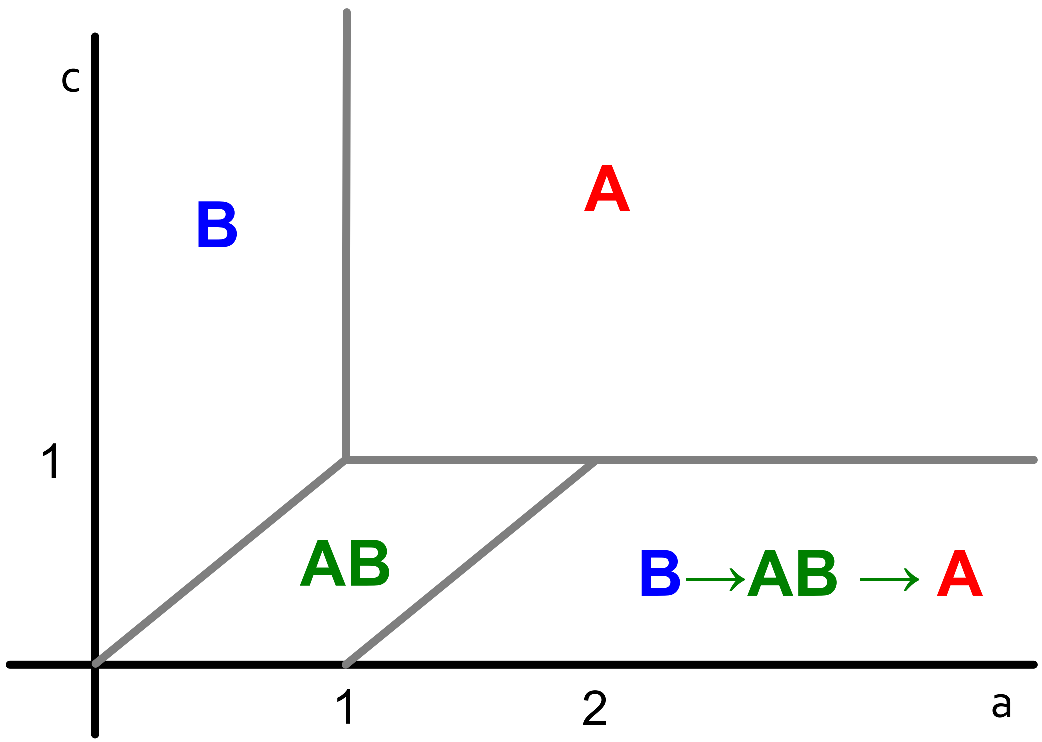
- **Same reward structure as before but now payoffs for *w* change: A**:a, **B**:1+1, **AB**:a+1-c
- **Notice: Now also AB spreads**
- What does node ***w*** in **AB-*w*-B** do?





For what pairs (c,a) does A spread?

- **Joining the two pictures:**



Lesson

- ***B* is the default throughout the network until new/better *A* comes along. What happens?**
  - **Infiltration:** If **B** is too compatible then people will take on both and then drop the worse one **(B)**
  - **Direct conquest:** If **A** makes itself not compatible – people on the border must choose. They pick the better one (**A**)
  - **Buffer zone:** If you choose an optimal level then you keep a static "buffer" between **A** and **B**