Naïve Learning in Social Networks and the Wisdom of Crowds

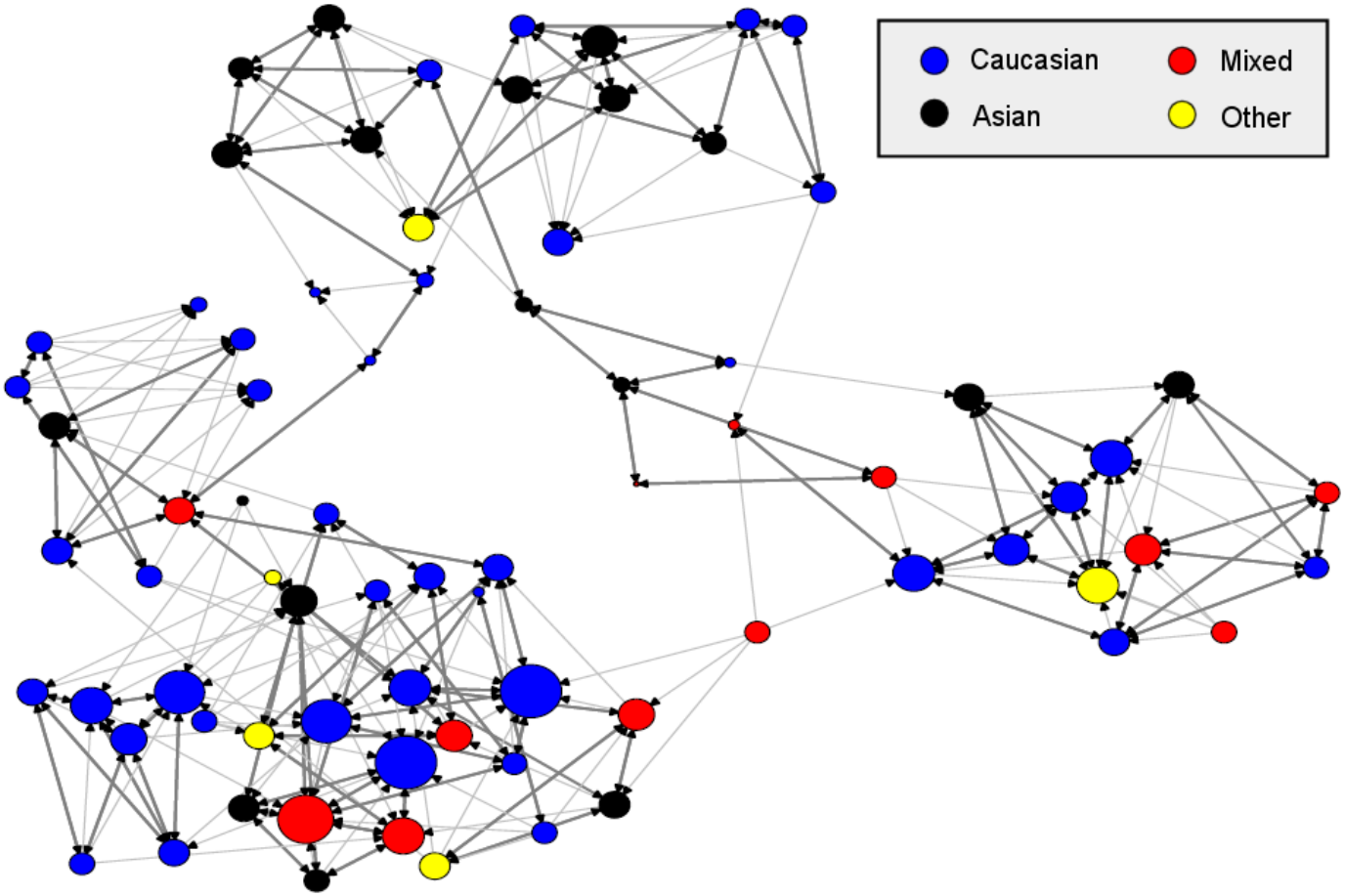
B. Golub and M.O. Jackson (2010)

Matteo Camboni
Leonardo Nini

Spring 2014 IGIER Visiting Students presentations

March 28, 2014

What is a network?



Learning in social networks

- The model investigates the ability of a population to aggregate information in an efficient way through social interaction.

Research question

Which are the social **network structures** that allow a society made up of individuals who communicate and update their beliefs naïvely to aggregate decentralized information **completely and correctly**?

Learning in social networks

- The model investigates the ability of a population to aggregate information in an efficient way through social interaction.

Research question

Which are the social **network structures** that allow a society made up of individuals who communicate and update their beliefs naïvely to aggregate decentralized information **completely and correctly**?

Main findings

1. **Individuals' beliefs converge to a consensus**, provided that the network is strongly connected and satisfies a weak aperiodicity condition.
2. Even with a naïve updating process, growing societies will eventually come close to the truth, as long as no agent holds a nonvanishing influence as the population grows.
   A society which converges to the truth is called **"wise".**

The DeGroot Model (1974)

1. The model, introduced by Morris De Groot in 1974, serves as a starting point for the authors' analysis.
2. There are n agents, indexed by the set $N = \{1, ..., n\}$ who interact in a network $\Gamma(n)$.
3. Each individual has to estimate an uknown parameter $\mu \in \mathbb{R}$.
4. Time is assumed to be discrete, i.e $.t = 0, 1, 2 \ldots$ we can think of periods as days.
5. Each agent receives at t=0 a noisy signal about the true value of the parameter $p_i^{(0)} = \mu + \varepsilon_i$, where ε_i is an randomly distributed error term with expected value of 0.

Behavioral Assumption

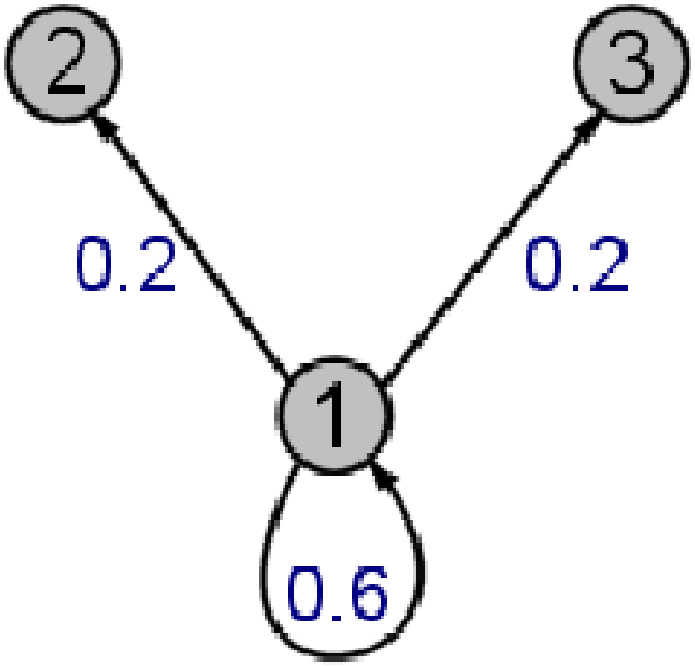
Attention

Agents are characterized by bounded rationality: they continue to use the same updating rule throughout the evolution, failing to adjust correctly for repetition and dependencies in information that they hear multiple times.

Updating

1. Each agent i communicates with other individuals, assigning precision π_{ij} to agent j. Its belief in period $t+1$ is assumed to be a weghted average of the beliefs held in period t by all agents.
2. Define the vector $\mathbf{p^{(t)}} = (p_1^{(t)}, p_2^{(t)}, p_3^{(t)}, ..., p_n^{(t)})$ as the vector including the belief of every individual at time t.
3. Hence, $\mathbf{p}^{(t+1)} = \mathbf{T}\mathbf{p}^{(t)}$ where $\mathbf{T}$ is a n x n stochastic interaction matrix such that
$$T_{ij} = \frac{\pi_{ij}}{\sum_{k=1}^{n} \pi_{ik}}$$
4. Since the updating rule is constant across periods,
$$\mathbf{p}^{(t)} = \mathbf{T}\mathbf{p}^{(t-1)} = \mathbf{T}^t\mathbf{p}^{(0)}$$

Example 1



$$p_1^{(t+1)} = 0.6p_1^{(t)} + 0.2p_2^{(t)} + 0.2p_3^{(t)}$$

Convergence

Definition 1

A matrix $\mathbf{T}$ is convergent if $\lim_{t\to\infty} \mathbf{T}^t\mathbf{p}$ exists for all vectors $\mathbf{p} \in [0,1]^n$

- An obvious question which may arise is: under which conditions does an interaction matrix $\mathbf{T}$ converge?
- In order to answer such question, we need some further definitions.

Convergence

Definition 1

A matrix $\mathbf{T}$ is convergent if $\lim_{t\to\infty} \mathbf{T}^t\mathbf{p}$ exists for all vectors $\mathbf{p} \in [0,1]^n$

- An obvious question which may arise is: under which conditions does an interaction matrix $\mathbf{T}$ converge?
- In order to answer such question, we need some further definitions.

Convergence

Definition 2

A *walk* in $\mathbf{T}$ is a sequence of nodes $i_1, i_2, i_3, ..., i_K$, not necessarily distinct, such that $\mathbf{T}_{\mathbf{i_k},\mathbf{i_{k+1}}} > \mathbf{0} \, \forall \, k \, \in \{1, 2, ..., K-1\}$.
The *lenght* of the walk is defined to be $K-1$.

Definition 3

A *cycle* in $\mathbf{T}$ is a walk $i_1, i_2, i_3, ..., i_K$ such that $i_1 = i_K$.
The *lenght* of a cycle with K (not necessarly distinct) entries is $K-1$.
A cycle is *simple* if the only node appearing twice in the sequence is $i_1 = i_K$.

Convergence

Definition 2

A *walk* in $\mathbf{T}$ is a sequence of nodes $i_1, i_2, i_3, ..., i_K$, not necessarily distinct, such that $\mathbf{T}_{\mathbf{i_k},\mathbf{i_{k+1}}} > \mathbf{0} \, \forall \, k \, \in \{1, 2, ..., K-1\}$.
The *lenght* of the walk is defined to be $K - 1$.

Definition 3

A *cycle* in $\mathbf{T}$ is a walk $i_1, i_2, i_3, ..., i_K$ such that $i_1 = i_K$.
The *lenght* of a cycle with K (not necessarly distinct) entries is $K - 1$.
A cycle is *simple* if the only node appearing twice in the sequence is $i_1 = i_K$.

Convergence

Definition 4

The matrix $\mathbf{T}$ is *strongly connected* if there is a path in $\mathbf{T}$ going from every node to any other node.

Definition 5

The matrix $\mathbf{T}$ is *aperiodic* if the gratest common divisor of the lenght of its simple cycles is 1.

Convergence

Definition 4

The matrix $\mathbf{T}$ is *strongly connected* if there is a path in $\mathbf{T}$ going from every node to any other node.

Definition 5

The matrix $\mathbf{T}$ is *aperiodic* if the gratest common divisor of the lenght of its simple cycles is 1.

Example 2: failure of aperiodicity

$$\mathbf{T} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

In this case,

$$\mathbf{T}^t = \begin{cases} \mathbf{I} & \textit{if } t \textit{ is even} \\ \mathbf{T} & \textit{if } t \textit{ is odd} \end{cases}$$

Result 1: Convergence of Beliefs in the De Groot model

Proposition 1

If $\mathbf{T}$ is a strongly connected matrix, the following propositions are equivalent

(a) $\mathbf{T}$ is convergent;

(b) $\mathbf{T}$ is aperiodic;

(c) $\exists!$ left eigenvector $\mathbf{s}$ of $\mathbf{T}$ corresponding to eigenvalue 1 whose entries sum to 1, s.t. $\forall\, \mathbf{p} \in [0,1]^n$,

$$\lim_{t\to\infty} \mathbf{p_i^{(t)}} = \left(\lim_{t\to\infty} \mathbf{T}^t \mathbf{p^{(0)}}\right)_i = \mathbf{s}\mathbf{p^{(0)}} \quad \forall_i$$

The influence vector

- The i_{th} element of the vector $\mathbf{s} = (s_1, ..., s_n) \in [0,1]^n$ can be interpreted as the influence of agent i.
- Since $\mathbf{s}$ is a left eigenvector of $\mathbf{T}$, the influence of each agent i can be computed as a weighted average of the influences of all agents with whom he is connected, where the weights are given by the importance that agent j assigns to agent i

$$s_i = \sum_{j=1}^{n} s_j T_{ji}$$

Extending the model

- Now, we examine a sequence of networks, in which the number n of agents grows.
- A *society* can be described by a sequence of interaction matrices $(\mathbf{T}(n))_{n=1}^{\infty}$ and of vectors $(\mathbf{p}^{(t)}(n))_{n=1}^{\infty}$ indexed by n, where $p_i^{(t)}(n)$ is the belief of agent i in network $\Gamma(n)$ at time t.

Definition 6: Wisdom

The sequence $(\mathbf{T}(n))_{n=1}^{\infty}$ is *wise* if

$$plim_{n\to\infty} \max_{i\leq n} \left| p_i^{(\infty)}(n) - \mu \right| = 0$$

Result 2: Convergence to the truth

Lemma 1

[Application of the Law of large numbers] if $(\mathbf{s}(n))_{n=1}^{\infty}$ is any sequence of influence vectors, then

$$plim_{n\to\infty} s(n)\mathbf{p}^{(0)}(n) = \mu$$

if and only if $s_1(n) \to 0$, where $s_1(n) = \max_{1\leq i\leq n} s_i(n)$

Proposition 2

If $(\mathbf{T}(n))_{n=1}^{\infty}$ is a sequence of convergent stochastic matrices, then it is wise if and only if the associated influence vectors are such that $s_1(n) \to 0$.

- Naïve agents can be misled by highly influential individuals.
- The limiting belief of all agents will converge to the truth if nobody's idiosyncratic error gets a positive weight as the society grows arbitrarily large.

Case of undirected networks with equal weights

- We assume that the network can be described by a (symmetric) matrix $\mathbf{G}$ which ij_{th} element is equal to 1 if and only if there is an (indirect) link between agents i and j, and equal to 0 otherwise.
- We assume also that each individual values equally the opinion of every agent with who he is in contact.
- Therefore, the stochastic Matrix $\mathbf{T}(\mathbf{G})$ is defined by $T_{ij} = G_{ij}/d_i(\mathbf{G})$, where $d_i(\mathbf{G}) = \sum_{j=1}^{n} G_{ij}$, is the *degree* of agent i, i.e. the number of individuals who have an (indirect) link with agent i.

Case of undirected networks with equal weights

- Define a sequence $(\mathbf{G}(n))_{n=1}^{\infty}$ of symmetric adjacency matrices.
- For each n, the degree of agent i is: $d_i(\mathbf{G}(n)) = \sum_{i=1}^{n} G_{ij}(n)$.

Corollary 1

The sequence $(\mathbf{G}(n))_{n=1}^{\infty}$ is wise if and only if

$$\max_{1 \leq i \leq n} s_i(n) = \max_{1 \leq i \leq n} \frac{d_i(\mathbf{G}(n))}{\sum_{i=1}^{n} d_i(\mathbf{G}(n))} \to 0 \text{ as } n \to \infty$$

Wisdom in terms of social structure

- We move now to a more general setting, in which links are not necessarily reciprocal and individuals may assign different weights to different agents.
- In order to identify some structural conditions that can ensure wisdom, we introduce some concepts.

Definition 7

A *family* is a sequence of groups B_n such that $B_n \subset \{1, 2, ...n\} \ \forall \ n$

Prominent family

$$T_{B,C} = \sum_{i \in B.\ j \in C} T_{ij}$$

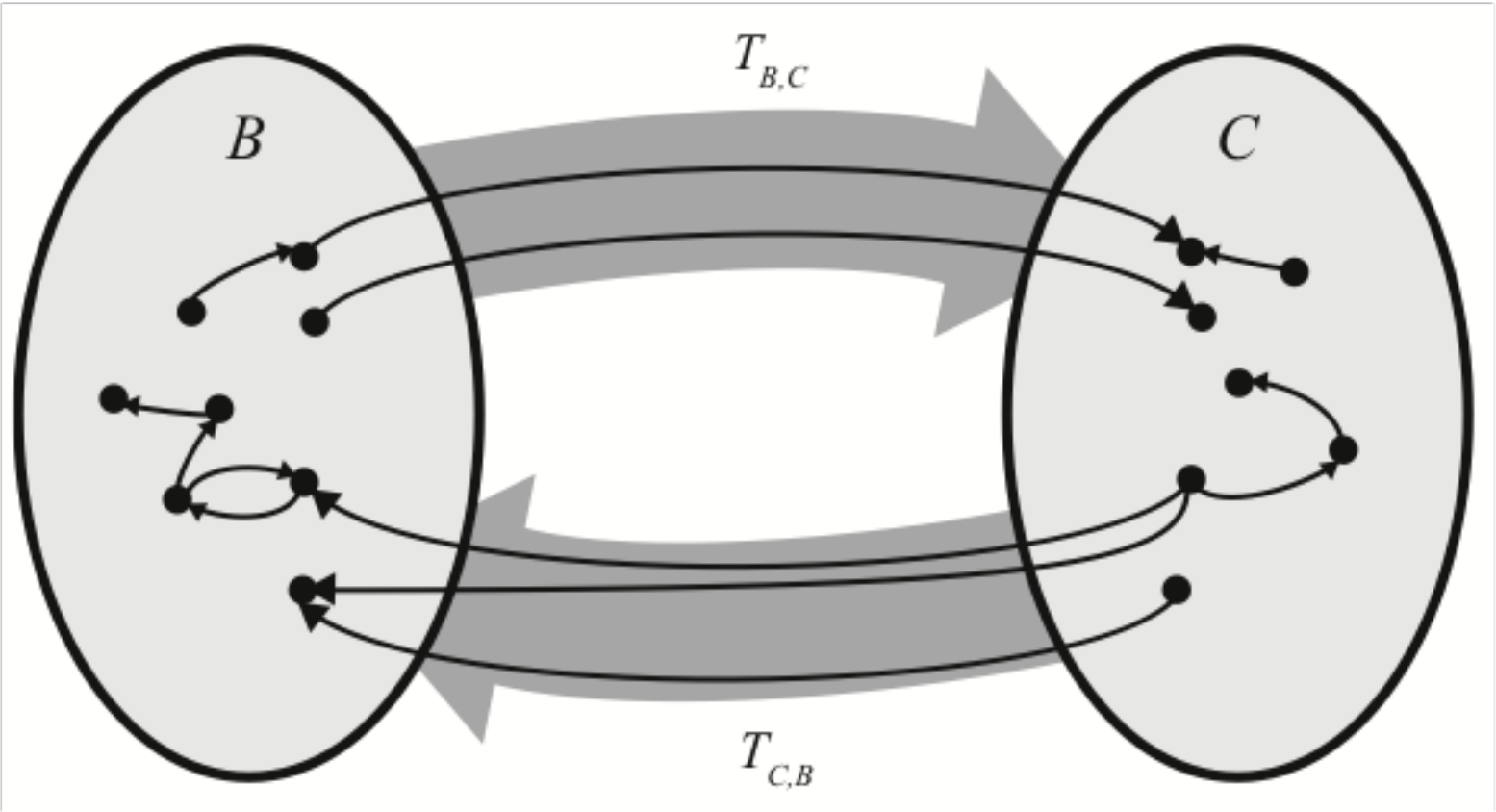


FIGURE 1

Prominent family

Definition 8

A group of agents B is *prominent in t steps* relative to $\mathbf{T}$ if $(\mathbf{T}^t)_{i,B} > 0 \ \ \forall i \notin B$.
Call $\pi_B(\mathbf{T}; t) := \min_{i \notin B} (\mathbf{T}^t)_{i,B}$ *the $t-$step prominence* of B relative to $\mathbf{T}$.

Definition 9

The family B_n is **uniformly prominent** relative to $(\mathbf{T}(n))_{n=1}^{\infty}$ if $\exists\, \alpha > 0 \ s.t. \ \ \forall\ n$ there is a t so that B_n is prominent in t steps relative to $\mathbf{T}(n)$ with $\pi_{B_n}(\mathbf{T}(n); t) \geq \alpha$.

Definition 10

The family B_n is **finite** if $\exists\ q \ s.t. \ \sup_n |B_n| < q$.

Prominent family

Definition 8

A group of agents B is *prominent in t steps* relative to $\mathbf{T}$ if $(\mathbf{T}^t)_{i,B} > 0 \ \ \forall i \notin B$.
Call $\pi_B(\mathbf{T}; t) := \min_{i \notin B} (\mathbf{T}^t)_{i,B}$ *the $t-$step prominence* of B relative to $\mathbf{T}$.

Definition 9

The family B_n is **uniformly prominent** relative to $(\mathbf{T}(n))_{n=1}^{\infty}$ if $\exists\, \alpha > 0 \ s.t. \quad \forall\ n$ there is a t so that B_n is prominent in t steps relative to $\mathbf{T}(n)$ with $\pi_{B_n}(\mathbf{T}(n); t) \geq \alpha$.

Definition 10

The family B_n is **finite** if $\exists\ q \ s.t. \ \sup_n |B_n| < q$.

Prominent family

Definition 8

A group of agents B is *prominent in t steps* relative to $\mathbf{T}$ if $(\mathbf{T}^t)_{i,B} > 0 \ \ \forall i \notin B$.
Call $\pi_B(\mathbf{T}; t) := \min_{i \notin B} (\mathbf{T}^t)_{i,B}$ *the $t-$step prominence* of B relative to $\mathbf{T}$.

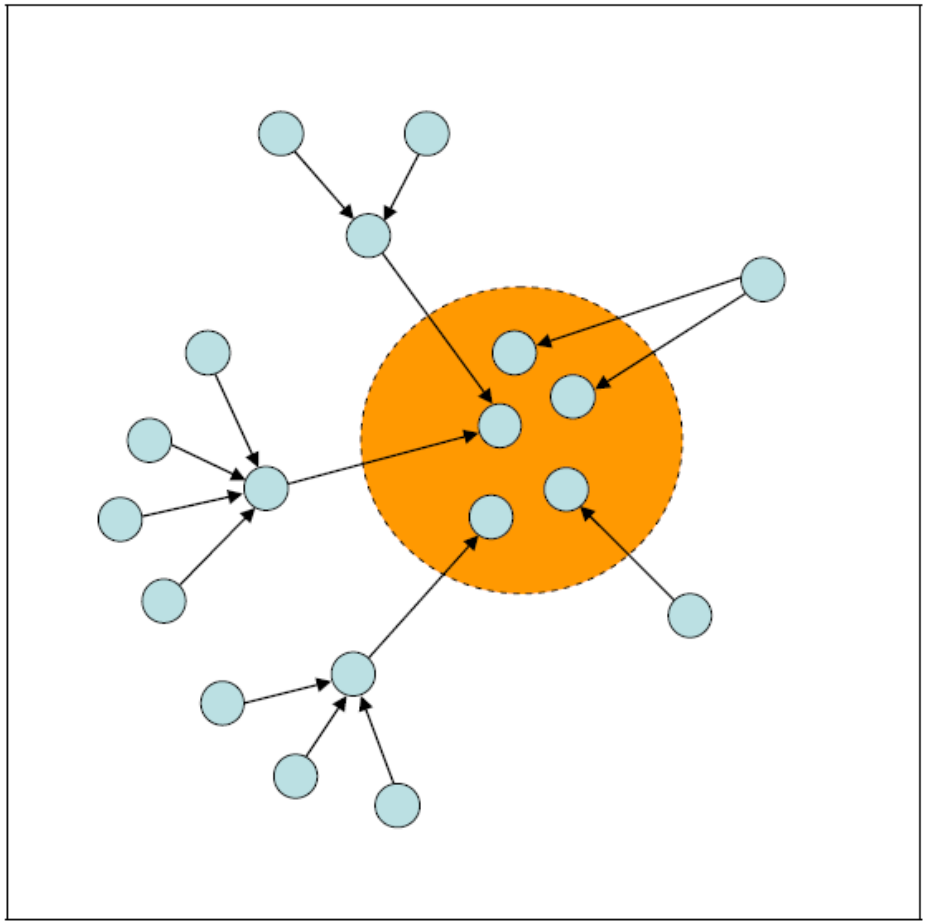
Definition 9

The family B_n is **uniformly prominent** relative to $(\mathbf{T}(n))_{n=1}^{\infty}$ if $\exists\, \alpha > 0 \ s.t. \quad \forall\ n$ there is a t so that B_n is prominent in t steps relative to $\mathbf{T}(n)$ with $\pi_{B_n}(\mathbf{T}(n); t) \geq \alpha$.

Definition 10

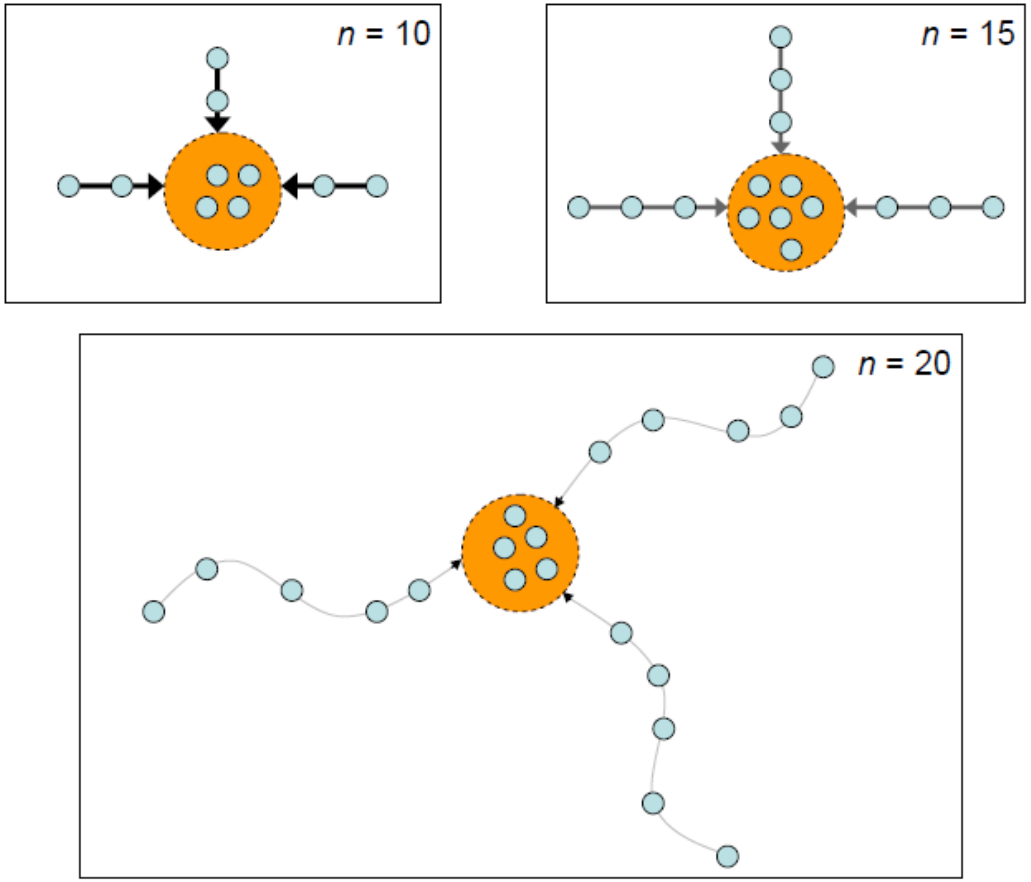
The family B_n is **finite** if $\exists\ q \ s.t. \ \sup_n |B_n| < q$.

Prominent group: intuition



The group inside the dashed circle is prominent in 2 steps.

Prominent family: intuition



Result 3: Prominent family

Proposition 3

If there is a finite, uniformly prominent family B_n wrt $\mathbf{T}(n)$, then the sequence $(\mathbf{T}(n))_{n=1}^{\infty}$ is not wise.

- Prominence is thus an hindrance to wisdom: families which are heavily influential towards the rest of the society impede convergence to the truth.

Example 3

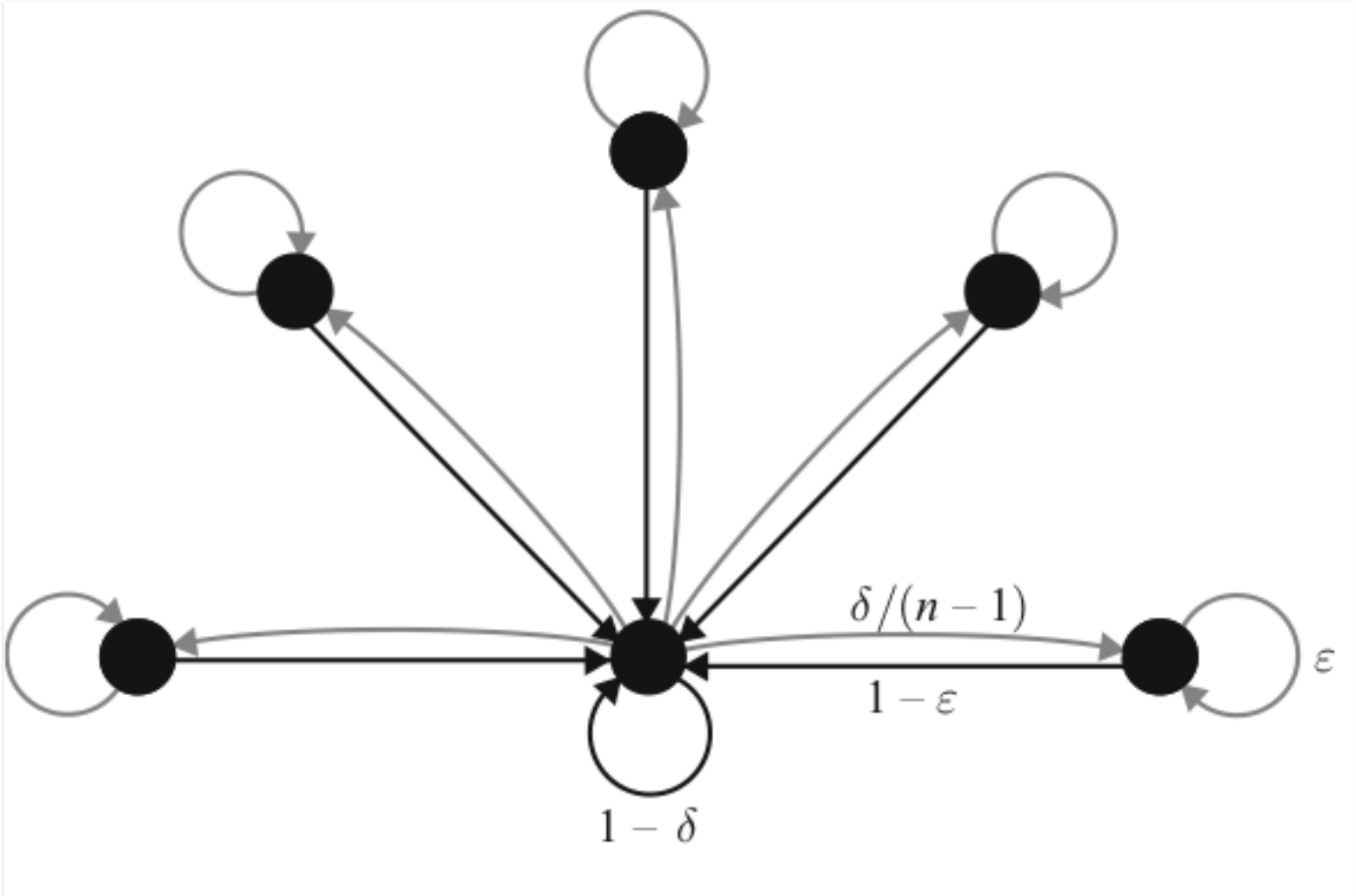


FIGURE 2

Example 3

$$\mathbf{T}(n) := \begin{bmatrix} 1-\delta & \frac{\delta}{n-1} & \frac{\delta}{n-1} & \cdots & \frac{\delta}{n-1} \\ 1-\varepsilon & \varepsilon & 0 & \cdots & 0 \\ 1-\varepsilon & & \varepsilon & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1-\varepsilon & 0 & 0 & \cdots & \varepsilon \end{bmatrix}.$$

$$s_i(n) = \begin{cases} \frac{1-\epsilon}{1-\epsilon+\delta} & \text{if } i = 1 \\ \frac{\delta}{(1-\epsilon+\delta)(n-1)} & \text{if } i > 1 \end{cases}$$

Example 4

- $\lim_{n\to\infty} s_1(n)$ can be made arbitrarly close to 1 by choosing a small $\delta \in (0, 1/2)$.

$$s_i(n) = \left(\frac{\delta}{1-\delta}\right)^{i-1} \frac{1-\left(\frac{\delta}{1-\delta}\right)}{1-\left(\frac{\delta}{1-\delta}\right)^n}$$

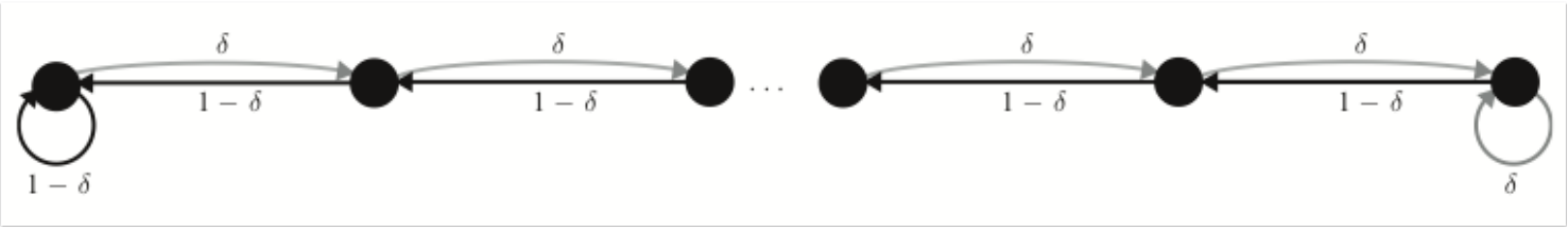


FIGURE 3

Result 4: Structural sufficient conditions for wisdom

Theorem 1

If $(\mathbf{T}(n))_{n=1}^{\infty}$ is a sequence of convergent stochastic matrices satisfying **balance** and **minimal out-dispersion**, then it is wise.

Balance Property

If there exists a sequence $j(n) \to \infty$ s.t. if $|B_n| \leq j(n)$, then

$$\sup_n \frac{T_{B_n^c, B_n}(n)}{T_{B_n B_n^c,}(n)} < \infty$$

Minimal Out- Dispersion Property

There exists a $q \in \mathbb{N}$ and $r > 0$ s.t. if B_n is finite, $|B_n| \geq q$ and $|C_n| / n \to 1$, then $T_{B_n C_n}(n) > r$ for large enough n.

Result 4: Structural sufficient conditions for wisdom

Theorem 1

If $(\mathbf{T}(n))_{n=1}^{\infty}$ is a sequence of convergent stochastic matrices satisfying **balance** and **minimal out-dispersion**, then it is wise.

Balance Property

If there exists a sequence $j(n) \to \infty$ s.t. if $|B_n| \leq j(n)$, then

$$\sup_n \frac{T_{B_n^c, B_n}(n)}{T_{B_n B_n^c,}(n)} < \infty$$

Minimal Out- Dispersion Property

There exists a $q \in \mathbb{N}$ and $r > 0$ s.t. if B_n is finite, $|B_n| \geq q$ and $|C_n| / n \to 1$, then $T_{B_n C_n}(n) > r$ for large enough n.

Result 4: Structural sufficient conditions for wisdom

Theorem 1

If $(\mathbf{T}(n))_{n=1}^{\infty}$ is a sequence of convergent stochastic matrices satisfying **balance** and **minimal out-dispersion**, then it is wise.

Balance Property

If there exists a sequence $j(n) \to \infty$ s.t. if $|B_n| \leq j(n)$, then

$$\sup_n \frac{T_{B_n^c, B_n}(n)}{T_{B_n B_n^c,}(n)} < \infty$$

Minimal Out- Dispersion Property

There exists a $q \in \mathbb{N}$ and $r > 0$ s.t. if B_n is finite, $|B_n| \geq q$ and $|C_n|/n \to 1$, then $T_{B_n C_n}(n) > r$ for large enough n.

The speed of convergence

- Issue: how long does it take $\mathbf{T}^t$ to approach its limit (if it exists)? In general, there is no relationship between speed of convergence and wisdom
- Consider the case in which agents weight each other equally: such society will converge immediately to wisdom.
- If instead all agents assign a weight equal to 1 to the same agent, then we would have immediate convergence but without wisdom.
- Lastly, consider a society in which all individuals place $1 - \varepsilon$ weight on themselves and distribute the rest equally: such society is wise but converge will happen arbitrarily slowly for small enough ε.

Related literature

- De Groot Model (and related literature in physics, computer science)
  - Sociology: centrality and prestige analysis, relationship between social structure and social learning (Katz, 1953, Bonacich, 1987)
- Herding Models (ex. Abhijit and Banerjee, 1992): agents converge to same belief.
- Observational Learning Models
  - Rosenberg et al., (2009): convergence and efficieny learning does not depend on the network architecture.
  - Bala and Goyal (1998): similar questions, but different setting (observational learning) and not precise calculations of the influence of each agent in the network.
  - Acemoglu et al. (2008)
- Network-based explanation of political opinion (DeMarzo et al., 2003): different questions, but similar learning results.

Conclusions

1. Small groups of opinion makers who attract a large share of attention without weighting considerably the rest of society are a great obstacle to correct common knowledge.
2. Such peculiarity may be interesting in formulating marketing or electoral strategies.
3. We can identify networks which are approximately line with the two structural sufficient conditions identified in the model, in which therefore agents will be facilitated in reaching a reasonable consensus.

Discussion

1. Strong behavioral assumptions: real agents are likely to employ more sophisticated belief updating methods.
2. If the society never aknowledges the true value of the parameter, there can be no valuation or revision of an agent's reliability.
3. The model assumes also that individuals have no cost in getting information.
4. Speed of convergence: if the process of beliefs updating converges too slowly, will it reach a steady state in "useful" time? What if μ is not fixed but changes over time?

Although the model does not answer to these issues, it conveys some reasonable results building on a simple framework.