Collective Belief, Naive Learning and Wisdom of the Crowd

Chenwei Shi Tsinghua University, Beijing

This is a joint work with Alexandru Baltag, Fenrong Liu and Sonja Smets

Reason to Believe

理 ◦ 信

Chenwei Shi

- relation between belief and evidence/justification/argument;
- consistent belief and inconsistent evidence;
- more information to believe, harder to maintain the consistency.

©Fan Ho

©Alex Webb

Is it possible to strike a balance between
believing more and believing more consistently?

Individual and Collective Belief

- Topological semantics for argumentation and belief;
- Indeterministic binary DeGroot model for collective belief.

Overview

Different ways of understanding collective belief

Some notions of collective belief studied in episemic logic

- Common belief;
- Everyone believes that ... (consensus);
- Distributed belief

Some rules of aggregating judgements or merging beliefs

- Majority rule;
- Distance-based rule;

Different ways of understanding collective belief

Some notions of collective belief studied in episemic logic

- Common belief;
- Everyone believes that ... (consensus);
- Distributed belief

Some rules of aggregating judgements or merging beliefs

- Majority rule;
- Distance-based rule;

Some concrete examples

- The Chinese investors **believe** that the slowdown of China's economy is temporary.
- The Chinese parents **believe** that baby formula from Europe is better than the local products.

Social interaction and influence on social networks

A group is not simply a set of agents. It is also about how people connect, interact and thus influence each other. Thus the notion of collective belief should also take the social interaction and influence on social networks into consideration.

Potential group belief

A rough definition of potential group belief

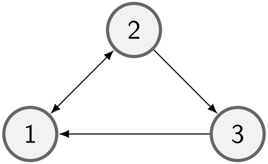
A group tends to believe p if and only if it is more probable that each group member, under other group members' influence, would finally believe p and keep believing p.

Three features of our definition

- binary belief
- high probability of finally reaching and keeping a consensus;
- social influence on social network;

Trust matrix

$$\mathbf{T} = \begin{bmatrix} 0 & 1 & 0 \\ 0.4 & 0 & 0.6 \\ 1 & 0 & 0 \end{bmatrix}$$



Belief change under social influence

Group members' belief about a given proposition P:

$$\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Group members' belief change under social influence:

$$\mathbf{Tb} = \begin{bmatrix} 0 & 1 & 0 \\ 0.4 & 0 & 0.6 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \color{red}{0.4} \\ 1 \end{bmatrix} .$$

Perspective shift

Group members' belief about a given proposition P:

$$\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Group members' tendency to believe P under social influence:

$$\mathbf{Tb} = \begin{bmatrix} 0 & 1 & 0 \\ 0.4 & 0 & 0.6 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.4 \\ 1 \end{bmatrix}.$$

Under the social influence, the second agent tends to believe P with probability 0.4.

Assume the trust matrix $\mathbf{T}$ is given as in the previous slide

$P(B'_1 = 1|B = 110) = 1, P(B'_2 = 1|B = 110) = 0.4, P(B'_3 = 1|B = 110) = 1$

Since the set of variables B'_i are mutually independent:

$$P(B' = 111|B = 100) = \prod_i P(B'_i = 1|B = 110) = 0.4$$

Question

$$P(B' = 101|B = 110) =?$$

Given a trust matrix

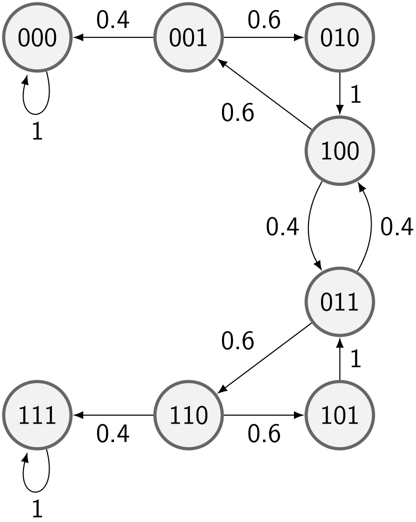
$$\mathbf{T} = \begin{bmatrix} 0 & 1 & 0 \\ 0.4 & 0 & 0.6 \\ 1 & 0 & 0 \end{bmatrix}$$

we can compute its corresponding transition matrix

$$\mathbb{T} = \begin{array}{c} \\ 111 \\ 110 \\ 101 \\ 100 \\ 011 \\ 010 \\ 001 \\ 000 \end{array} \begin{array}{c} \begin{array}{cccccccc} 111 & 110 & 101 & 100 & 011 & 010 & 001 & 000 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.4 & 0 & 0.6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.4 & 0 & 0.6 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Transition diagram for Markov chain

Figure: The transition diagram for $\mathbb{T}$



$$\mathbb{T}^2 = \begin{array}{c} \\ 111 \\ 110 \\ 101 \\ 100 \\ 011 \\ 010 \\ 001 \\ 000 \end{array} \begin{array}{c} \begin{array}{cccccccc} 111 & 110 & 101 & 100 & 011 & 010 & 001 & 000 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.4 & 0 & 0 & 0 & 0.6 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 & 0 & 0 & 0 \\ 0 & 0.24 & 0 & 0.16 & 0 & 0.36 & 0 & 0.24 \\ 0.24 & 0 & 0.36 & 0 & 0.16 & 0 & 0.24 & 0 \\ 0 & 0 & 0 & 0 & 0.4 & 0 & 0.6 & 0 \\ 0 & 0 & 0 & 0.6 & 0 & 0 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$\mathbb{T}^n$

$\mathbb{T}^n_{110,111}$: the probability of arriving in the state 111 after n step given the initial state 110 and trust matrix **T**

Potential group belief

Definition (Group's potential belief)

Given the group G's initial belief state $\mathbf{b}$ and trust matrix $\mathbf{T}$, the group G tends to believe the given proposition if and only if there exists a natural number N such that for any $n \geq N$,

$$\mathbb{T}^n_{\mathbf{b}\mathbf{1}} > 0.5$$

where $\mathbf{1}$ is a constant vector with the entry 1.

- Under what conditions does the sequence converge?
- What is its relationship with $\mathbf{T}, \mathbf{T}^2, \mathbf{T}^3, \ldots$?

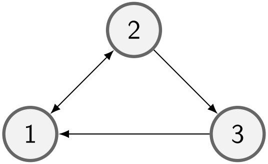
Strongly connectedness

$$\mathbf{T_1} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.2 & 0.8 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{T_2} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.2 & 0.8 & 0 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$$

Strongly connectedness

$$\mathbf{T} = \begin{bmatrix} 0 & 1 & 0 \\ 0.4 & 0 & 0.6 \\ 1 & 0 & 0 \end{bmatrix}$$



Aperiodicity

Definition

Given a trust matrix $\mathbf{T}$, the period of an agent i in the model is the greatest common divisor of the members in the set $\{n \in \mathbb{N} \mid \mathbf{T}^n_{ii} > 0\}$:

$$g(i) = \gcd\{n \in \mathbb{N} \mid \mathbf{T}^n_{ii} > 0\} \ .$$

i is *aperiodic* if $g(i) = 1$ and periodic if $g(i) > 1$. $\mathbf{T}$ is aperiodic if and only if all the agents in it are aperiodic.

Convergence

Theorem

Given a trust matrix $\mathbf{T}$, *if it is strongly connected and aperiodic, then the powers of its corresponding transition matrix* $\mathbb{T}$ *converge to a limiting matrix, that is,* $\mathbb{T}^\infty$ *exists.*

Proof

step 1

Given a trust matrix **T**, if it is strongly connected and aperiodic, then its corresponding transition matrix $\mathbb{T}$ is **a transition for an absorbing Markov chain**.

step 2

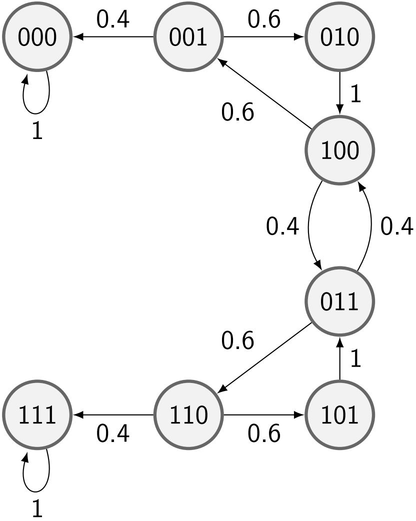
Given a trust matrix **T**, if its corresponding transition matrix $\mathbb{T}$ is **a transition for an absorbing Markov chain**, then the powers of its corresponding transition matrix $\mathbb{T}$ converge to a limiting matrix, that is, $\mathbb{T}^\infty$ exists.

The definition of an absorbing Markov chain

Definition

A Markov chain is absorbing if there exists in it at least one absorbing state, and if for every state the probability of reaching an absorbing state (not necessarily in one step) is strictly positive.

An absorbing Markov chain



Theorem

Given a trust matrix $\mathbf{T}$*, if it is strongly connected and aperiodic, then its corresponding transition matrix* $\mathbb{T}$ *is a transition for an absorbing Markov chain, moreover, there are only two absorbing states* $\mathbf{1}$ *and* $\mathbf{0}$*.*

$$\mathbb{T}^{100} \approx \begin{array}{c} \\ 111 \\ 110 \\ 101 \\ 100 \\ 011 \\ 010 \\ 001 \\ 000 \end{array} \begin{array}{c} \begin{array}{cccccccc} 111 & 110 & 101 & 100 & 011 & 010 & 001 & 000 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.77 & 0 & 0 & 0 & 0 & 0 & 0 & 0.23 \\ 0.62 & 0 & 0 & 0 & 0 & 0 & 0 & 0.38 \\ 0.38 & 0 & 0 & 0 & 0 & 0 & 0 & 0.62 \\ 0.62 & 0 & 0 & 0 & 0 & 0 & 0 & 0.38 \\ 0.38 & 0 & 0 & 0 & 0 & 0 & 0 & 0.62 \\ 0.23 & 0 & 0 & 0 & 0 & 0 & 0 & 0.77 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Proposition

Given a Kripke-DeGroot model, if the trust matrix $\mathbf{T}$ is strongly connected and aperiodic, and the group G's initial belief state is $\mathbf{b}^Q$, then the group G tends to believe Q if and only if

$$\mathbb{T}^{\infty}_{\mathbf{b}^Q\mathbf{1}} > \mathbb{T}^{\infty}_{\mathbf{b}^Q\mathbf{0}}$$

where $\mathbb{T}$ is the transition matrix generated from $\mathbf{T}$.

A summary of the results on the convergence

Theorem

Given an IBDM, if the trus matrix $\mathbf{T}$ *is strongly connected, then the following four statements are equivalent:*

1. *the* **trust** *matrix* $\mathbf{T}$ *is* **aperiodic***;*
2. *the* **transition** *matrix generated from* $\mathbf{T}$ *is a transition matrix for an* **absorbing** *Markov chain;*
3. *the powers of the* **transition** *matrix generated from* $\mathbf{T}$ *converge;*
4. *the powers of the* **trust** *matrix* $\mathbf{T}$ *converge.*

Kripke-DeGroot frame

Let G be a group of agents and $\mathbf{T}$ be its trust matrix.

$$(W, \{R_i\}_{i \in G}, \mathbf{T})$$

where R_i is world-invariant and serial.

$$\mathbf{b}^Q, \text{ where } Q \subseteq W$$

$$\mathbf{b}_i^Q = 1 \text{ if } R_i(w) \subseteq Q, \quad \text{otherwise, } \mathbf{b}_i^Q = 0$$

Definition

Given the group G's initial belief state $\mathbf{b}^Q$ about Q and trust matrix $\mathbf{T}$, the group G tends to believe Q if and only if there exists a natural number N such that for any $n \geq N$,

$$\mathbb{T}^n_{\mathbf{b}^Q \mathbf{1}} > 0.5$$

Potential group belief is closed upwards and pairwise consistent

Given a Kripke-DeGroot model, assuming that the trust matrix is strongly connected and aperiodic,

- if the group tends to believe Q and $Q \subseteq S$, then the group tends to believe S;
  Reason: if $Q \subseteq S$, then $\mathbf{b}^Q \leq \mathbf{b}^S$.
- if the group tends to believe Q and $S \subseteq W - Q$, then the group does not tend to believe S;
  Reason: $\mathbf{b}^{W-Q} \leq \mathbf{1} - \mathbf{b}^Q$

Why closure under conjunction fails

Example

Given a Kripke-DeGroot model $(\{pq, p\overline{q}, \overline{p}q, \overline{pq}\}, \mathbf{T}, f, V)$, where

- $\mathbf{T} = \begin{bmatrix} 0.3 & 0.3 & 0.4 \\ 0.3 & 0.3 & 0.4 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$
- $f(1) = \{pq\}, f(2) = \{pq, p\overline{q}\}, f(3) = \{pq, \overline{p}q\}$
- $V(p) = \{pq, p\overline{q}\}, V(q) = \{pq, \overline{p}q\}$

Because

$$\mathbf{T}^{\infty} = \mathbf{T}$$

and

$$\mathbf{b}^{[\![p]\!]} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{b}^{[\![q]\!]} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{b}^{[\![p \wedge q]\!]} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

According to the definition of the regular community's potential belief, $B_G p$ and $B_G q$ hold in the model while $B_G(p \wedge q)$ does not.

An equivalent way of defining potential belief for a community

We call a group with a strongly connected and aperiodic trust matrix "regular community". To compute $\mathbb{T}^\infty$ for a regular community, we only need to compute $\mathbf{T}^\infty$.

Theorem

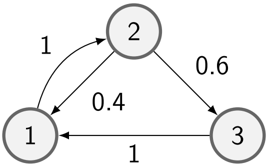
Given a Kripke-DeGroot model, if its trust matrix $\mathbf{T}$ is strongly connected and aperiodic, and the group G's initial belief state is $\mathbf{b}^Q$, then the group G tends to believe Q if and only if

$$\mathbf{T}^\infty_{i*}\mathbf{b}^Q > 0.5$$

where $\mathbf{T}^\infty_{i*}$ is the ith row vector of $\mathbf{T}^\infty$ and $\mathbf{T}^\infty_{i*} = \mathbf{T}^\infty_{j*}$ for any $i, j \in G$.

What does $\mathbf{T}^\infty$ mean?

$$\mathbf{T} = \begin{bmatrix} 0 & 1 & 0 \\ 0.4 & 0 & 0.6 \\ 1 & 0 & 0 \end{bmatrix}$$



$$\mathbf{T}^2 = \begin{bmatrix} 0.4 & 0 & 0.6 \\ 0.6 & 0.4 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{T}^{100} \approx \begin{bmatrix} 0.3846 & 0.3846 & 0.2308 \\ 0.3846 & 0.3846 & 0.2308 \\ 0.3846 & 0.3846 & 0.2308 \end{bmatrix}$$

$\mathbf{T}^{\infty}_{*j}$: group member j's influence on the whole group.

Language

Let G be a finite set of agents, At be a set of atomic propositions. The language $\mathcal{L}_{B\succcurlyeq}$ is given by the following grammar:

$$\varphi ::= p \mid C \succcurlyeq D \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_i\varphi$$

where $p \in \mathrm{At}$ and $C, D \subseteq G$.

Semantics

Given a Kripke-DeGroot model $\mathcal{M} = (\mathcal{KD}, V)$ where $\mathcal{KD}$ is a Kripke-DeGroot frame and $V : At \to 2^W$ is a valuation function, and a possible world w in W,

$\mathcal{M}, w \models p$	iff	$w \in V(p)$
$\mathcal{M}, w \models C \succcurlyeq D$	iff	$\sum_{i \in C} \mathbf{u}_i \geq \sum_{i \in D} \mathbf{u}_i$
$\mathcal{M}, w \models \neg\varphi$	iff	$\mathcal{M}, w \not\models \varphi$
$\mathcal{M}, w \models \varphi \wedge \psi$	iff	$\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
$\mathcal{M}, w \models B_i\varphi$	iff	$R_i(w) \subseteq [\![\varphi]\!]$

Community's Potential Belief

$$B_G\varphi := \bigvee_{C\subseteq G}((C \succ \overline{C}) \wedge \bigwedge_{i\in C} B_i\varphi)$$

where $\overline{C} = G \setminus C$ and $C \succ \overline{C} := (C \succcurlyeq \overline{C}) \wedge \neg(\overline{C} \succcurlyeq C)$.

Community's Stable Belief

$$S_G := \bigvee_{C \subseteq G, C \neq \emptyset} \bigwedge_{i \in C} ((\{i\} \succ \overline{C}) \wedge B_i \varphi)$$

Axiom System

Propositional tautologies and Modus Ponens
K: $B_i\varphi \to (B_i(\varphi \to \psi) \to B_i\psi)$
Necessitation rule: if φ is derivable, then $B_i\varphi$ is derivable
$\succcurlyeq$**4**: $C \succcurlyeq D \to B_i(C \succcurlyeq D)$ $\succcurlyeq$**5**: $\neg(C \succcurlyeq D) \to B_i\neg(C \succcurlyeq D)$
SP: $C \succ \emptyset$ for $C \neq \emptyset$
CO: $(C \succcurlyeq D) \vee (D \succcurlyeq C)$
Scott: if $|\{k \mid i \in C_k \text{ and } 0 \leq k \leq n\}| = |\{l \mid i \in D_l \text{ and } 0 \leq l \leq n\}|$ for all $x \in G$, then $\bigwedge_{i=0}^{n-1}(C_i \succcurlyeq D_i) \to (D_n \succcurlyeq C_n)$ is derivable.

Agents as Random Variables/Test Function

- For each proposition H, an (individual or collective) agent can be seen as a Bernoulli random variable X^H;

example

Consider three agents. They try to test the hypothesis "the coin is not biased" (H^c) by observing an experiment of flipping the coin ten times. The first agent would reject the hypothesis ($X_1 = 1$) if either tail or head appears more than **6** times; the second agent would reject the hypothesis ($X_2 = 1$) if either tail or head appears more than **7** times; the third agent ($X_3 = 1$) would reject the hypothesis if either tail or head appears more than **8** times.
Moreover, assume that the influence vector for these three agents is $(0.3, 0.3, 0.4)^\top$. Then we can also take the community as a random variable $G_3 = 0.3X_1 + 0.3X_2 + 0.4X_3$. The community would reject the hypothesis ($G_3 = 1$) if $0.3X_1 + 0.3X_2 + 0.4X_3 > 0.5$.

Truth Tracking Test Functions

A test function X is **truth tracking** if and only if

$$\inf_{\theta \in H} P_\theta(X = 1) > \beta_H \geq 0.5 \quad .$$

Note that X can be either an individual agent's test function or a collective agent's test method.

Question 1

What can ensure a community's test method to be truth tracking?

Wisdom of the Crowd

- What is a crowd? A growing community. A sequence of growing communities $\{G_n\}$.
- What is wisdom of the crowd? As $n \to \infty$, G_n tends to be truth tracking.

Question 2

What can ensure the crowd to be wise?

A rough answer

- the influence vector w
- the expectation of the distribution of the test methods needs to be larger than 0.5 for all $\theta \in H$.