Collective Learning in Games through Social Networks

Nina Gierasimczuk
joint work with Sanne Kosterman

Institute for Logic, Language and Computation
University of Amsterdam

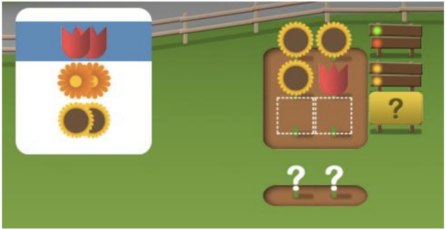
LABEX CIMI Pluridisciplinary Workshop on Game Theory
Toulouse, November 20th, 2015

Recent projects on multi-agent learning

- Learning semantics via coordination on model-checking games.
- Learning roles through voting behavior and mass functions.
- **Collective learning in games through social networks**.

Recent projects on multi-agent learning

- Learning semantics via coordination on model-checking games.
- Learning roles through voting behavior and mass functions.
- **Collective learning in games through social networks**.



Serious Games: Examples of Teams Learning in Games

Military Training: US Army

Screen from Serious Games by Harun Farocki (2011)

Practical Motivation: Serious Games and Social Networks

Merging techniques of serious games and social networks.

Constructing an **active** (games) and **social** (networks) learning environments.

Decisions one needs to make

1. **Game Structure**: players, actions, individual payoffs.
2. **Learning Goal**: what the players should eventually learn.
3. **Social Network Learning (SNL) Update**: how agents update their private information after communicating in the network.
4. **Gameplay**: defines how agents determine what strategy they will play in the game after communicating with neighbors in the network.
5. **Game Learning (GL) Update**: how agents use information about strategies and corresponding rewards to reinforce their strategy and learn from previous actions in the game.

The Game-Network Learning Model: Main Idea

Cooperative games:

- Grand coalition
- Group-rational players

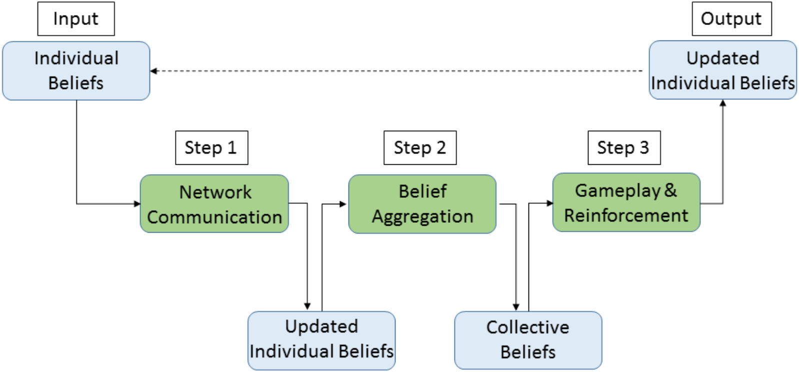
Collective learning:

- Goal: learn towards **social optimum**
- Update beliefs about **joint** strategies

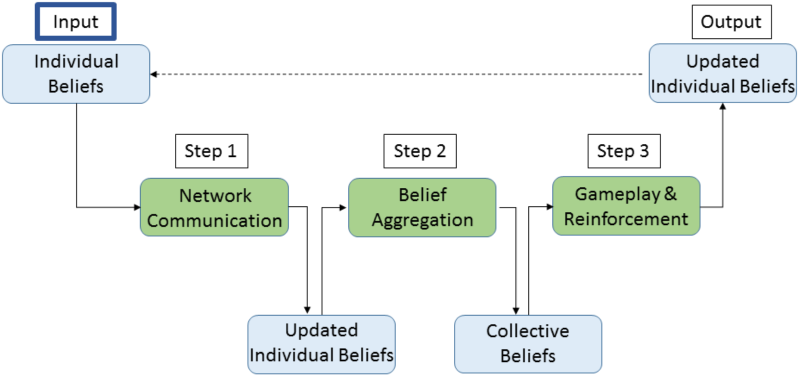
Two belief updates:

- After network communication (feedback from neighbors)
- After gameplay (feedback from payoffs)

The Game-Network Learning Model: Main Idea



The Game-Network Learning Model: Main Idea



Input: Individual Beliefs

Stochastic belief matrix:

$$B = \begin{pmatrix} b_{11} & \dots & b_{1k} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nk} \end{pmatrix}$$

Input: Individual Beliefs

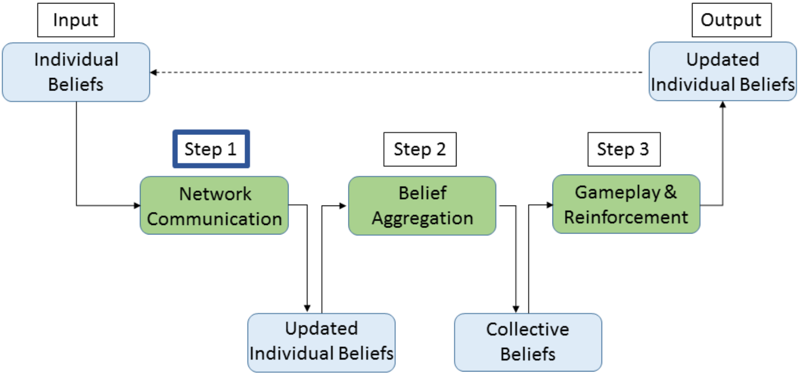
Stochastic belief matrix:

$$B = \begin{pmatrix} b_{11} & \dots & b_{1k} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nk} \end{pmatrix}$$

For example:

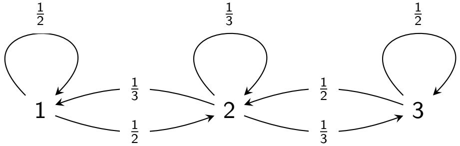
$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Rows are agents, columns are joint strategies.
The entries are strengths of belief in a strategy being the optimal one.


Input
Individual Beliefs
Output
Updated Individual Beliefs
Step 1
Network Communication
Step 2
Belief Aggregation
Step 3
Gameplay & Reinforcement
Updated Individual Beliefs
Collective Beliefs

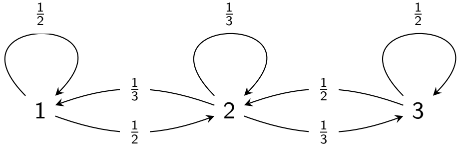
Step 1: Network Communication

1. Weights of trust:



Step 1: Network Communication

1. Weights of trust:

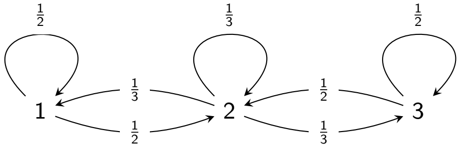


2. Belief update:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \rightsquigarrow \quad \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

Step 1: Network Communication

1. Weights of trust:

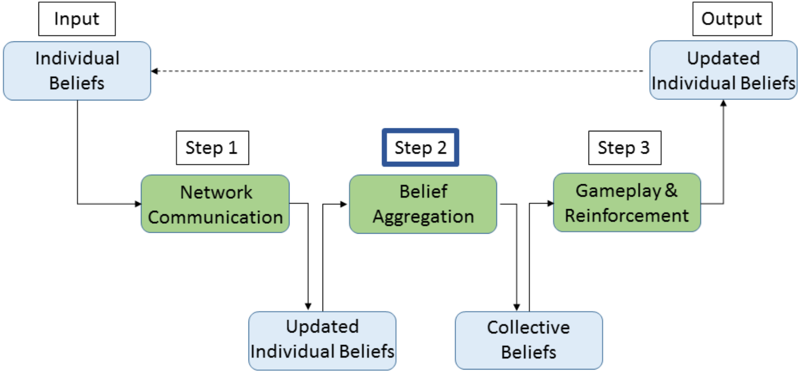


2. Belief update:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \rightsquigarrow \quad \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

Player's perspective: **weighted average** of beliefs of neighbors.

The Game-Network Learning Model



Step 2: Belief Aggregation

- Purpose: deciding collectively which joint strategy to play
- Method: **probabilistic social choice function** (PSCF)

$$B = \begin{pmatrix} b_{11} & \dots & b_{1k} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nk} \end{pmatrix} \quad \mapsto \quad \vec{b} = \begin{pmatrix} b_1 & \dots & b_k \end{pmatrix}$$

Step 2: Belief Aggregation

- Purpose: deciding collectively which joint strategy to play
- Method: **probabilistic social choice function** (PSCF)

$$B = \begin{pmatrix} b_{11} & \dots & b_{1k} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nk} \end{pmatrix} \quad \mapsto \quad \vec{b} = \begin{pmatrix} b_1 & \dots & b_k \end{pmatrix}$$

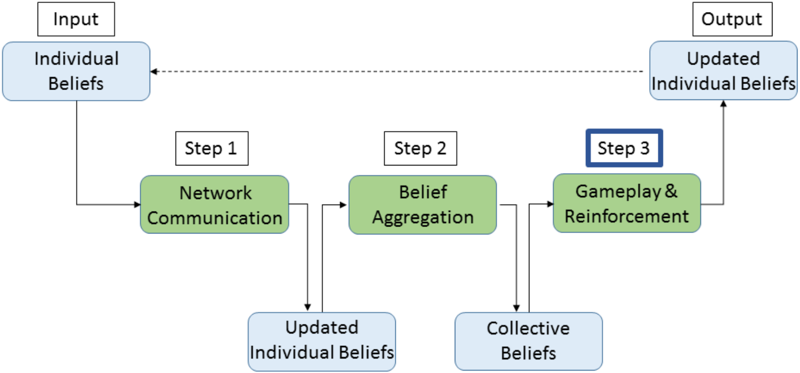
- **Averaged** probabilistic social choice function (aPSCF): takes (weighted) average of beliefs and outputs it a societal probability distribution rather than an ordering.

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 5/18 & 8/18 & 5/18 \end{pmatrix}$$

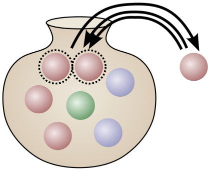
Axiomatic Properties for aPSCF and wPSCF

Independence of Alternatives	✓	✓
Unanimity	✓	✓
Neutrality	✓	✓
Anonymity	✓	✓
Pareto Optimality	✓	✓
Social Rationality	✓	✓
Non-Dictatorship	✓	-
Consistency	✓	-
Strategy-Proofness	-	-

The Game-Network Learning Model



Step 3: Gameplay and Reinforcement



Reinforcement learning:

- Stochastic games
- Law of Effect
- Learn towards high reward

Collective reinforcement learning:

- Reinforce *joint* strategies
- With *social welfare* fraction

In NBA stats earn you money.
Why would you fire the best players?

'The secret of basketball is that it's not about basketball.'

'And that's what Isiah learned while following those Lakers and Celtics teams around: it wasn't about basketball. Those teams were loaded with talented players, yes, but that's not the only reason they won. They won because they liked each other, knew their roles, ignored statistics, and valued winning over everything else. They won because their best players sacrificed to make everyone else happy. They won as long as everyone remained on the same page.'

Reinforcement in the Game-Network model

Remember belief aggregation:

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 5/18 & 8/18 & 5/18 \end{pmatrix}$$

Reinforcement in the Game-Network model

Remember belief aggregation:

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 5/18 & 8/18 & 5/18 \end{pmatrix}$$

Now suppose the players:

- Play joint strategy $s(2)$ with probability $8/18$
- Receive average social welfare $U(s) = 1/6$

Reinforcement in the Game-Network model

Belief update about joint strategies:

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \rightsquigarrow \quad \begin{pmatrix} 5/12 & 7/12 & 0 \\ 5/18 & 8/18 & 5/18 \\ 0 & 7/12 & 5/12 \end{pmatrix}$$

Reinforcement in the Game-Network model

Belief update about joint strategies:

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \rightsquigarrow \quad \begin{pmatrix} 5/12 & 7/12 & 0 \\ 5/18 & 8/18 & 5/18 \\ 0 & 7/12 & 5/12 \end{pmatrix}$$

By **Bush-Mosteller** reinforcement:

$$\begin{aligned} s(1): &\quad b_{21} = 1/3 - 1/6 \cdot 1/3 = 5/18 \\ s(2): &\quad b_{22} = 1/3 + 1/6 \cdot 2/3 = 8/18 \\ s(3): &\quad b_{23} = 1/3 - 1/6 \cdot 1/3 = 5/18 \end{aligned}$$

More on reinforcement

After aggregation by aPSCF, the coalition holds a societal probability distribution over the set of joint strategies. A joint strategy chosen to be played with certain probability. Players get to know the corresponding social welfare and calculate an average fraction. The fraction is used by each player for reinforcement.

More on reinforcement

After aggregation by aPSCF, the coalition holds a societal probability distribution over the set of joint strategies. A joint strategy chosen to be played with certain probability. Players get to know the corresponding social welfare and calculate an average fraction. The fraction is used by each player for reinforcement.

- Probability of played strategy is increased by a payoff-dependent fraction of the distance between the original probability the max. probability 1.
- The probability of other strategies are decreased proportionally.

More on reinforcement

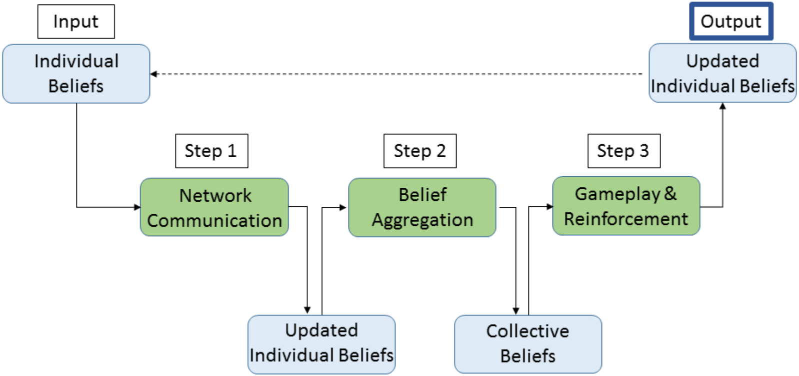
After aggregation by aPSCF, the coalition holds a societal probability distribution over the set of joint strategies. A joint strategy chosen to be played with certain probability. Players get to know the corresponding social welfare and calculate an average fraction. The fraction is used by each player for reinforcement.

- Probability of played strategy is increased by a payoff-dependent fraction of the distance between the original probability the max. probability 1.
- The probability of other strategies are decreased proportionally.

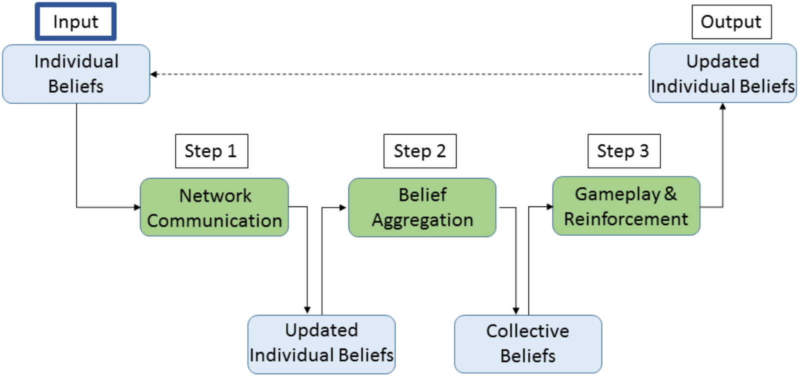
Two advantages of Bush-Mosteller reinforcement:

- makes use of utility values that are scaled in the interval from 0 to 1.
- violates *Law of Practice*: learning does not slow down.

The Game-Network Learning Model



The Game-Network Learning Model



Learning Effect

Can network communication in a game have a **positive influence** on the learning outcome?

Can network communication **speed up learning** towards social optimum?

Can network communication **increase the probability** for playing the social optimum?

Learning Effect at a Given Round t

Dependent on the *network expertise* and *network structure*

Learning Effect at a Given Round t

Dependent on the **network expertise** and *network structure*

Definition (Expert for round t)

We say an agent $i_e \in N$ is an **expert for round** t if his belief for the social optimum is higher than the average belief of society for the social optimum.

Learning Effect at a Given Round t

Dependent on the **network expertise** and *network structure*

Definition (Expert for round t)

We say an agent $i_e \in N$ is an **expert for round** t if his belief for the social optimum is higher than the average belief of society for the social optimum.

Definition (Maximal Expert for round t)

We say an agent $i_m \in N$ is a **maximal expert for round** t if his belief for the social optimum is maximal.

Learning Effect at a Given Round t

Dependent on the *network expertise* and **network structure**

Definition (Weight Centrality)

Let $w_i = \sum_{m \in N} w_{mi}$ be the total weight that agent i receives from his neighbours. The **weight centrality** of some agent $i \in N$ is then given by the fraction $C_i^w = \frac{w_i}{n}$.

Learning Effect at a Given Round t

Dependent on the *network expertise* and *network structure*

Definition (Weight Centrality)

Let $w_i = \sum_{m \in N} w_{mi}$ be the total weight that agent i receives from his neighbours. The **weight centrality** of some agent $i \in N$ is then given by the fraction $C_i^w = \frac{w_i}{n}$.

Theorem

If $C_{i_m}^w > \frac{1}{n} \geq C_i^w$ for all maximal experts $i_m \in \mathcal{E}_{\max}^t$ and other players $i \in N \backslash \mathcal{E}_{\max}^t$, then the probability for playing the social optimum at round t after network communication is higher than before network communication.

Learning Effect in the Long Run

Definition (Stable Expert)

Let $\mathcal{E}^t$ be the set of experts for round t. We say an agent $i \in N$ is a **stable expert** if i is an expert for any round.

Learning Effect in the Long Run

Definition (Stable Expert)

Let $\mathcal{E}^t$ be the set of experts for round t. We say an agent $i \in N$ is a **stable expert** if i is an expert for any round.

Theorem

Let $\mathcal{E}^1$ be the set of initial experts for round $t = 1$. If

(i) $\mathcal{E}^1$ *is* **maximally closed**; *and*

(ii) $\mathcal{E}^1$ *is in* **agreement** *at round $t = 1$,*

then $\mathcal{E}^1 = \mathcal{E}^t_{\text{max}} \subseteq \mathcal{E}^t$ for all $t \geq 1$, so that stable experts exist.

Learning Effect in the Long Run

Definition (Stable Expert)

Let $\mathcal{E}^t$ be the set of experts for round t. We say an agent $i \in N$ is a **stable expert** if i is an expert for any round.

Theorem

Let $\mathcal{E}^1$ be the set of initial experts. If

(i) $\mathcal{E}^1$ *is* **maximally closed**;

(ii) $\mathcal{E}^1$ *is in* **agreement** *at round* $t = 1$; *and*

(iii) $C^w_{i_e} > \frac{1}{n} \geq C^w_i$ *for all* $i_e \in \mathcal{E}^1$ *and* $i \in N \backslash \mathcal{E}^1$,

then the probability of playing the social optimum after communication is higher than before network communication at every round $t \geq 1$.

Conclusions

Theoretical conclusions:

- Interdisciplinary computational model
- For learning in *cooperative games* through *social networks*
- Communication positively influences learning:
  (i) Experts with high weight centrality: *given round t*
  (ii) Stable experts with high weight centrality: *every round*
- Applications in Airline Safety Heros in CAG

Perspectives

Directions for future work:

- Simulations in progress
- Competitive games (between coalitions?)
- Dynamic networks and changing trusts
- Epistemology for modeling knowledge
- Psychological experiments: serious games & social networks

Thank you!

Appendices

Recommendations towards Serious Games

1. Include independent experts with central position
2. Include game elements enhancing experts' reliability
3. Include accessible non-human resources in network
4. Include (group) rewards for cooperation
5. Include private chats rather than blogs for non-experts

Case Study: Airline Safety Heroes

Game players: employees of an airline company

Game objective: learn how to prevent and solve unsafe situations

Game type: digital cards game

(I) Three types of cards:
- (1).1 unsafe situation
- (2).2 prevention tool
- (3).3 solution

(II) Forming a set:
- (1).1 exchange with stock
- (2).2 exchange with players

Case Study: Airline Safety Heroes

Expand the game with social network communication:

1. Card exchange via online chat functions
2. Communication about correctness of sets
3. Managers / employees of safety department as experts
4. Provide experts with central position
5. Group rewards for helping co-players

ALGORITHMS (I)

Algorithm 1 Network Communication at round t

Input: Weight matrix W; probability matrix B^t

1: for all $i \in N$, $s(j) \in S$: $b_{ij}^{t^+} := \sum_{m \in N} w_{im} b_{mj}^t$

2: $B^{t^+} := \left(b_{ij}^{t^+}\right)_{n \times k} = W \cdot B^t$

Output: Probability matrix B^{t^+}

Algorithm 2 Belief Aggregation at round t

Input: Probability matrix B^{t^+}

1: for all $s(j) \in S$: $b_j^{t^+} := \frac{1}{n} \sum_{i \in N} b_{ij}^{t^+}$

2: $\vec{b}^{t^+} := (b_1^{t^+}, \ldots, b_k^{t^+})$

Output: Probability vector $\vec{b}^{t^+}$

Algorithms (II)

Algorithm 3 Gameplay and Reinforcement at round t

Input: Probability vector $\vec{b}^{t^+}$; probability matrix B^{t^+}

$s^t := s(j)$, s.t. $s(j)$ is drawn from S with probability $b_j^{t^+}$ ▷ Gameplay

$U(s^t) := \frac{1}{n} \sum_{i \in N} u_i(s^t)$ ▷ Average s.w.

for all $i \in N$, $s(j) \in S$: ▷ Reinforcement

$$b_{ij}^{t+1} := \begin{cases} b_{ij}^{t^+} + \lambda \cdot U(s^t)(1 - b_{ij}^{t^+}) & \text{if } \; s(j) = s^t \\ b_{ij}^{t^+} - \lambda \cdot U(s^t) b_{ij}^{t^+} & \text{if } \; s(j) \neq s^t \end{cases}$$

$B^{t+1} := \left(b_{ij}^{t+1}\right)_{n \times k}$

Output: Probability matrix B^{t+1}

Hidden Assumptions (I)

Game structure

(I) In each round each player i can choose from the same set S_i of possible strategies to play in the stage game and this set is equal for all players. Utility functions u_i are scaled between 0 and 1.

(II) Players do not know their own or others' utility functions. The utility after each round of gameplay, together with the average social welfare, is revealed to all players separately, before the next round starts.

(III) Players have a bounded memory: at each round t players only know their probabilistic beliefs and received payoffs from the previous round $t-1$. (Not needed in case players do not know actions of others.)

(IV) Players are group-rational. They act as a single grand coalition and play honestly.

Hidden Assumptions (II)

Network structure:

(I) The weighted edges in E_W represent how much an agent trusts his neighbour in the network with regard to his expertise about the game being played.

(II) The network structure and corresponding weights of the directed edges are determined beforehand and do not change while the game is played.

(III) Players in the network are only able to directly communicate about the game with their neighbours.

(IV) Players are not aware of the entire network structure, they only know who their neighbours are. They do know however, how many players exist in the entire network.

Networks, Beliefs and Rationality

Belief:

- How to interpret *degrees* of belief?
- Threshold: i believes θ is true iff $b_i(\theta) \geq p$, for $p \in [0,1]$
- But then: how to define consensus/agreement?

Trust:

- Drawback: changing beliefs about others' beliefs not allowed
- Changing trusts avoids duplication of information
- In fact: matrix calculation 'changes' trusts by associativity

Rationality:

- Taking into account opinions of untrusted neighbours
- Agents rationally agree with method $\Rightarrow$ consensus is rational
- Initial assignment of weights without bias

Axiomatic Properties for wPSCFs

Notes:

- **Unanimity**: bi-implication for *winner certainty* and *zero unanimity* only holds for aPSCFs, not for wPSCFs.
- **Anonimity**: if $b'_{i\bullet} = b_{\sigma(i)\bullet}$ then $F(B) = \sigma F(B')$, where σF is the wPSCF obtained by permuting the weight vector that defines F, using σ.
- **Pareto optimality** and **Monotonicity**: special notion for only wPSCFs that requires the individual $\hat{\imath} \in N$ to have a positive weight.
- **Consistency**: one could define a wPSCF F_{n+m} for two merged groups N and M by the normalized weight vector $\vec{w} = (aw_1, \ldots, aw_n, (1-a)w'_1, \ldots, (1-a)w'_m)$ for any $a \in [0, 1]$.