Opinion dynamics under out-group discrimination

Steffen Eger
UKP Lab, Technical University Darmstadt

September 19, 2016
Opinion dynamics

- Models how agents form beliefs and opinions about social/economic/everyday/... issues
- A lot of research from various disciplines such as economics, complexity science/physics, etc.
  - Golub and Jackson 2012/2014, Acemoglu et al. 2010/2012
Example
Negative relationships

- More recent works allow negative relationships between agents
- Also named anti-conformity, opposition, out-group discrimination, ...
- Relevant works are:
  - Cao et al. 2013: Rebels lead to the doctrine of the mean
  - Altafini 2013: Consensus Problems on Networks With Antagonistic Interactions
  - Javarone 2014: Social Influences in Opinion Dynamics: the Role of Conformity
  - etc.
Contributions

- Provide a generalized model of ‘opposition’
- Provide utility function based motivations
- Provide a few mathematical results
A set of \( n \) agents, denoted by \([n] = \{1, \ldots, n\}\)

A set \( S \), the opinion spectrum

- \( S \) is either continuous, e.g., \( S = [0, 1] \) (CONTINUOUS MODEL), or
- \( S \) is discrete, e.g., \( S = \{\text{Yes, No}\} \) (DISCRETE MODEL)

Each agent \( i \) has an in-group (his ‘friends’) and an out-group (his ‘enemies’)

Each agent \( i \) also has two associated functions

- \( \mathcal{I} \) is the identity function on \( S \) — for his in-group
- \( \mathcal{D}_i \) is not the identity function on \( S \) — for his out-group. Call \( \mathcal{D}_i \) also deviation or inversion function
Each agent $i \in [n]$ has a utility function for the opinions $b_1, \ldots, b_n$ of all other agents

$$u_i(b_1, \ldots, b_n) = - \sum_{j \in \text{In}(i)} W_{ij} (b_i - \mathcal{F}(b_j))^2 - \sum_{j \in \text{Out}(i)} W_{ij} (b_i - \mathcal{D}_i(b_j))^2$$

Similar utility function for the DISCRETE MODEL, but omitted here.
Agents are utility-maximizers, they choose their opinions $b_i$ to maximize their utility functions given the opinions of their peers.

In the CONTINUOUS MODEL, this leads to the dynamics:

$$b_i(t + 1) = \sum_{j \in \text{in}(i)} W_{ij} \tilde{\xi}(b_j(t)) + \sum_{j \in \text{Out}(i)} W_{ij} \mathcal{D}_i(b_j(t))$$

In the DISCRETE MODEL, this leads to the dynamics:

$$b_i(t + 1) = \text{weighted majority (inverted) opinion of peers}$$
Example

- Let $S = \{0, 1\}$ be a binary opinion space.
- Let for all $i \in [n]$: $D_i(x) = D(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x = 1 \end{cases}$
- If $\text{In}(i) = [n]$: agent $i$ always chooses (weighted) majority opinion of his peers $\rightarrow$ “conformist”
- If $\text{In}(i) = \emptyset$: agent $i$ always chooses (weighted) minority opinion of his peers $\rightarrow$ “anti-conformist”
Results 1 — Long-run disagreement

- Under (a few, weak, technical conditions) agents disagree in the long-run whenever

\[
\text{Fix}(\mathcal{O}_i) \cap \text{Fix}(\mathcal{O}_j) = \emptyset
\]

for two agents \(i, j\)

- Here \(\text{Fix}(\mathcal{O}) = \{x \mid \mathcal{O}(x) = x\}\)

- The result means that to agree in the long-run there must exist an opinion \(x\) which is “neutral” for all agents \(i \in [n]\) in the sense that \(\mathcal{O}_i(x) = x\)

- \(\rightarrow\) long-run/persistent disagreement is “likely” under our model of out-group discrimination
Opposition bipartite network
Results 2 — Definitions

Reverse opposition bipartite network
Consider the following special case

- \( S = \mathbb{R} \)
- \( \mathcal{D}_i(x) = \mathcal{D}(x) = -x \)
- Also assume some technical conditions: symmetry of out-group relationship, aperiodicity of graphs wherein agents interact, connectedness of graphs, etc.
- Then, there are only three possible outcomes in the long-run: \textbf{consensus, polarization} and \textbf{non-convergence} (oscillation of beliefs)
Results 2 — Opinion polarization

- Agents’ opinions polarize in the long-run iff network wherein they interact is **opposition bipartite**
- Agents’ opinions do not converge in the long-run iff **reverse opposition bipartite**
- Agents’ opinions reach a consensus in the long-run iff neither **opposition bipartite** nor **reverse opposition bipartite**
Conclusion

- Presented a generalized model of opinion dynamics under out-group discrimination (negative relationships)
- Based on the idea of “deviation functions” and in-group/out-group structure
- Provided utility function based motivation
- Also shown two mathematical results:
  - long-run agreement in society is difficult under our model
  - we also gave network conditions of long-run polarization, consensus and non-convergence in a special case
Relevant papers:

Thank You!