



Utrecht University

Dominik Klein

# AI for Open Societies

Session IV - Emergence of Collective Patterns

# Today's lecture



- **Content:** Collective patterns
- **Methodological:** Agent-based modelling

# Emergence of Collective Patterns



- Using AI systems to understand, improve on and design institutions for open societies

# Emergence of Collective Patterns



- Using AI systems to understand, improve on and design institutions for open societies
  - Communication and Participation
    - Expert networks, citizen initiatives (→ participative budgeting in Amsterdam,...)
    - Fairness, avoid bubbles and polarization...
  - Resilient Institutions
    - Economic Shocks, (Foreign?) malign actors
  - Organization of Public Spaces (Security, Interaction...)

⇒ Collective perspective

⇒ Agent based Computer Simulations as central tool.

# Emergence of Collective Patters

Some aspects of **open societies** are discernible at individual level, others only at the **collective level**.



# Emergence of Collective Patterns



Some aspects of **open societies** are discernible at individual level, others only at the **collective level**.

- Polarization  
(Hegselmann and Krause 2002, Douven and Hegselmann 2021)
- Segregation  
(Schelling 1969)
- Inequality  
(Klein et al. 2020)
- (Social) Norms  
(Bicchieri 2005, Lisciandra et al. 2013)
- (Gendered) Division of Labor  
(o'Connor 2019)

# Emergence of Collective Patterns



Some aspects of **open societies** are discernible at individual level, others only at the **collective level**.

- Polarization  
(Hegselmann and Krause 2002, Douven and Hegselmann 2021)
- Segregation  
(Schelling 1969)
- Inequality  
(Klein et al. 2020)
- (Social) Norms  
(Bicchieri 2005, Lisciandra et al. 2013)
- (Gendered) Division of Labor  
(o'Connor 2019)

**General Picture:** Individuals act, politics defines the playing field...  
but some effects only show on social level. ⇒ **Emergence**

# Plan and Goals for Today



- 1 Theoretical apparatus to talk about emergence
- 2 Some important concepts
- 3 Case study: opinion dynamics
- 4 Agent-Based Modelling as AI tool

## Take Home

- Familiarity with the emergence way of thinking
- Insights to what can happen and language to discuss this
- Agent-based modelling as a tool

# An Example



Part of the [https://ccl.northwestern.edu/netlogo/ Models library](https://ccl.northwestern.edu/netlogo/Models%20library)  
(freely downloadable)

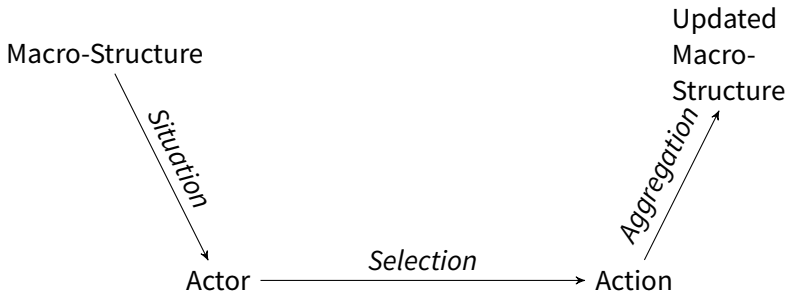
# Emergence



- **Emergence:**  
Complex social patterns arise from the interactions of individuals. The resulting social patterns may have properties that do not immediately follow from individuals and their properties.
  
- In a slogan “The sum is more than its parts”

# Coleman's bathtub

Explanation in the Social Sciences



## Methodological Individualism:

Every explanation of social level phenomena needs to refer to individual agents, e.g. how their perception of the environment and their (inter)actions give rise to the social level phenomenon.

# Some toy Explananda



- Walking on the right hand side
- Panic buying...
- Using the same messenger

# Plan and Goals for Today



- 1 Theoretical apparatus to talk about emergence models ✓
- 2 Some important concepts
- 3 Case study: opinion dynamics
- 4 Agent-Based Modelling as AI tool

# Important concepts (and Observations)



- Emergence
- Path Dependencies
- Tipping Points
- (Non-Monotonicity)
- Direction of Effect

# A Second Example



Let's play a game

# A Second Example



Let's play a game

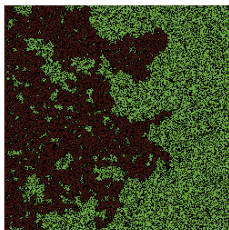
## Path Dependencies

Small scale random effects early on in a model run may have significant and lasting effects on long term behavior.

## Example

- Emergence of Firms
- Construction of Social Networks
  - ⇒ Preferential Attachment

# A third example



## Tipping Points:

Miniscule changes in model parameters can have drastic influences on the output.

Mathematically speaking, the function from input to output values need not be continuous

- Climate Change
- Emergence of Trust/Revolutions
- ...

# Direction of Effect - do they (have to) want it?



Question:

How are individual motives and emergent effect related?

# Direction of Effect - do they (have to) want it?



## Question:

How are individual motives and emergent effect related?

- Collective Pattern can be in line with individual motivations or competences
  - ! Can even be stronger than individual motivations (Segregation)
- Collective Pattern can be accidental (Bar Example)
- Collective Pattern can be opposed to individual motivation (Adam Smith and the invisible hand)
- Collective Pattern may not have an individual counterpart (Trends, Flocking...)

# Important concepts



- Emergence ✓
- Path Dependencies ✓
- Tipping Points ✓
- (Non-Monotonicity)
- Direction of Effect ✓

# Plan and Goals for Today



- 1 Theoretical apparatus to talk about emergence models ✓
- 2 Some important concepts ✓
- 3 Case Study: Opinion Dynamics
- 4 Agent-Based Modelling as AI tool

# Plan and Goals for Today



- 1 Theoretical apparatus to talk about emergence models ✓
- 2 Some important concepts ✓
- 3 Case Study: Opinion Dynamics
- 4 Agent-Based Modelling as AI tool

# Hegselmann-Krause: The mechanism



- Stance on a  $[0; 1]$  scale
- Agents
  - Have some private starting position
  - Try to learn from each other
    - ⇒ Have to determine who is trustworthy / who to learn from.
- Every agent has a **confidence level**  $\epsilon$ .  
Takes only opinions into account that are at most  $\epsilon$  from their own.
- New opinion: Average over all opinions the agent takes into account

R. Hegselmann, U.Krause. *Opinion dynamics and bounded confidence models*. Journal of Artificial Social Societies and Social Simulation 2002.

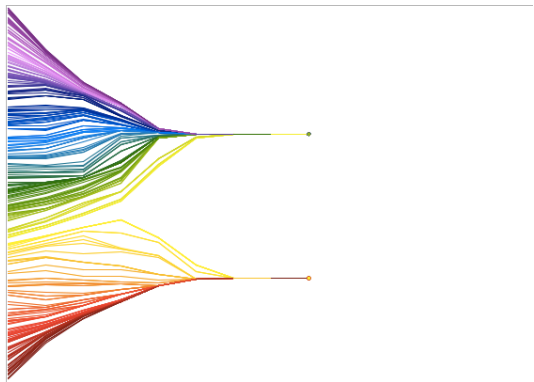
# Hegselmann-Krause: The mechanism



- Stance on a  $[0; 1]$  scale
- Agents
  - Have some private starting position
  - Try to learn from each other
    - ⇒ Have to determine who is trustworthy / who to learn from.
- Every agent has a **confidence level**  $\epsilon$ .  
Takes only opinions into account that are at most  $\epsilon$  from their own.
- New opinion: Average over all opinions the agent takes into account

$$pos_{new}^j = \frac{1}{\#\{j \mid |pos_{old}^j - pos_{old}^i| < \epsilon\}} \sum_{j: |pos_{old}^j - pos_{old}^i| < \epsilon} pos_{old}^j$$

R. Hegselmann, U.Krause. *Opinion dynamics and bounded confidence models*. Journal of Artificial Social Societies and Social Simulation 2002.



# Findings



## ■ Base Model

No assumption of correct value

→ Either questions without truth value (beauty, moral value)  
or where information is hard to get.

## ■ Polarization

can form, even though everybody is unbiased and rational  
(=Clusters of agents too far apart to talk to each other)

## ■ Non-monotonicity:

Changes in confidence interval  $\epsilon$  can sometimes increase the  
number of clusters (though usually decreases)

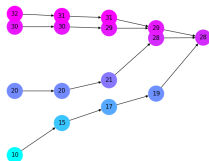
# Openness Monotonicity



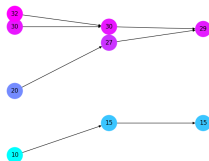
Changes in confidence interval  $\epsilon$  can sometimes increase the number of clusters

For simplicity: opinions on interval  $[0,100]$  instead of  $[0;1]$

Positions: 10, 20, 30, 32, 32, 32, 32



$\epsilon = 11$



$\epsilon = 12$

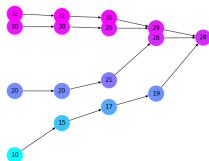
# Openness Monotonicity



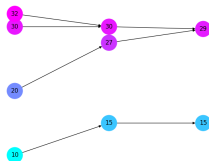
Changes in confidence interval  $\epsilon$  can sometimes increase the number of clusters

For simplicity: opinions on interval  $[0,100]$  instead of  $[0;1]$

Positions: 10, 20, 30, 32, 32, 32, 32



$\epsilon = 11$



$\epsilon = 12$

- And can also move the position of consensus (Duijf and Klein, in progress)

# Findings



## ■ Base Model

No assumption of correct value

→ Either questions without truth value (beauty, moral value)  
or where information is hard to get.

## ■ Polarization

can form, even though everybody is unbiased and rational  
(=Clusters of agents too far apart to talk to each other)

## ■ Non-monotonicity:

Changes in confidence interval  $\epsilon$  can sometimes increase the  
number of clusters (though usually decreases)

↑ missing concept from last section

# The extended model



- There is a correct state  $\tau$
- Next to social signals there is also (rare) signals about  $\tau$ .
- Three types of agents:
  - Free Riders: Only Social Signal Updating as before
  - Truth Seekers: update position
$$pos_{new}^{i, truthseeking} = (1 - \alpha)pos_{new}^i + \alpha \cdot \tau$$
  - Campaigners: Fixed position  $\rho$

Igor Douven and Rainer Hegselmann. *Mis-and disinformation in a bounded confidence model*. Artificial Intelligence 2021.



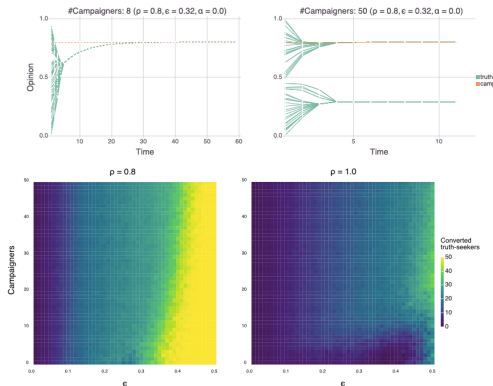
...Hold a fixed position  $\rho$ , independent of the truth  $\tau$ .

- **Misinformation** Aim that the public believes a falsehood (about a given topic)
- **Disinformation** Aim to impede/distract the public from believing a truth (about a given topic)

**Note:** For logically consistent agents Misinformation implies Disinformation, but not vice versa.

# Observations

## i) Without influence of the truth

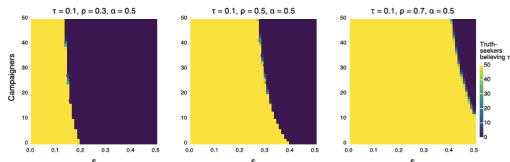


- Having a more extreme campaign position hurts the campaigners
- Having more/stronger campaigners may impede the campaign

# Observations



*ii) With influence of the truth, no free riders*

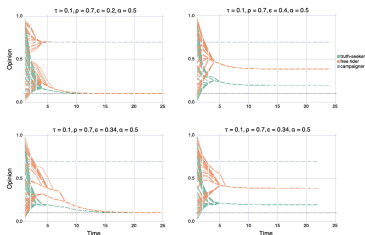


Note: graphic changes from mis- to disinformation

- Light deviation (small distance between  $\tau$  and  $\rho$ ) more successful than strong deviations
- Having more/stronger campaigners can help the campaign (contrary to before)

# Observations

## iii) With influence of the truth and free riders



- If anything, free riders have a negative influence on truth seekers (additional graph in the article)
- Disinformation against truth seekers can increase when free riders are present (misinformation can never)
- Number of free riders not too important for the effect
- Free riders themselves can be drawn to campaigners, truth seekers, or trapped in the middle.

# Observations



## *iv) Variation: changing confidence*

- Updating confidence intervall from peers
- Free riders have (even) less impact on truth seekers
- Stronger distance between  $\tau$  and  $\alpha$  can increase effect on free riders (contrary to before...)

# Interpretations of the model

- Offers some insights into complex interplay of iterated social learning and campaigning.
- Iterated interaction produces arguably counter-intuitive findings



# Interpretations of the model



- Offers some insights into complex interplay of iterated social learning and campaigning.
  - Iterated interaction produces arguably counter-intuitive findings
    - (for campaigners it can help to be few and subtle)
  - Few general lessons, effects can revert in model variations
  - Open to various interpretations.
  - Qualitative, but not quantitative interpretations
  - Heavily idealized
    - Absence of social networks
    - No substantial arguments exchanged
    - Everybody has same  $\epsilon$
    - Double counting
- ⇒ Questions of validity and transferability

# Some extensions



- Having underlying network structure: Can only hear those that I am connected to (Wu et al 2022)
- Multiple discussion topics at once (Lorenz 2008)
  - Closeness on topic  $x$  can allow me to listen to somebody with whom I disagree on  $y$ )
  - Often: One central cluster, some around
- Even more AI: Hegselmann Krause meets LLM (e.g. Betz 2023)
  - Substantial Argument exchange, positions derived through LLM

# Emergent Phenomena



- Social Level Phenomena emerging through individual interaction
- Can be subjected to tipping points and path dependencies
  - Makes analytic (mathematical) solutions often intractable
  - May be hard to have valid intuitions about.
- Suitable AI tool: Agent based computer simulations
- By now a number of free and easy to use software packages exist.

D. Klein, J.Marx, K. Fischbach. *Agent-Based Modeling in Social Science, History, and Philosophy: An Introduction* . HSR 2018.

# Agent-Based Modelling as AI tool

- Agent-Based explanations require
  - A set of agents
  - Their interaction



# Agent-Based Modelling as AI tool



- Agent-Based explanations require
  - A set of agents
  - Their interaction
- **Ideal for computational AI models.**
  - ⇒ Require a formal characterization of
    - Landscapes
    - Agents
    - ...their behavior and interaction
    - Allow for high degrees of heterogeneity
- Can be done in python, but specialized languages (NetLogo...) exist
- Usually, old-fashioned symbolic AI. But there are integrations with LLM.

# Agent based Modelling



- Strengths
  - Highly versatile (psychologic mechanisms, rationalistic, multitude of agents)
  - Allow to study and discover emergent phenomena
- Weaknesses
  - Sometimes difficult to test for validity.
  - Inherent computational complexity
- Opportunities
  - Can help in policy design, applicable to diverse set of agents
- Threats
  - Over-interpretation of results, depending on idealization
  - Simplification

# Integration with other parts of AIOS



- Hortensius
  - Specify psychologic decision rules for agents
  - Specify aspects of network building behavior (who connects to whom...)
- van der Vegt
  - Specify criteria and influence rules for communication
  - Specify linguistic component of a model.
- van Maanen:
  - Use insights for calibration of agents and their behavioral rules
  - Specify Models for how agents decide



01-05-2024



Utrecht  
University

Security in Open Societies

Future of Work

Bottom-up Initiatives  
for Societal Change

Open Cities

Behaviour and Institutions

Transitions and well-being

Democracy and  
good governance

Fair Transitions

Contesting Governance

Longtermism and Institutional Change

Futures of Democracy

Markets and Corporations

The Transactional State  
as an Institution for Good

Gender, Diversity and Global Justice

Openness challenged: the university at risk?

In/Equality

Equality and diversity

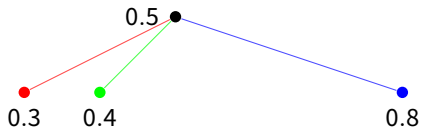
## IOS-Themes: 15 platforms



Questions?

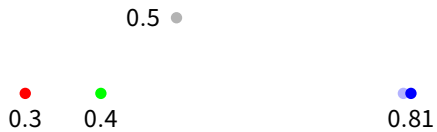


$\epsilon = 0.5$





$\epsilon = 0.5$





$\epsilon = 0.5$

