Computational modelling for social research

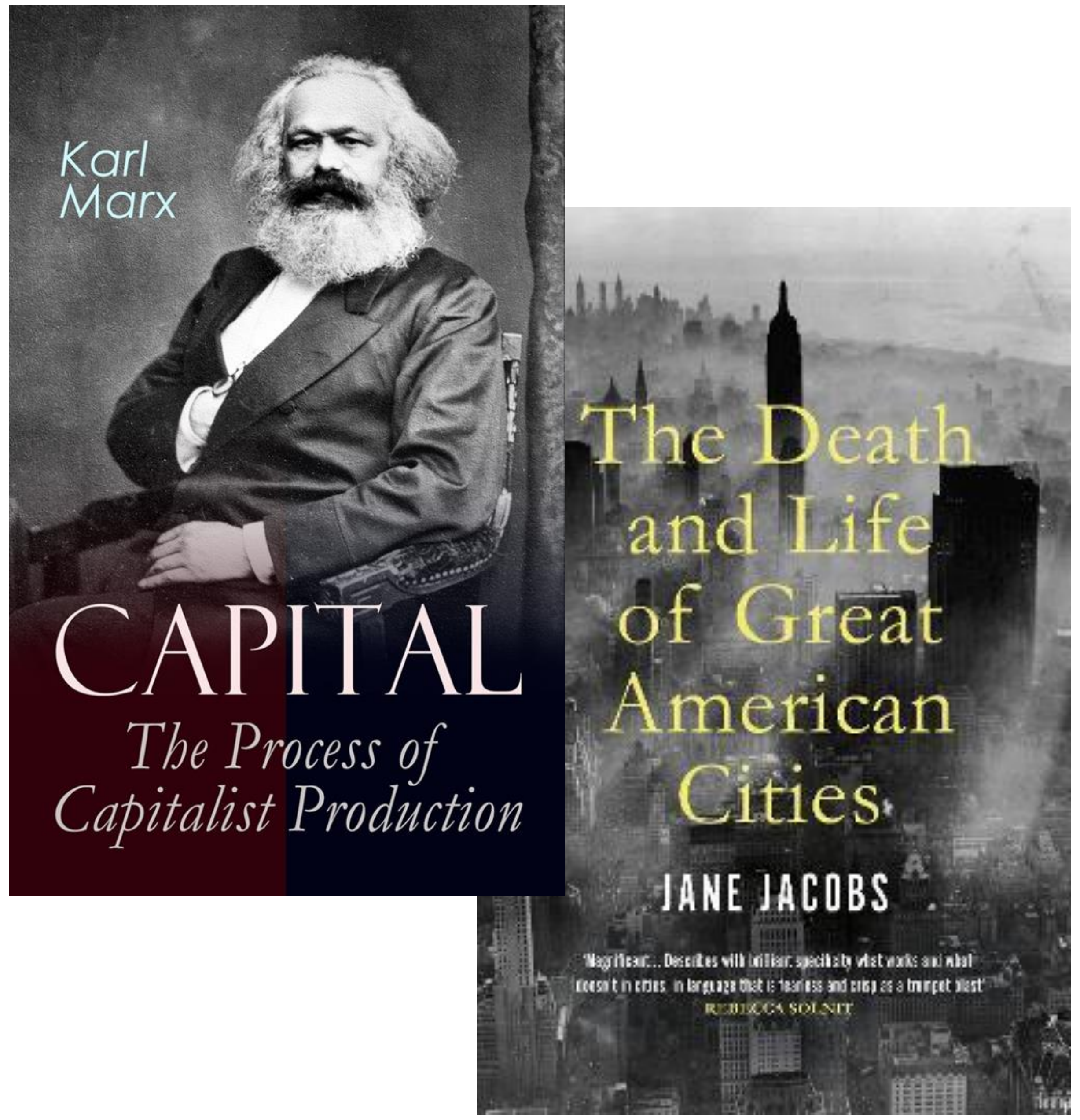
Part II: Introduction to formal modelling

Michele Tizzoni

January 28, 2025

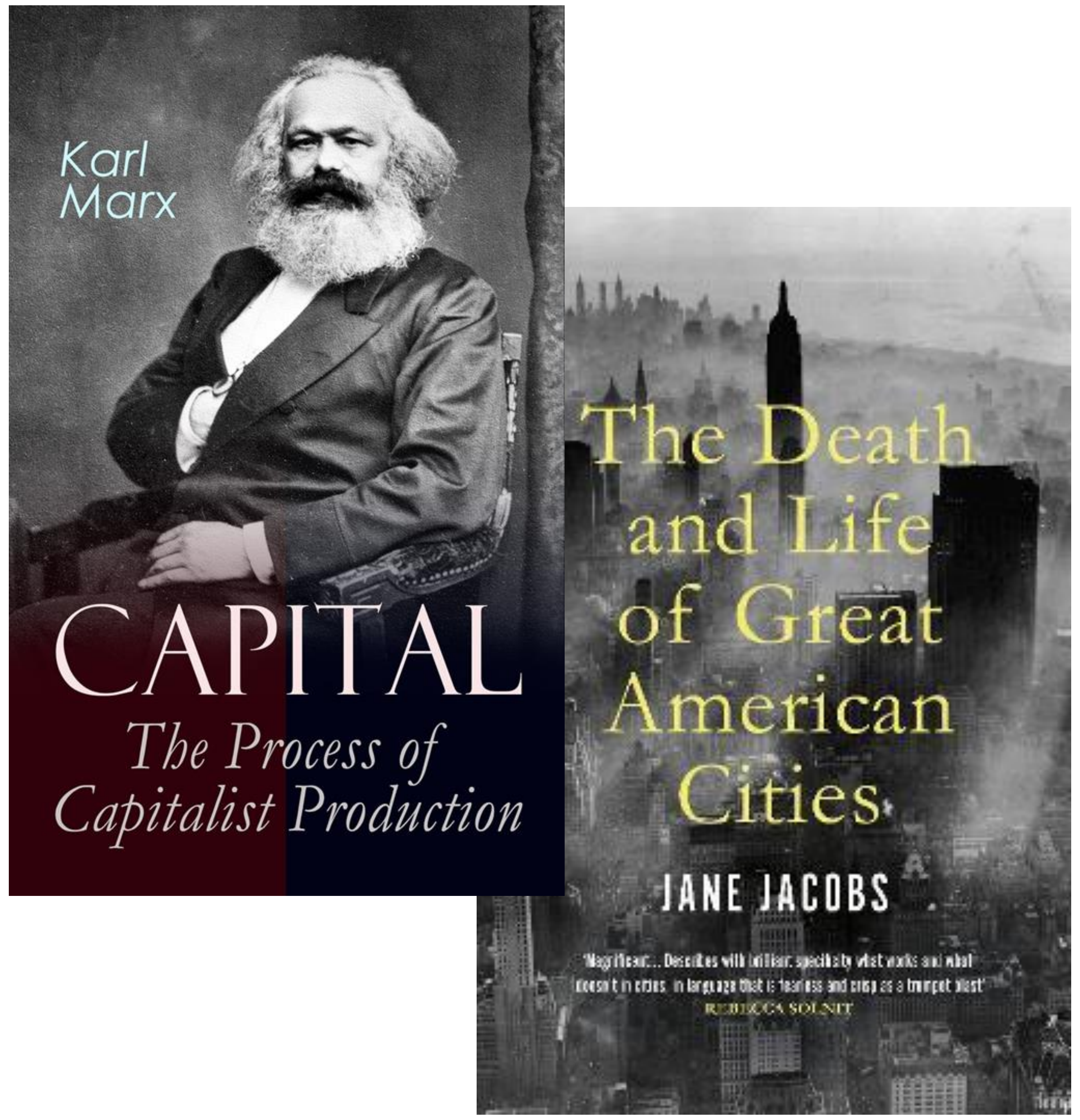
What is a model in social research?

What is a model in social research?



Verbal / theoretical model

What is a model in social research?

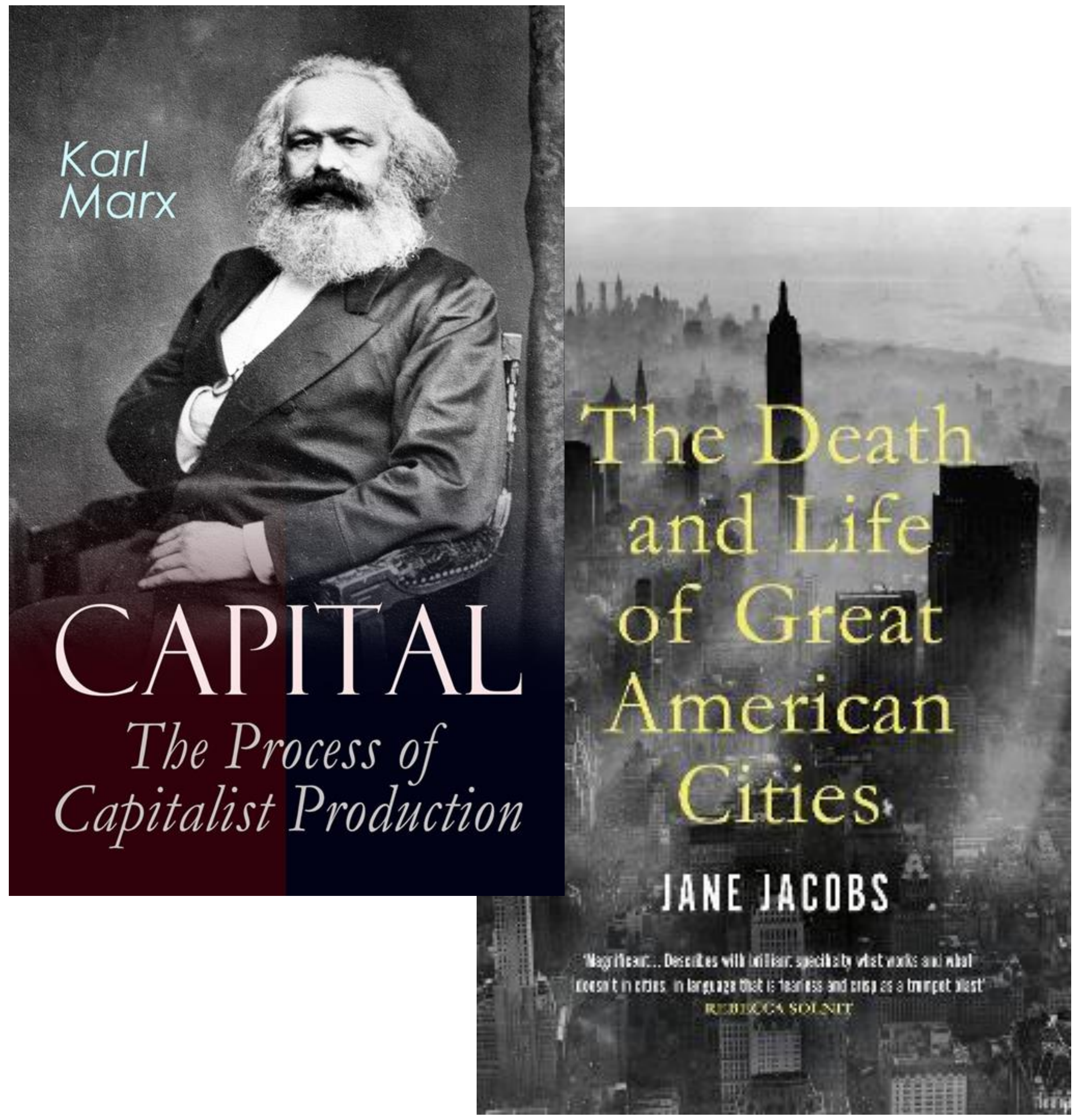


Verbal / theoretical model

$$\hat{Y} = \beta x + \epsilon$$

Statistical model

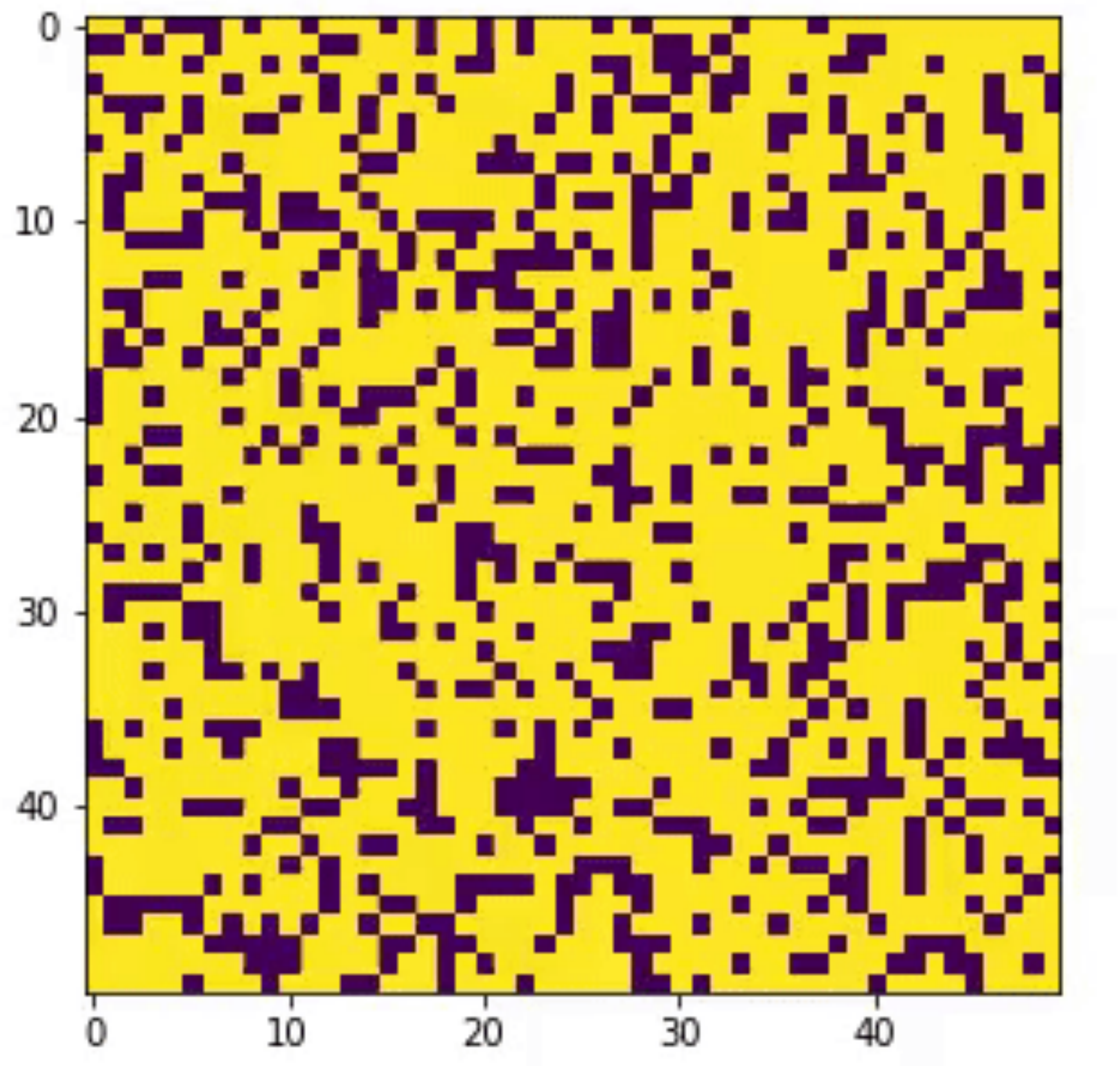
What is a model in social research?



Verbal / theoretical model

$$\hat{Y} = \beta x + \epsilon$$

Statistical model



Formal/dynamical model

Why do we need formal models?

> All models are wrong.
> Some are useful.

-George E. P. Box

Why do we need formal models?

reality

abstraction,
simplification

Why do we need formal models?

- A formal model is a **simplified version** of reality, written in **mathematical equations** or **computer code**.

- Formal models are useful because **reality is complex**. No model is ever a complete recreation of reality.

- A formal model contains only those **few elements and processes** that the modeller suspects to be important.

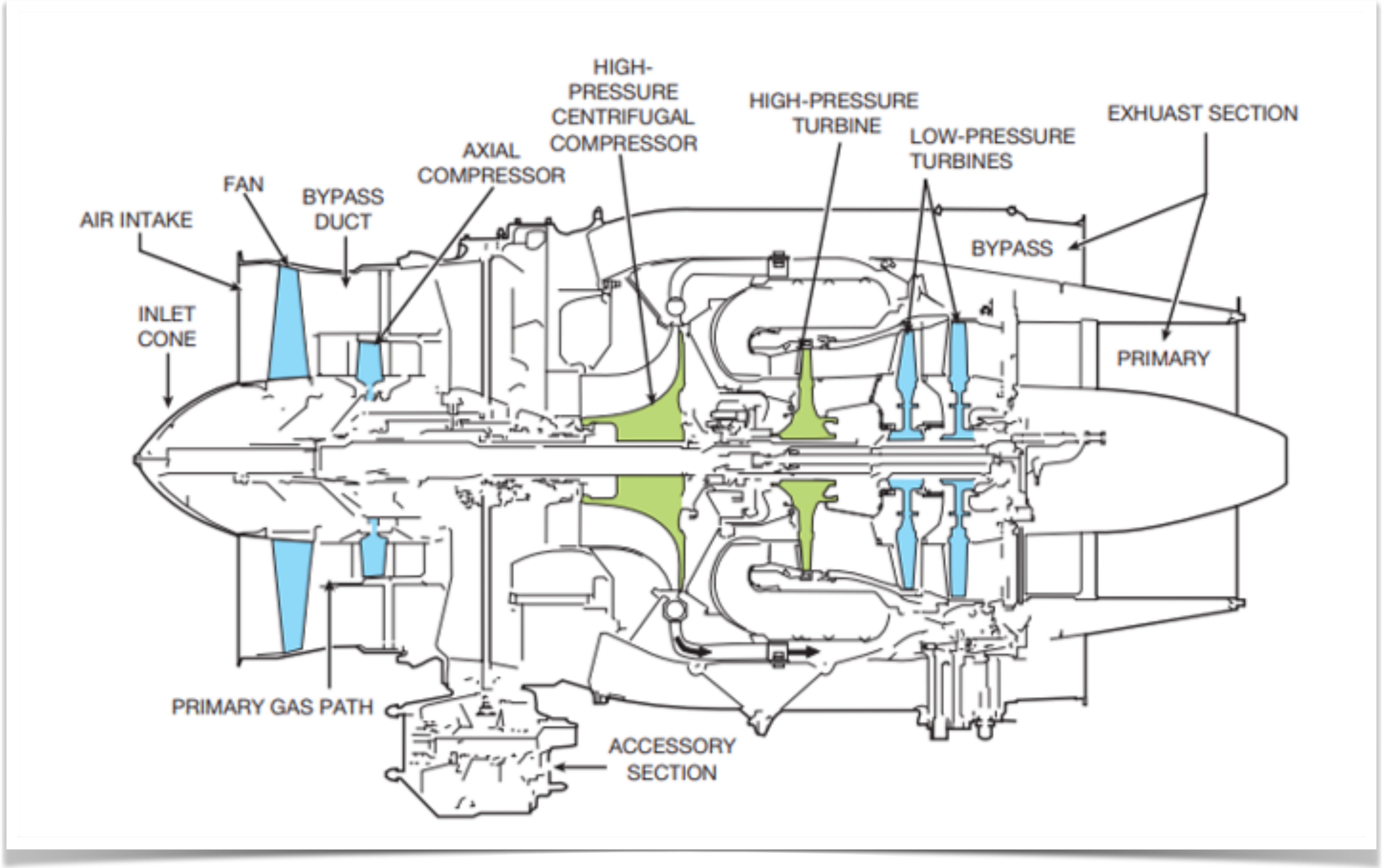
Why do we need formal models?

- The **advantage** of formal modelling is that we are forced to **precisely specify every element** and process that we propose.

- Models can also help to **understand the consequences of our theories**.

- With formal models we can set up systems where **different forces or mechanisms interact** and observe the dynamics of their interaction: we can study **complex systems.**

Complex systems



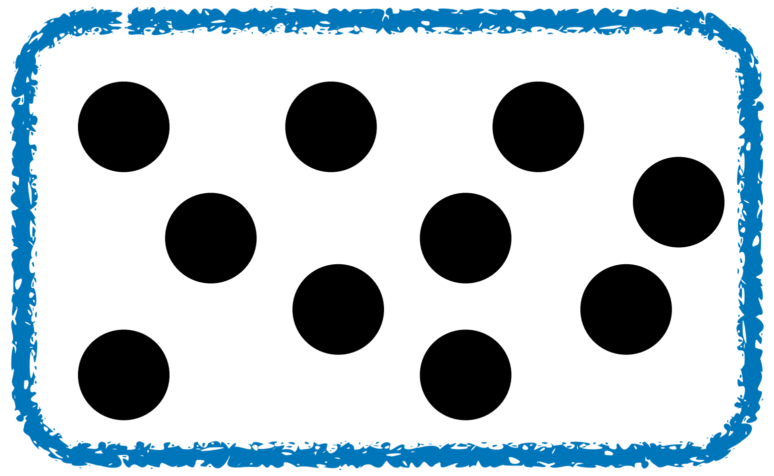
Not complex

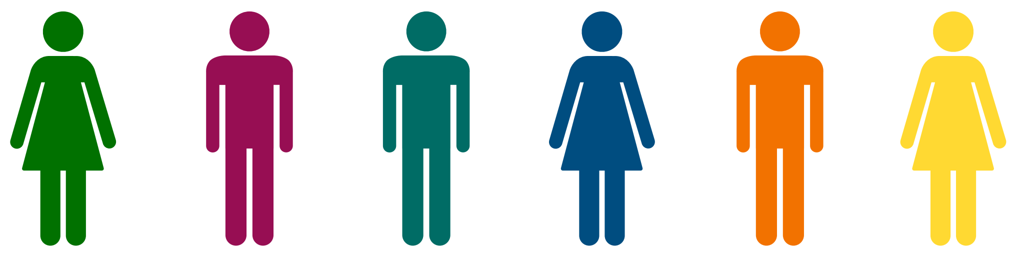
Complex

Types of formal models

Population level

Individual-based

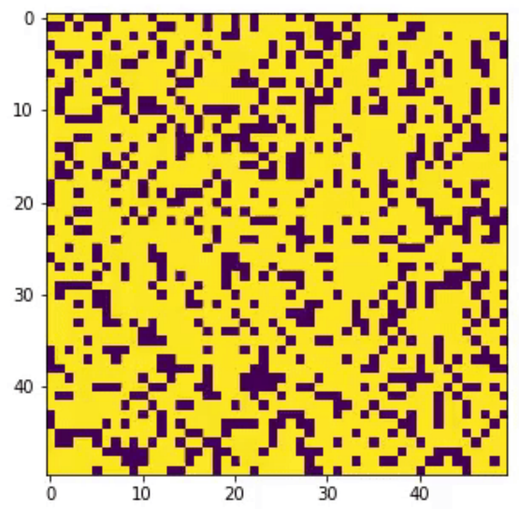




Analytically tractable

Computer simulations

$$\begin{cases} \dfrac{di}{dt} = -\lambda si \\ \dfrac{ds}{dt} = \lambda si - \alpha s(s+r) \\ \dfrac{dr}{dt} = \alpha s(s+r) \end{cases}$$



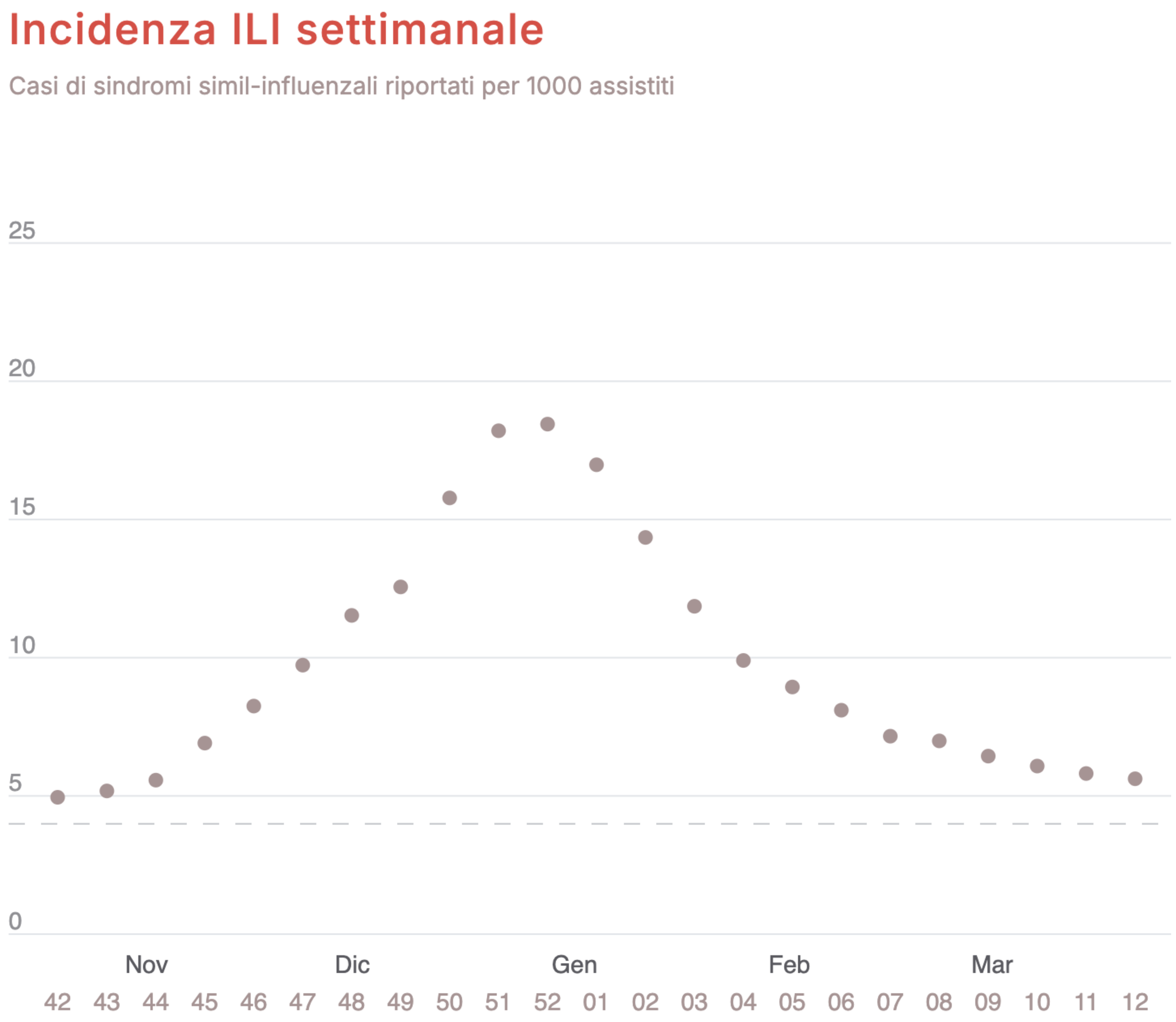
Formal models in the natural sciences

- **Biology**: predator-prey model (Lotka-Volterra equations)
  - *Models interactions between predators and preys in an ecosystem*
- **Evolutionary biology**: the replicator equation
  - *Models the frequency change of a strategy in an evolving population*
- **Genetics**: Hardy-Weinberg equation
  - *Models the allele frequency distribution in a non-evolving population*
- **Epidemiology**: the SIR model
  - *Models disease spread in a closed population*
- All of physics, chemistry, geology…

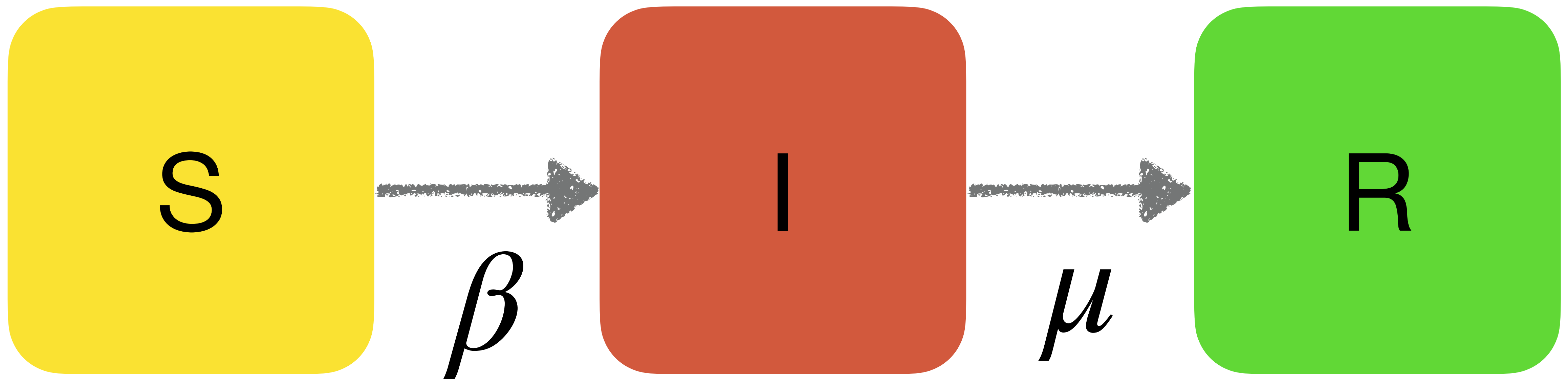
A classic example: epidemics

Seasonal influenza

in Italy: 2023-2024



The SIR model (1924)



- Population is closed and constant (no demographics) $S + I + R = N$
- Population is "well mixed" (no heterogeneities)
- **Recovery** takes place at constant rate: μ
- **Transmission** rate is constant: β

The SIR model (1924)

$$
\begin{cases}
\dfrac{dS}{dt} = -\beta\dfrac{SI}{N} \\
\dfrac{dI}{dt} = \beta\dfrac{SI}{N} - \mu I \\
\dfrac{dR}{dt} = \mu I
\end{cases}
$$

Epidemic threshold

$$R_0 = \frac{\beta}{\mu} > 1$$

The disease will infect a fraction of the population

The SIR model (1924)

$$\begin{cases} \dfrac{dS}{dt} = -\beta\dfrac{SI}{N} \\ \dfrac{dI}{dt} = \beta\dfrac{SI}{N} - \mu I \\ \dfrac{dR}{dt} = \mu I \end{cases}$$

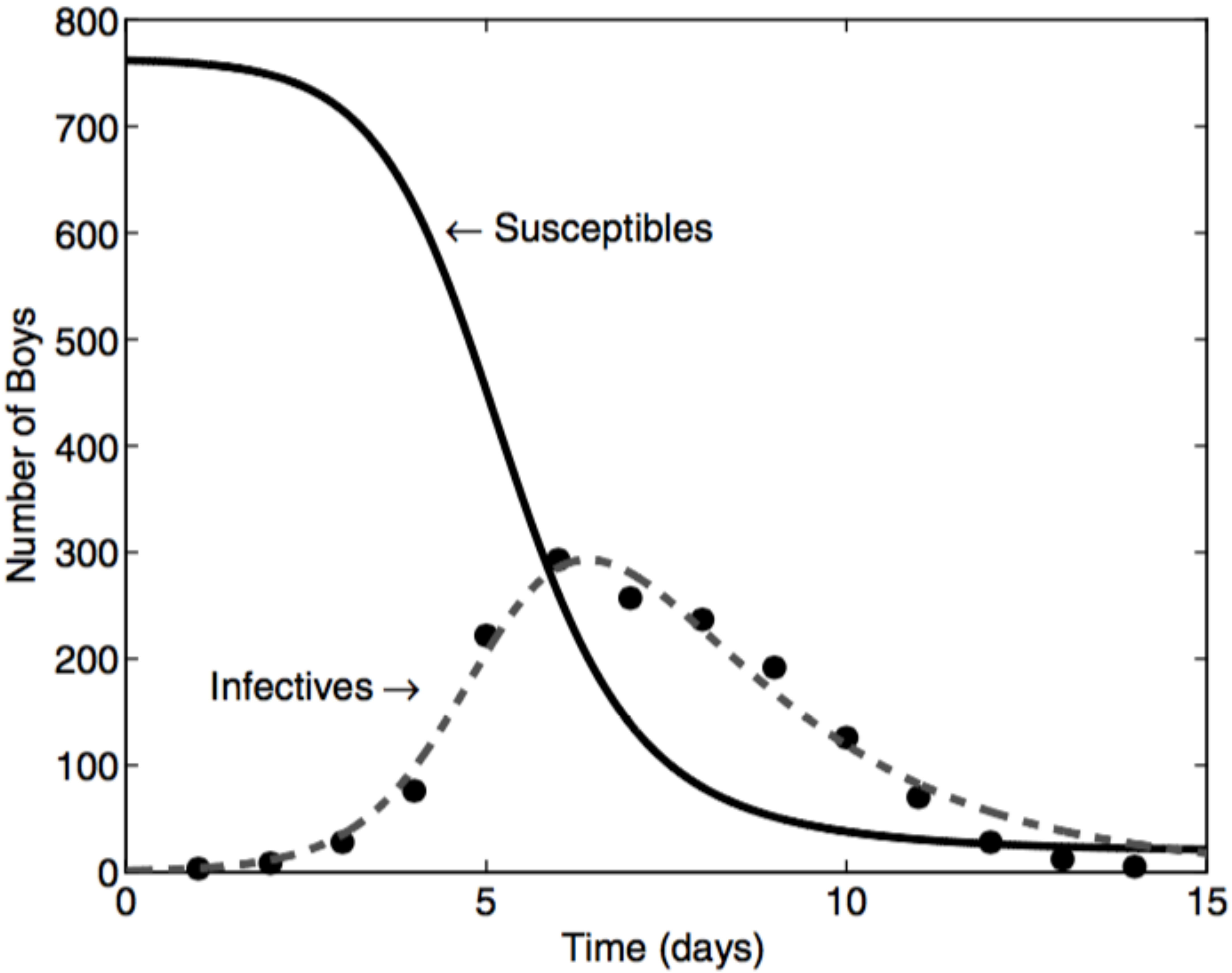
Epidemic threshold

$$R_0 = \frac{\beta}{\mu} > 1$$

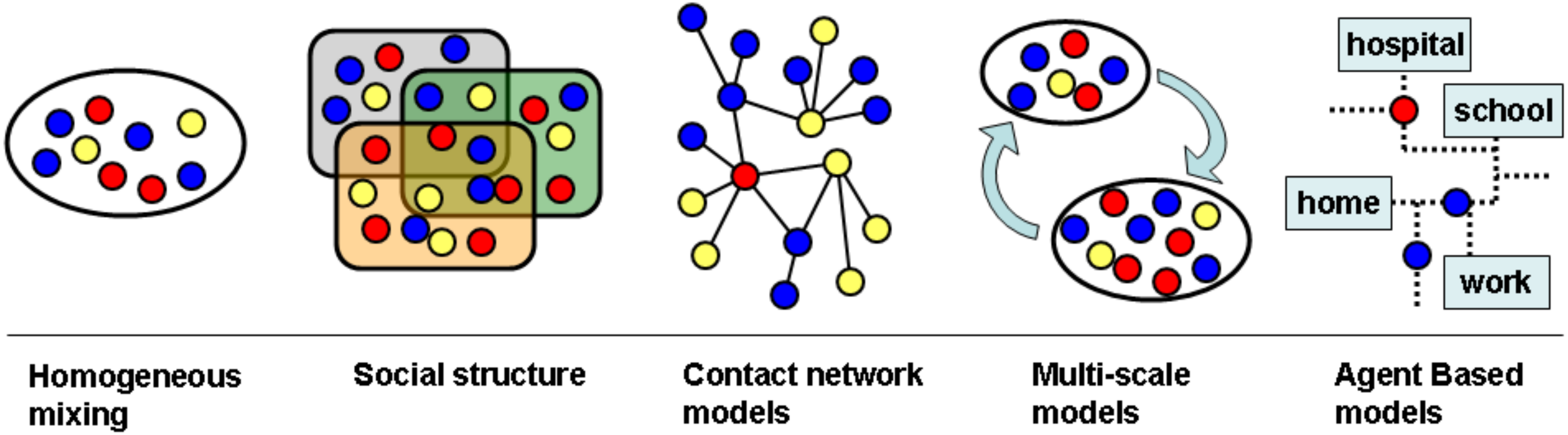
The disease will infect a fraction of the population

The SIR model (1924)

- Flu epidemic in a boarding school in England, 1978 (data from BMJ)



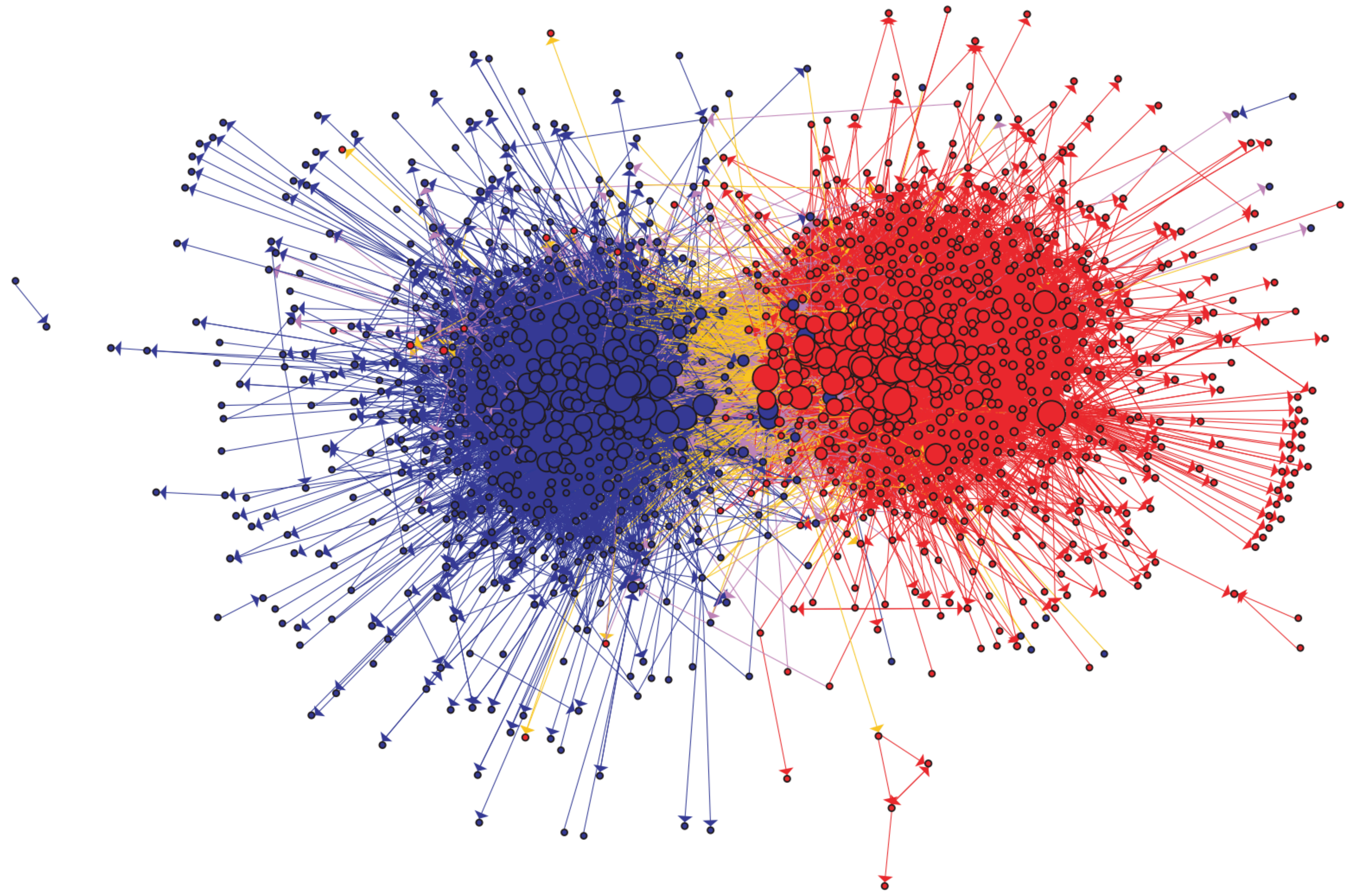
A taxonomy of epidemic models



Population level ↔ Individual-based

Formal models in social research

- Social influence
- Rumor propagation
- Opinion/consensus formation
- Cooperative phenomena
- Conflict resolution
- Spatial segregation
- Cultural evolution
- Scientific production
- Economic complexity
- and many others…



Adamic & Glance "Divided they blog" 2005

Social dynamics models

- **Rumor/gossip model**
- The voter model
- The Axelrod’s model
- The Schelling’s model of segregation
- The naming game

Information spreading

- Rumors and information spreading phenomena are the prototypical examples of social contagion processes in which the infection mechanism can be considered of psychological origin.
- Social contagion phenomena refer to different processes that depend on the individual propensity to adopt and diffuse knowledge, ideas or simply a habit.
- The similarity between social contagion processes and epidemiological models was recognized quite a long ago.

Information spreading

- Rumors and information spreading phenomena are the prototypical examples of social contagion processes in which the infection mechanism can be considered of psychological origin.
- Social contagion phenomena refer to different processes that depend on the individual propensity to adopt and diffuse knowledge, ideas or simply a habit.
- The similarity between social contagion processes and epidemiological models was recognized quite a long ago.

Social and physiological contagion processes **differ in some important features**:

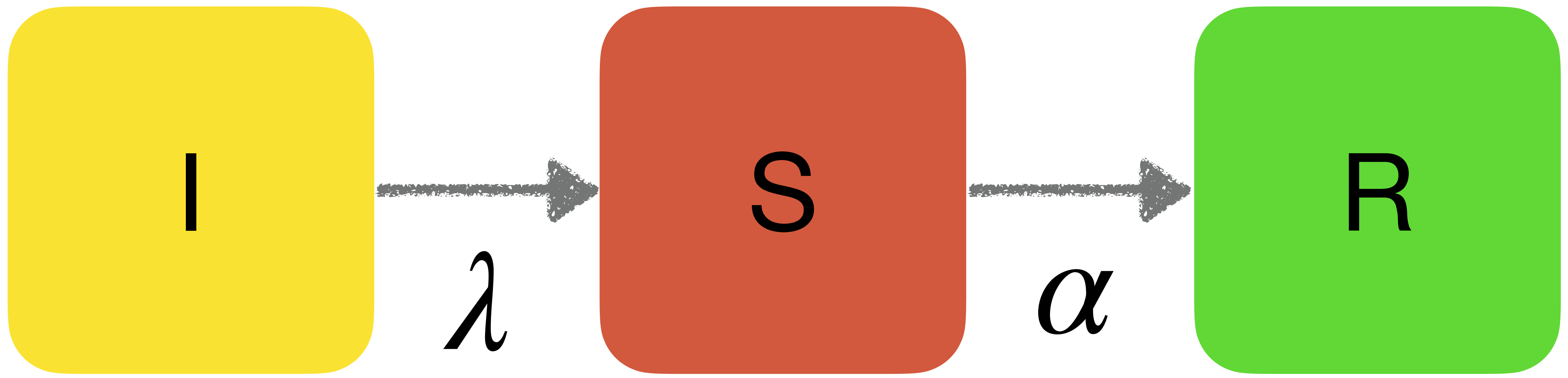
- spread of information is an intentional act, unlike a pathogen contamination.
- it is usually advantageous to access a new idea or information, so being infected is no longer just a passive process.
- acquiring a new idea or information may need **time and exposure to more than one source of information**, thus the importance of models in which memory has an important role.

A model for gossip

Dailey & Kendall, Stochastic rumors (1965)

Epidemics	Rumor
Susceptible	Ignorant
Infected	Spreader
Recovered	Stifler

A model for gossip



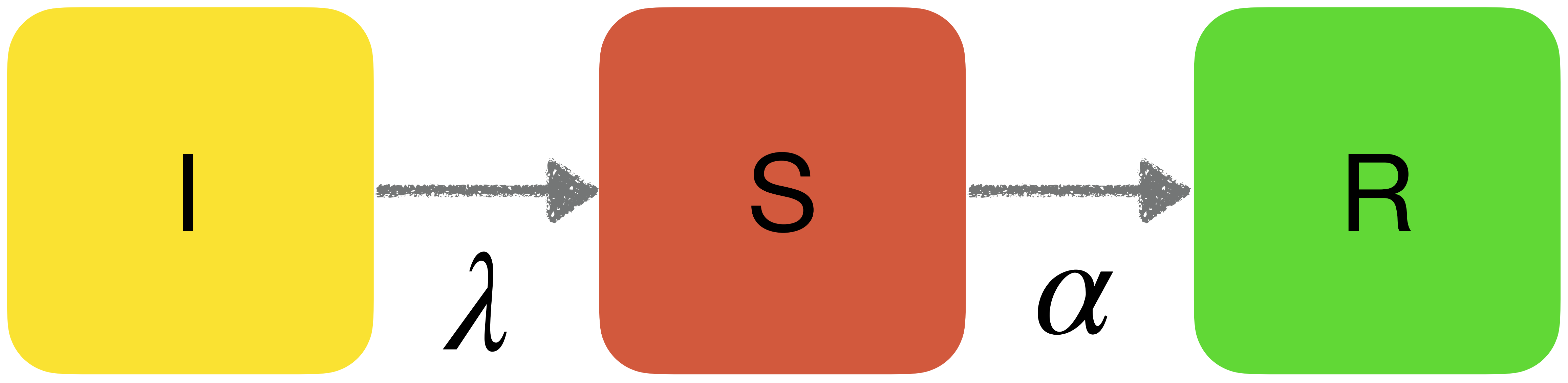
Ignorant + Spreader → 2 Spreaders

Spreader + Spreader →Stifler + Spreader

Spreader + Stifler → 2 Stiflers

A model for gossip

$$\begin{cases} \dfrac{di}{dt} = -\lambda si \\ \dfrac{ds}{dt} = \lambda si - \alpha s(s+r) \\ \dfrac{dr}{dt} = \alpha s(s+r) \end{cases}$$



A model for gossip

$$\frac{ds}{dt} = (\lambda + \alpha)si - \alpha s$$

Equivalent to the epidemic SIR

With $\lambda + \alpha = \beta$ and $\alpha = \mu$

$$\boxed{R_0 = \frac{\lambda + \alpha}{\alpha} > 1}$$

$$\boxed{\frac{\lambda}{\alpha} > 0}$$

No threshold in homogeneous systems

Social dynamics models

- Rumor/gossip model
- **The voter model**
- The Axelrod’s model
- The Schelling’s model of segregation
- The naming game

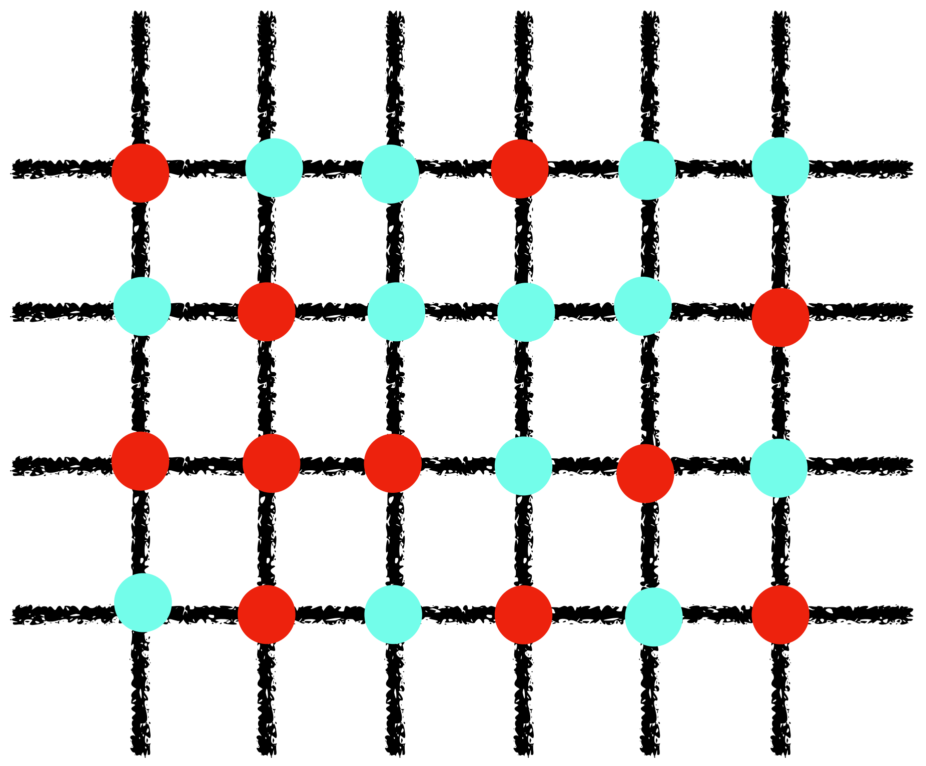
The voter model

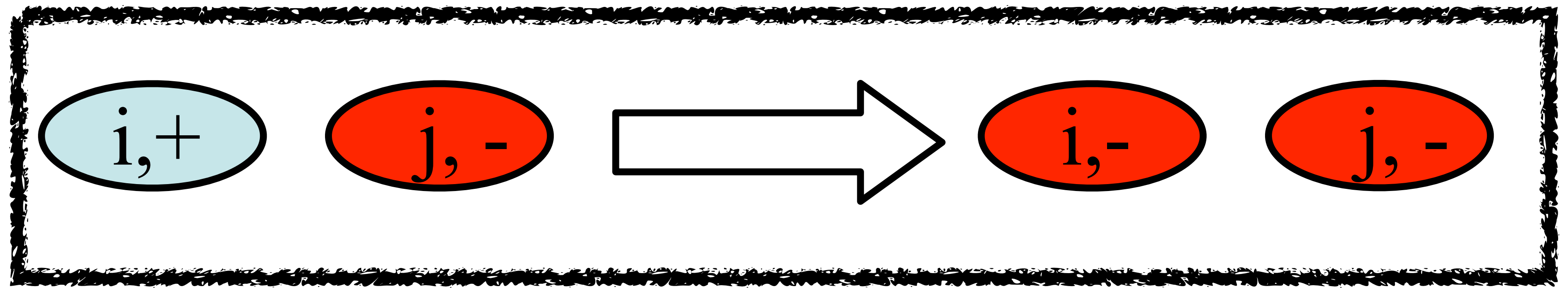
Voter model on a grid

- N agents (k = 1, …, N)
- Opinion: s = +1, -1

At each time step:

- Choose one agent i
- i chooses at random one of his neighbours j
- Agent i adopts the opinion of agent j

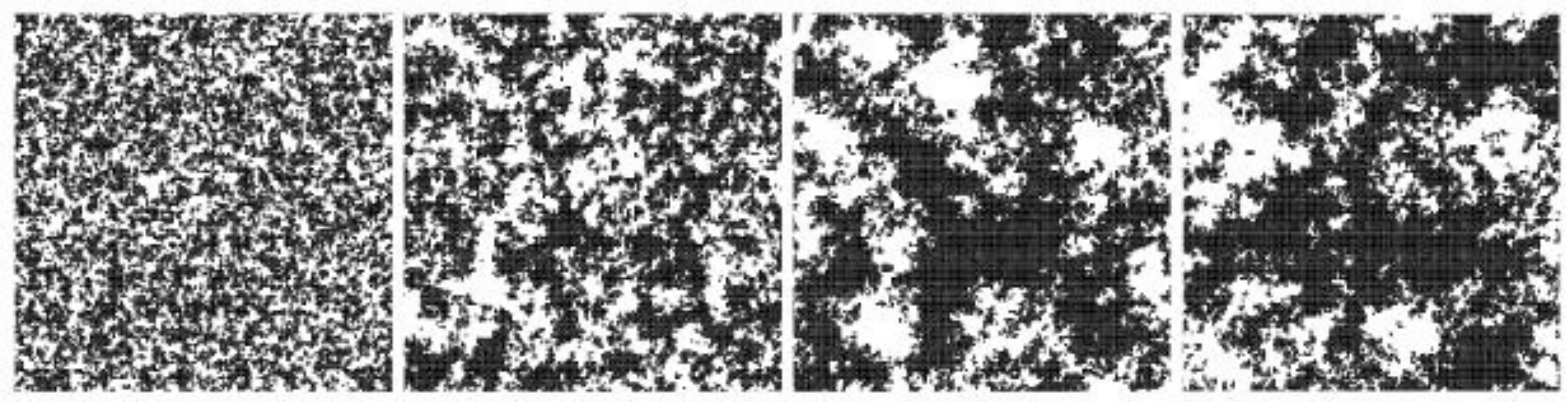




Clifford & Sudbury, A model for spatial conflict (1975)

The voter model

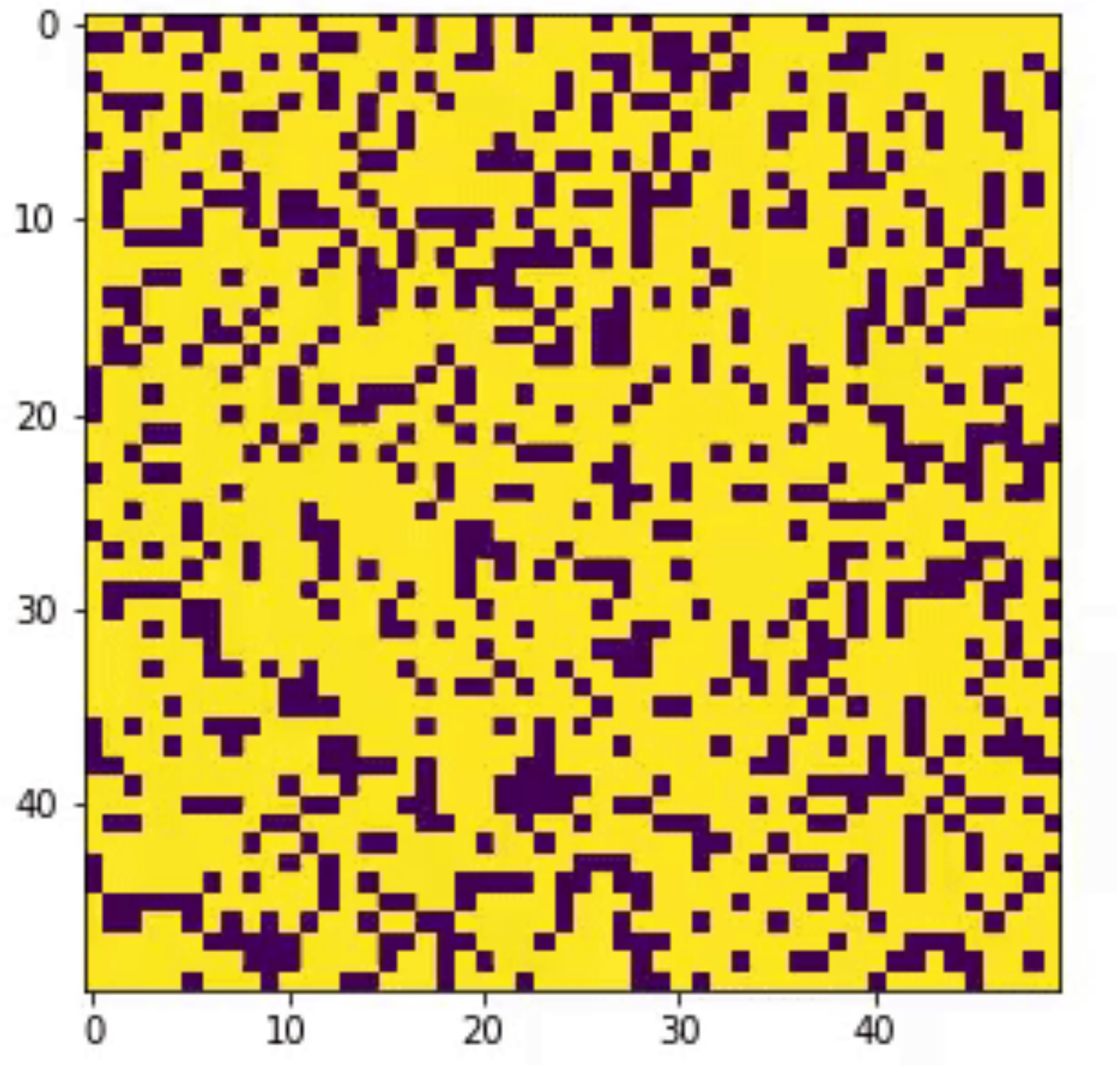
“Coarsening” of domains of similar opinion agents



Questions:

- Convergence to consensus? **Always** for dimensions D=1 and D=2
- How? In how much time?

The voter model



- This is an individual based model that can be "solved" analytically.
- It is possible to show that for dimensionality **D>2**, the system **does not reach consensus** and domains of different opinions coexist indefinitely in time.
- **On (scale-free) networks**, **consensus is achieved quickly**. This rapid consensus is facilitated by "hubs- nodes of the largest degrees that influence their very many neighbours.

Social dynamics models

- Rumor/gossip model
- The voter's model
- **The Axelrod's model**
- The Schelling's model of segregation
- The naming game

The Axelrod's model

Robert Axelrod. The dissemination of culture (1997)

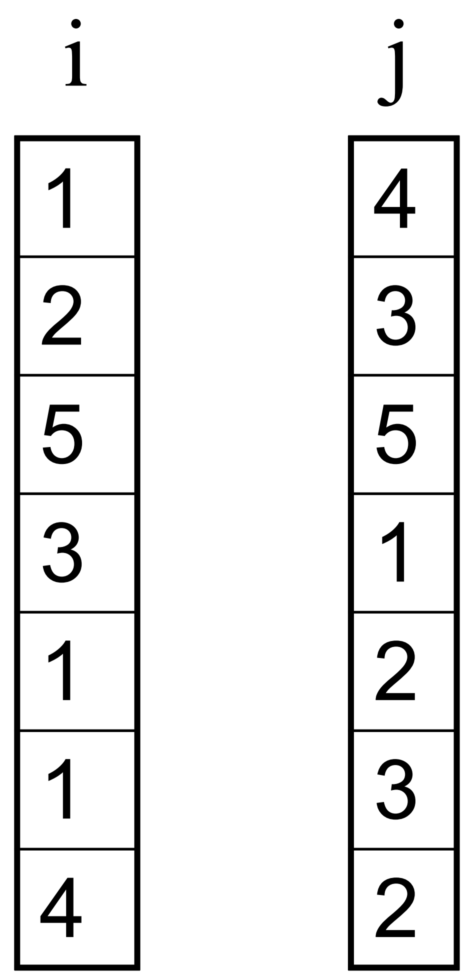
"If people tend to become more alike in their beliefs, attitudes and behaviour when they interact, why do not such differences eventually all disappear?"

Why don't we become all alike?

The Axelrod's model

N agents i=1,..,N on a lattice

- Each agent has **F** attributes
- Each attribute can take **q** values

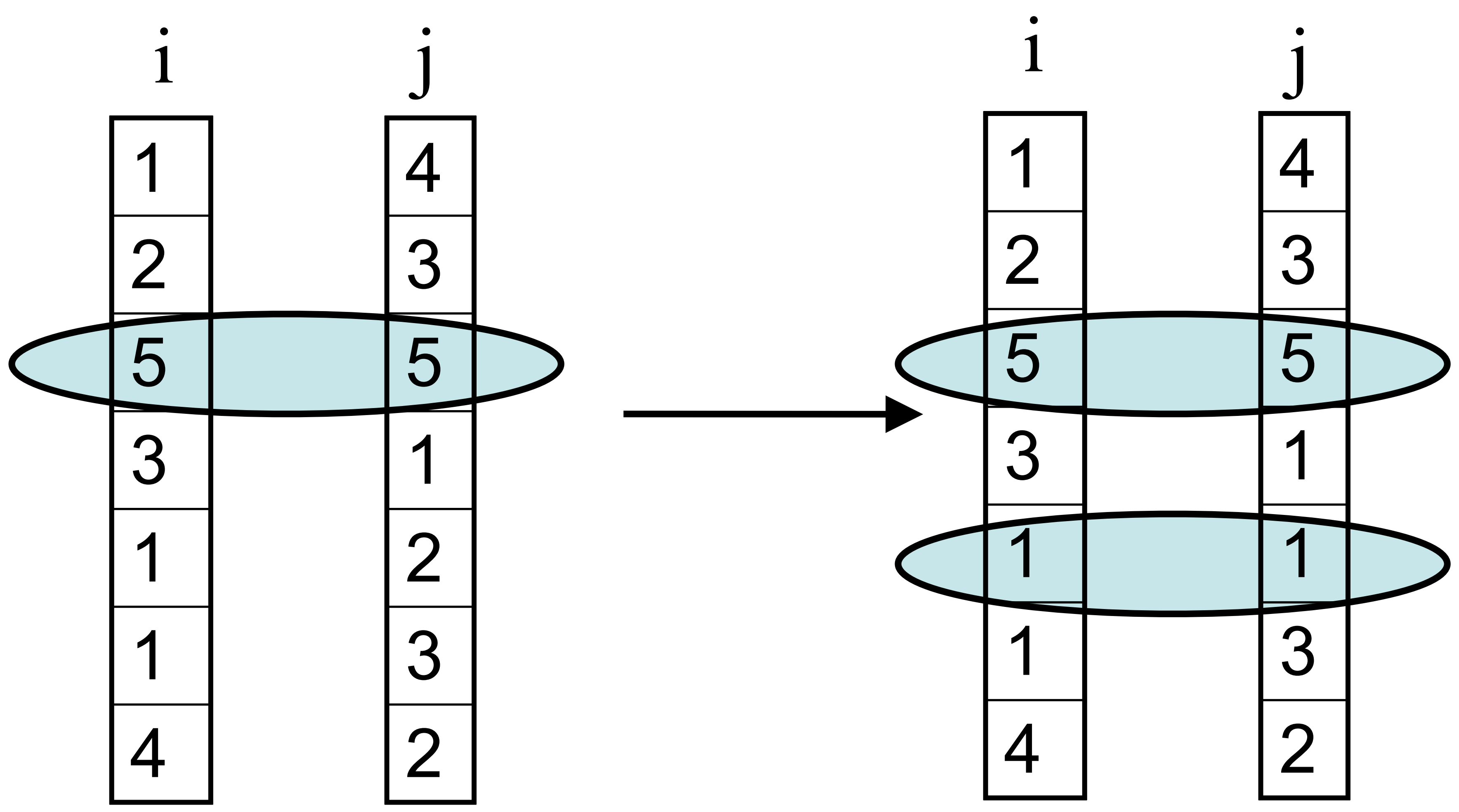


- If i and j have no common attribute:

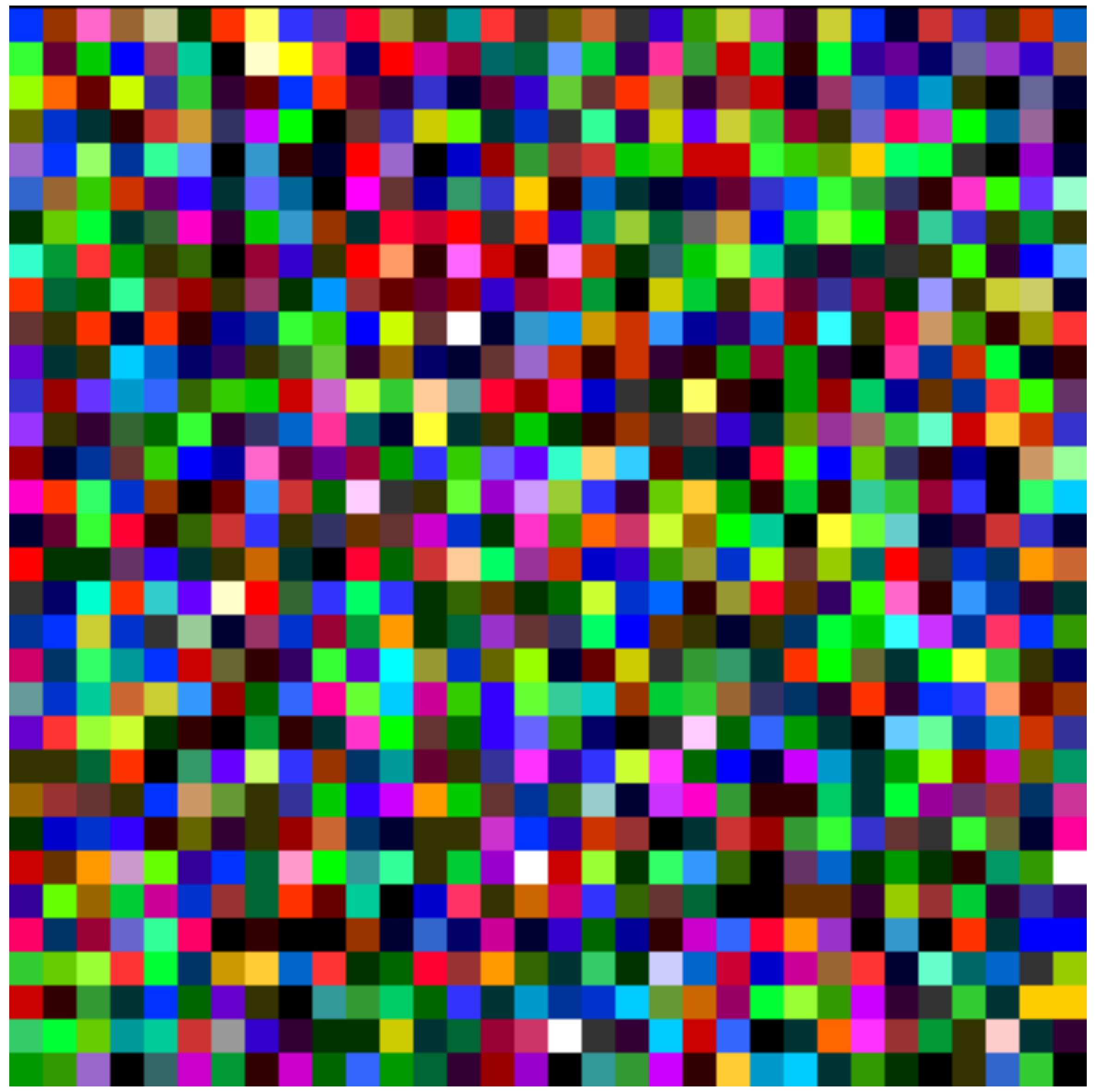
No interaction possible

- If i and j have at least one common attribute:

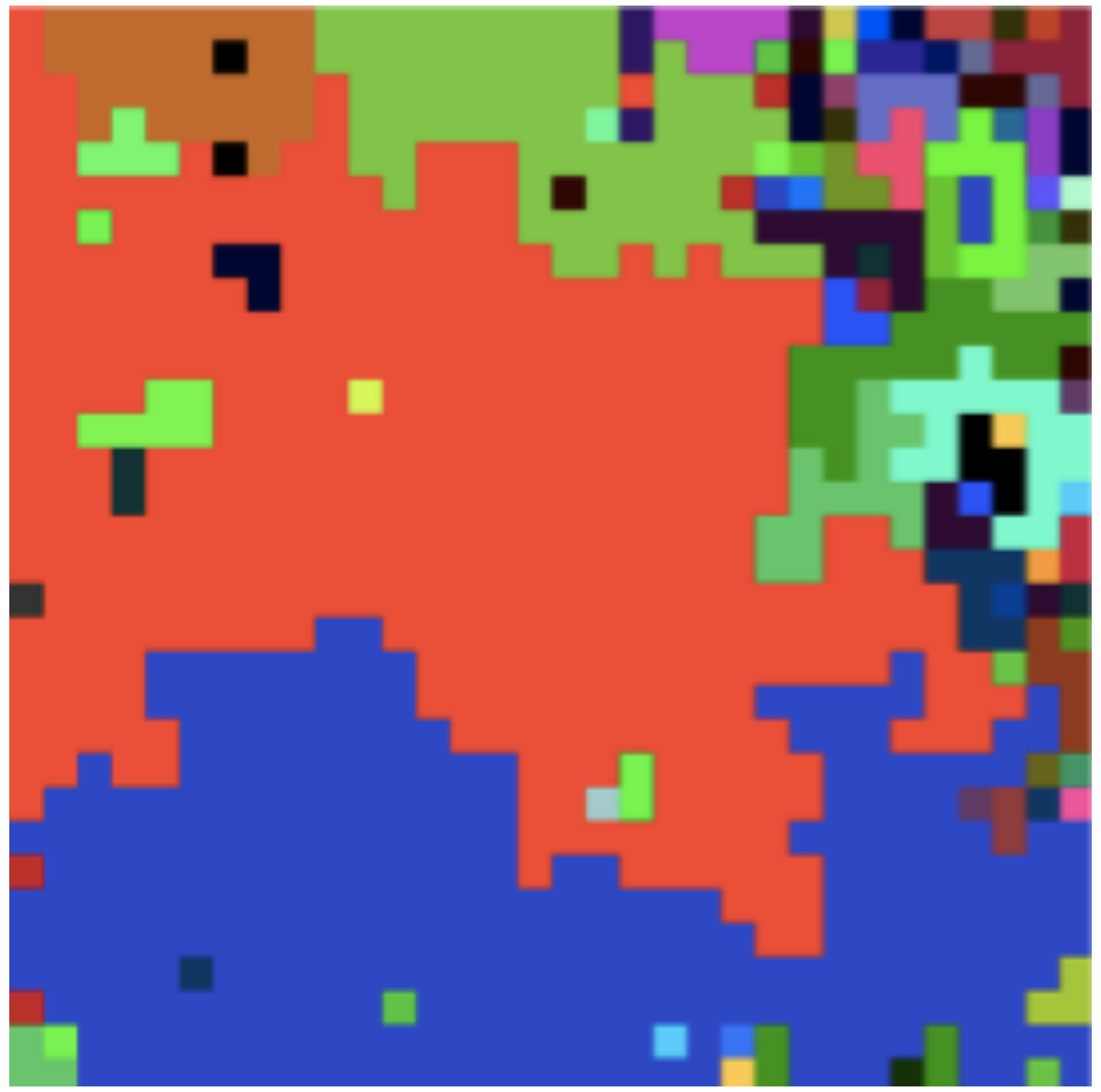
i chooses one of the other attributes and adopt j's value



The Axelrod's model

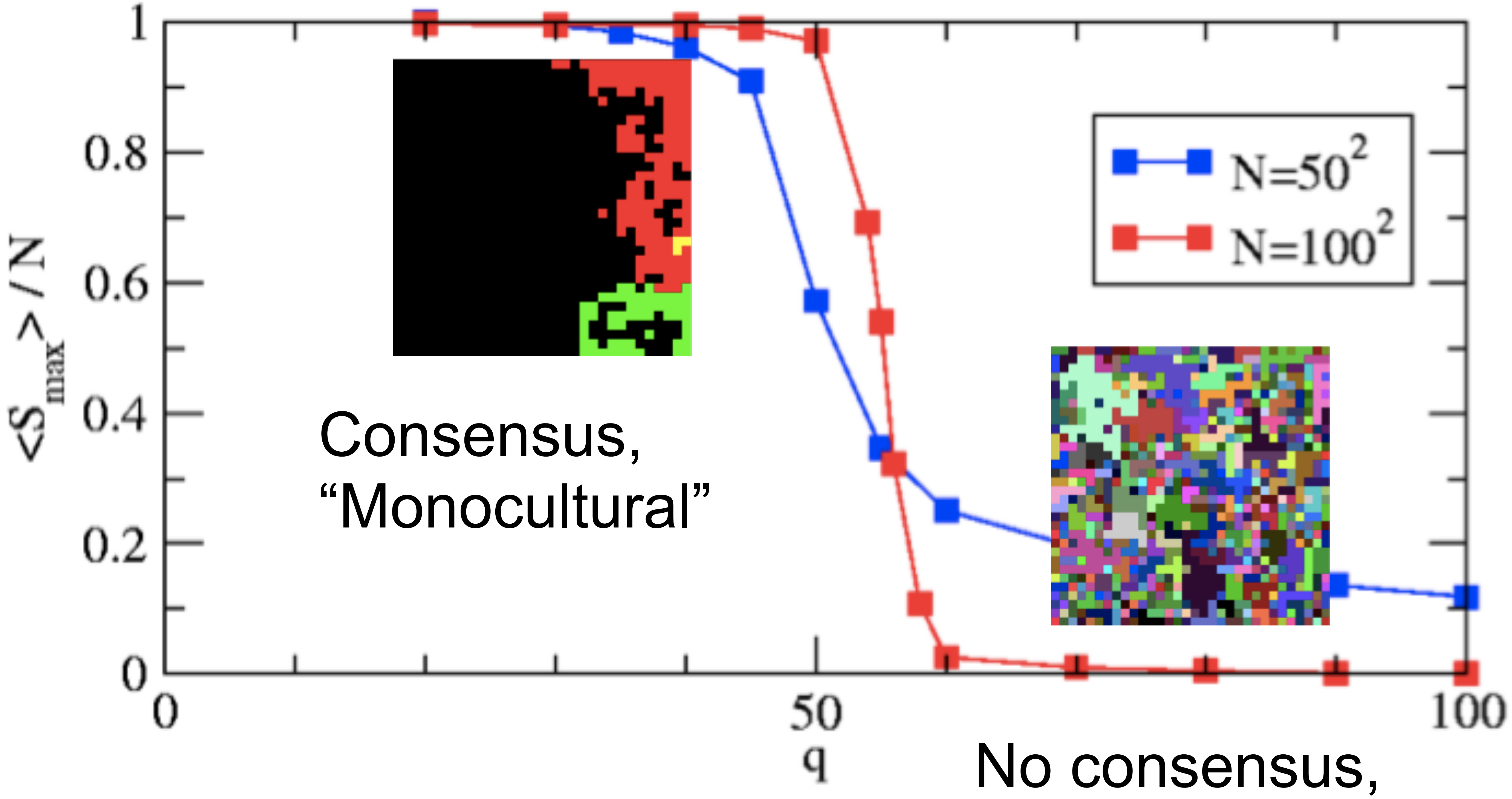


t=0: initial disordered state



Large t: freezing

The Axelrod's model



No consensus,
"Multicultural"

- Transition from consensus to fragmented state as q increases.

Castellano, Marsili, Vespignani, Physical Review Letters 2003

Social dynamics models

- Rumor/gossip model
- The voter's model
- The Axelrod's model
- **The Schelling's model of segregation**
- The naming game

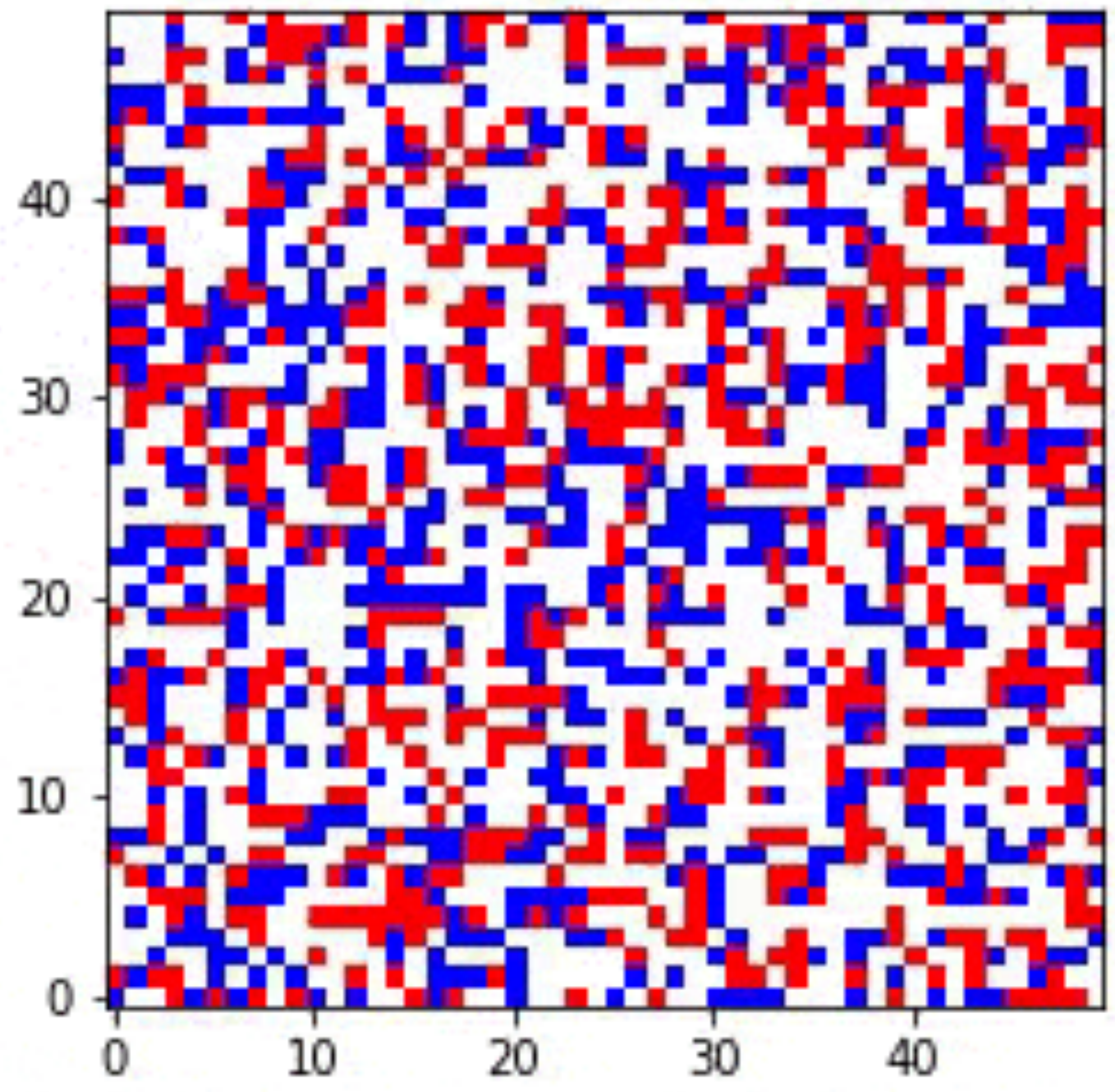
The Schelling's model

Agents are placed on a grid

- N agents (k = 1, …, N)
- Opinion/race: s = +1, -1

At each time step:

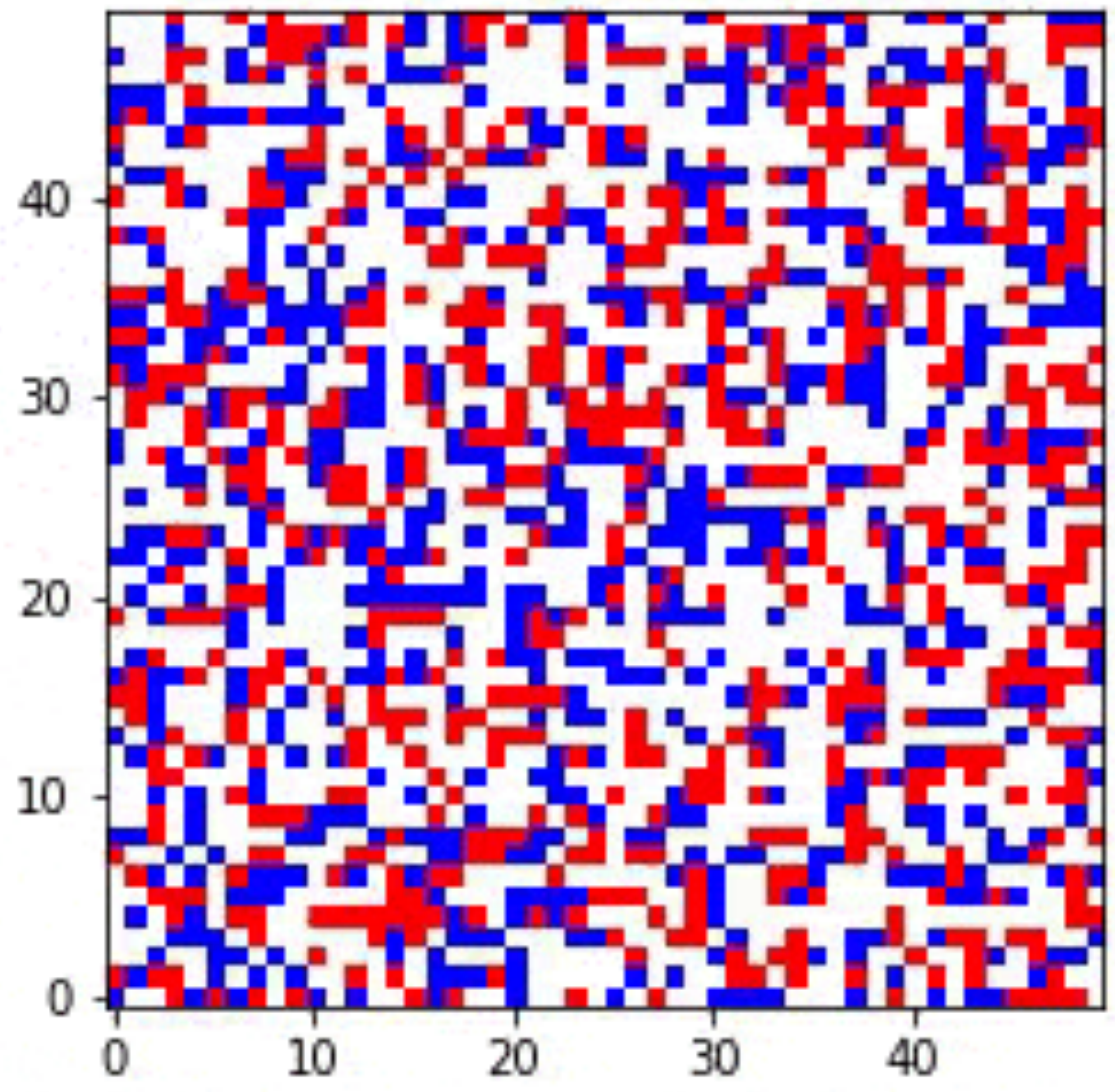
- All agents i check their neighbourhood
- If the fraction of **different** neighbours exceeds a **threshold X**, i will relocate to a different **empty** spot where their preference is met.



Schelling T., Dynamic models of segregation (1971)

The Schelling's model

- Even with **mild preferences** (X = 30%) the grid eventually becomes **highly segregated**.
- This happens without any explicit law or strong biases, **just from individual choices**.
- Segregation can arise naturally only from our innate preference to find similar neighbours.
- Eliminating explicit discriminatory policies may not be sufficient to counteract segregation.



Schelling T., Dynamic models of segregation (1971)

Social dynamics models

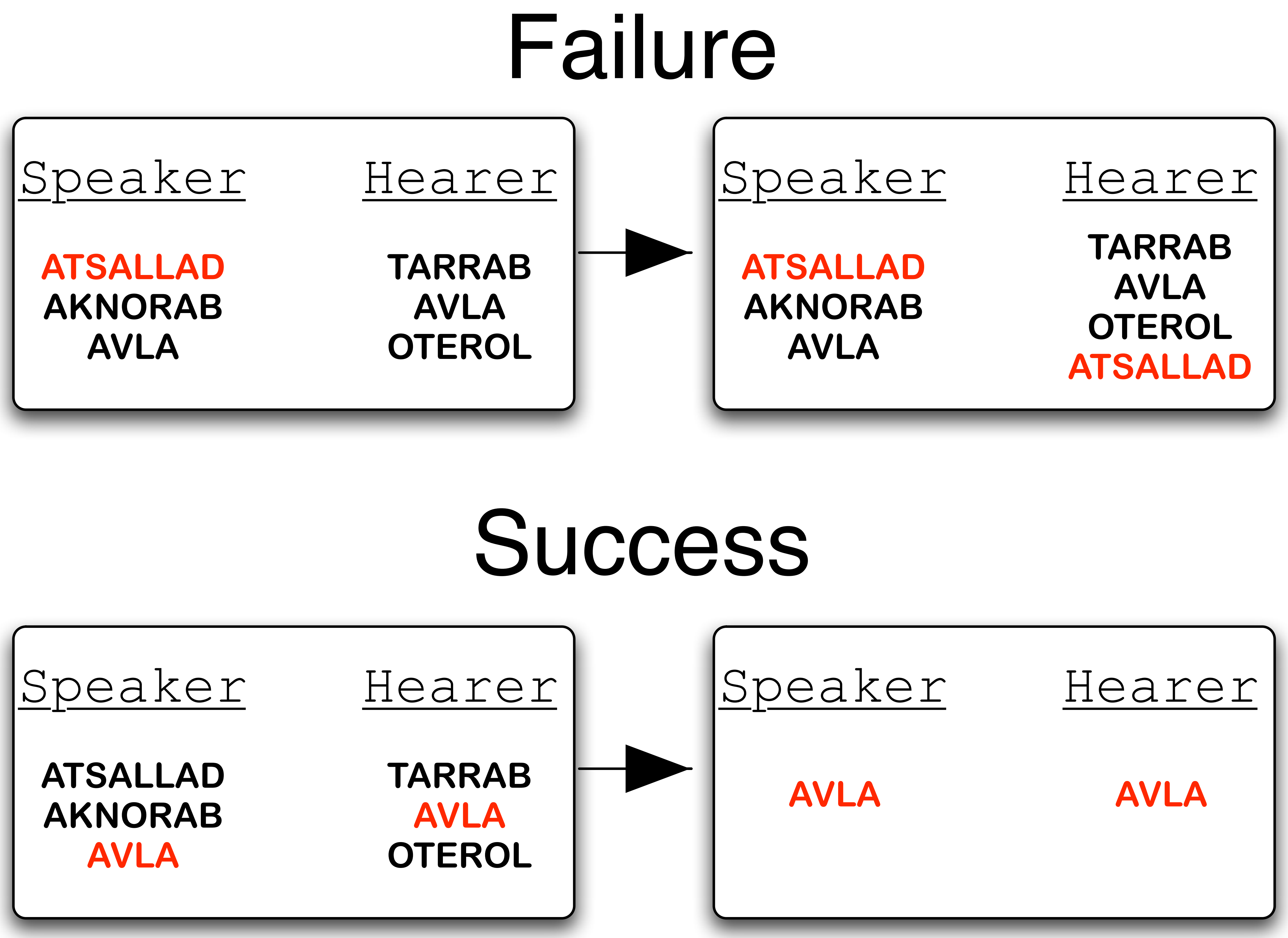
- Rumor/gossip model
- The voter's model
- The Axelrod's model
- The Schelling's model of segregation
- **The naming game**

The naming game

- N agents (k = 1, …, N)
- Each agent starts with an empty dictionary

At each time step:

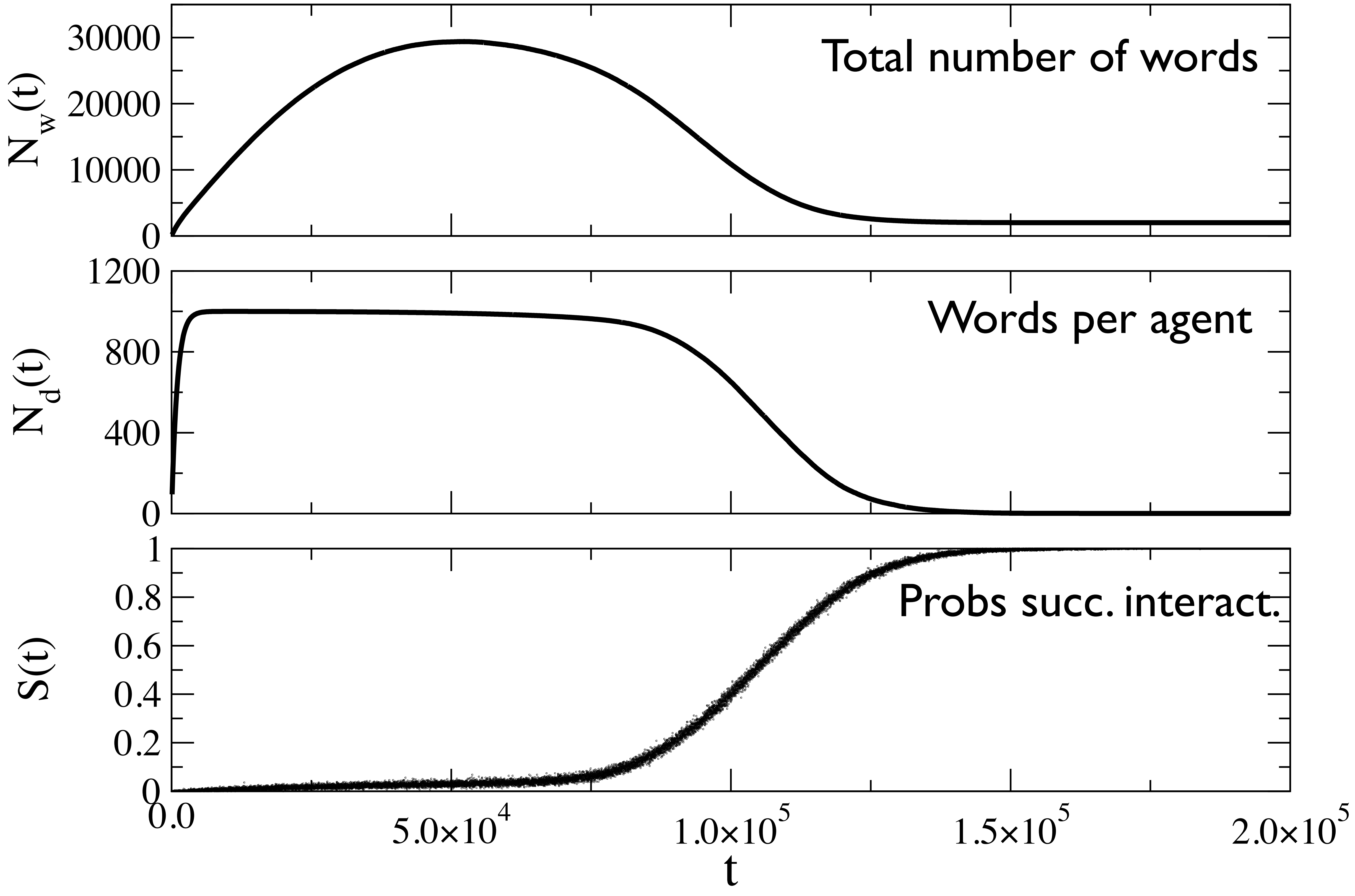
- One agent k (speaker) chooses at random one of her neighbours j (the hearer)
- Speaker randomly selects one of its words and conveys it to the hearer;
- if the hearer's inventory contains such a word, the two agents update their inventories so as to keep only the word involved in the interaction (success);
- otherwise, the hearer adds the word to those already stored in its inventory (failure).



The naming game

- Convergence is reached with a quite abrupt disorder/order transition that starts approximately just after the peak of the total number of words curve has disappeared.

- **Initially, agents hardly understand each others (S(t) is very low); then the inventories start to present significant overlaps, so that S(t) increases until it reaches 1.**



Baronchelli, Andrea. "A gentle introduction to the minimal Naming Game." Belgian Journal of Linguistics 30.1 (2016): 171-192.

The committed minority

SOCIAL SCIENCE

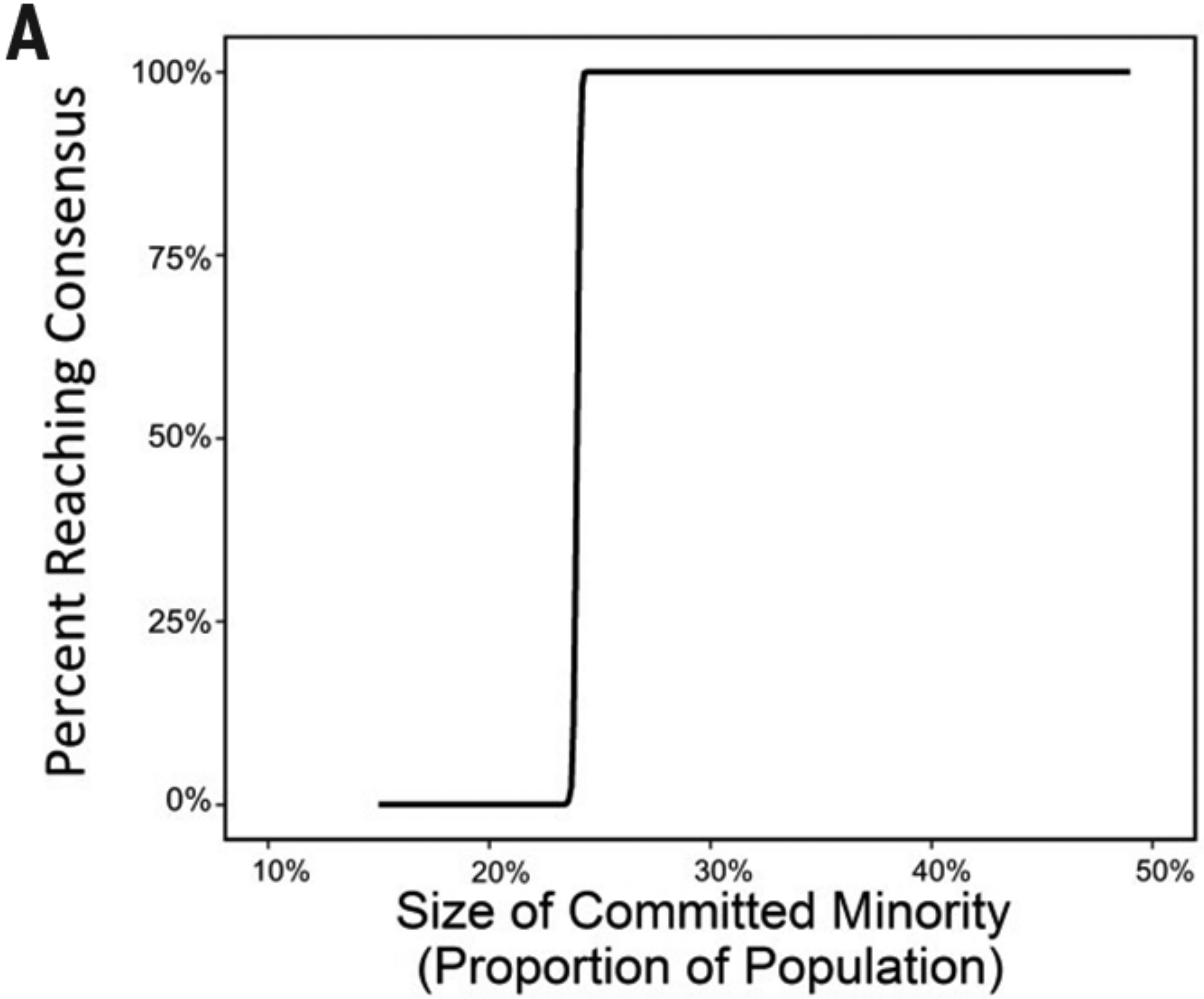
Experimental evidence for tipping points in social convention

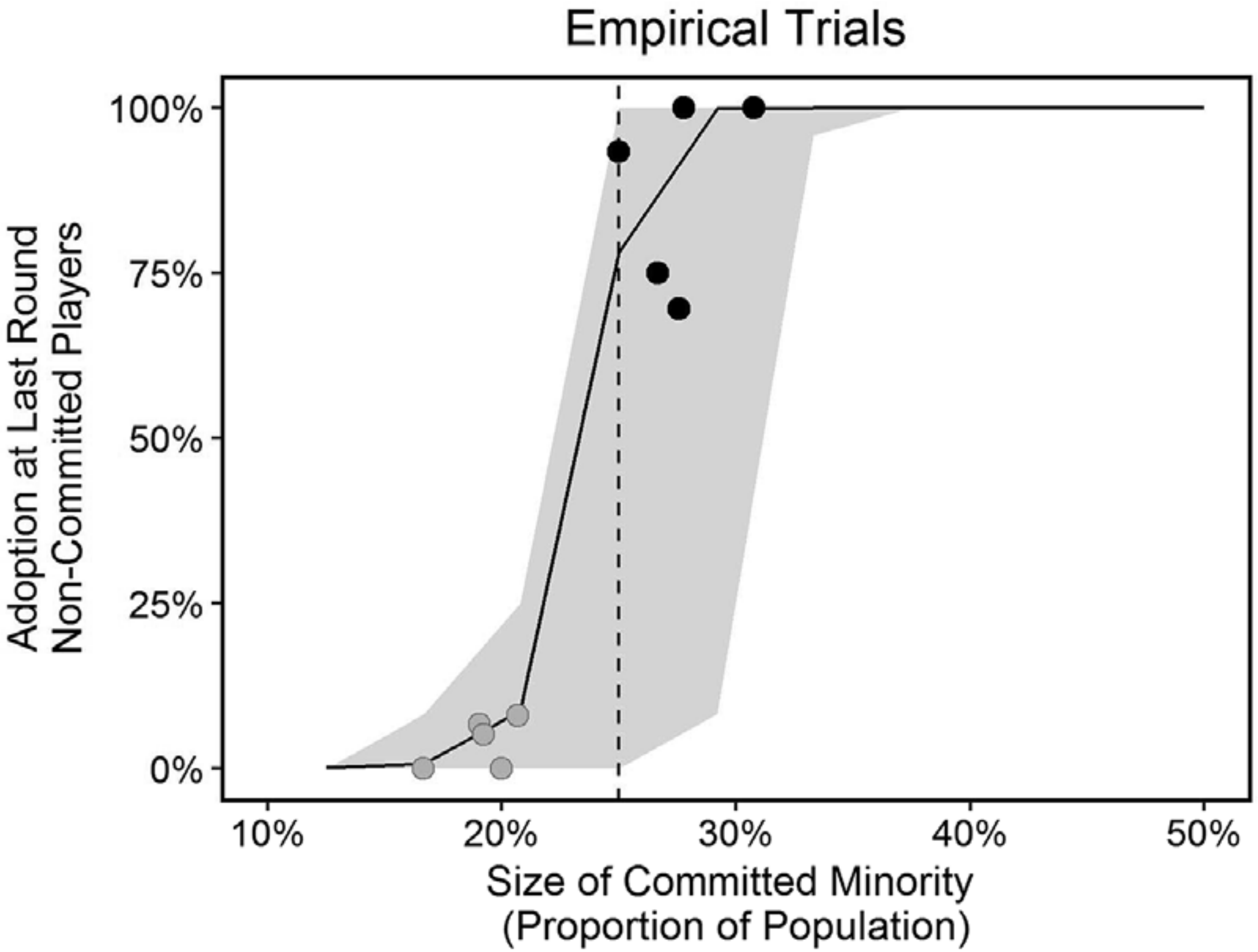
Damon Centola[1,2]*, Joshua Becker[1], Devon Brackbill[1], Andrea Baronchelli[3]

Theoretical models of critical mass have shown how minority groups can initiate social change dynamics in the emergence of new social conventions. Here, we study an artificial system of social conventions in which human subjects interact to establish a new coordination equilibrium. The findings provide direct empirical demonstration of the existence of a tipping point in the dynamics of changing social conventions. When minority groups reached the critical mass—that is, the critical group size for initiating social change—they were consistently able to overturn the established behavior. The size of the required critical mass is expected to vary based on theoretically identifiable features of a social setting. Our results show that the theoretically predicted dynamics of critical mass do in fact emerge as expected within an empirical system of social coordination.

- Real world experiment
- Individuals are assigned to groups and coordinate online to name a given object (face)
- Once consensus is reached, a team of confederates (a committed minority) is introduced in the group with the goal of overturning the consensus.
- Results show that in all cases, the minority manages to change the convention.

The committed minority





Centola, Damon, et al. "Experimental evidence for tipping points in social convention." *Science* 360.6393 (2018): 1116-1119.

The committed minority

Many more examples

- Game theory (Nash equilibrium)
- Threshold models for behaviour adoption (Schelling and Granovetter)
- Models for social networks (small-world, preferential attachment)
- Crowd behaviour (flocking, pedestrians, applause dynamics)
- Formation of hierarchies (Bonabeau model)

Food for thought

- Since the early XX century, many formal models have been introduced in social research.

- **Interdisciplinary approaches**: contribution from math, physics, economics, statistics, etc.

- Until 1990s most of the formal models were **theoretical studies**. Since the early 2000s, **empirical studies** have grown thanks to **large-scale data**.

Food for thought

- Large-scale data have prompted the development of large-scale models (typically individual-based with several parameters).

- **Trade-off between model realism and complexity**: too realistic often means too specific. Too many parameters do not allow a meaningful comparison with empirical data.

- Formal models are **most useful** when they capture **key features** of the system under study with **few assumptions** and parameters.

Suggested readings (a biased list)

- Castellano, Fortunato, Loreto “Statistical physics of social dynamics” Rev. Mod. Phys. 2009
- Mark Buchanan “The social atom”.
- Alex Pentland “Social physics”.