香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

人工智能學院
School of Artificial Intelligence

# Large Language Models for Social Simulations

(AIE1902)

## Unit 2: Introduction to Agent-Based Modeling (ABM) I

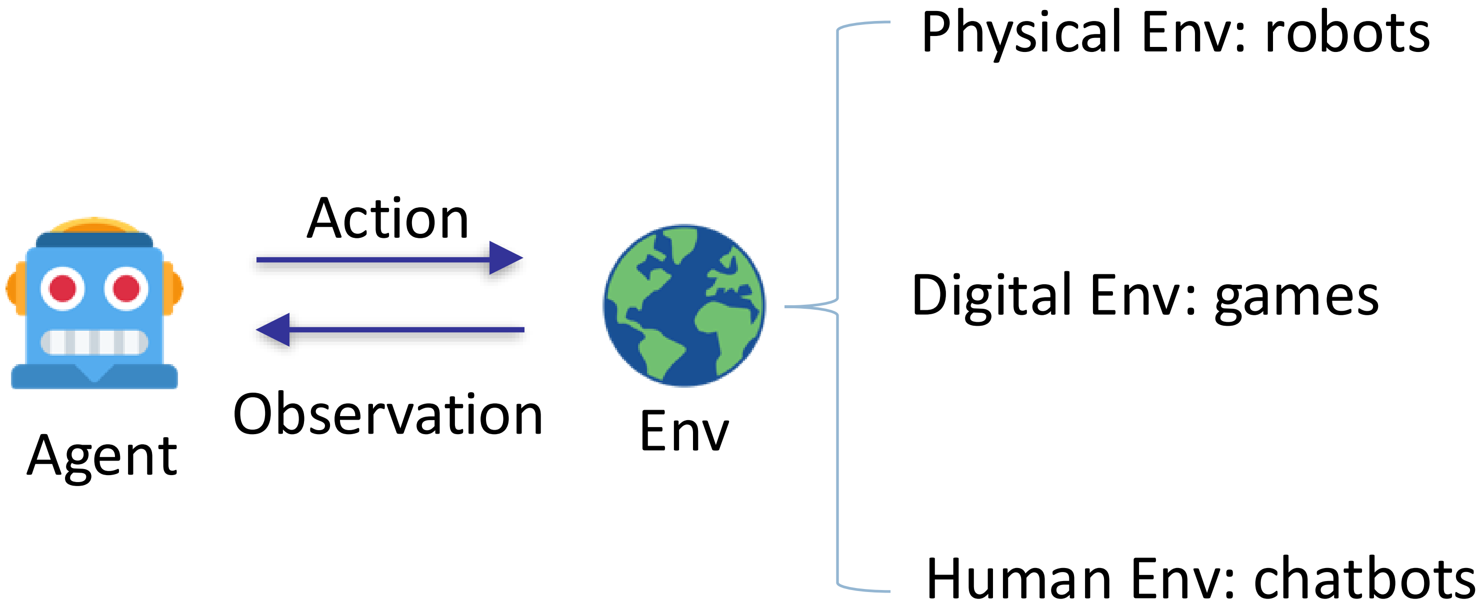

Assistant professor, Zhao Zhanzhan 赵展展 (Instructor)

PhD student, Jia Xiao 贾晓 (Teaching assistant)

School of Artificial Intelligence

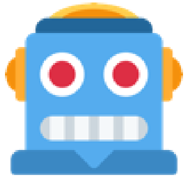
# Part 1: What is ABM？

# What is “agent”?



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
人工智能學院
School of Artificial Intelligence

# What is "agent"?




Action

Observation


- Physical Env: robots

- Digital Env: games

- Human Env: chatbots

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

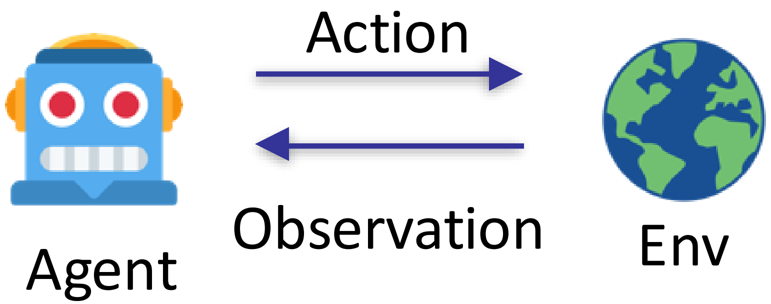
人工智能學院
School of Artificial Intelligence

# What is “agent”?

Memory

Computation/
think

Actions



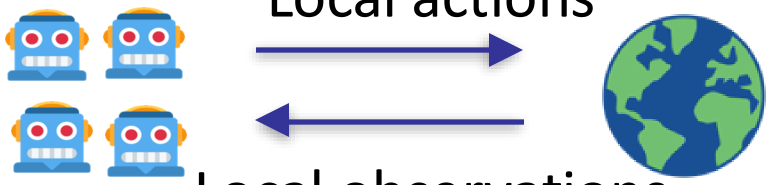

Physical Env: robots

Digital Env: games

Human Env: chatbots

# What is a “multi-agent system”?

Local actions



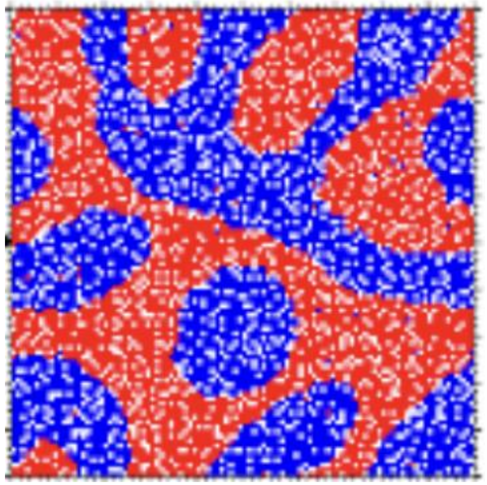
Local observations

Multi-agents

Env

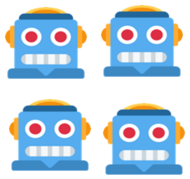
Physical Env:
Swarm robotics

Digital Env:
Agent-based Modeling (ABM)
Complex systems
Game theory
Multi-agent RL
Comp social science



Human Env:
Sociology
psychology

# What is a "multi-agent system"?



Actions

Observations

Env

Multi-agents

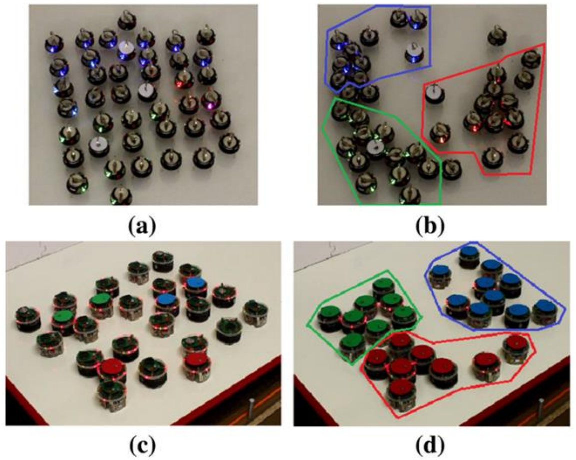
Physical Env:
Swarm robotics

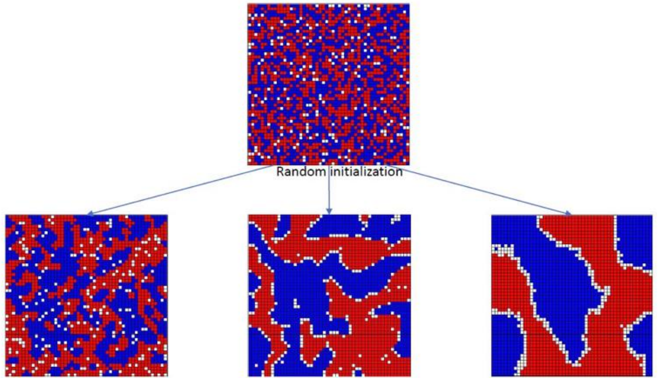
Model

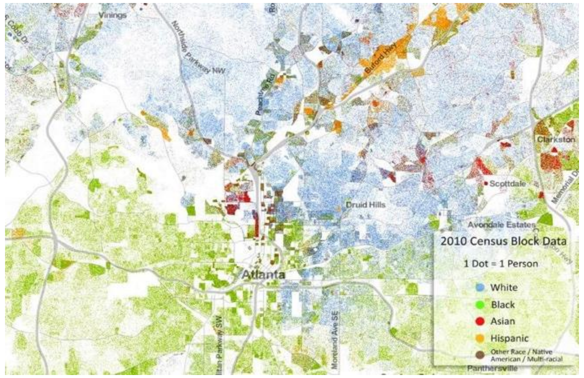
Digital Env:
Agent-based Modeling (ABM)
Complex systems
Game theory
Multi-agent RL
Comp social science

Inform / Explain

Human Env:
Sociology
psychology



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
人工智能學院
School of Artificial Intelligence

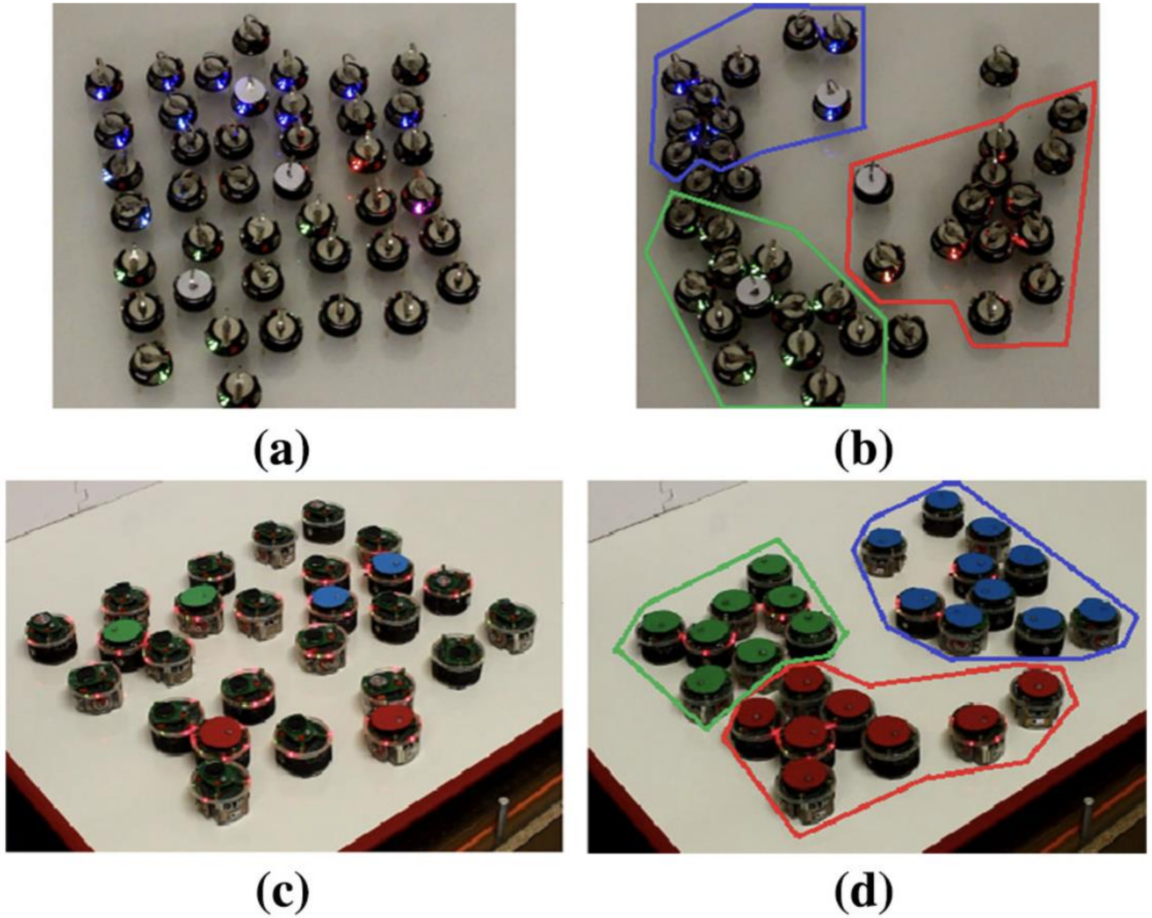
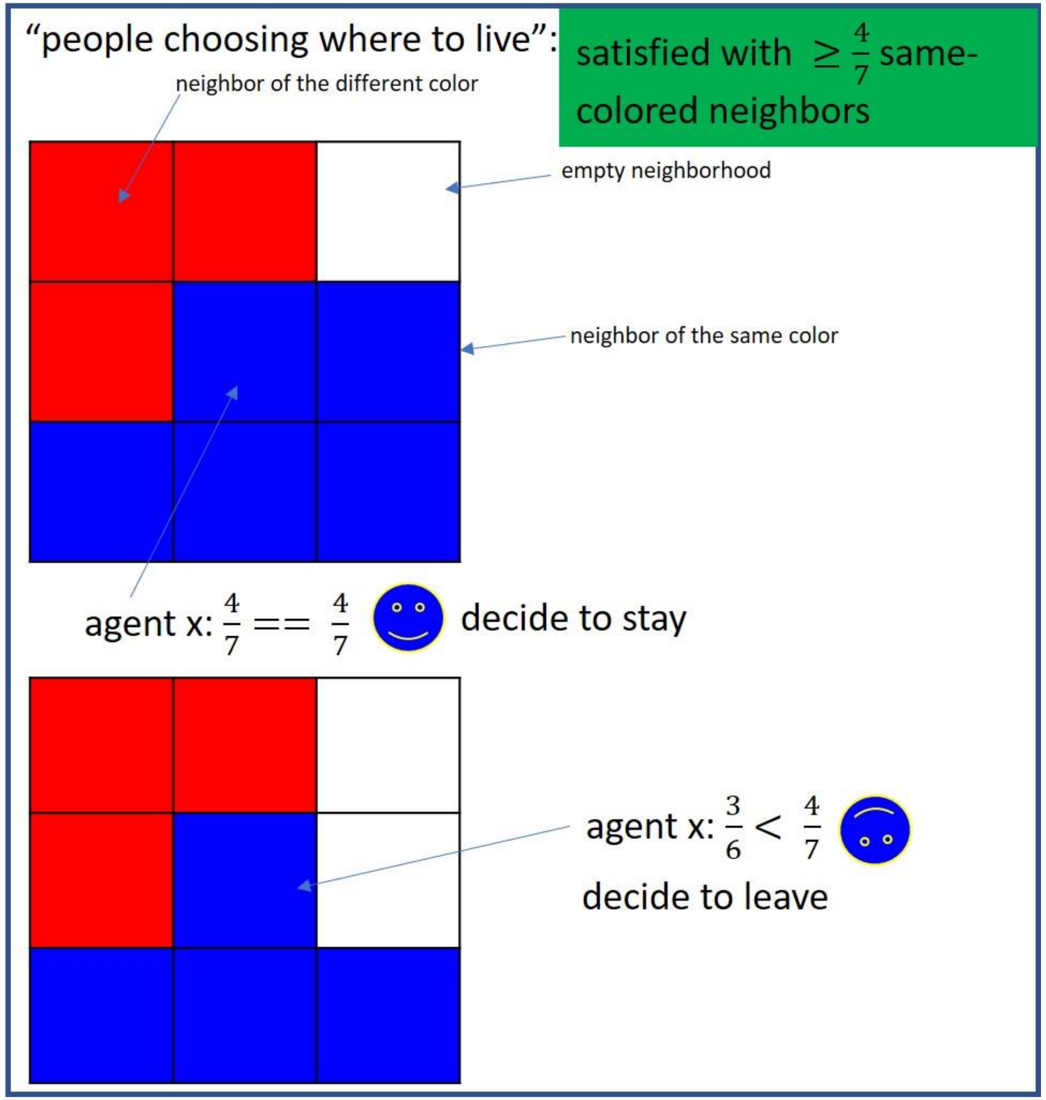
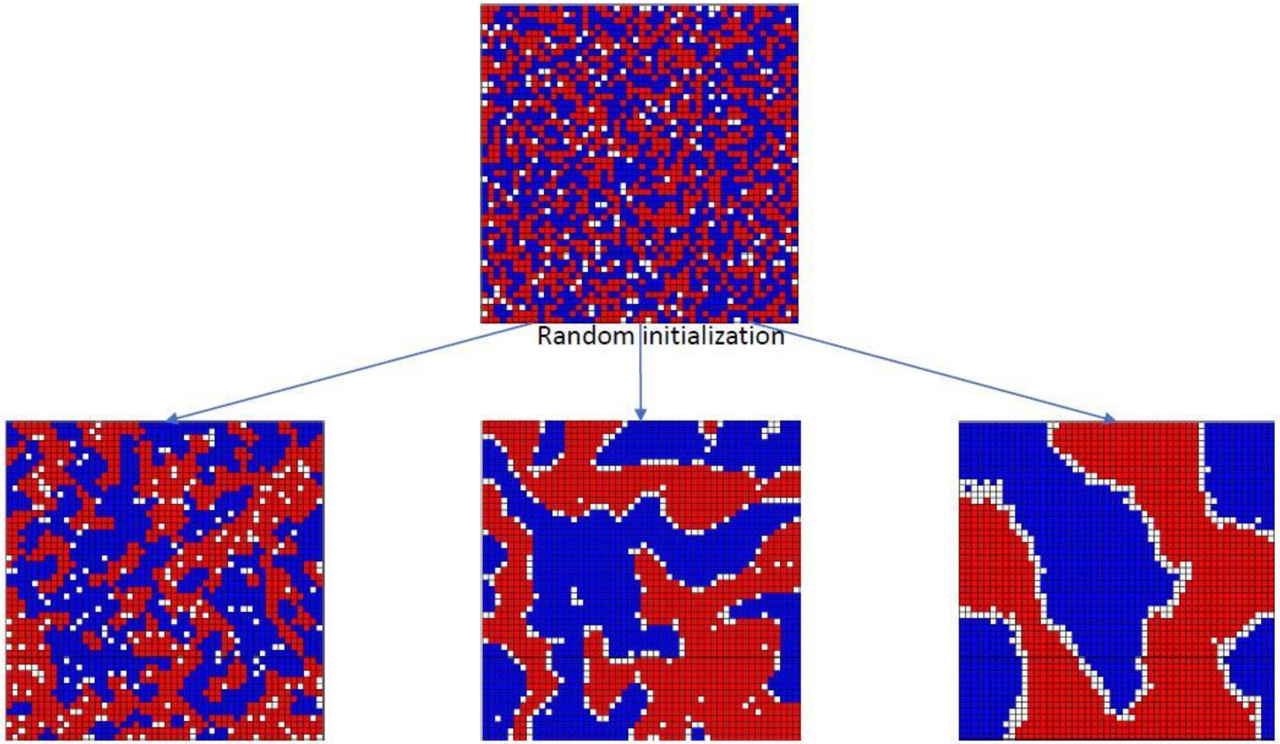


Fig. 15 Snapshots from a segregation trial with Kilobots: **a** initial grid formation with three leaders, marked with white tags; **b** result after segregation occurred. Trial with e-pucks: **c** initial grid formation with three pairs of leaders marked with tags; **d** result after segregation occurred, tags were added after the experiments for visualisation

https://link.springer.com/article/10.1007/s11721-016-0119-0

# Schelling Model




*Being satisfied with 30%* or more same-colored neighbors, an average agent *ends up in the neighborhood with 72%* same-colored neighbors.

*Being satisfied with 50%* or more same-colored neighbors, an average *agent ends up in the neighborhood with 93%* same-colored neighbors.

*Being satisfied with 75%* or more same-colored neighbors, an average *agent ends up in the neighborhood with 100%* same-colored neighbors.

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen
人工智能學院 School of Artificial Intelligence

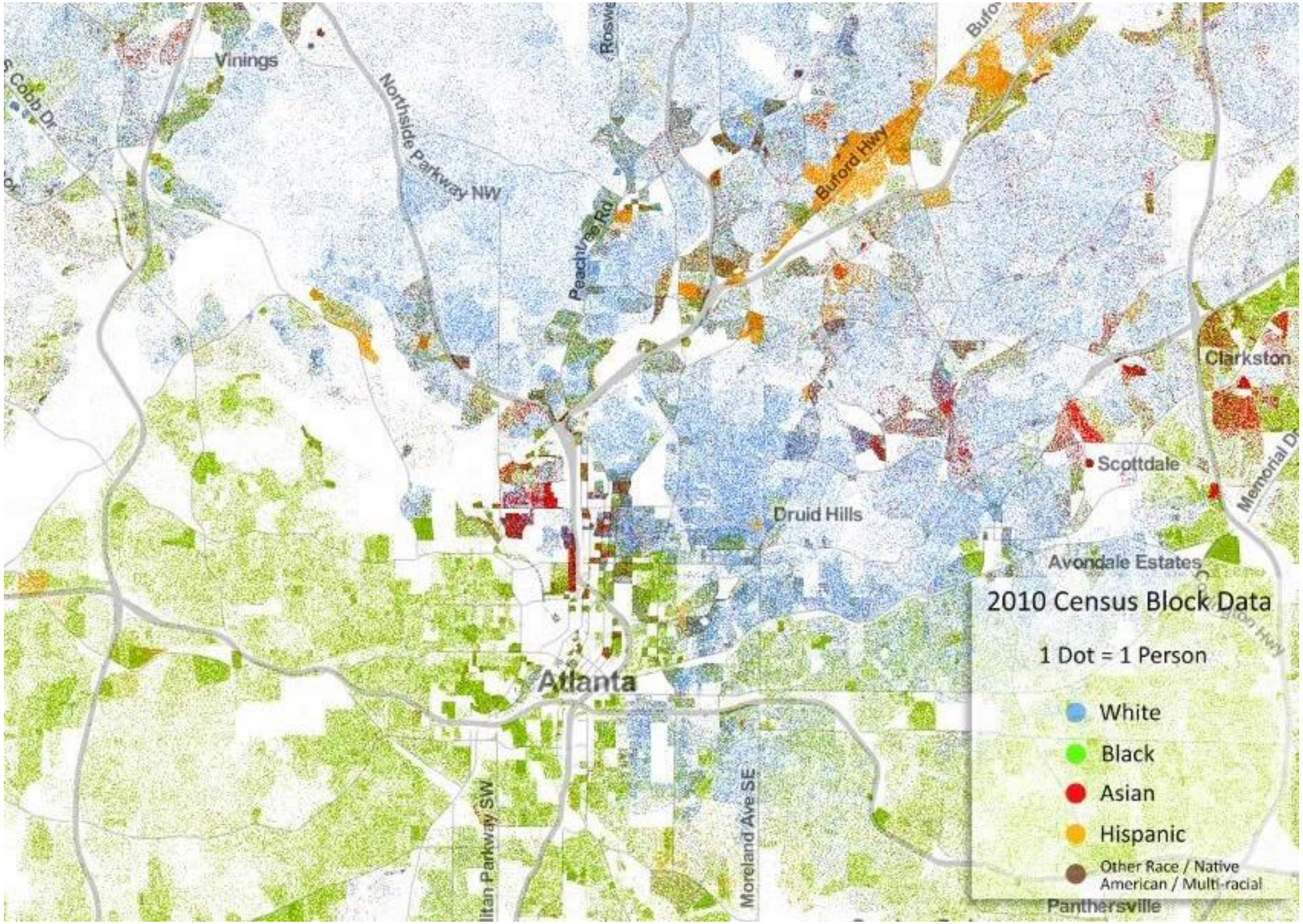
# Thomas Schelling (1921–2016)

- American economist and political scientist
- 2005 Nobel Prize in Economics
- Known for:

  **Micromotives and Macrobehavior** (1978)

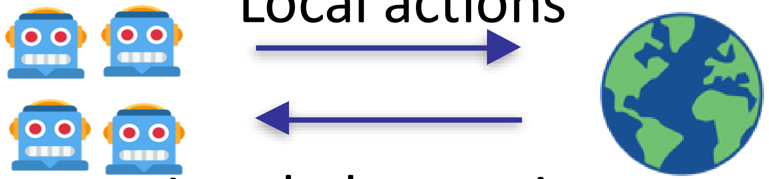
  Early models of **segregation dynamics**

  Illustrating how **simple individual rules** can lead to **complex social patterns**


Vinings
S Cobb Dr
Northside Parkway NW
Roswe
Peachtree Rd
Bufo
Buford Hwy
Clarkston
Scottdale
Druid Hills
Avondale Estates
Memorial Dr
Atlanta
Moreland Ave SE
Panthersville
2010 Census Block Data
1 Dot = 1 Person
White
Black
Asian
Hispanic
Other Race / Native American / Multi-racial

# What is a “multi-agent system”?

Local actions



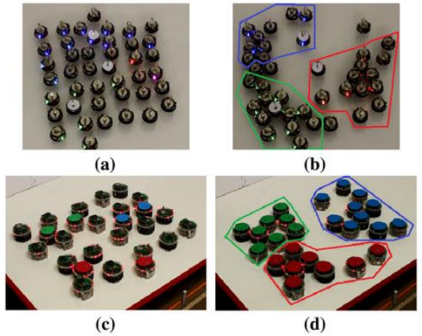
Local observations

Multi-agents

Env

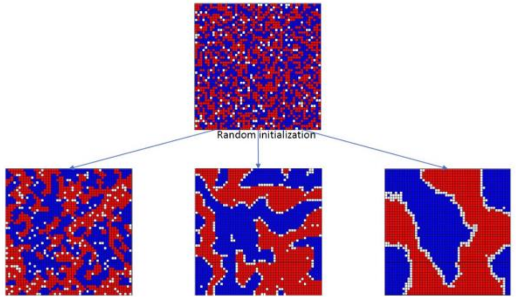
Physical Env:
Swarm robotics

More controllable agents
less complex




Model

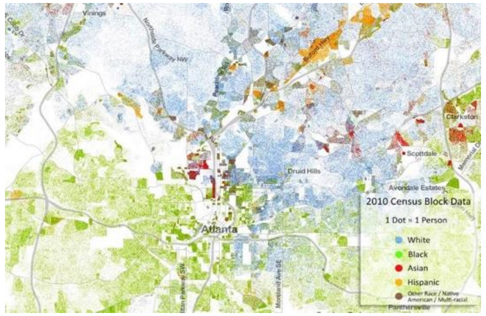
Digital Env:
Agent-based Modeling (ABM)
Complex systems
Game theory
Multi-agent RL
Comp social science




Inform / Explain

Human Env:
Sociology
psychology

Less controllable agents
More complex



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

人工智能學院
School of Artificial Intelligence

Individual
Shared
Memory

Decentralized
Centralized
Computation/
think

Actions

Communication
(a network)

Synchronized
Asynchronized
Coordination

Roles
(heterogeneity / diversity)

Local actions

Local observations

Multi-agents

Env

Physical Env:
Swarm robotics

Model

Digital Env:
Agent-based Modeling
Complex systems
Game theory
Multi-agent RL
Comp social science

Inform / Explain

Human Env:
Sociology
psychology

Individual
Shared
Memory

Decentralized
Centralized
Computation/ think

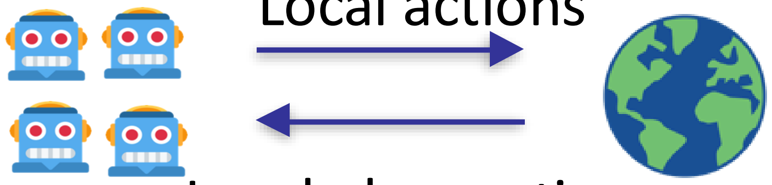
Actions

Communication (a network)

Synchronized
Asynchronized
Coordination

Roles (heterogeneity / diversity)

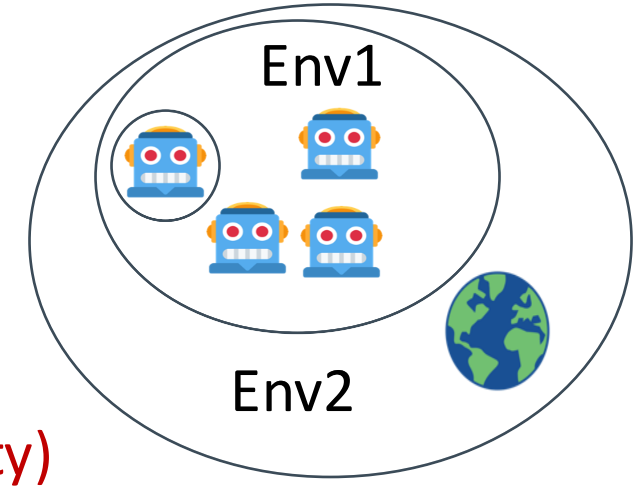
Local actions



Local observations

Multi-agents

Env




Physical Env:
Swarm robotics

Model

Digital Env:
Agent-based Modeling
Complex systems
Game theory
Multi-agent RL
Comp social science

Inform / Explain

Human Env:
Sociology
psychology

# What is multi-agent system's Emergence??



Individual

Memory

Decentralized

Computation/ think

Tool use

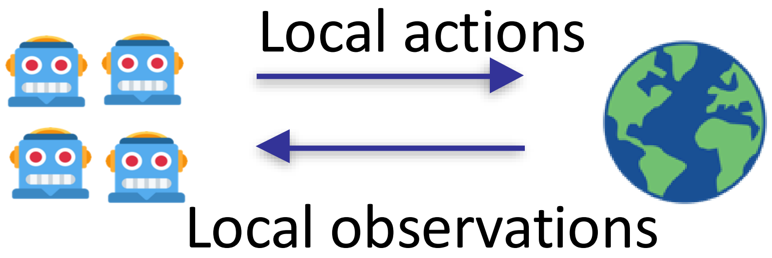
Synchronized

Asynchronized

Communication (a network)

Coordination

Roles (heterogeneity / diversity)



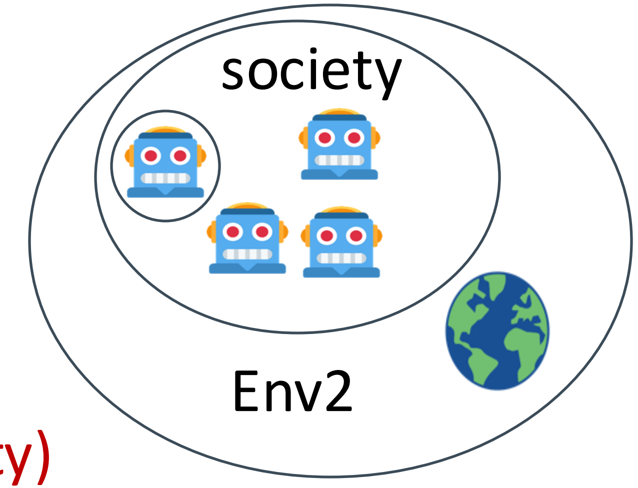

Multi-agents

Env

Physical Env:

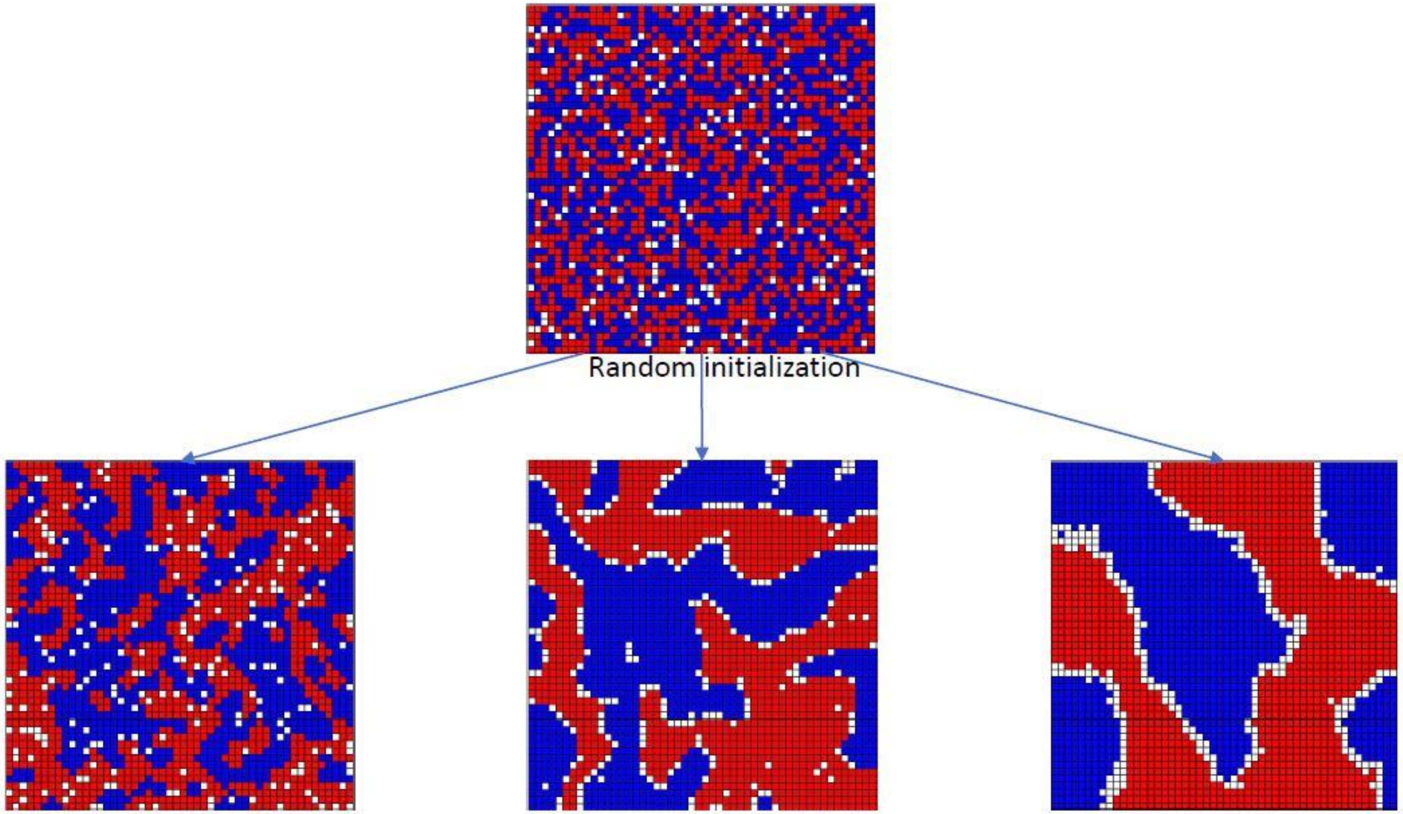
Digital Env:

Human Env:




Emergence: A system-level surprise happened.

**Emergence** refers to a phenomenon where **complex patterns, behaviors, or properties arise from simple interactions among individual components** of a system, without any central control or explicit design. The whole system exhibits properties that **cannot be easily predicted** just by looking at its individual parts.




*Being satisfied with 30%* or more same-colored neighbors, an average agent *ends up in the neighborhood with 72%* same-colored neighbors.

*Being satisfied with 50%* or more same-colored neighbors, an average *agent ends up in the neighborhood with 93%* same-colored neighbors.

*Being satisfied with 75%* or more same-colored neighbors, an average *agent ends up in the neighborhood with 100%* same-colored neighbors.

# Key Characteristics of Emergence

- Bottom-Up Formation:
  Global patterns emerge from **local interactions** among simple units.
  There is **no central controller**—the system self-organizes.

- Unpredictability:
  The outcome is often **unexpected or more than the sum of its parts**.
  The system's behavior **cannot be easily reduced** to the properties of individual components.

- Self-Organization:
  The system **adjusts dynamically** without external intervention.
  Adaptation and evolution often play a role.

- Multi-Level Effects:
  Emergent behavior is usually **observed at a higher level** than the interactions of individual components.

# Multi-agent System Coordination

**Synchronized vs. Asynchronous Execution:**

•**Synchronized:** Agents take turns or follow a predefined execution schedule.

•**Asynchronous:** Agents act independently, without waiting for others, leading to more realistic and dynamic simulations.

Eg. To mimic this uniformly random activation sequence in a local way, we assume each particle has its own Poisson clock with mean 1 and activates after a delay $t$ drawn with probability $e^{-t}$. After completing its activation, a new delay is drawn to its next activation, and so on.

| Year | Agent-Based Modeling | Game Theory | Multi-Agent RL (MARL) |
|---|---|---|---|
| 1940s | Cellular Automata (Von Neumann & Ulam) | Minimax theorem (Von Neumann, 1928) | - |
| 1950s | - | Nash Equilibrium (Nash, 1950) | - |
| 1970s | Schelling’s Segregation Model (1971) | Evolutionary Game Theory (Maynard Smith, 1973) | - |
| 1980s | Boids flocking model (Reynolds, 1986) | Iterated Prisoner’s Dilemma (Axelrod, 1984) | Markov Games (Littman, 1994, later in the 90s) |
| 1990s | Sugarscape Model (Epstein & Axtell, 1996) | Expanding economic & social science applications | Independent Q-learning for Multi-Agent Systems (Tan, 1993) |
| 2000s | Large-scale agent simulations (e.g., traffic, economics) | Game Theory applied in AI, auctions, and mechanism design | Cooperative MARL (Claus & Boutilier, 1998), Nash Q-learning (Hu & Wellman, 2003) |
| 2010s | AI-driven ABMs, integration with ML | AI + Game Theory (AlphaGo, 2016) | MADDPG (Multi-Agent Deep Deterministic Policy Gradient, Lowe et al., 2017), Self-Play in AlphaZero (2017) |
| 2020s | Large-scale AI-driven simulations (e.g., social simulations) | Game-Theoretic AI (AlphaZero, OpenAI Five) | MARL for real-world applications (e.g., traffic control, robotics, LLM multi-agent systems) |

Agent-based Modeling (ABM): rule-based individual modeling

Complex systems: rule-based individual modeling and focus on emergence

Game theory: rational individual assumption and focus on strategic interactions

Multi-agent RL: rational individual assumption and focus on learning the best strategic interactions

LLM-based Multi-agent System: human-like, generality, the closest to human reactions than any rules or strategic interactions.

# How to Build an ABM？

- **Problem Definition**
  - Clearly identify the research question or phenomenon to simulate
- **Agent Design**
  - Define agent
  - Specify behavior rules (decision-making, interaction protocols)
- **Environment Design**
  - Determine env
  - Define how agents perceive and interact with the environment
- **Simulation Execution**
  - Set parameters and initial conditions
  - Run simulations with varying scenarios
- **Analysis & Validation**
  - Observe emergent patterns at the macro level
  - Compare simulation results with real-world data or theory

# Part 2:
# Competition Between Collective and Individual Dynamics

**“Sometimes, everyone chasing their own happiness can make the whole society worse off.”**

—today, we’ll see exactly why, and when, that happens.

# Individual Actions vs. Collective Outcomes – Why It Matters?

Different academic fields take very different approaches to connecting the **micro** (individual) and **macro** (collective) levels:

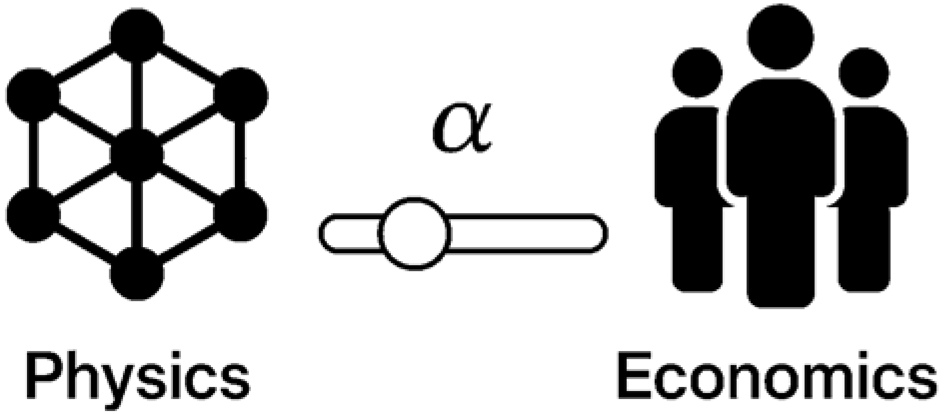
- **Physics:** Focuses on *global optimization* – systems move toward the state that minimizes total free energy.
- **Economics:** Focuses on *individual optimization* – each agent seeks to maximize their own utility, leading to a **Nash equilibrium**.

# Individual Actions vs. Collective Outcomes – Why It Matters?

These approaches can produce **radically different results**:

- Global optimization → highest overall "happiness" or efficiency.
- Individual optimization → everyone looks after their own interest, but the collective outcome can be far from ideal.

This paper builds a simple, adjustable model with a cooperation **parameter** $\alpha$, allowing us to move smoothly between the two perspectives and study the differences in outcomes.



# A Simplified City Model

The “city” is divided into $Q$ blocks, each containing $H$ housing units (cells).

Each housing unit can hold at most **one resident**.

The **density** of block $q$ is: $\rho_q = \frac{n_q}{H}$

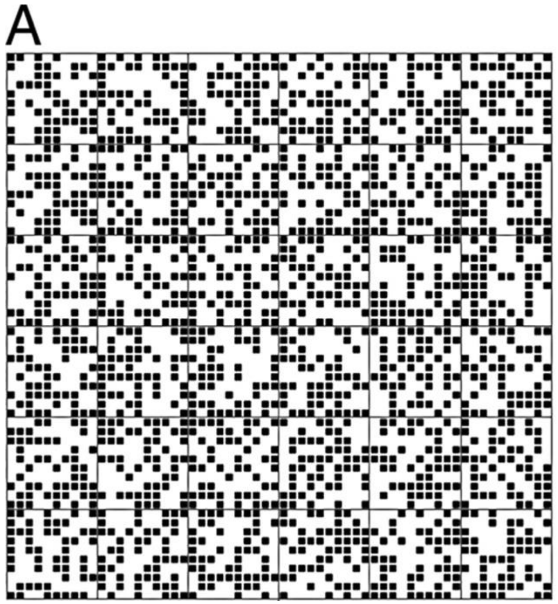
where $n_q$, is the number of residents in block $q$.

All residents share the **same utility function** $u(\rho_q)$,

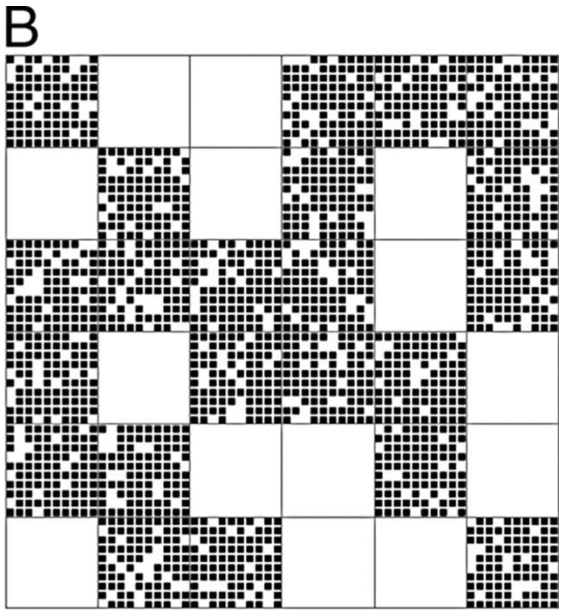
which measures how satisfied they are with the density of their block.

**Global utility** is the total satisfaction of all residents: $U = H \sum_q \rho_q \, u(\rho_q)$

This setup lets us connect the *micro-level* (individual preferences) with the *macro-level* (overall city pattern).



Mixed state – residents evenly distributed across blocks.



Segregated state – some blocks overcrowded, others empty.

# Move or Stay? Considering Myself and Others

At each step:

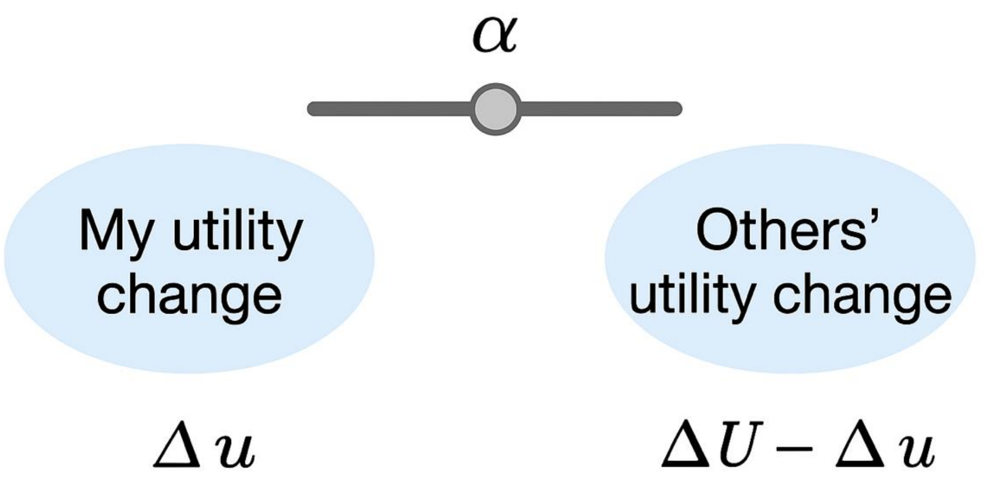
1. Randomly pick **one resident** and **one vacant unit**.
2. Calculate the “gain” G from moving.

Formula for gain: $G = \Delta u + \alpha(\Delta U - \Delta u)$

- $\Delta u$ : Change in *my* own utility if I move.
- $\Delta U - \Delta u$ : Change in *everyone else’s* utility.
- $A$ : Global *cooperation level* shared by all residents.

Special cases:

- $\alpha = 0$: Purely selfish – only care about my own utility.
- $\alpha = 1$: Fully cooperative – only care about total utility.




This setup allows us to smoothly adjust between **economic behavior** (self-interest) and **physical-system-like behavior** (global optimization).

# Residents Prefer "Half-Full" Blocks

The model uses an **asymmetrically peaked utility function**:

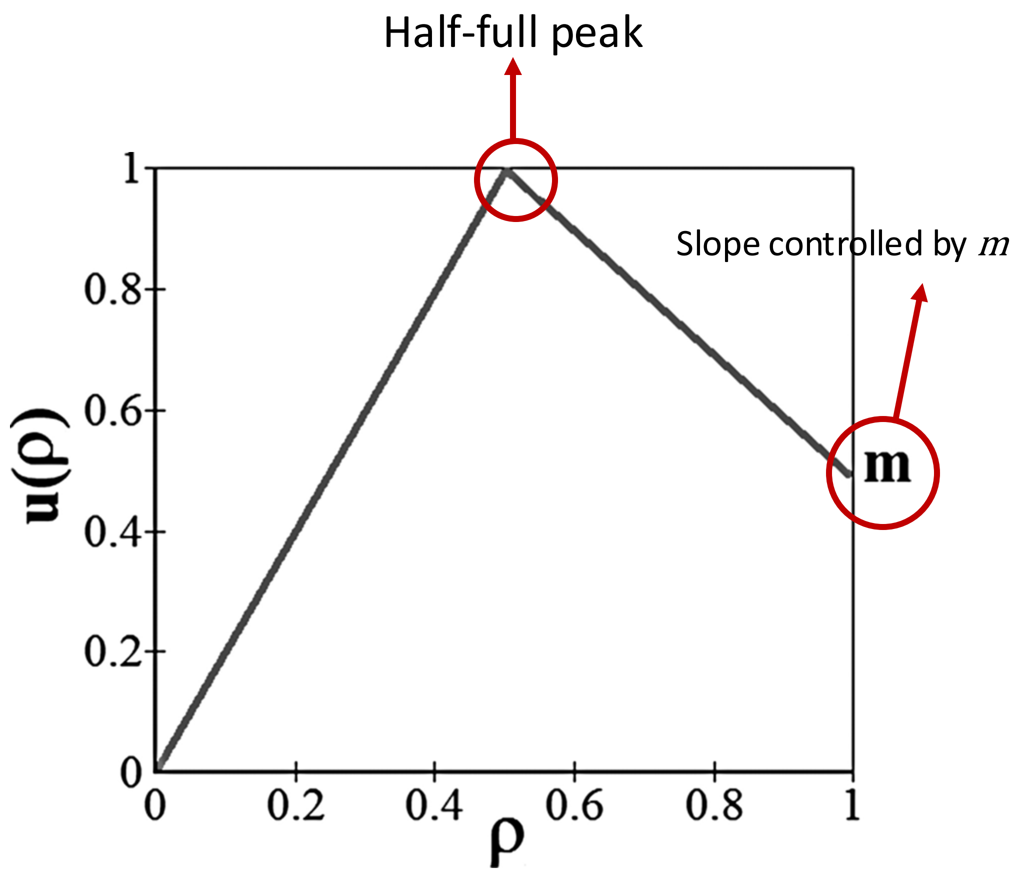
- Utility is **highest** when the block density is $\rho=1/2$ (half-full).
- This reflects a preference for mixed neighborhoods rather than overcrowded or empty ones.

A shape parameter $m$ controls **how fast utility declines** on the crowded side ($\rho>1/2$):

- Smaller $m\rightarrow$ utility drops quickly as crowding increases.
- Larger $m\rightarrow$ residents tolerate higher density before utility declines.

For the example analysis:

- **Global density** fixed at $\rho_0=1/2$.
- **Temperature** $T\rightarrow0$ to focus purely on utility-driven decisions without random noise.



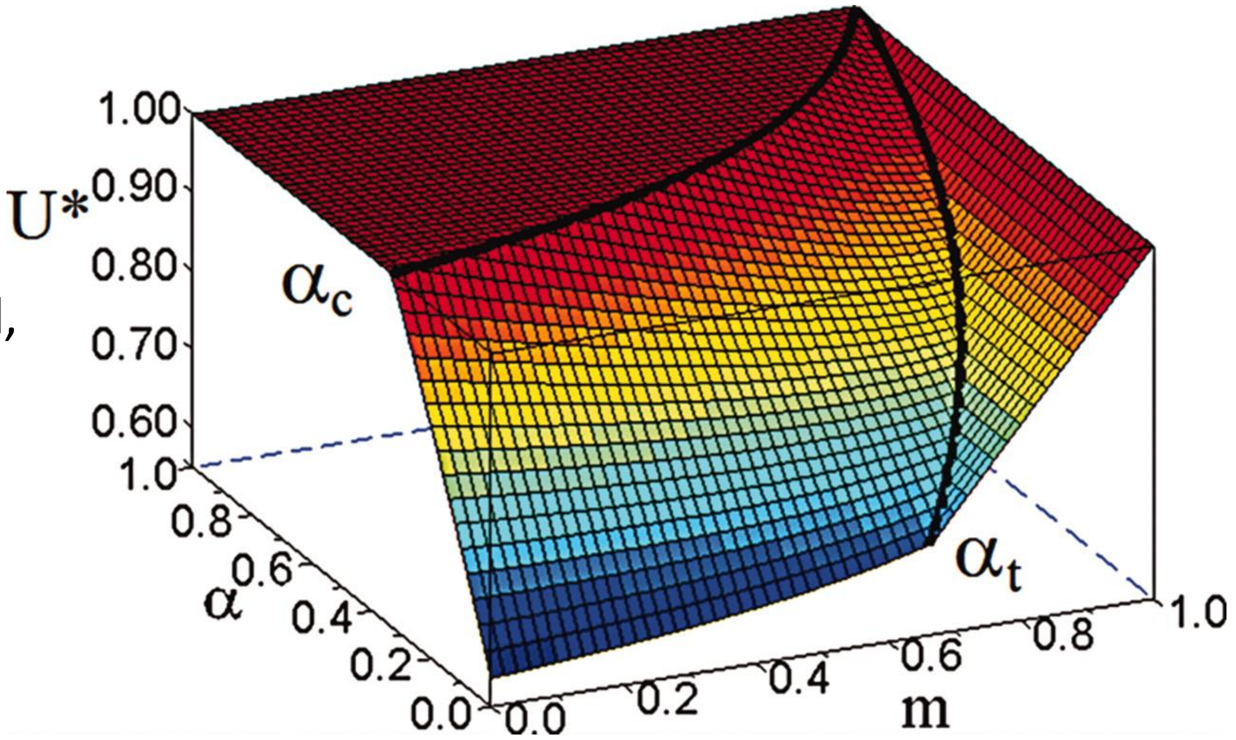
# From Mixed to Segregated: The Critical Point of $\alpha$

**High cooperation** ($\alpha$ large):

- **Mixed state** – each block has the same density.
- Highest possible global utility.

**Low cooperation** ($\alpha$ small):

- **Segregated state** – some blocks are overcrowded, others are empty.
- Lower global utility despite individuals trying to improve their own satisfaction.




There is a **critical value**: $\alpha_c = \dfrac{1}{3 - 2m}$

- If $\alpha > \alpha_c$: mixed state.
- If $\alpha < \alpha_c$: segregated state.

💡As $m$ increases (more tolerance for crowding), the required $\alpha_c$ for a mixed state becomes higher.

# The Role of Noise and Model Robustness

**Low temperature** (*T* small):

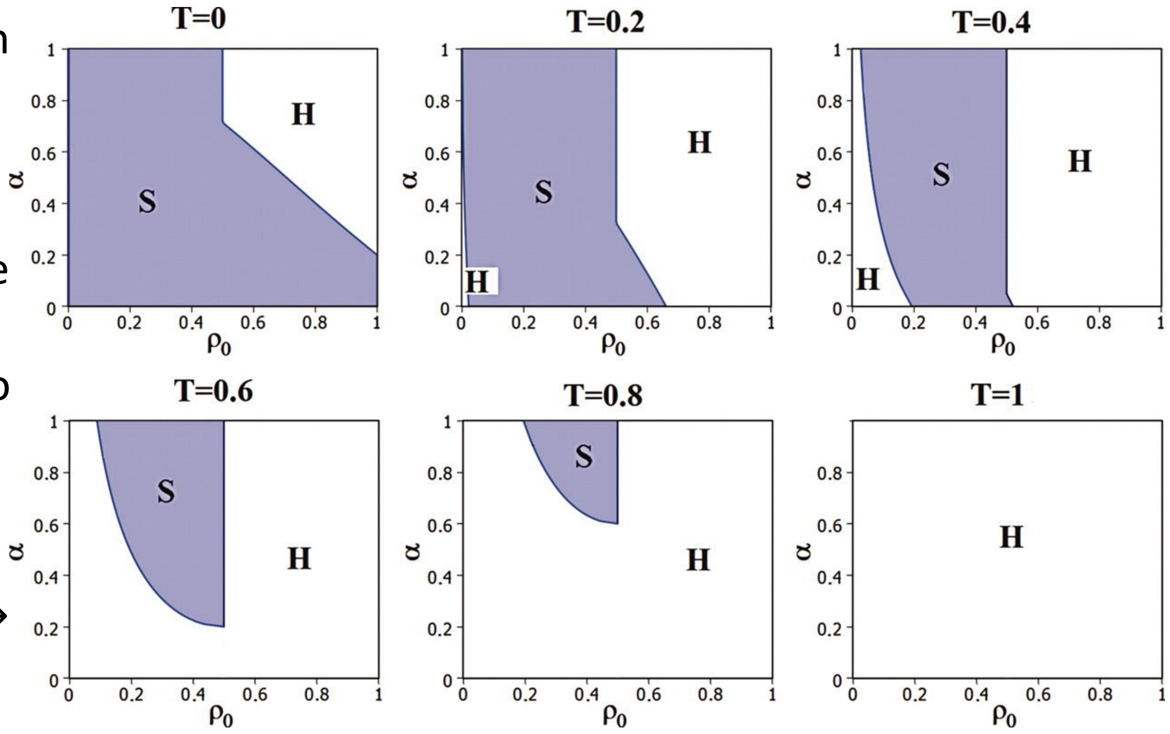
- Noise is weak; the phase diagram and conclusions remain unchanged.

**High temperature** (*T* large):

- Randomness dominates → city becomes more homogeneous.
- This homogeneity is **not** due to utility optimization but to random movement (“noise-driven mixing”).

**Robustness checks**:

- Adding two types of residents (e.g., red and green) → qualitative results remain the same.
- Changing neighborhood definition (nearest neighbors vs. predefined blocks) → qualitative results unchanged.




💡**Analogy:** Just like stirring coffee☕ — constant stirring (noise) evens out local differences in concentration.

# Economic and Policy Implications

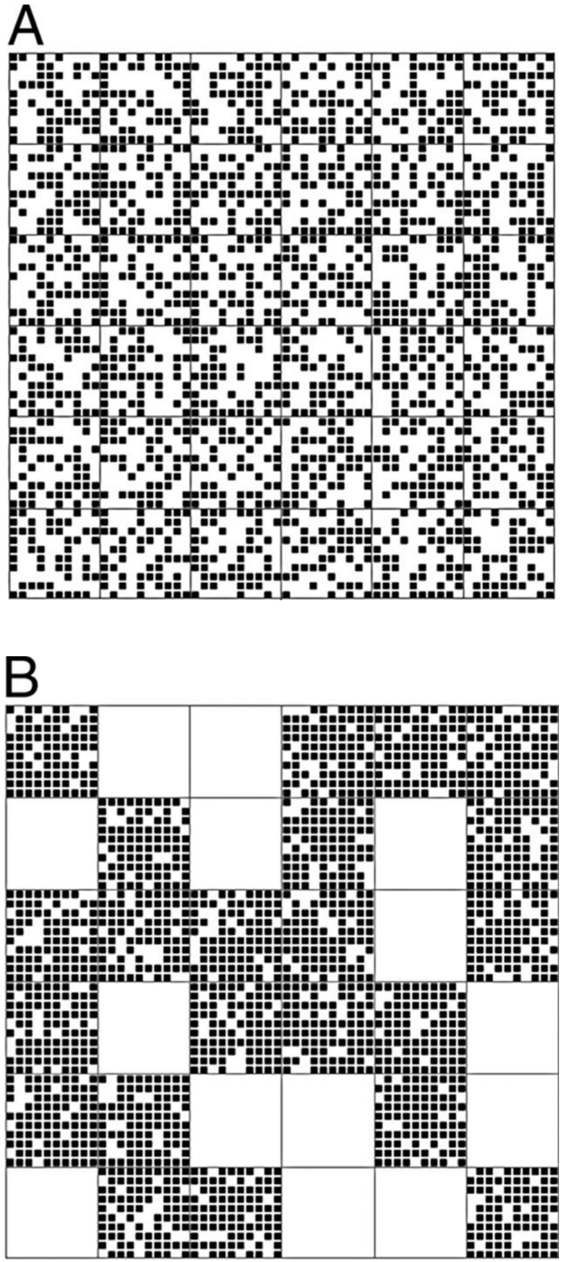
From an **economics perspective**, the cooperation parameter $\alpha$ can be seen as the **strength of policy mechanisms** that internalize externalities (e.g., a *Pigouvian tax*).

**Traditional view:**

- Only when $\alpha = 1$ are all externalities removed, achieving the optimal outcome.

**Finding of this paper:**

- As long as $\alpha$ exceeds the critical threshold $\alpha_c$, the system can avoid inefficient segregation and reach an efficient mixed state.
- This means society can achieve **significant improvement** even without completely eliminating all externalities.



# Thank You!

Thank you for your attention and participation.

📫 **Contact Information:**

zhanzhanzhao@cuhk.edu.cn (Dr. Zhao Zhanzhan)
225080011@link.cuhk.edu.cn (Mr. Jia Xiao )