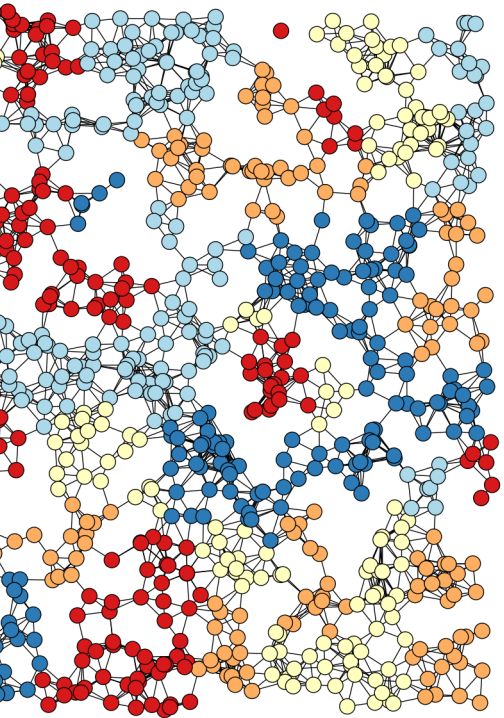


# Computational Learning Theory and Beyond

Open HPI

## Excursion: Game Theory, Segregation and Potential Functions

Algorithm Engineering Group

Pascal Lenzner
Louise Molitor

# Game Theory

**Game Theory** is the study of mathematical models of strategic interaction between rational decision-makers. [Wikipedia; Meyerson '91]

Game Theory

Strategic Agents — choose → Strategies — to maximize → Utility Function

Outcome of the Game

vector of all chosen strategies

is in

Equilibrium

Solution Concept

- subset of all outcomes
- stable: agents don't want to change strategies

Nash Equilibrium

Core

Pairwise Stability

. . .

**Rock-Paper-Scissors**

| Player 2 \ Player 1 | Rock | Paper | Scissors |
|---|---|---|---|
| Rock | 0 / 0 | +1 / −1 | −1 / +1 |
| Paper | −1 / +1 | 0 / 0 | +1 / −1 |
| Scissors | +1 / −1 | −1 / +1 | 0 / 0 |

# Game Theory

**Game Theory** is the study of mathematical models of strategic interaction between rational decision-makers. [Wikipedia; Meyerson ’91]

Game Theory

Strategic Agents — choose — Strategies — to maximize — Utility Function

Core questions:

- Do equilibria exist?
- Properties of equilibria?
- How to design the game to enforce certain equilibria?

Rock-Paper-Scissor

| Player 2 \ | Rock | Paper | Scissors |
|---|---|---|---|
| Rock | 0 | −1 | +1 |
| | 0 | +1 | −1 |
| Paper | +1 | 0 | −1 |
| | −1 | 0 | +1 |
| Scissors | −1 | +1 | 0 |
| | +1 | −1 | 0 |

Nash Equilibrium

Core

Pairwise Stability

. . .

# Algorithmic Game Theory

**Game Theory** is the study of mathematical models of strategic interaction between rational decision-makers. [Wikipedia; Meyerson ’91]

Algorithmic Game Theory

Study games via the ”Algorithmic Lens”

Core questions:

- Do equilibria exist?
- Properties of equilibria?
- How to design the game to enforce certain equilibria?

How can players find equilibria?

Is equilibrium detection hard?

Approximate equilibria?

Impact of selfishness?

How to design algorithms for a strategic environment?

# Algorithmic Game Theory in AI

**Game Theory** is the study of mathematical models of strategic interaction between rational decision-makers. [Wikipedia; Meyerson ’91]

Algorithmic Game Theory in AI

Study games via the ”Algorithmic Lens”

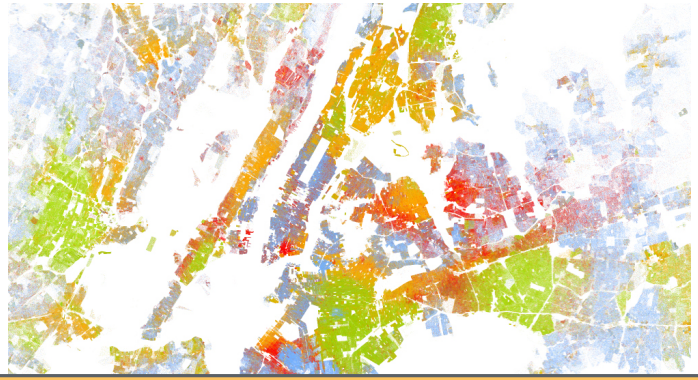
Study multi-agent systems

Core questions:

- Do equilibria exist?
- Properties of equilibria?
- How to design the game to enforce certain equilibria?

- How can players find equilibria?
- Is equilibrium detection hard?
- Approximate equilibria?
- Impact of selfishness?
- How to design algorithms for a strategic environment?

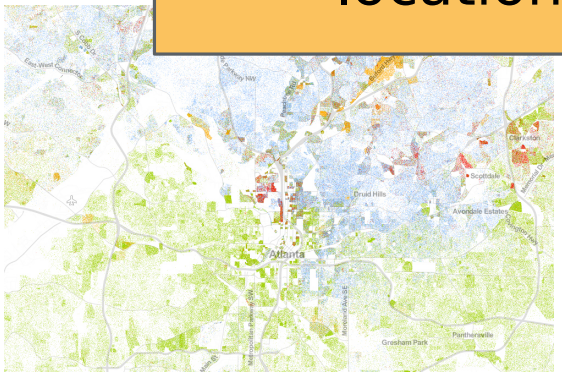
# Motivation: Residential Segregation



Residential Segregation

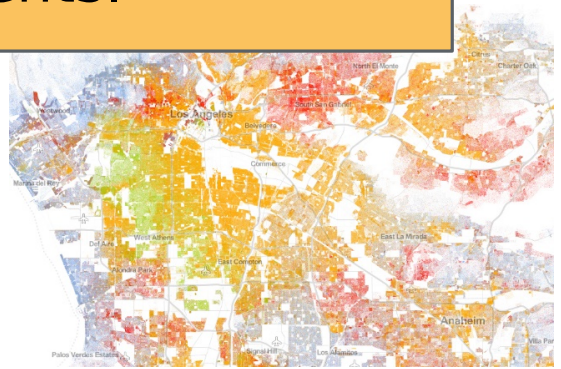
[Plots from: The racial dot map, https://demographics.coopercenter.org/Racial-Dot-Map/]

Segregation is the outcome of strategic location choices by selfish agents.




Atlanta

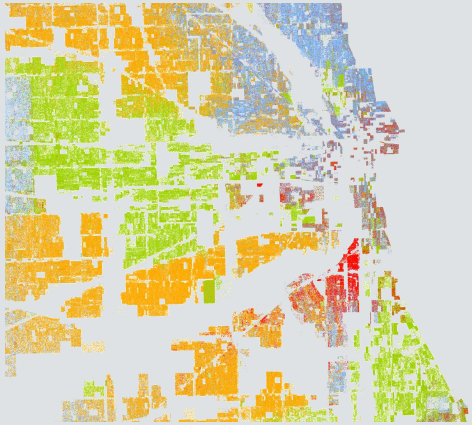
Chicago




Los Angeles

# Schelling's Segregation Model

Real-World Segregation:



Residential Segregation in Chicago

## Schelling's Model: [J. Math. Soc. 1(2), '71]

- two types of agents on a line or grid.
- agents have tolerance parameter $\tau \in [0, 1]$
- agent is content if at least $\tau$-fraction of neighbors is of own type
- discontent agents swap or jump randomly

Phenomenon: From initial random placement the process reaches a severely segregated state even if $\tau < \frac{1}{2}$.

https://ncase.me/polygons/

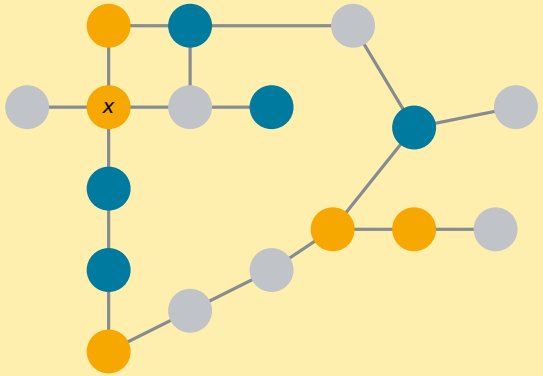
- explains why even a population of tolerant agents can end up in a segregated state

"Micromotives vs. Macrobehavior"

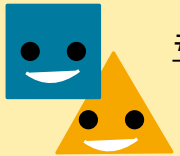
# Game-Theoretic Schelling Segregation

Game-theoretic version:

Chauhan et al. [SAGT'18]




- model with strategic agents on an arbitrary host graph
- agents only care about their neighborhood; agent is happy if



$$\frac{\#\text{agents of same type in neighborhood}}{\#\text{agents in neighborhood}} \geq \tau$$

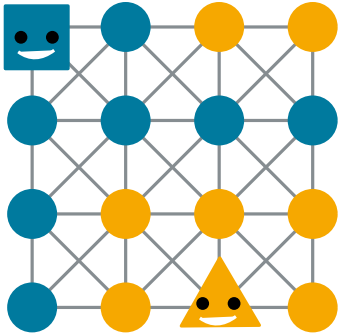
cost function of an agent:

$$cost = \max\left(0, \tau - \frac{\#\text{agents of same type in neighborhood}}{\#\text{agents in neighborhood}}\right)$$

## Swap Game

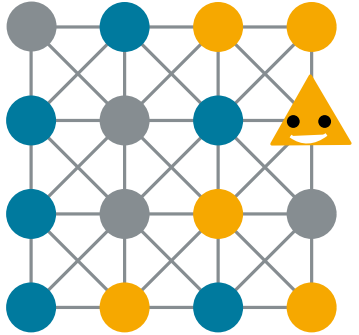
- pairs of unhappy agents can swap positions

$\tau = \frac{1}{2}$ :



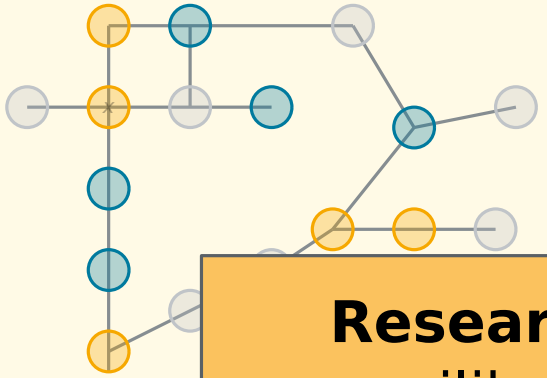
## Jump Game

- unhappy agents can jump to empty positions

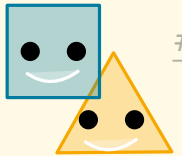


# Game-Theoretic Schelling Segregation

Game-theoretic version:
Chauhan et al. [SAGT'18]



- model with strategic agents on an arbitrary host graph
- agents only care about their neighborhood; agent is happy if



$$\frac{\#\text{agents of same type in neighborhood}}{\#\text{agents in neighborhood}} \geq \tau$$

cost function of an agent:

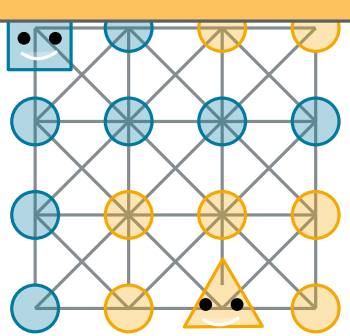
$$\left(\frac{\alpha}{\text{d}}\right)$$

**Research question:** Can agents find an equilibrium via iterative improving moves?

## Swap Game

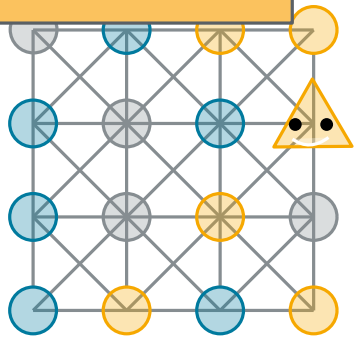
- pairs of unhappy agents can swap positions

$\tau = \frac{1}{2}$ :



## Jump Game

- unhappy agents can jump to empty positions



# Finding Equilibria via Improving Moves

**Research question:** Can agents find an equilibrium via iterative improving moves?

- model with strategic agents on host graph
- cost function of an agent:

$$cost = \max\left(0, \tau - \frac{\#\text{agents of same type in neighborhood}}{\#\text{agents in neighborhood}}\right)$$

**Answers:**

## Swap Game

- YES, if $\tau \leq \frac{1}{2}$ ← Proof via potential function!
- NO, if $\tau > \frac{1}{2}$ ← Proof via improving response cycle!

## Jump Game

- NO, for all $0 < \tau < 1$ ← Proof via improving response cycle!

# Potential Functions

**Research question:** Can agents find an equilibrium via iterative improving moves?

- model with strategic agents on host graph
- cost function of an agent:

$$cost = \max\left(0, \tau - \frac{\#\text{agents of same type in neighborhood}}{\#\text{agents in neighborhood}}\right)$$

**Answers:**

Swap Game

- YES, if $\tau \leq \frac{1}{2}$ ← Proof via potential function!

**Potential Function $\Phi$:**

- maps state of a game to a real number
- if agent improves, then potential decreases by at least $\varepsilon > 0$
- minimum potential value is lower bounded
- every local minimum of $\Phi$ corresponds to equilibrium state

**Proof:** Choose $\Phi = \frac{1}{2} \sum_{\text{agents}} \#\text{neighbors of other type}$

$= \#\text{edges between agents with different type}$

# Improving Response Cycles

**Research question:** Can agents find an equilibrium via iterative improving moves?

- model with strategic agents on host graph
- cost function of an agent:

$$cost = \max\left(0, \tau - \frac{\#\text{agents of same type in neighborhood}}{\#\text{agents in neighborhood}}\right)$$

**Answers:**

Jump Game

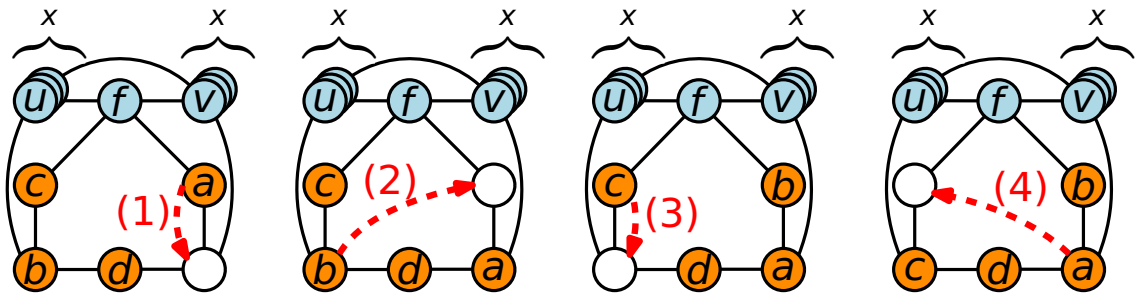
- NO, for all $0 < \tau < 1$

← Proof via improving response cycle!

**Improving Response Cycle:**

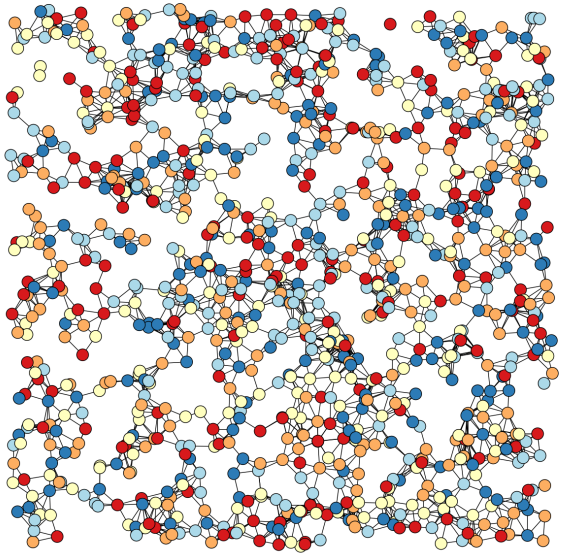
- cyclic sequence of states of the game
- next state can be reached by an improving move of some agent

**Proof:**




$$x > \max\left(\frac{2}{\tau}, \frac{1}{1-\tau}\right)$$

# Wrap-Up



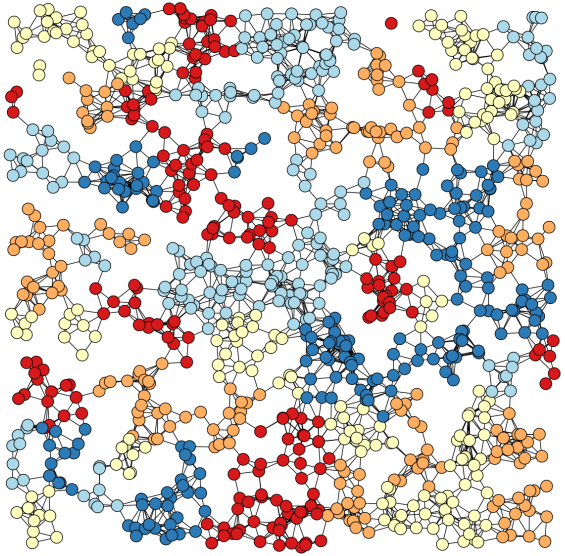
**(Algorithmic) Game Theory**

- suited for analysing interactions of (selfish) agents
- recent trend in foundational AI research

**Schelling's Model**

- residential segregation as agent-based system

selfish location choice



**Potential Functions**

- tool for analysing dynamic systems
- existence guarantees convergence to equilibrium