Multi-agent learning

Real-valued spatial games

Gerard Vreeswijk, Intelligent Systems Group, Computer Science Department,
Faculty of Sciences, Utrecht University, The Netherlands.

Robert Axelrod: "The Evolution of Cooperation"

Robert Axelrod: "The Evolution of Cooperation" (1984)

Round 1 (1980):

- Solicited entries for a round-robin tournament:

  Programs would play the IPD for 200 rounds against every other contestant.
- Fourteen participants from social sciences, math, CS, economy.
- Programs were of all sorts and kinds.
- Winner: TFT. (TFT was also the shortest program!)

Round 2 (1981):

- Sixty-two participants from six different countries.
- In this second round all participants knew that TFT was the winner.
- Publication in Science! (Vol. 211, No. 4489. March 27, pp. 1390-1396.)

Observation & reflection (1984):

- Book: "The Evolution of Cooperation" (1984)

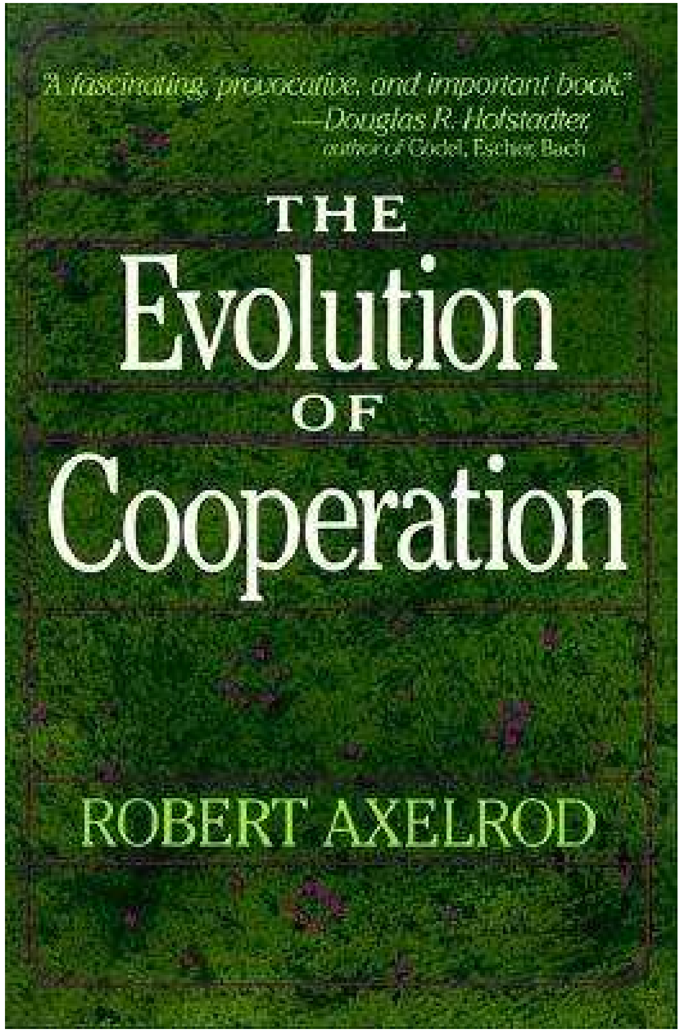
Robert Axelrod: “The Evolution of Cooperation” (1984)

Questions:

- How can cooperation emerge in a population of self-interested individuals?
- Which mechanisms lie at the basis of successful cooperation?

Hypotheses:

- Reciprocity (I scratch your back, you’ll scratch mine).
- Clusters of cooperation.
- Protection against invasion by inferior strategies.



General motivation

Biological perspective: cooperation is observed in a population.

- Under what conditions will cooperation and altruism emerge in a world of egoistic actors without central authority?
- Which mechanisms ensure cooperation?
- What origins does such mechanisms have?

  (Imitation? Learning? Evolution? Combinations of such mechanisms?)

Multi-agent system perspective: cooperation is sometimes desired in a multi-agent system.

Problem. How can we get artificial agents that are selfish and self-interested to cooperate just on the basis of their own self-interests, without any form of centralised control?

Plan for today

Part I: Spatial Iterated Prisoners' Dilemma (SIPD)

- Evolution of cooperation with pure memoryless strategies.
- Demo.

Part II: Spatial Iterated Prisoners' Dilemma with Continuous Actions

- Demo.
- Analysis. Explain (analytically) the following.
  1. How investors obtain a benefit from clustering.
  2. How first clusters of investors are formed.
  3. How the asymptotic investment value can be deduced.

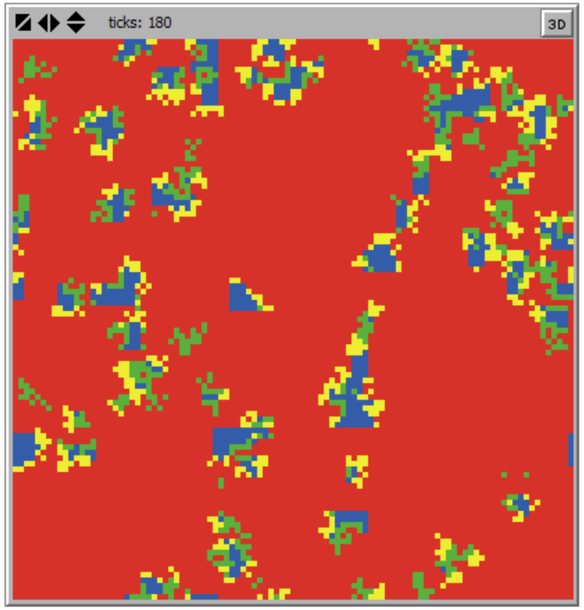
Part I: Spatial Iterated Prisoners' Dilemma

Spatial Iterated Prisoners' Dilemma (SIPD)

- Payoff matrix, with reward to defect $0 \leq \alpha \leq 3$:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 1,1 & 0,\alpha \\ \alpha,0 & 0,0 \end{pmatrix} \end{array}$$

- A strategy is a pure, memoryless action.
- Per tick, all agents interact with their neighbours, and adopt strategy of best of $3 \times 3 - 1$.
- Interesting when initial cooperation is low (e.g., 5%), and reward to defect is "balanced" (e.g., $\alpha = 1.6$).

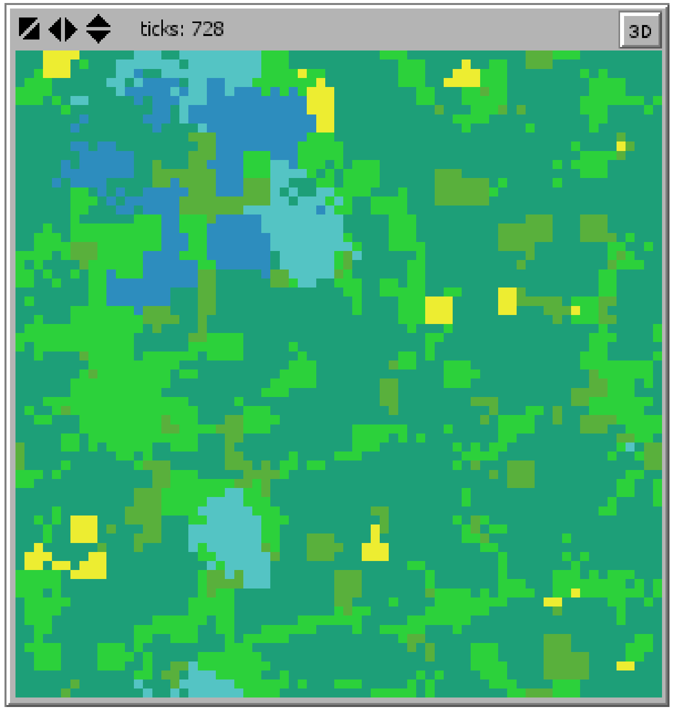


	Previous	*Current*
Blue	C	C
Red	D	D
Green	C	D
Yellow	D	C

Part II: Spatial Iterated Prisoners' Dilemma with Continuous Actions

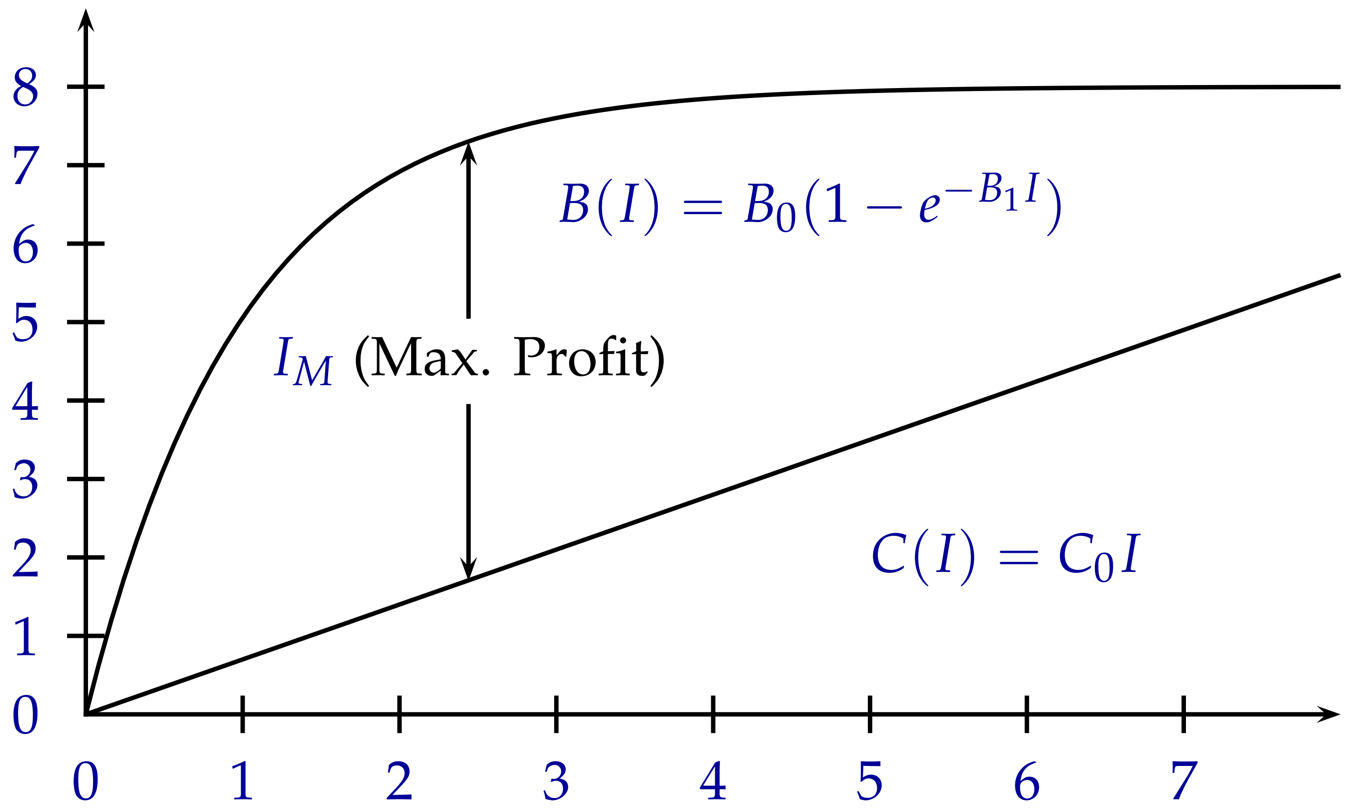
Spatial Iterated Prisoners' Dilemma with Continuous Actions

- This time, actions are investments $0 \leq I \leq I_{\max}$. Investments are there to please your neighbours, and only your neighbours.
  - You invest I.
    * This wil cost you $C(I)$ per neighbour.
    * Every neighbour will benefit from your investment with $B(I)$.
  - Your neighbour j invests I_j. Same story, other way around.
- $P(I, \{I_j\}_j) = -8\,C(I) + \sum_{j=1}^{8} B(I_j)$.



Colours indicate the level of investment (not the fitness). Red (< 0.0001), Orange, Brown, Yellow, Green, Lime Turquoise, Cyan, Sky, Blue, Violet, Magenta, Pink ($I_{\max}$).

Cost and Benefit as function of Investment



- In Killingback *et al.*: $B_0 = 8$, $B_1 = 1$, and $C_0 = 0.7$.

Relation with one stage discrete Prisoners' Dilemma

- Suppose $I_1 < I_2 < I_M$. With the functions B and C defined as above,

$$P(I_1, I_2) > P(I_2, I_2) > P(I_1, I_1) > P(I_2, I_1)$$

- $P(I_1, I_2)$ Corresponds to (D, C) (Free rider.)
  $P(I_2, I_2)$ Corresponds to (C, C) (Cooperation.)
  $P(I_1, I_1)$ Corresponds to (D, D) (Every man for himself and God for us all.)
  $P(I_2, I_1)$ Corresponds to (C, D) (Sucker.)

- If two neighbours invest $I_1 = 0.3$ and six invest $I_2 = 0.5$, then

$$\begin{array}{cc} & \begin{array}{cc} 6\times\mathrm{C} & 2\times\mathrm{D} \end{array} \\ \begin{array}{c} \mathrm{C} \\ \mathrm{D} \end{array} & \begin{pmatrix} P(0.5, 0.5) & P(0.3, 0.5) \\ P(0.5, 0.3) & P(0.3, 0.3) \end{pmatrix} \end{array} = \begin{array}{cc} & \begin{array}{cc} 6\times\mathrm{C} & 2\times\mathrm{D} \end{array} \\ \begin{array}{c} \mathrm{C} \\ \mathrm{D} \end{array} & \begin{pmatrix} 22, 22 & 14, 24 \\ 24, 14 & 15, 15 \end{pmatrix} \end{array}$$

Process

- There are 70×70 cells in the land of Torusota.
- All cells start with an extremely low level of investment I_{init}, uniformly distributed over $(0, 10^{-4})$.
- Loop.
  - All cells interact with their eight neighbours $j = 1, \dots, 8$ and compute the profit $B(I_j) - C(I)$ for each interaction.

$$\text{Fitness} = -8\, C(I) + \sum_{j=1}^{8} B(I_j).$$

  - All cells adopt the fitness of their fittest neighbour (including the cell itself).
  - All cells mutate their level of investment I with probability 0.01. If mutated, the mutation rate is normally distributed with variance $I/10$.

Mutation can be seen as "random act of kindness"

- From Wikipedia:

  "A selfless act performed by a person or persons wishing to either assist or cheer up an individual or in some cases an animal. There will generally be no reason other than to make people smile, or be happier. Either spontaneous or planned in advance, random acts of kindness are encouraged by various communities."

  `http://en.wikipedia.org/wiki/Random_acts_of_kindness`

- Generally takes place for an entire week during the earlier months of the year. In Canada, it is held either in February or in March

SIRE (Foundation for Idealistic Advertising)

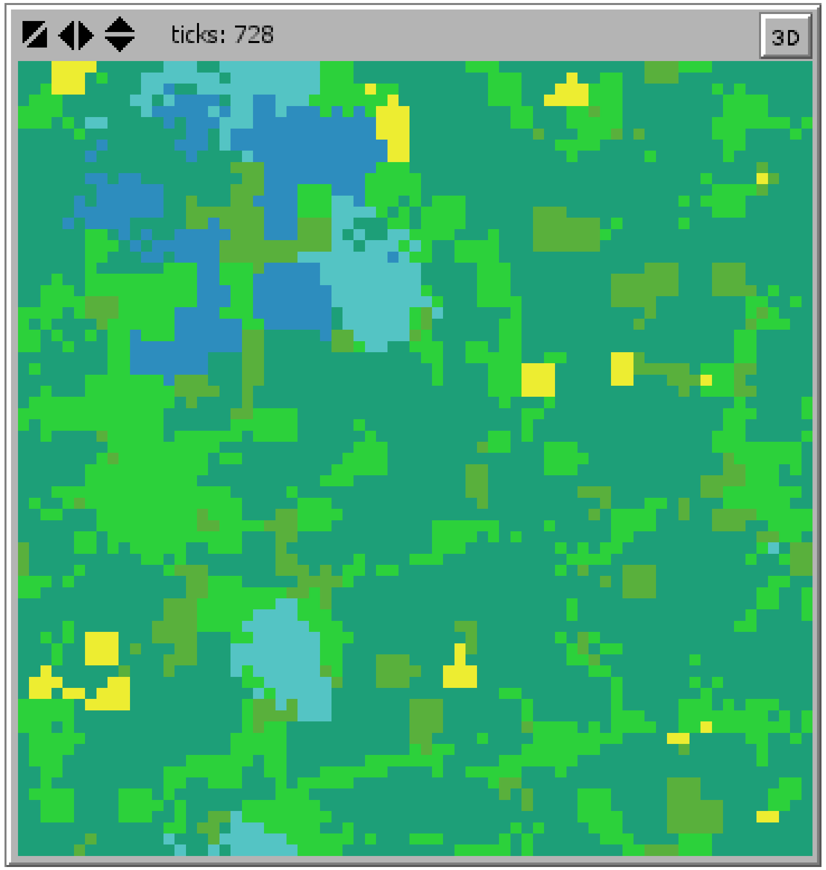
- "Aardige mensen, hoe gaan we er mee om". (How to deal with kind people.)
- "Pas op—/! Aardig!" (Watch out—/! Kind person!)

How to deal with kind people http://www.pasopaardig.nl

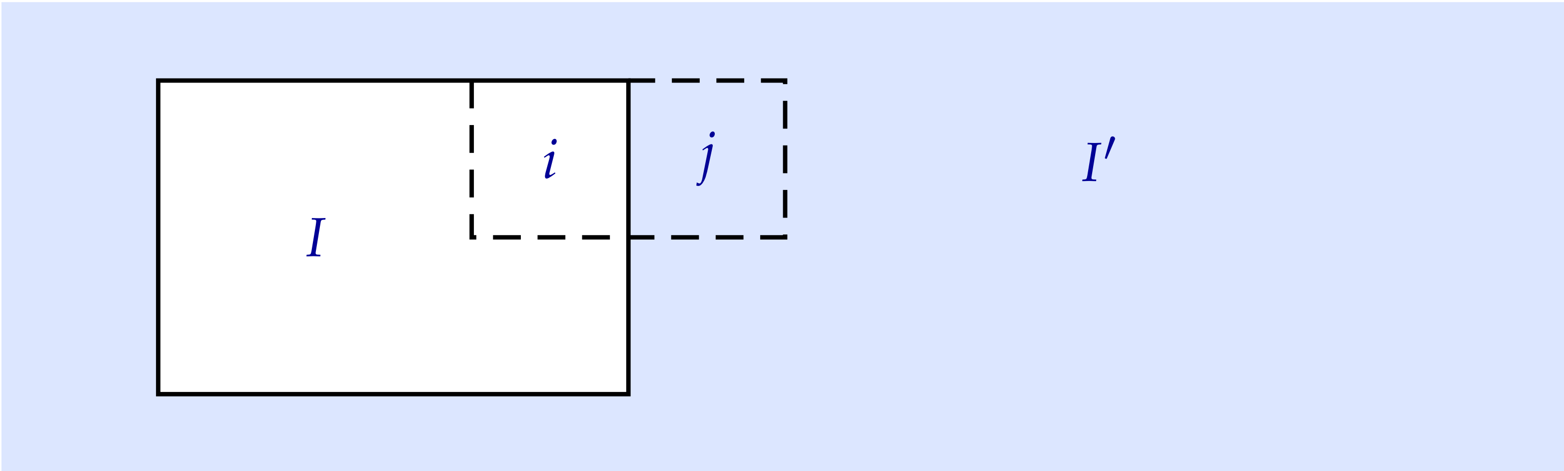
Analysis

The article of Killingback *et al.* concludes with three observations, that are "explained".

1. **Expansion:** clusters of high investors help each other.
2. **Bootstrap:** clusters are formed in the first place.
3. **Long term:** investment converges to
$$\frac{1}{B_1} \ln \left(\frac{B_0 B_1}{4 C_0} \right).$$



First mechanism: I-investors mutually benefit from clustering



Suppose $I > I'$, and $n = \#I$-neighbours of i, and $m = \#I$-neighbours of j.

A necessary condition for cell i to "take over" cell j is

$$nB(I) + (8-n)B(I') - 8\,C(I) > mB(I) + (8-m)B(I') - 8\,C(I')$$

Whether this is sufficient follows from applying similar arguments to neighbours of cell j.

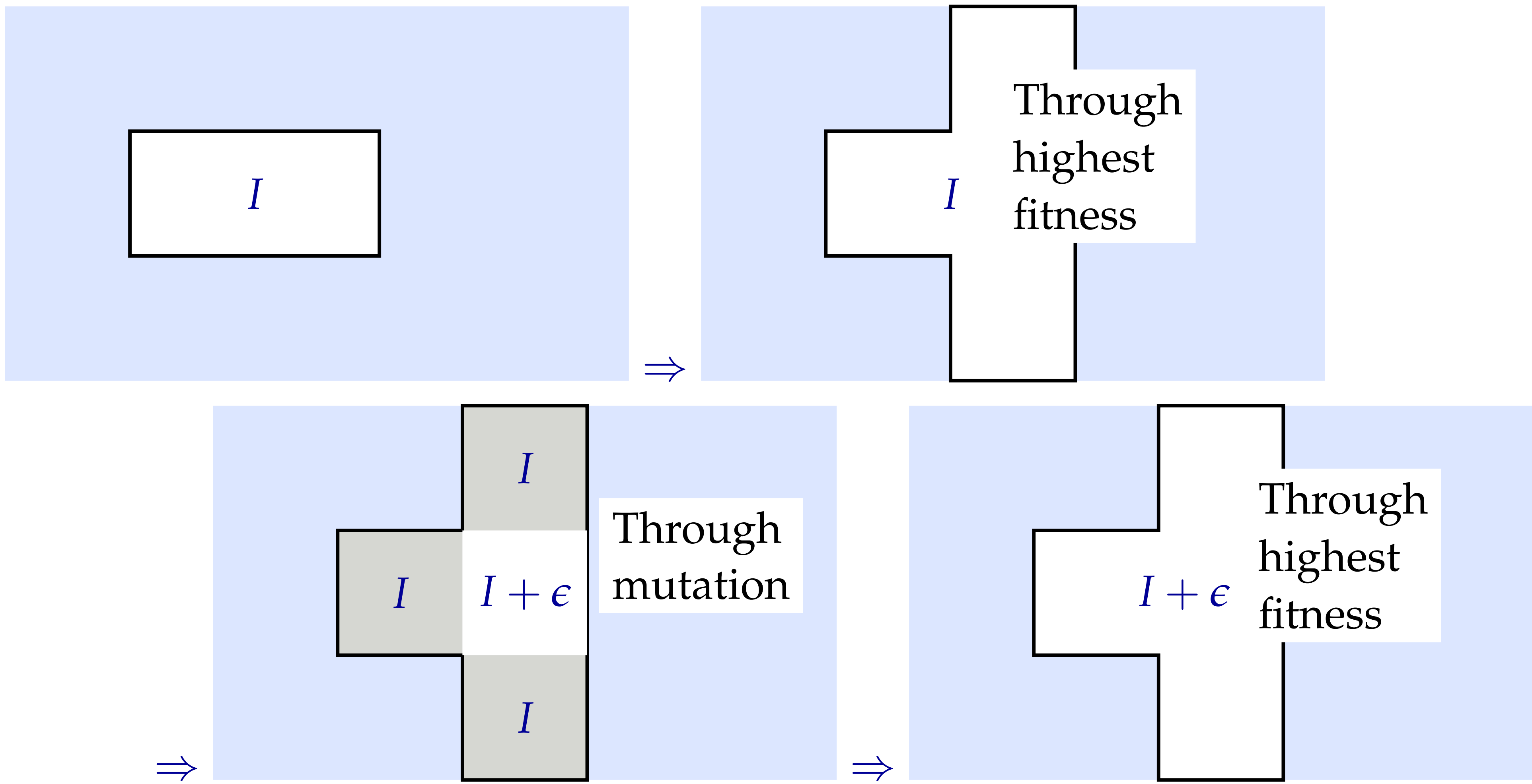
Sufficient conditions to "take over" a neighbour

$$
\begin{aligned}
& nB(I) + (8-n)B(I') - 8\,C(I) > mB(I) + (8-m)B(I') - 8\,C(I') \\
\Leftrightarrow\ & (n-m)B(I) - (n-m)B(I') - 8(C(I) - C(I')) > 0 \\
\Leftrightarrow\ & (n-m)[1-e^{-I}] - (n-m)[1-e^{-I'}] - C_0(I-I') > 0 \\
\Leftrightarrow\ & (n-m)[e^{-I'} - e^{-I}] - C_0(I-I') > 0 \\
\Leftrightarrow\ & \frac{e^{-I} - e^{-I'}}{I-I'} < \frac{-C_0}{n-m} \quad (\text{provided } n-m>0\text{, that is, provided } n-m \in \{1,2\}) \\
\Leftarrow\ & -e^{-I} < \frac{-C_0}{n-m} \quad \left(\text{since } \frac{e^{-I} - e^{-I'}}{I-I'} < \frac{d}{dI}e^{-I} = -e^{-I}\right) \\
\Leftrightarrow\ & e^{-I} > \frac{C_0}{n-m} \ \Leftrightarrow\ I < \ln\left(\frac{n-m}{C_0}\right).
\end{aligned}
$$

Sufficient conditions to "take over" a neighbour

	$n - m$	constraint on investment I
adjacent to cluster of free riders	-4	no constraint
more free riders than investors	-3	no constraint
	-2	no constraint
	-1	no constraint
	0	unsatisfiable
more investors than free riders	1	only if $I < \ln(1/0.7) = 0.36$
adjacent to cluster of investors	2	only if $I < \ln(2/0.7) = 1.05$

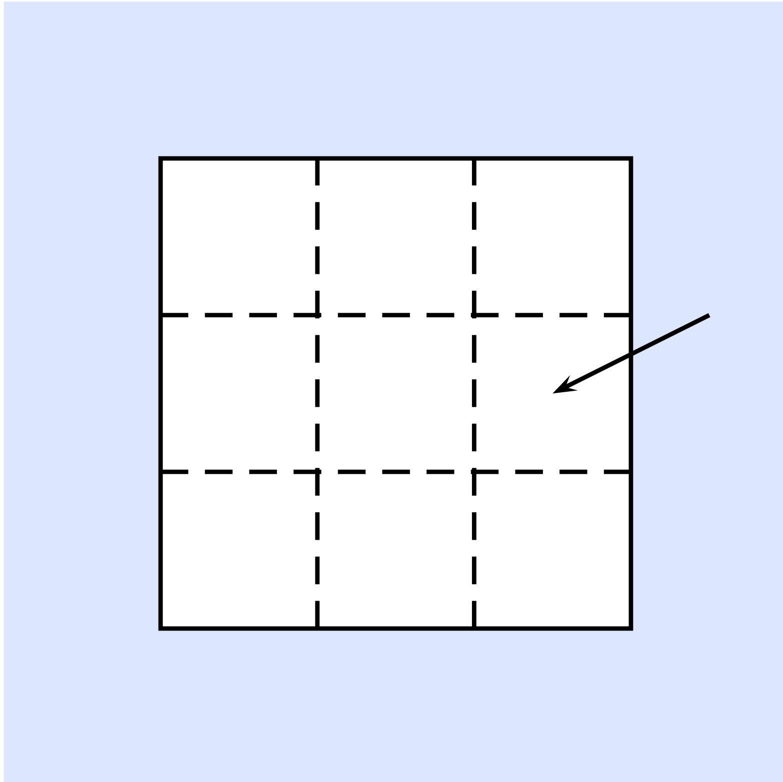
Second mechanism: the formation of basic clusters



Third analysis: asymptotic value of I: I^*

- Consider a cluster of high investors in a sea of of lower investors.
- For which values of I is it impossible to expand?
- By considering what happens at the edge of a cluster, one may derive that a cluster cannot expand as soon as

$$I > (1/B_1)\ln(B_0 B_1/4C_0).$$



Values for investment I that stop expansion

$$nB(I) + (8-n)B(I') - 8 \cdot C(I) < mB(I) + (8-m)B(I') - 8\,C(I')$$
$$\Leftrightarrow (n-m)B(I) - (n-m)B(I') - 8(C(I) - C(I')) < 0$$
$$\Leftrightarrow (n-m)[B_0(1-e^{-B_1 I})] - (n-m)[B_0(1-e^{-B_1 I'})] - 8C_0(I - I') < 0$$
$$\Leftrightarrow (n-m)[B_0 e^{-B_1 I'} - B_0 e^{-B_1 I}] - 8C_0(I - I') < 0$$
$$\Leftrightarrow \frac{B_0 e^{-B_1 I} - B_0 e^{-B_1 I'}}{I - I'} > -\frac{8C_0}{2} \quad \text{(since } n = 5 \text{ and } m = 3\text{)}$$
$$\Leftrightarrow -B_0 B_1 e^{-B_1 I} > -4C_0 \quad \left(\text{since } \frac{B_0 e^{-IB_1} - B_0 e^{-B_1 I'}}{I - I'} > -B_0 B_1 e^{-B_1 I}\right)$$
$$\Leftrightarrow B_0 B_1 e^{-B_1 I} < 4C_0$$
$$\Leftrightarrow I > (1/B_1) \ln\left(\frac{B_0 B_1}{4C_0}\right).$$

Joshua Epstein

Joshua Epstein, a professor in the Department of Emergency Medicine, is launching an advanced modeling center to develop simulations of chronic diseases, economic turmoil and sudden disasters, such as the toxic chemical cloud in Los Angeles model seen on the computer here. Photo: Keith Weller.

Problems with Agent-based modeling

Epstein:

"(. . .) among skeptics toward agent modeling, the central indictment is tripartite:

First, that in contrast to mathematical "hard" science, there are no equations for agent-based models.

Second, that agent models are not deductive (. . .)

Third, that they are *ad hoc*, not general.

I will argue that the first two claims are false and that, at this stage in the field's development, the third is unimportant." (p. 1590)

Joshua M. Epstein (2006). "Remarks on the Foundations of Agent-based Generative Social Science". Chapter 34 in: *Handbook of Computational Economics*, Vol. 2: "Agent-Based Computational Economics," Handbooks in Economics Series, pp. 1586-1602. Elsevier/North-Holland.

Further

- Agent-based Computational Economics (ACE).

Robert Axelrod (2006). "Agent-based Modeling as a Bridge Between Disciplines". Chapter 33 in: *Handbook of Computational Economics*, Vol. 2: "Agent-Based Computational Economics," Handbooks in Economics Series, pp. 1565-1586. Elsevier/North-Holland.

Robert Axelrod and Leigh Tesfatsion (2006). "A Guide for Newcomers to Agent-based Modeling in the Social Sciences". Appendix A in: *Handbook of Computational Economics*, Vol. 2: "Agent-Based Computational Economics," Handbooks in Economics Series, pp. 1648-1658. Elsevier/North-Holland.

Leigh Tesfatsion and Kenneth L. Judd (2006). *Handbook of Computational Economics*, Vol. 2: "Agent-Based Computational Economics". Handbooks in Economics Series. Elsevier/North-Holland.

- Other games with continuous actions: Cournot dynamics, Stackelberg dynamics.
- Computational Biology.