

VII. Cooperation & Competition

The Iterated Prisoner's Dilemma

The Prisoners' Dilemma

- Devised by Melvin Dresher & Merrill Flood in 1950 at RAND Corporation
- Further developed by mathematician Albert W. Tucker in 1950 presentation to psychologists
- It “has given rise to a vast body of literature in subjects as diverse as philosophy, ethics, biology, sociology, political science, economics, and, of course, game theory.” — S.J. Hagenmayer
- “This example, which can be set out in one page, could be the most influential one page in the social sciences in the latter half of the twentieth century.” — R.A. McCain

Prisoners' Dilemma: The Story

- Two criminals have been caught
- They cannot communicate with each other
- If both confess, they will each get 10 years
- If one confesses and accuses other:
 - confessor goes free
 - accused gets 20 years
- If neither confesses, they will both get 1 year on a lesser charge

Prisoners' Dilemma Payoff Matrix

		Bob	
		cooperate	defect
Ann	cooperate	-1, -1	-20, 0
	defect	0, -20	-10, -10

- defect = confess, cooperate = don't
- payoffs < 0 because punishments (losses)

Ann's "Rational" Analysis (Dominant Strategy)

		Bob	
		cooperate	defect
Ann	cooperate	-1, -1	-20, 0
	defect	0, -20	-10, -10

- if cooperates, may get 20 years
- if defects, may get 10 years
- \therefore , best to defect

Bob's "Rational" Analysis (Dominant Strategy)

		Bob	
		cooperate	defect
Ann	cooperate	-1, -1	-20, 0
	defect	0, -20	-10, -10

- if he cooperates, may get 20 years
- if he defects, may get 10 years
- \therefore , best to defect

Suboptimal Result of “Rational” Analysis

		Bob	
		cooperate	defect
Ann	cooperate	-1, -1	-20, 0
	defect	0, -20	-10, -10

- each acts individually rationally \Rightarrow get 10 years (dominant strategy equilibrium)
- “irrationally” decide to cooperate \Rightarrow only 1 year

Summary

- Individually rational actions lead to a result that all agree is less desirable
- In such a situation you cannot act unilaterally in your own best interest
- Just one example of a (game-theoretic) *dilemma*
- Can there be a situation in which it would make sense to cooperate unilaterally?
 - **Yes**, if the players can expect to interact again in the future

The Iterated Prisoners' Dilemma

and Robert Axelrod's Experiments

Assumptions

- No mechanism for enforceable threats or commitments
- No way to foresee a player's move
- No way to eliminate other player or avoid interaction
- No way to change other player's payoffs
- Communication only through direct interaction

Axelrod's Experiments

- Intuitively, expectation of future encounters may affect rationality of defection
- Various programs compete for 200 rounds
 - encounters each other and self
- Each program can remember:
 - its own past actions
 - its competitors' past actions
- 14 programs submitted for first experiment

IPD Payoff Matrix

		B	
		cooperate	defect
A	cooperate	3, 3	0, 5
	defect	5, 0	1, 1

N.B. Unless $DC + CD < 2 CC$ (i.e. $T + S < 2 R$),
can win by alternating defection/cooperation

Indefinite Number of Future Encounters

- Cooperation depends on expectation of **indefinite** number of future encounters
- Suppose a known finite number of encounters:
 - No reason to C on last encounter
 - Since expect D on last, no reason to C on next to last
 - And so forth: there is no reason to C at all

Analysis of Some Simple Strategies

- Three simple strategies:
 - **ALL-D**: always defect
 - **ALL-C**: always cooperate
 - **RAND**: randomly cooperate/defect
- Effectiveness depends on environment
 - **ALL-D** optimizes local (individual) fitness
 - **ALL-C** optimizes global (population) fitness
 - **RAND** compromises

Expected Scores

↓ playing ⇒	ALL-C	RAND	ALL-D	Average
ALL-C	3.0	1.5	0.0	1.5
RAND	4.0	2.25	0.5	2.25
ALL-D	5.0	3.0	1.0	3.0

Result of Axelrod's Experiments

- Winner is Rapoport's **TFT** (Tit-for-Tat)
 - cooperate on first encounter
 - reply in kind on succeeding encounters
- Second experiment:
 - 62 programs
 - all know **TFT** was previous winner
 - **TFT** wins again

Expected Scores

↓ playing ⇒	ALL-C	RAND	ALL-D	TFT	Avg
ALL-C	3.0	1.5	0.0	3.0	1.875
RAND	4.0	2.25	0.5	2.25	2.25
ALL-D	5.0	3.0	1.0	$1 + 4/N$	2.5+
TFT	3.0	2.25	$1 - 1/N$	3.0	2.3125-

Demonstration of Iterated Prisoners' Dilemma

Run NetLogo demonstration
PD N-Person Iterated.nlogo

Characteristics of Successful Strategies

- *Don't be envious*
 - at best TFT ties other strategies
- *Be nice*
 - i.e. don't be first to defect
- *Reciprocate*
 - reward cooperation, punish defection
- *Don't be too clever*
 - sophisticated strategies may be unpredictable & look random; be clear

Tit-for-Two-Tats

- More forgiving than **TFT**
- Wait for two successive defections before punishing
- Beats **TFT** in a noisy environment
- E.g., an unintentional defection will lead **TFTs** into endless cycle of retaliation
- May be exploited by feigning accidental defection

Effects of Many Kinds of Noise Have Been Studied

- Misimplementation noise
- Misperception noise
 - noisy channels
- Stochastic effects on payoffs
- General conclusions:
 - sufficiently little noise \Rightarrow generosity is best
 - greater noise \Rightarrow generosity avoids unnecessary conflict but invites exploitation

More Characteristics of Successful Strategies

- Should be a generalist (robust)
 - i.e. do sufficiently well in wide variety of environments
- Should do well with its own kind
 - since successful strategies will propagate
- Should be cognitively simple
- Should be evolutionary stable strategy
 - i.e. resistant to invasion by other strategies

Kant's Categorical Imperative

“Act on maxims that can at the same time have for their object themselves as universal laws of nature.”

Ecological & Spatial Models

Ecological Model

- What if more successful strategies spread in population at expense of less successful?
- Models success of programs as fraction of total population
- Fraction of strategy = probability random program obeys this strategy

Variables

- $P_i(t)$ = probability = proportional population of strategy i at time t
- $S_i(t)$ = score achieved by strategy i
- $R_{ij}(t)$ = relative score achieved by strategy i playing against strategy j over many rounds
 - fixed (not time-varying) for now

Computing Score of a Strategy

- Let n = number of strategies in ecosystem
- Compute score achieved by strategy i :

$$S_i(t) = \sum_{k=1}^n R_{ik}(t)P_k(t)$$

$$\mathbf{S}(t) = \mathbf{R}(t)\mathbf{P}(t)$$

Updating Proportional Population

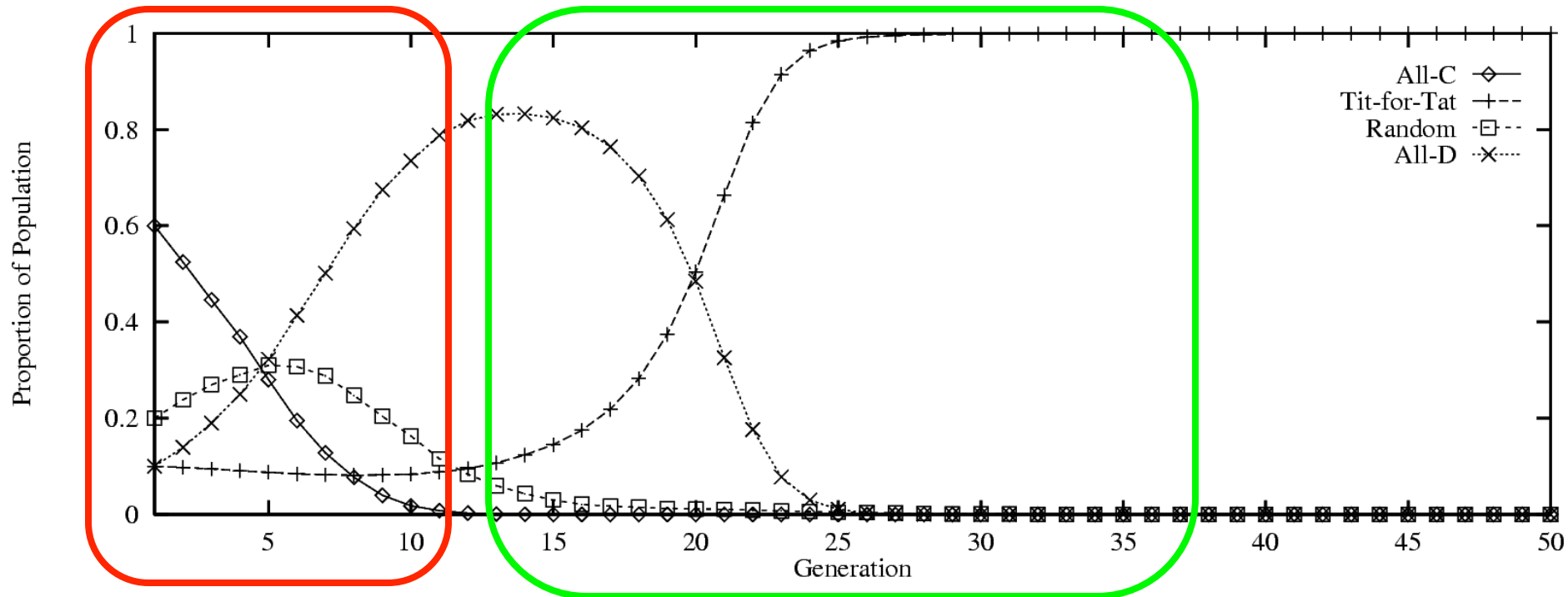
$$P_i(t + 1) = \frac{P_i(t)S_i(t)}{\sum_{j=1}^n P_j(t)S_j(t)}$$

Some Simulations

- Usual Axelrod payoff matrix
- 200 rounds per step

Demonstration Simulation

- 60% ALL-C
- 20% RAND
- 10% ALL-D, TFT



NetLogo Demonstration of Ecological IPD

Run EIPD.nlogo

Collectively Stable Strategy

- Let w = probability of future interactions
- Suppose cooperation based on reciprocity has been established
- Then no one can do better than **TFT** provided:

$$w \geq \max\left(\frac{T - R}{R - S}, \frac{T - R}{T - P}\right)$$

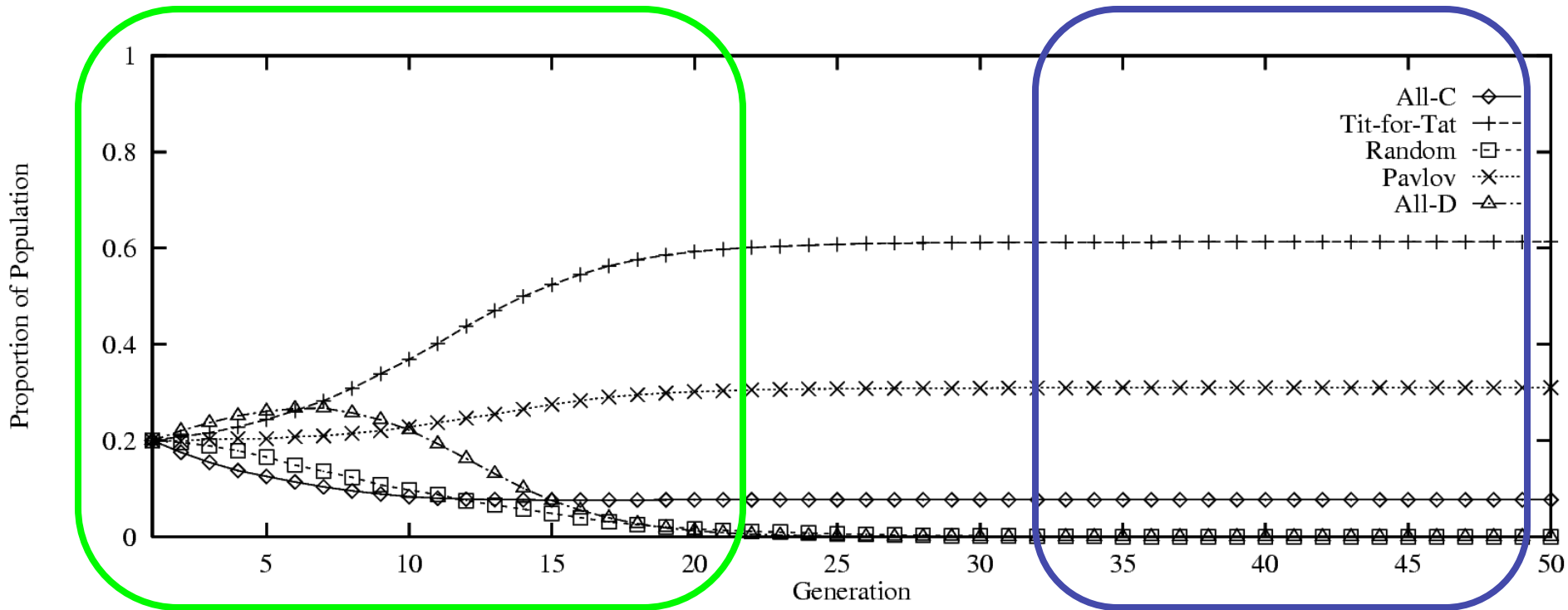
- The **TFT** users are in a Nash equilibrium

“Win-Stay, Lose-Shift” Strategy

- Win-stay, lose-shift strategy:
 - begin cooperating
 - if other cooperates, continue current behavior
 - if other defects, switch to opposite behavior
- Called **PAV** (because suggests Pavlovian learning)

Simulation without Noise

- 20% each
- no noise



Effects of Noise

- Consider effects of noise or other sources of error in response
- **TFT:**
 - cycle of alternating defections (CD, DC)
 - broken only by another error
- **PAV:**
 - eventually self-corrects (CD, DC, DD, CC)
 - can exploit **ALL-C** in noisy environment
- Noise added into computation of $R_{ij}(t)$

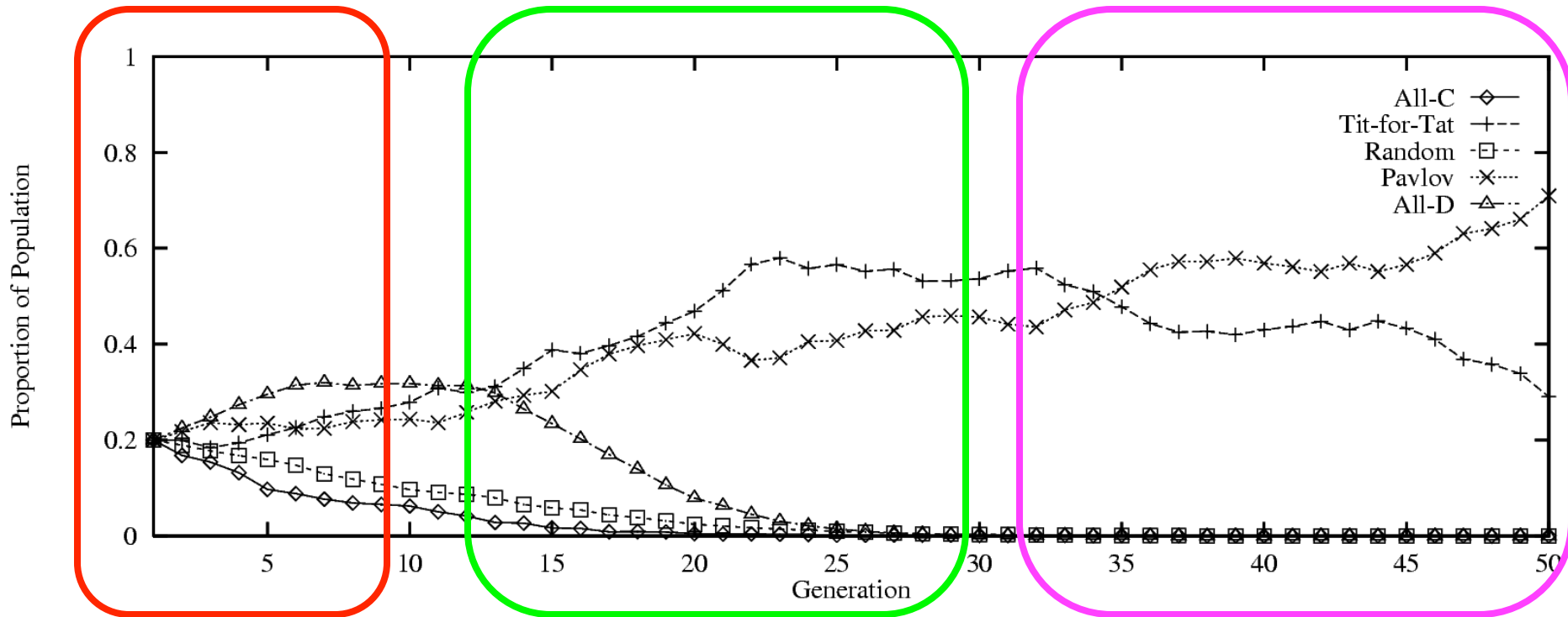
Flake's Simulation with Noise

- $R(t)$ is computed over r rounds
- $A_{ik}(j)$ = action of strategy i playing against strategy k in round j
 - Normal strategy i action with probability $1 - p_n$
 - Random C/D with probability p_n
- Note that this overestimates effects of noise

$$R_{ik}(t) = \sum_{j=1}^r \text{payoff} [A_{ik}(j) A_{ki}(j)]$$

Simulation with Noise

- 20% each
- 0.5% noise



Run Flake's EIPD with Noise

[EIPD-cbn.nlogo](#)

Spatial Effects

- Previous simulation assumes that each agent is equally likely to interact with each other
- So strategy interactions are proportional to fractions in population
- More realistically, interactions with “neighbors” are more likely
 - “Neighbor” can be defined in many ways
- Neighbors are more likely to use the same strategy

Spatial Simulation

- Toroidal grid
- Agent interacts only with eight neighbors
- Agent adopts strategy of most successful neighbor
- Ties favor current strategy

NetLogo Simulation of Spatial IPD

Run SIPD-async-alter.nlogo

Typical Simulation ($t = 1$)



Colors:

ALL-C

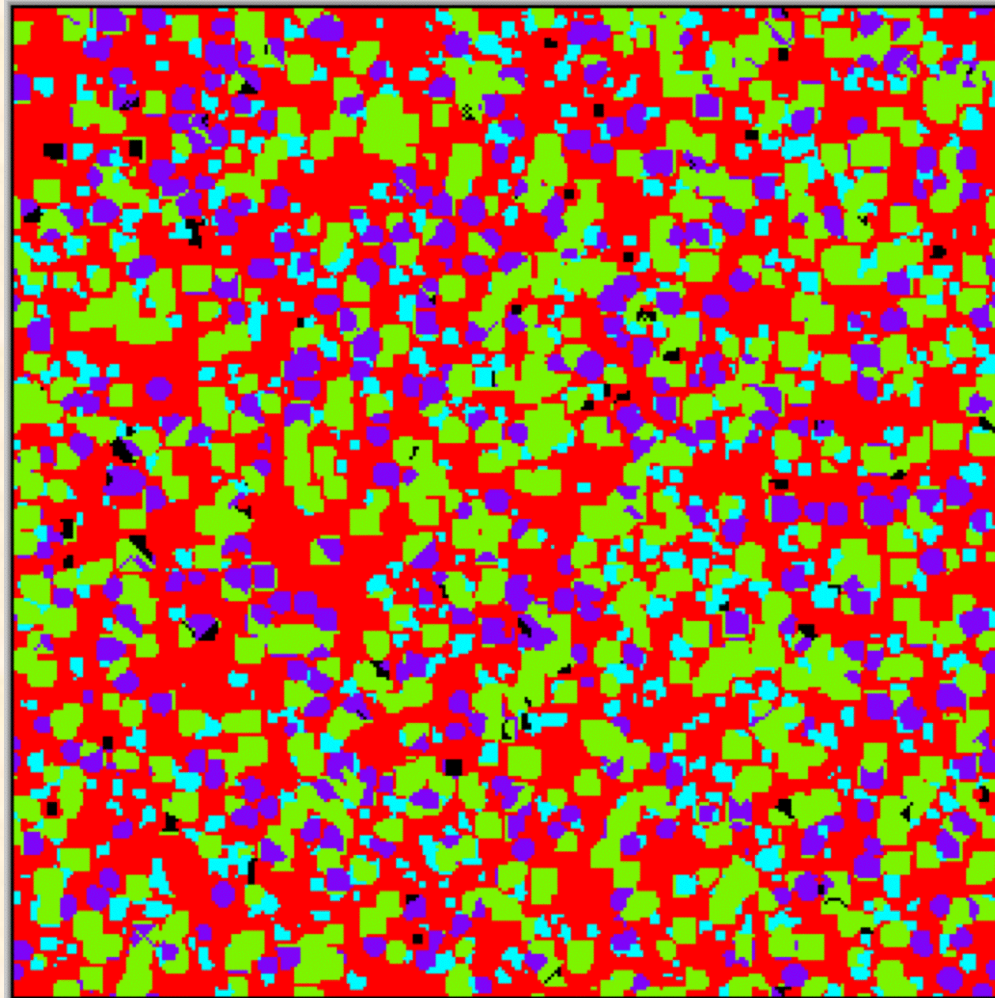
TFT

RAND

PAV

ALL-D

Typical Simulation ($t = 5$)



Colors:

ALL-C

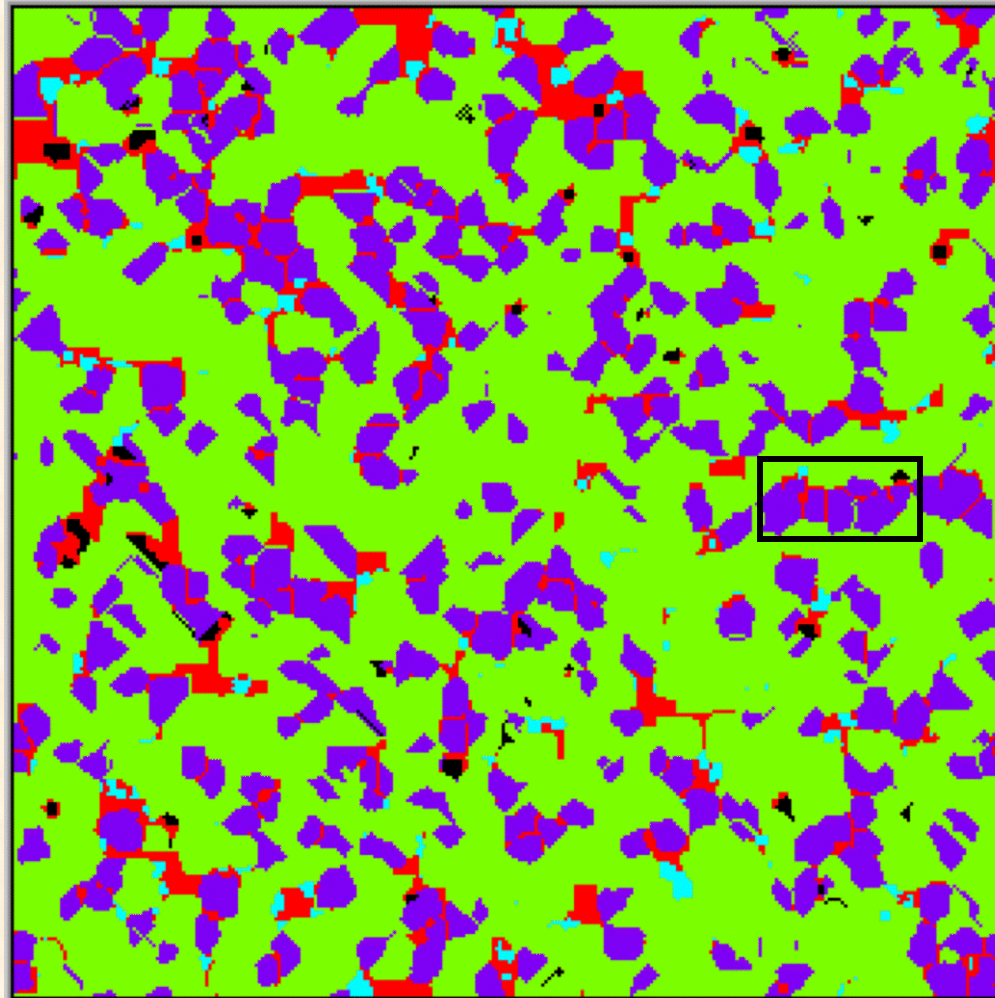
TFT

RAND

PAV

ALL-D

Typical Simulation ($t = 10$)



Colors:

ALL-C

TFT

RAND

PAV

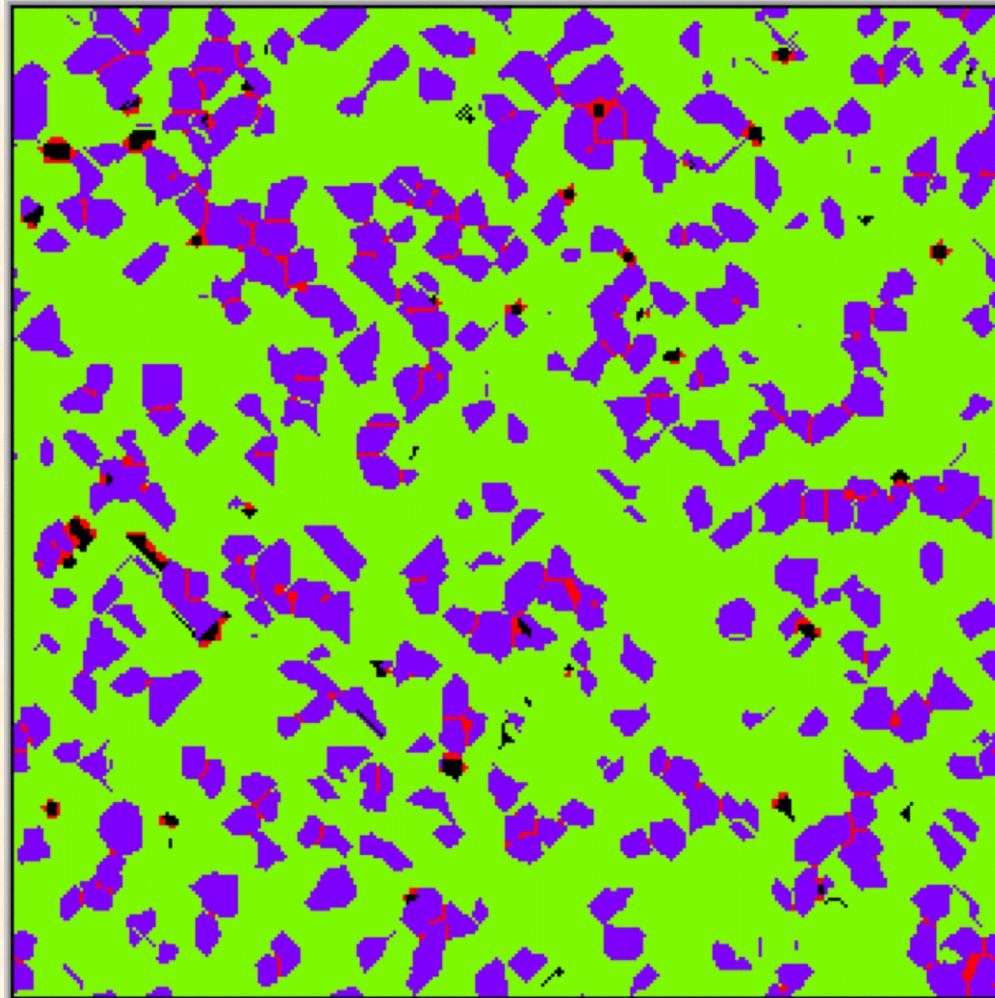
ALL-D

Typical Simulation ($t = 10$)

Zooming In



Typical Simulation ($t = 20$)



Colors:

ALL-C

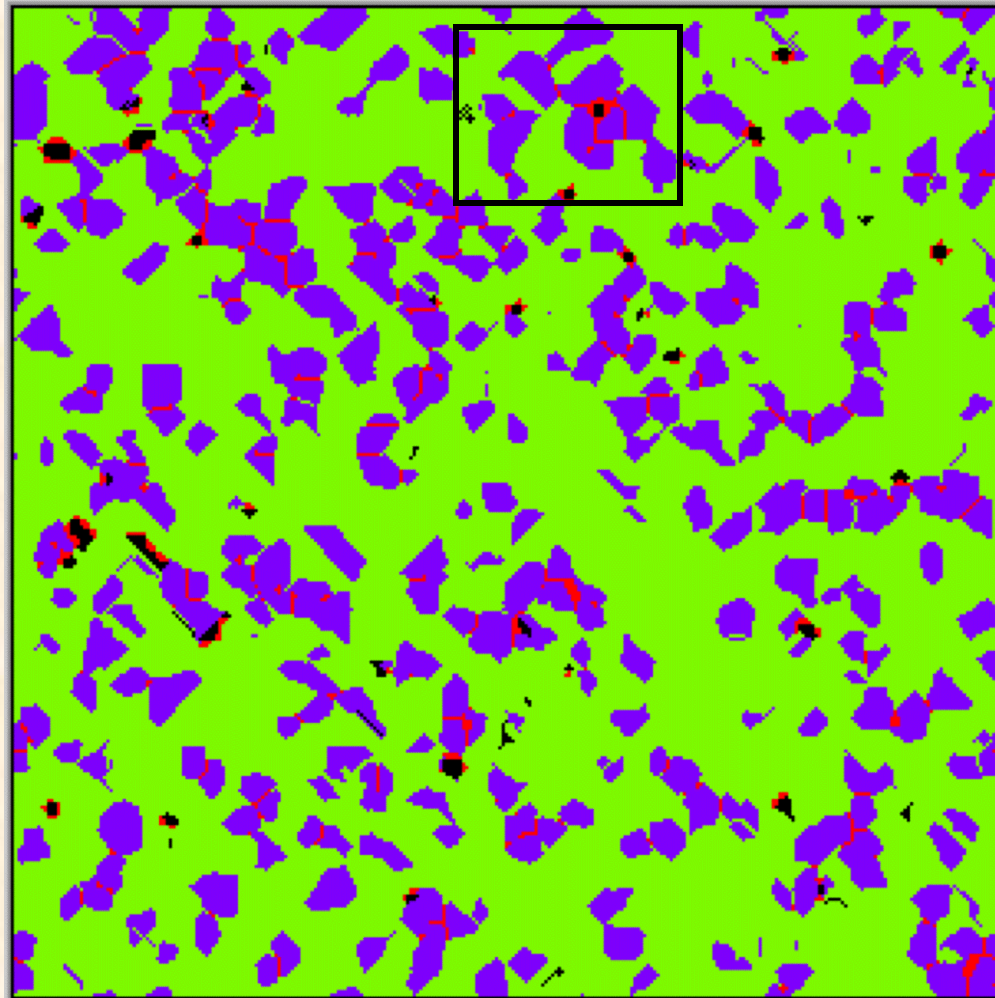
TFT

RAND

PAV

ALL-D

Typical Simulation ($t = 50$)



Colors:

ALL-C

TFT

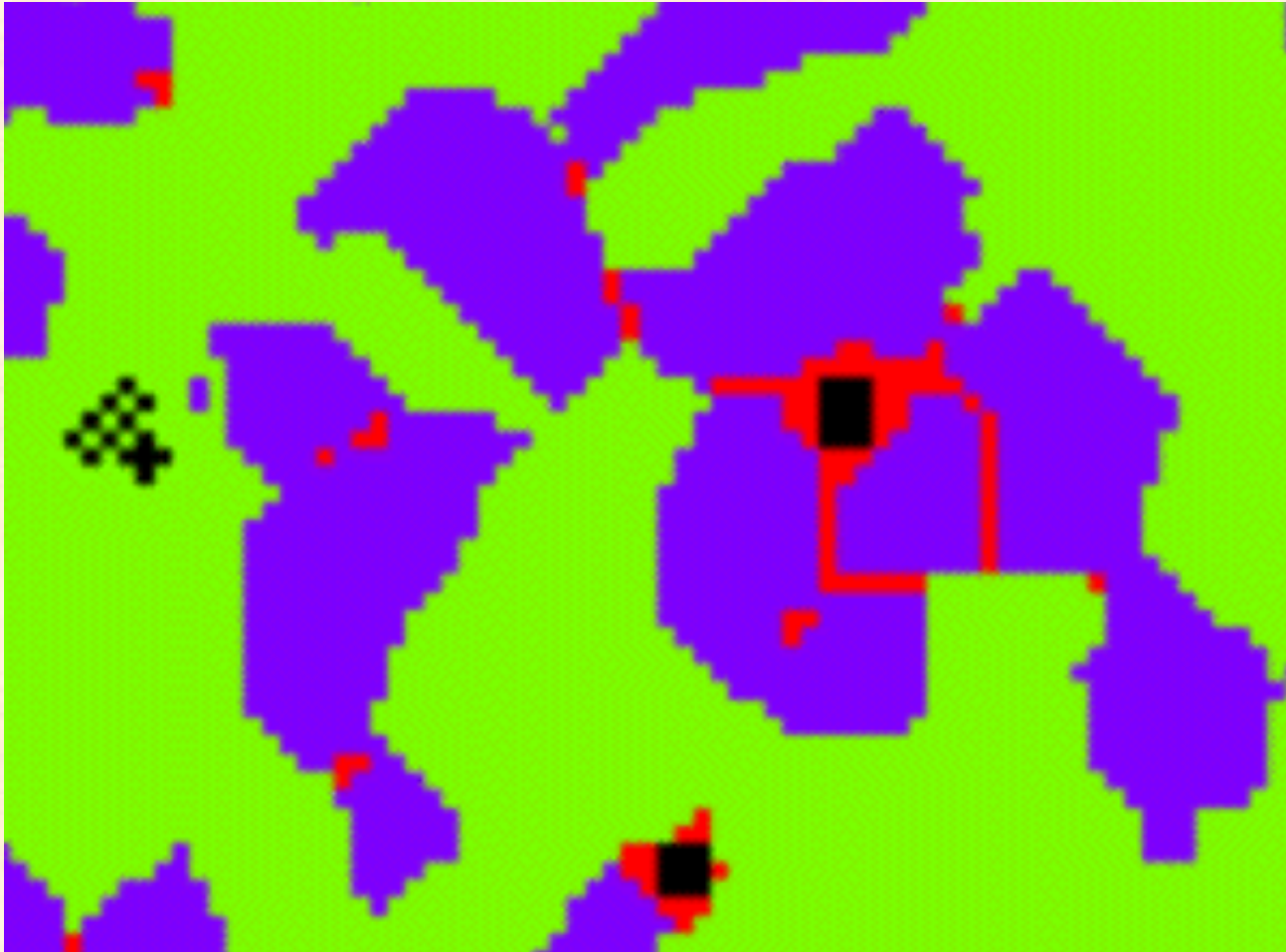
RAND

PAV

ALL-D






Typical Simulation ($t = 50$)

Zoom In



SIPD Without Noise

Legend

-  — All-C
-  — Tit-for-Tat
-  — Random
-  — Pavlov
-  — All-D

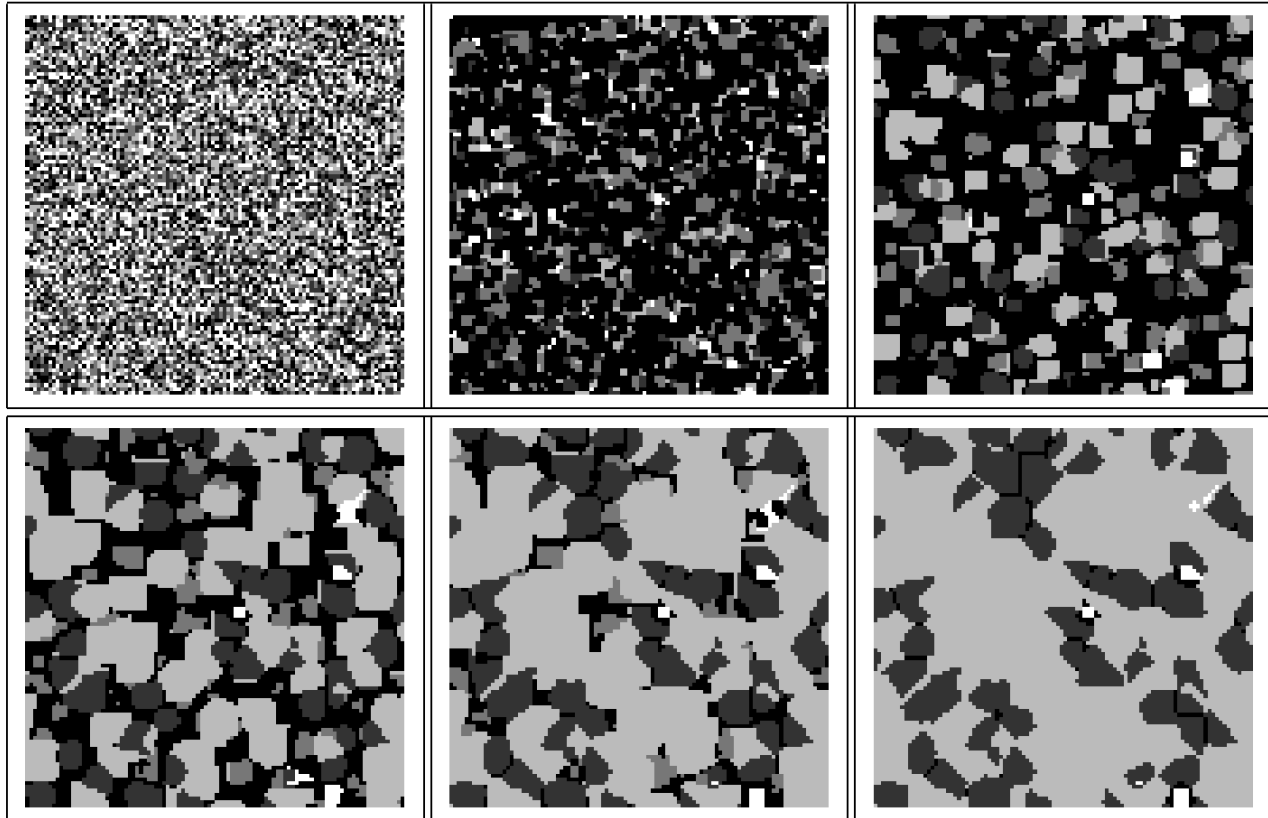


Figure 17.4 Competition in the spatial iterated Prisoner's Dilemma without noise

Figure from *The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation*. Copyright © 1998–2000 by Gary William Flake. All rights reserved. Permission granted for educational, scholarly, and personal use provided that this notice remains intact and unaltered. No part of this work may be reproduced for commercial purposes without prior written permission from the MIT Press.

Conclusions: Spatial IPD

- Small clusters of cooperators can exist in hostile environment
- Parasitic agents can exist only in limited numbers
- Stability of cooperation depends on expectation of future interaction
- Adaptive cooperation/defection beats unilateral cooperation or defection

NetLogo Simulation of Spatial IPD

[Run SIPD-async-alter.nlogo](#)

Additional Bibliography

1. von Neumann, J., & Morgenstern, O. *Theory of Games and Economic Behavior*, Princeton, 1944.
2. Morgenstern, O. “Game Theory,” in *Dictionary of the History of Ideas*, Charles Scribners, 1973, vol. 2, pp. 263-75.
3. Axelrod, R. *The Evolution of Cooperation*. Basic Books, 1984.
4. Axelrod, R., & Dion, D. “The Further Evolution of Cooperation,” *Science* **242** (1988): 1385-90.
5. Poundstone, W. *Prisoner’s Dilemma*. Doubleday, 1992.