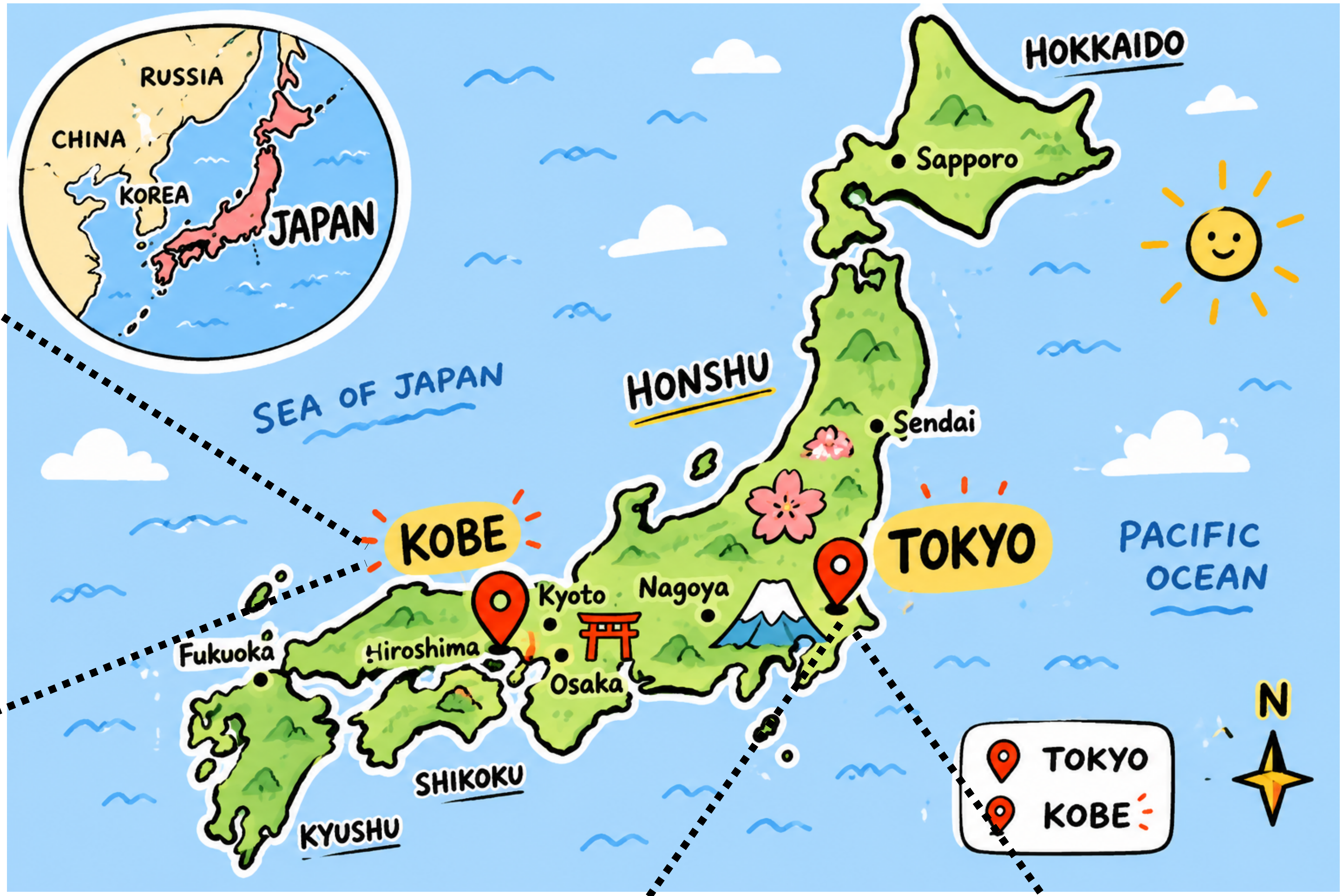
Computer Tournaments in Repeated Games: Past, Present, and Future

Nikoleta E. Glynatsi

2026 UBIQUITOUS TRUSTWORTHY INTELLIGENCE
INTERNATIONAL ACADEMIC CONFERENCE

RIKEN

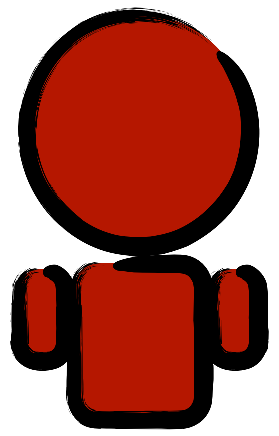
R-CCS

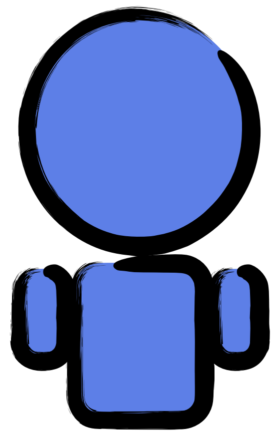


iTHEMS

https://ithems-math-soc.github.io/

Direct reciprocity





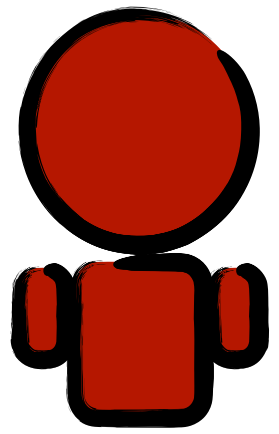
$$\begin{array}{c|cc} & C & D \\ \hline C & 3 & 0 \\ D & 5 & 1 \end{array}$$

Nash Equilibrium

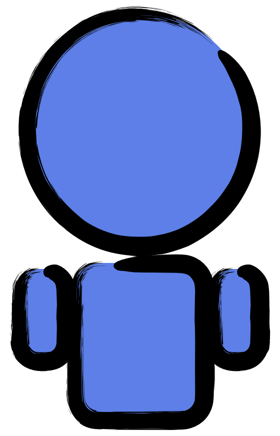
$$T > R > P > S$$

$$2R > T + S$$

$$\boxed{1}\ \begin{array}{ccc} & C & D \\ C & 3 & 0 \\ D & 5 & 1 \end{array} \qquad \boxed{2}\ \begin{array}{ccc} & C & D \\ C & 3 & 0 \\ D & 5 & 1 \end{array} \quad \cdots \quad \boxed{n}\ \begin{array}{ccc} & C & D \\ C & 3 & 0 \\ D & 5 & 1 \end{array} \qquad \boxed{n+1}\ \begin{array}{ccc} & C & D \\ C & 3 & 0 \\ D & 5 & 1 \end{array}$$



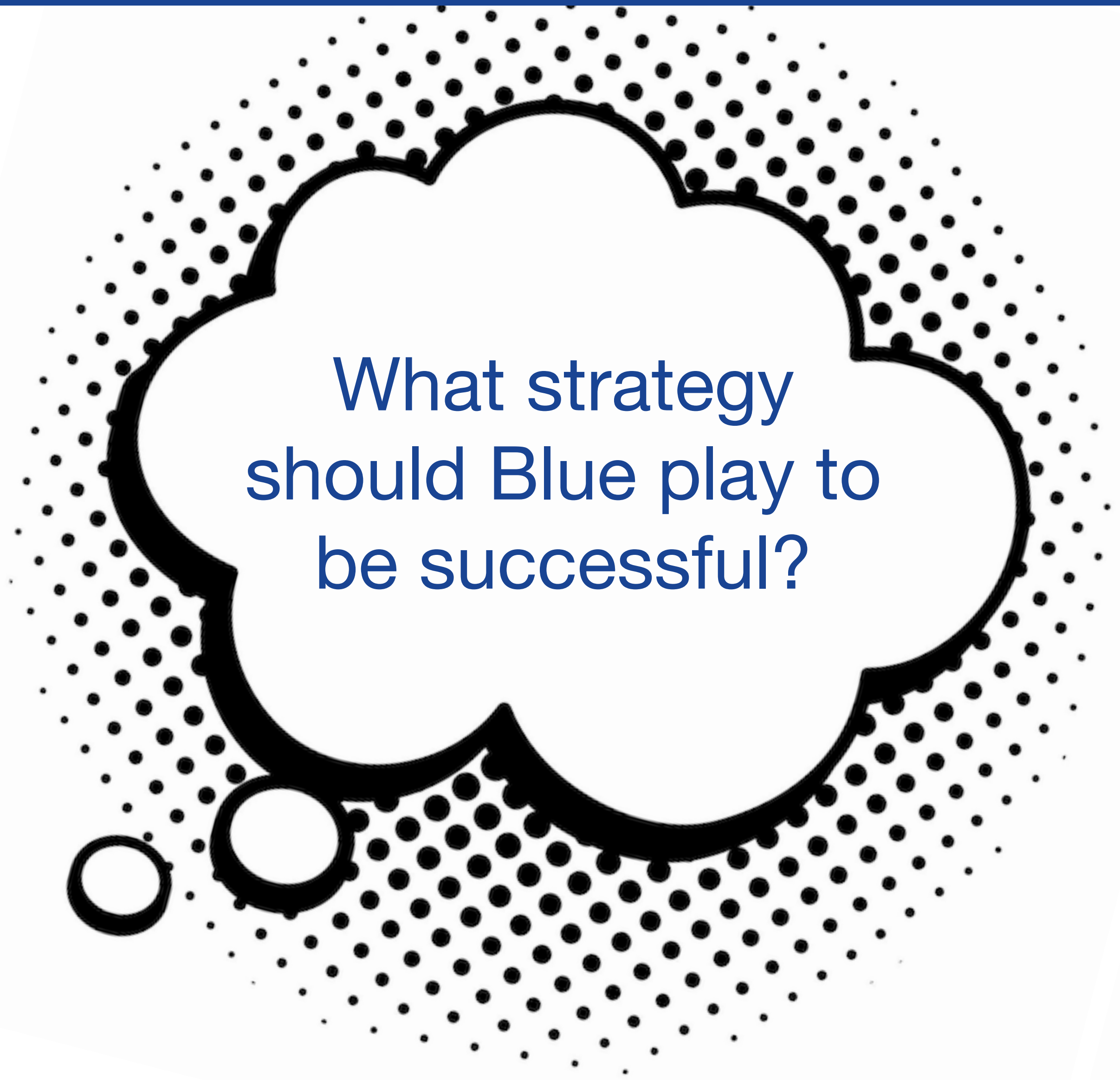
$C \quad D \quad \cdots \quad C \quad C/D$

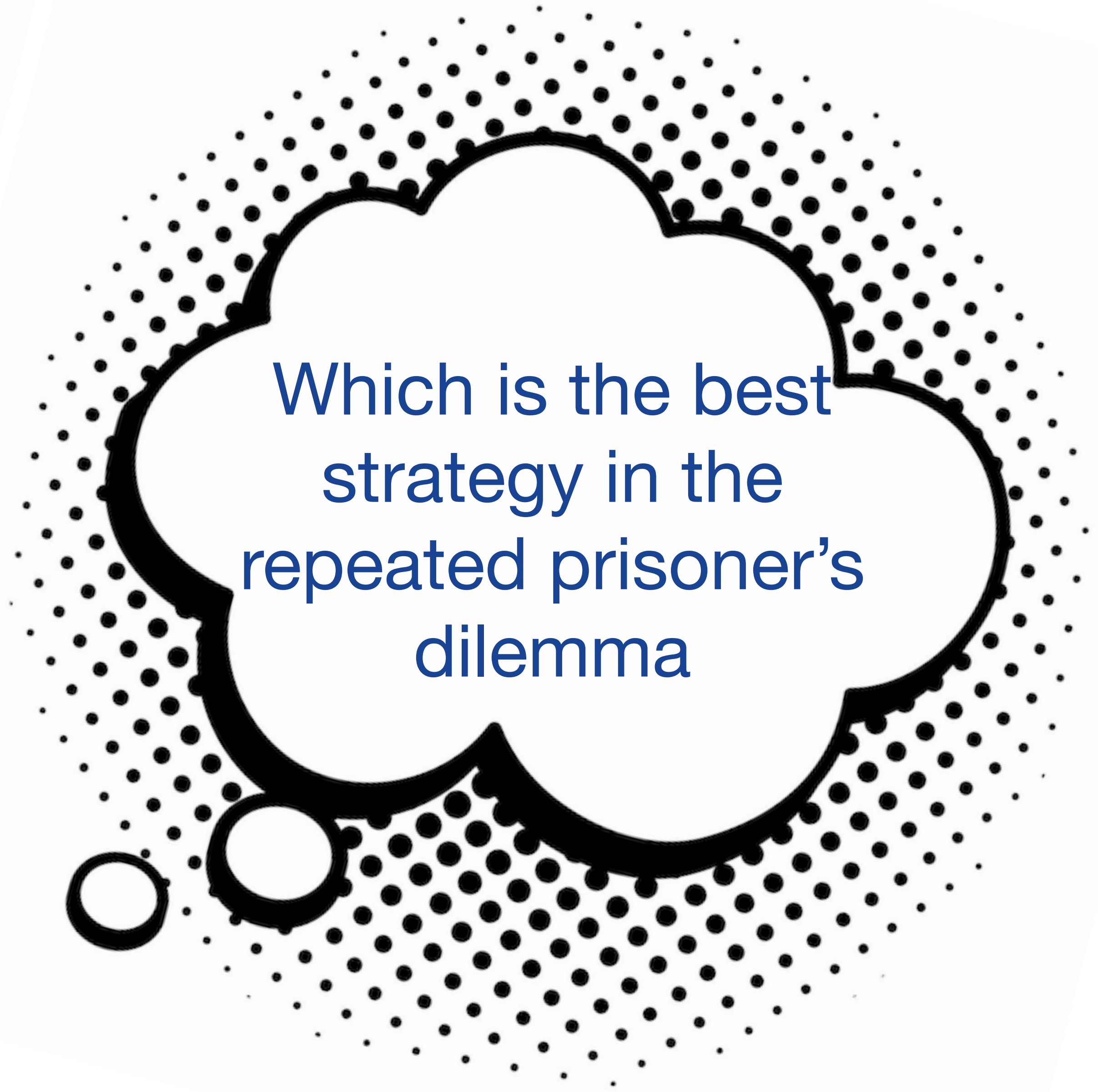


$D \quad C \quad \cdots \quad D$

Strategies: determine behavior from observations and past interactions.

Computer tournaments (1980)





Robert Axelrod

Axelrod’s First

RULES

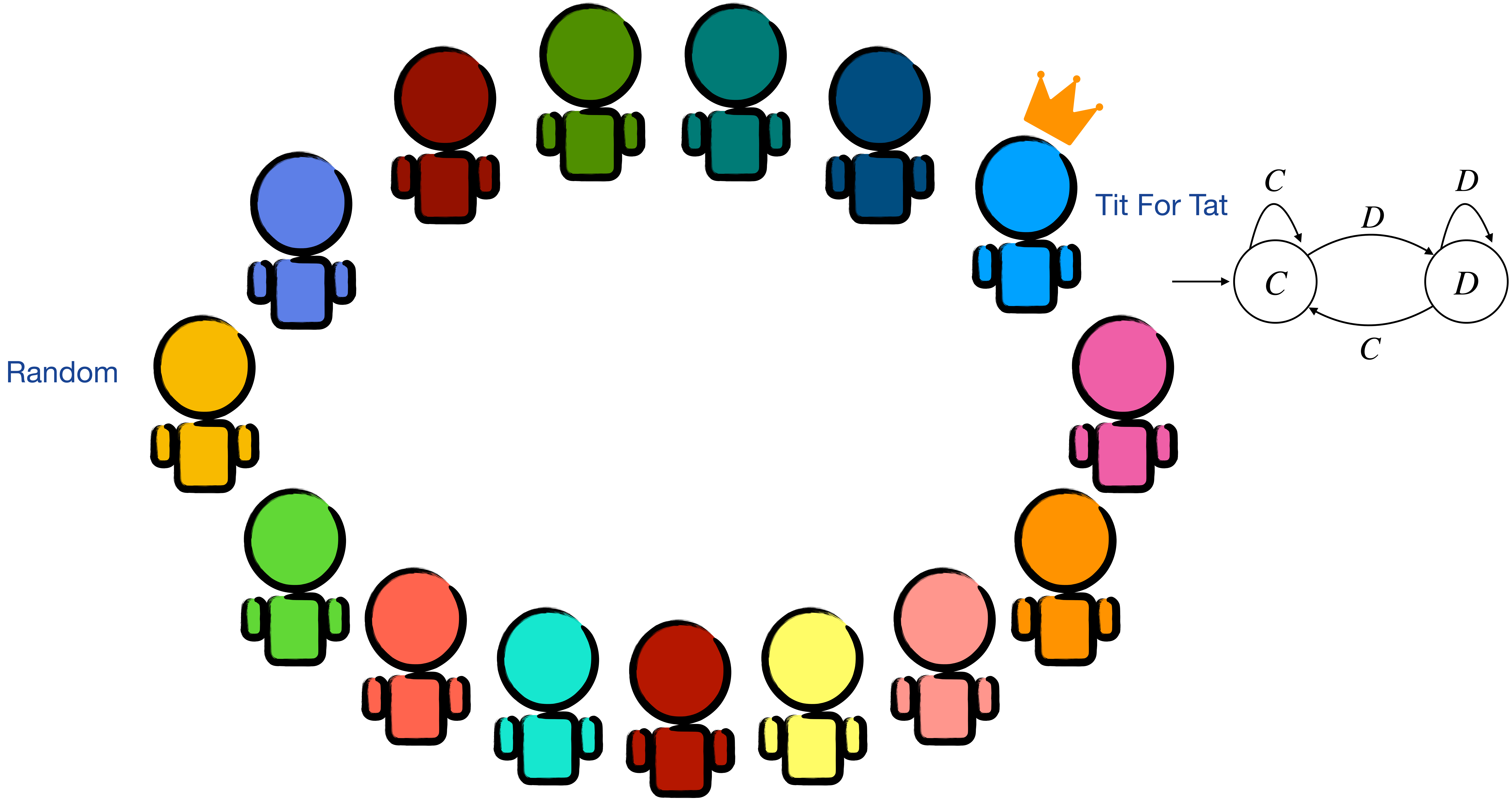
- Strategies are submitted as computer code.
- Each strategy plays a match against every other strategy, a copy of itself, and the Random strategy.
- Each match consists of 200 rounds.
- The tournament is repeated five times.
- The strategy with the highest average payoff across all matches is declared the winner.

Effective Choice in the Prisoner’s Dilemma

ROBERT AXELROD
Institute of Public Policy Studies
University of Michigan

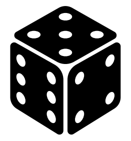
This is a “primer” on how to play the iterated Prisoner’s Dilemma game effectively. Existing research approaches offer the participant limited help in understanding how to cope effectively with such interactions. To gain a deeper understanding of how to be effective in such a partially competitive and partially cooperative environment, a computer tournament was conducted for the iterated Prisoner’s Dilemma. Decision rules were submitted by entrants who were recruited primarily from experts in game theory from a variety of disciplines: psychology, political science, economics, sociology, and mathematics. The results of the tournament demonstrate that there are subtle reasons for an individualistic pragmatist to cooperate as long as the other side does, to be somewhat forgiving, and to be optimistic about the other side’s responsiveness.

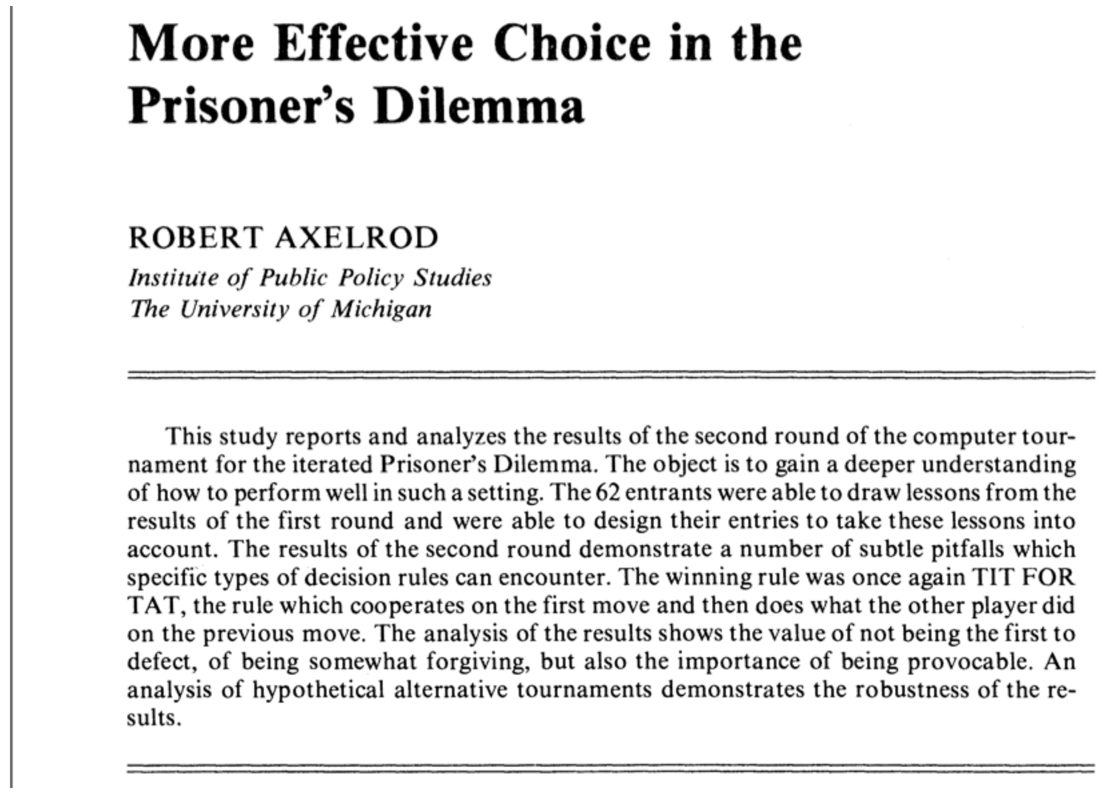
Axelrod’s First



Axelrod’s Second

SECOND TOURNAMENT RULES

- Strategies are submitted as computer code.
- Each strategy plays a match against every other strategy, a copy of itself, and the Random strategy.
- Each match consists of  rounds.
- The tournament is repeated five times.
- The strategy with the highest average payoff across all matches is declared the winner.



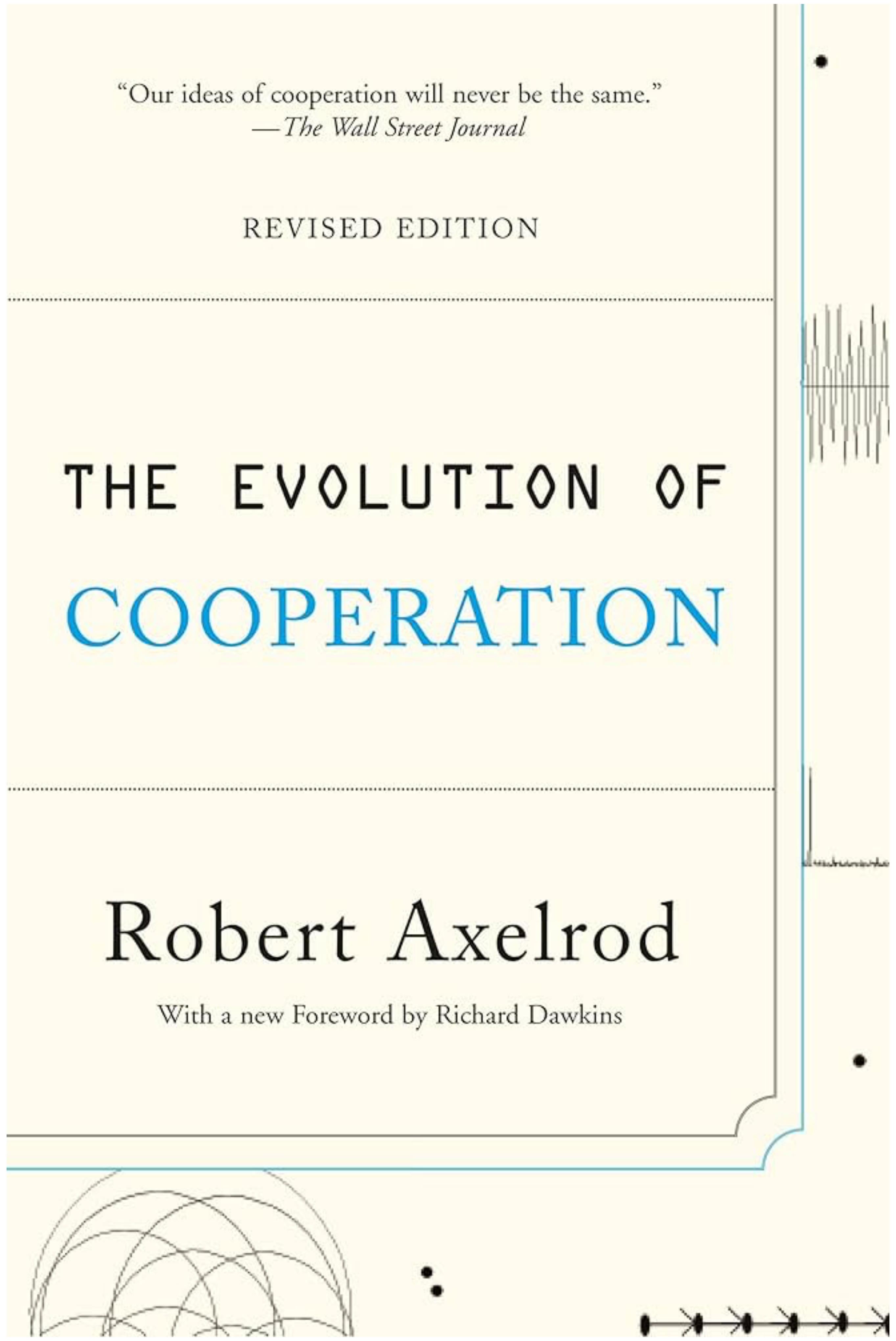
More Effective Choice in the Prisoner’s Dilemma

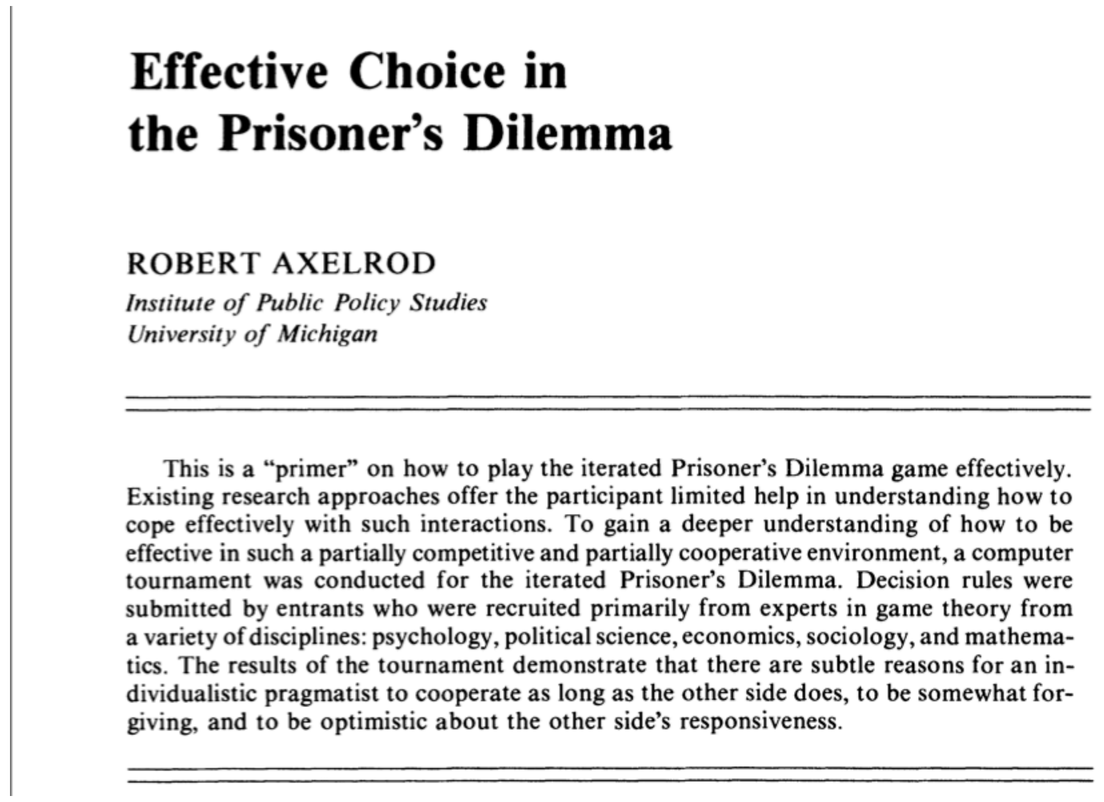
ROBERT AXELROD
Institute of Public Policy Studies
The University of Michigan

This study reports and analyzes the results of the second round of the computer tournament for the iterated Prisoner’s Dilemma. The object is to gain a deeper understanding of how to perform well in such a setting. The 62 entrants were able to draw lessons from the results of the first round and were able to design their entries to take these lessons into account. The results of the second round demonstrate a number of subtle pitfalls which specific types of decision rules can encounter. The winning rule was once again TIT FOR TAT, the rule which cooperates on the first move and then does what the other player did on the previous move. The analysis of the results shows the value of not being the first to defect, of being somewhat forgiving, but also the importance of being provocable. An analysis of hypothetical alternative tournaments demonstrates the robustness of the results.

Axelrod’s Second

Evolution of cooperation

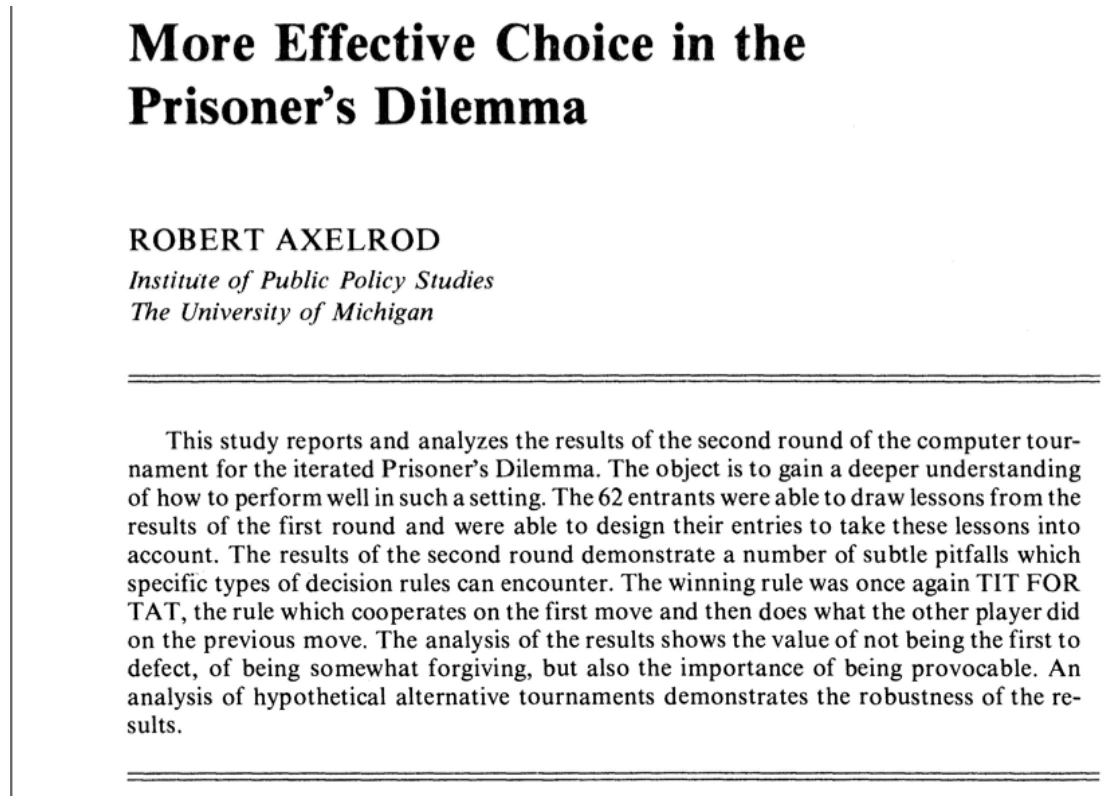




Effective Choice in the Prisoner's Dilemma

ROBERT AXELROD
Institute of Public Policy Studies
University of Michigan

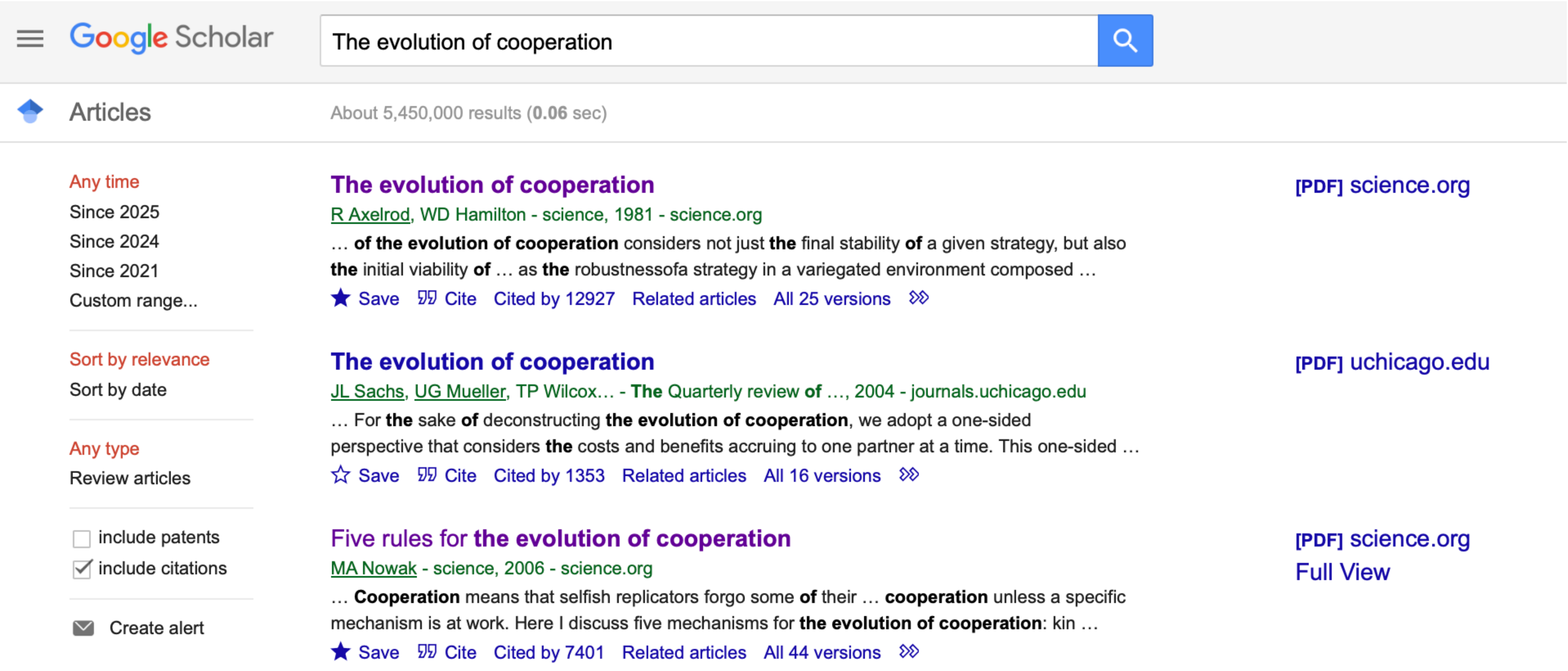
This is a "primer" on how to play the iterated Prisoner's Dilemma game effectively. Existing research approaches offer the participant limited help in understanding how to cope effectively with such interactions. To gain a deeper understanding of how to be effective in such a partially competitive and partially cooperative environment, a computer tournament was conducted for the iterated Prisoner's Dilemma. Decision rules were submitted by entrants who were recruited primarily from experts in game theory from a variety of disciplines: psychology, political science, economics, sociology, and mathematics. The results of the tournament demonstrate that there are subtle reasons for an individualistic pragmatist to cooperate as long as the other side does, to be somewhat forgiving, and to be optimistic about the other side's responsiveness.



More Effective Choice in the Prisoner's Dilemma

ROBERT AXELROD
Institute of Public Policy Studies
The University of Michigan

This study reports and analyzes the results of the second round of the computer tournament for the iterated Prisoner's Dilemma. The object is to gain a deeper understanding of how to perform well in such a setting. The 62 entrants were able to draw lessons from the results of the first round and were able to design their entries to take these lessons into account. The results of the second round demonstrate a number of subtle pitfalls which specific types of decision rules can encounter. The winning rule was once again TIT FOR TAT, the rule which cooperates on the first move and then does what the other player did on the previous move. The analysis of the results shows the value of not being the first to defect, of being somewhat forgiving, but also the importance of being provocable. An analysis of hypothetical alternative tournaments demonstrates the robustness of the results.



Evolution of cooperation

Strategies

Human-designed strategies

Evolutionary search

Machine learning

Success criteria

Evolutionary stability

Cooperation in equilibrium

<u>Tournament performance</u>

Robustness to noise

Properties

Strategy classification

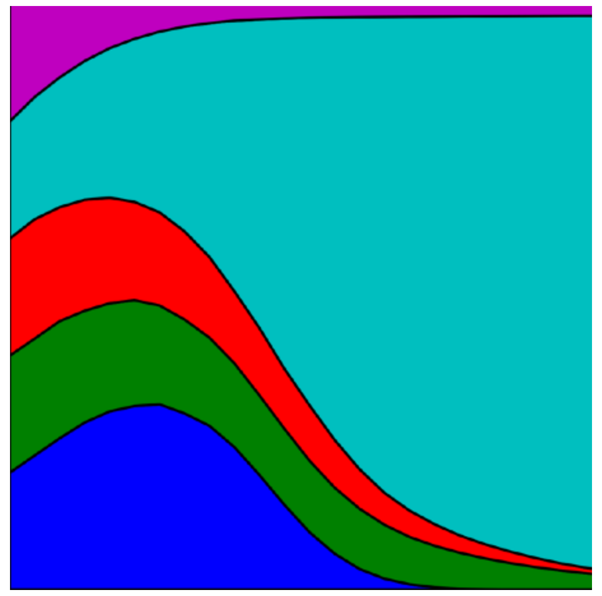
Behavioral fingerprints

Strategy complexity

Many questions are still studied in isolation

Limited access to source code

Axelrod-Python (mid-2010s)

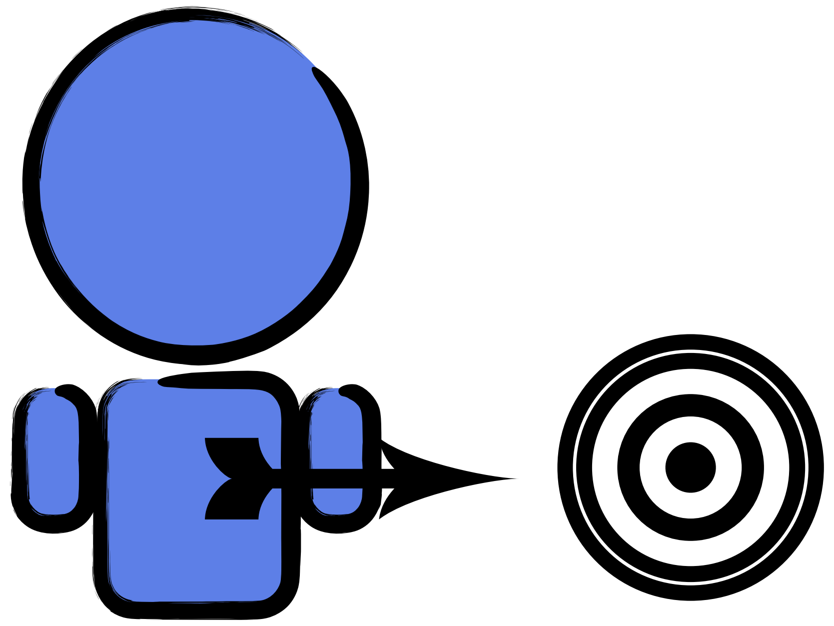


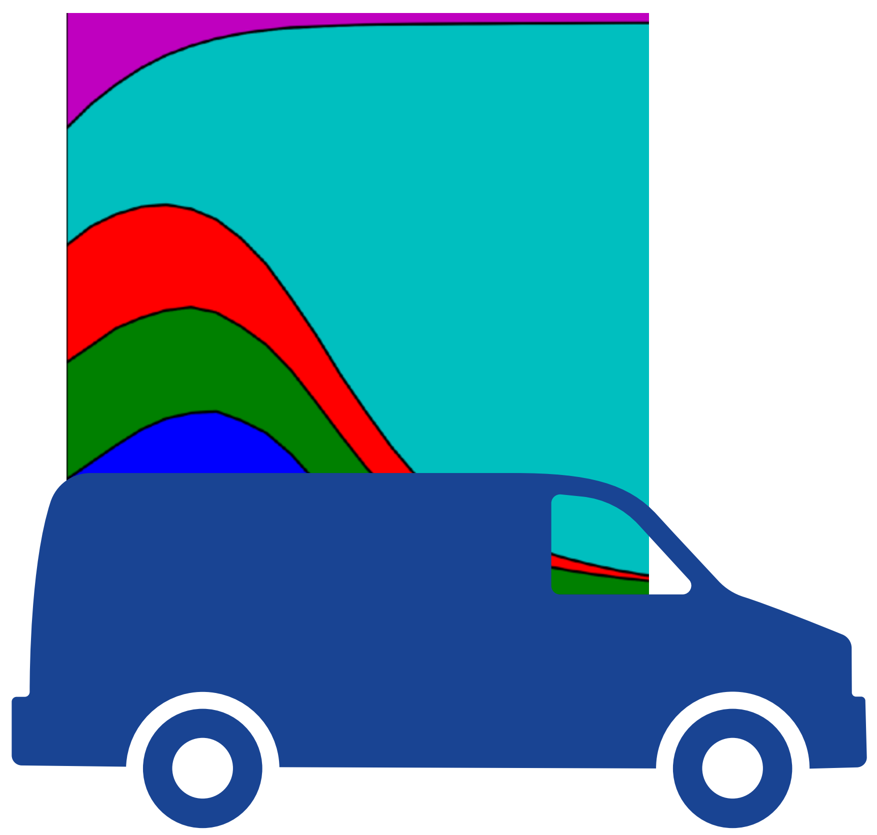
https://axelrod.readthedocs.io/en/stable/

- Open-source Python library for the Iterated Prisoner's Dilemma
- Reproduces and extends Axelrod's original tournaments
- Large collection of implemented strategies
- Reproducible and extensible computational experiments

[1] An open framework for the reproducible study of the iterated prisoner's dilemma. Journal of Open Research Software. https://doi.org/10.5334/jors.125

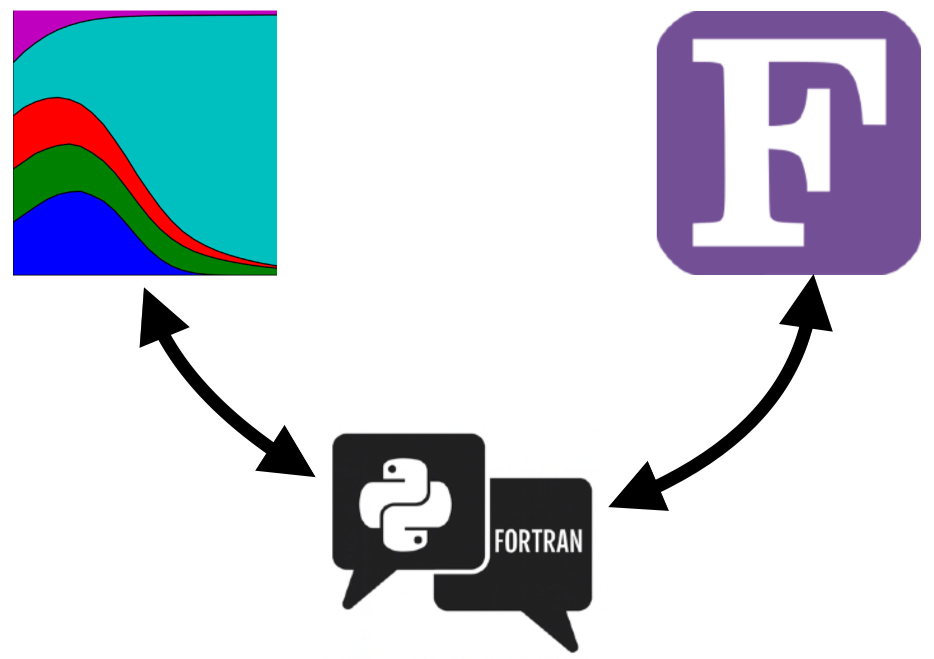
1. Training Strategies



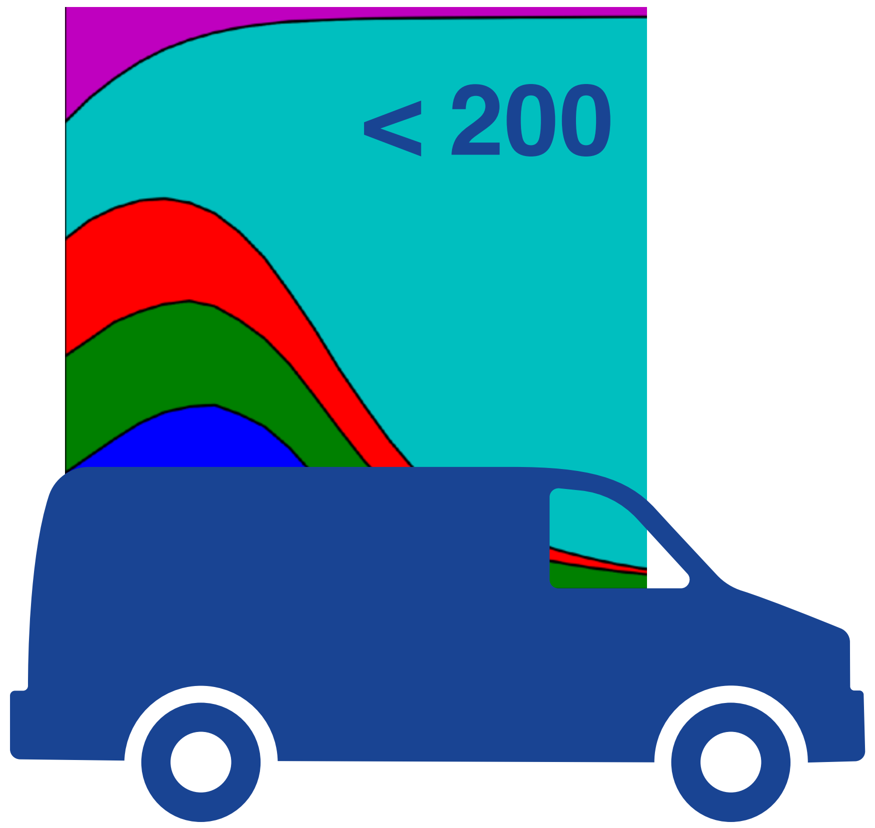


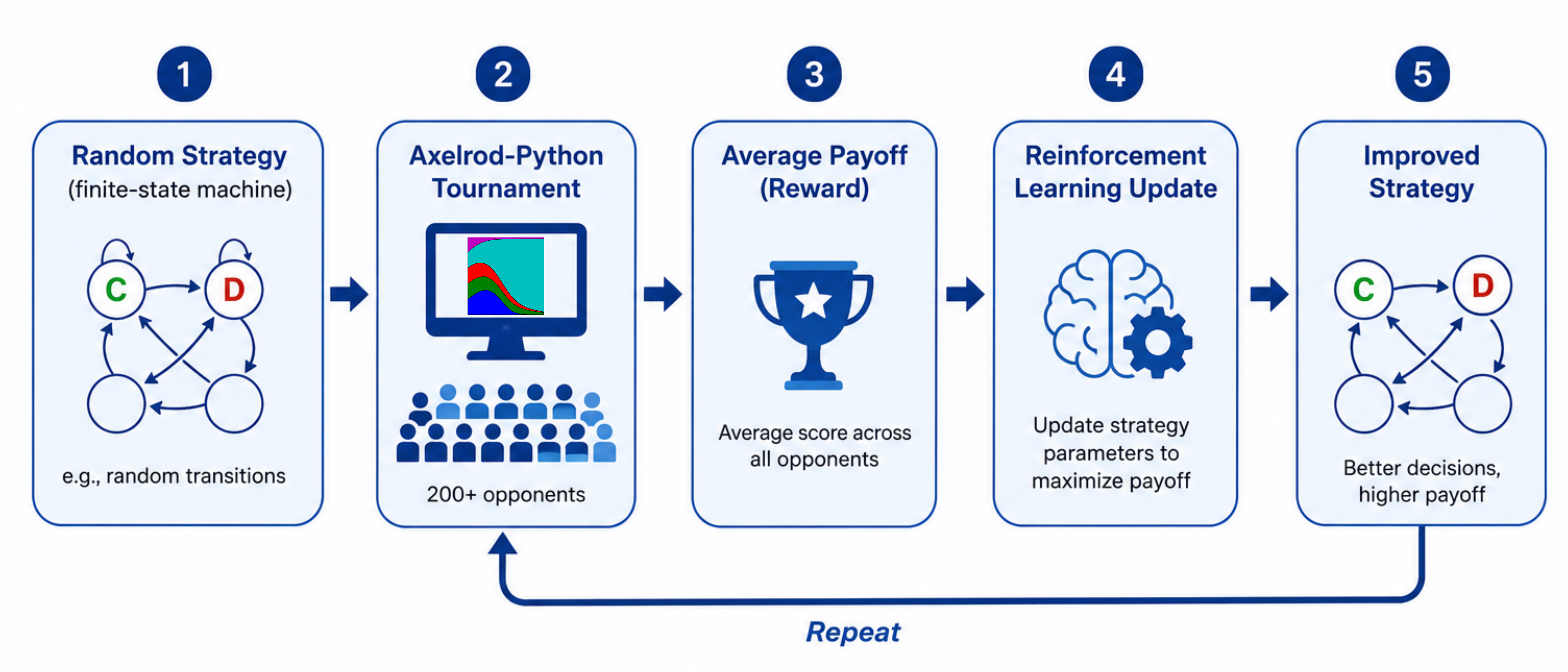
2. Successful Properties

3. Reproducing Axelrod's second



1. Training Strategies





Can we train strategies using reinforcement learning?

1. Training Strategies

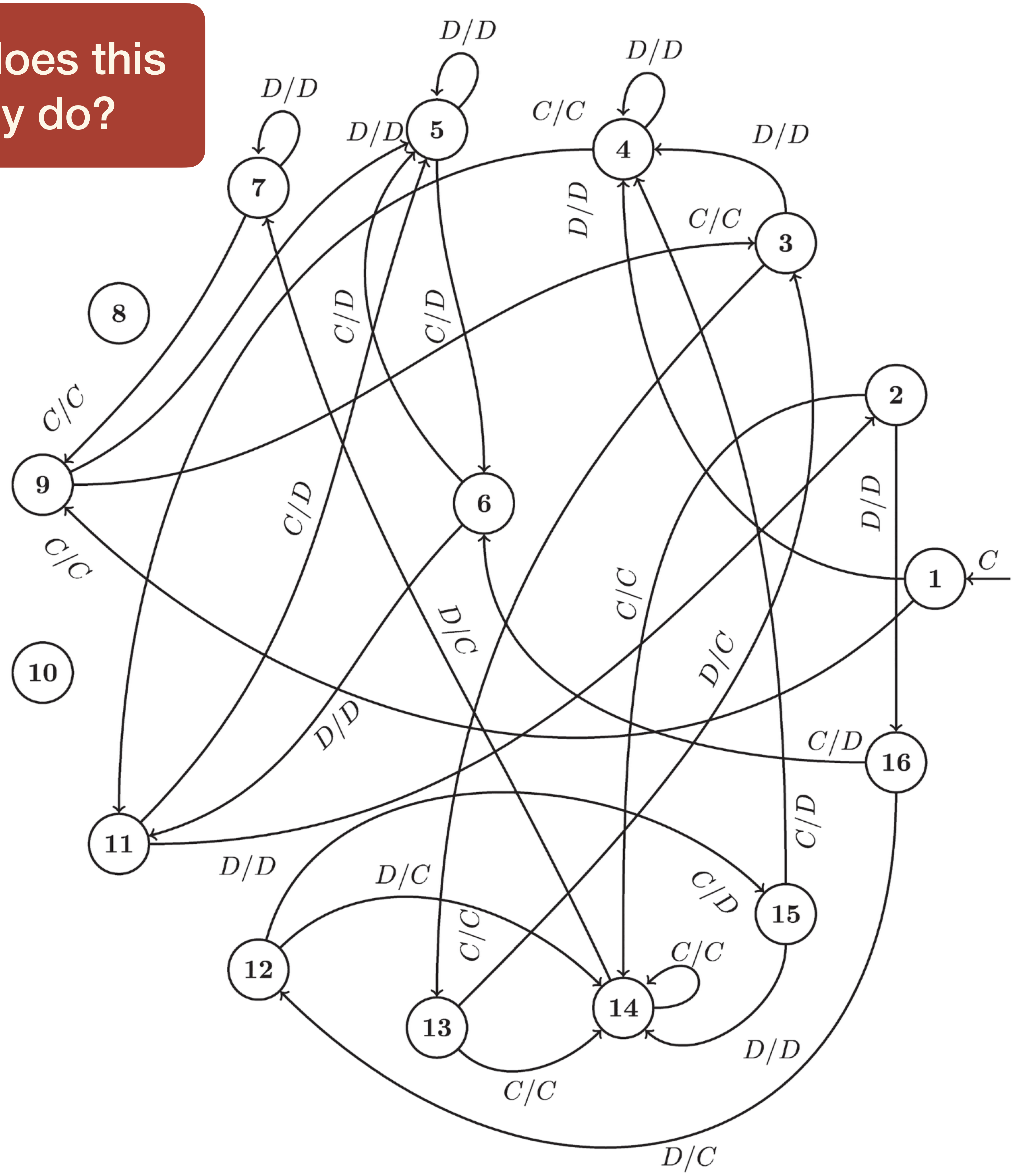
Top performing strategies in a tournament with over 200 strategies.

	mean	std	min	5%	25%	50%	75%	95%	max
EvolvedLookerUp2_2_2*	2.955	0.010	2.915	2.937	2.948	2.956	2.963	2.971	2.989
Evolved HMM 5*	2.954	0.014	2.903	2.931	2.945	2.954	2.964	2.977	3.007
Evolved FSM 16*	2.952	0.013	2.900	2.930	2.943	2.953	2.962	2.973	2.993
PSO Gambler 2_2_2*	2.938	0.013	2.884	2.914	2.930	2.940	2.948	2.957	2.972
Evolved FSM 16 Noise 05*	2.919	0.013	2.874	2.898	2.910	2.919	2.928	2.939	2.965
PSO Gambler 1_1_1*	2.912	0.023	2.805	2.874	2.896	2.912	2.928	2.950	3.012
Evolved ANN 5*	2.912	0.010	2.871	2.894	2.905	2.912	2.919	2.928	2.945
Evolved FSM 4*	2.910	0.012	2.867	2.889	2.901	2.910	2.918	2.929	2.943
Evolved ANN*	2.907	0.010	2.865	2.890	2.900	2.908	2.914	2.923	2.942
PSO Gambler Mem1*	2.901	0.025	2.783	2.858	2.884	2.901	2.919	2.942	2.994
Evolved ANN 5 Noise 05*	2.864	0.008	2.830	2.850	2.858	2.865	2.870	2.877	2.891
DBS	2.857	0.009	2.823	2.842	2.851	2.857	2.863	2.872	2.899
Winner12	2.849	0.008	2.820	2.836	2.844	2.850	2.855	2.862	2.874
Fool Me Once	2.844	0.008	2.818	2.830	2.838	2.844	2.850	2.857	2.882
Omega TFT: 3, 8	2.841	0.011	2.800	2.822	2.833	2.841	2.849	2.859	2.882

https://doi.org/10.1371/journal.pone.0188046.t002

1. Training Strategies

What does this strategy do?



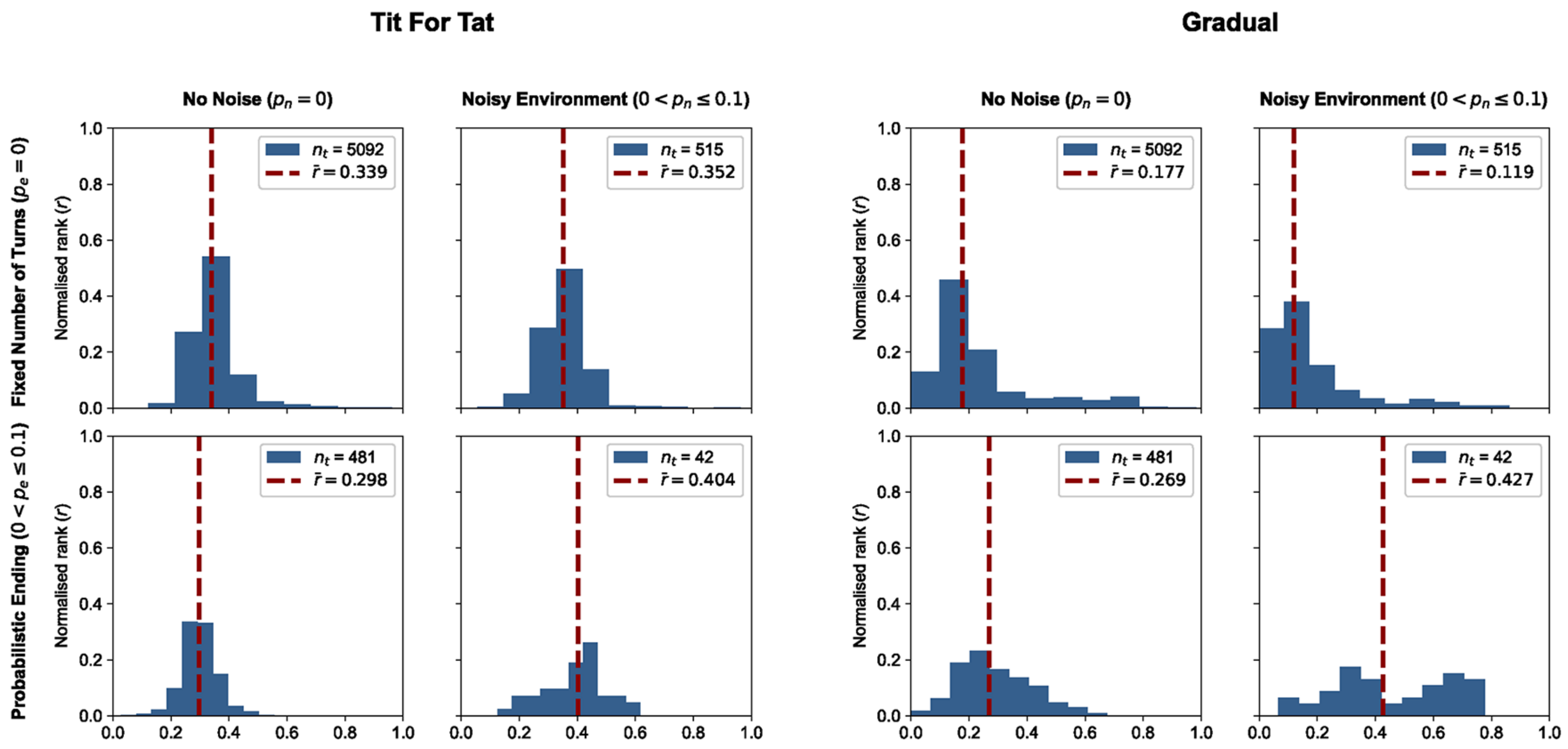
- From the training emerged strategies that were cooperative but also took advantage of simple strategies
- Strategies developed properties that were not hard coded, like handshakes

[2] Reinforcement learning produces dominant strategies for the iterated prisoner's dilemma. PloS one. https://doi.org/10.1371/journal.pone.0188046

[3] Evolution reinforces cooperation with the emergence of self-recognition mechanisms. PloS one. https://doi.org/10.1371/journal.pone.0204981

2. Successful Properties

Do successful strategies have some common properties?



2. Successful Properties

	Standard		Noisy $p_n \leq 0.1$		Probabilistic ending $p_e \leq 0.1$		Noisy probabilistic ending	
	r	**median score**	**r**	**median score**	**r**	**median score**	**r**	**median score**
CC to *C* rate	-0.501	0.501	-0.210	0.194	-0.336	0.348	0.087	0.015
CD to *C* rate	0.226	-0.199	0.337	-0.235	0.458	-0.352	0.609	-0.372
DC to *C* rate	0.127	-0.100	0.227	-0.111	0.164	-0.105	0.410	-0.203
DD to *C* rate	0.412	-0.396	0.549	-0.391	0.433	-0.378	0.615	-0.407
C_r	-0.323	0.383	0.298	-0.051	-0.060	0.160	0.595	-0.213
C_{max}	0.000	0.050	-0.000	0.244	-0.000	0.079	-0.000	0.296
C_{min}	0.000	0.085	0.000	-0.070	0.000	0.128	0.000	0.000
C_{median}	0.000	0.209	0.000	0.572	-0.000	0.324	0.000	0.667
C_{mean}	0.000	0.229	-0.000	0.583	-0.000	0.354	-0.000	0.689
C_r / C_{max}	-0.323	0.381	0.307	-0.076	-0.060	0.156	0.608	-0.246
C_{min} / C_r	0.109	-0.080	-0.141	-0.011	0.024	0.029	-0.335	0.092
C_r / C_{median}	-0.330	0.353	0.326	-0.258	-0.065	0.111	0.614	-0.464
C_r / C_{mean}	-0.331	0.357	0.325	-0.228	-0.066	0.114	0.617	-0.431
N	-0.000	-0.009	-0.000	-0.017	-0.000	0.011	0.000	0.139
k	-0.000	-0.002	-0.000	-0.003	-0.000	0.010	-0.000	0.035
n	-0.000	-0.125	-0.000	-0.392	-	-	-	-
p_n	-	-	0.000	-0.244	-	-	0.000	-0.272
p_e	-	-	-	-	0.000	0.257	0.000	0.568
Make use of game	-0.003	-0.022	-0.047	0.014	-0.046	0.022	-0.110	0.057
Make use of length	-0.158	0.124	-0.224	0.139	-0.173	0.128	-0.206	0.115
SSE	0.473	-0.452	0.589	-0.412	0.458	-0.418	0.571	-0.383
stochastic	0.006	-0.024	0.010	-0.007	-0.001	0.001	-0.001	0.002
memory usage	-0.098	0.108	-0.080	0.114	-	-	-	-

https://doi.org/10.1371/journal.pcbi.1012644.t004

2. Successful Properties

1. Do not be envious by striving for a payoff larger than the opponent’s payoff.
2. Be “nice” by not being the first to defect.
3. Reciprocate both cooperation and defection; Be provocable to retaliation and forgiveness.
4. Do not be too clever by scheming to exploit the opponent.

1. Be a little bit envious
2. Be “nice” in non-noisy environments or when game lengths are longer
3. Reciprocate both cooperation and defection appropriately; Be provocable in tournaments with short matches, and generous in tournaments with noise
4. It’s ok to be clever
5. Adapt to the environment; Adjust to the mean population cooperation

[4] Properties of winning iterated prisoner's dilemm
strategies. PLOS Computational Biology.
https://doi.org/10.1371/journal.pcbi.1012644

Summary *(so far)*

Reinforcement Learning (2017)

Learned strategies outperform many classical strategies.
Machine learning discovers novel cooperative behaviours.

Self-Recognition (2018)

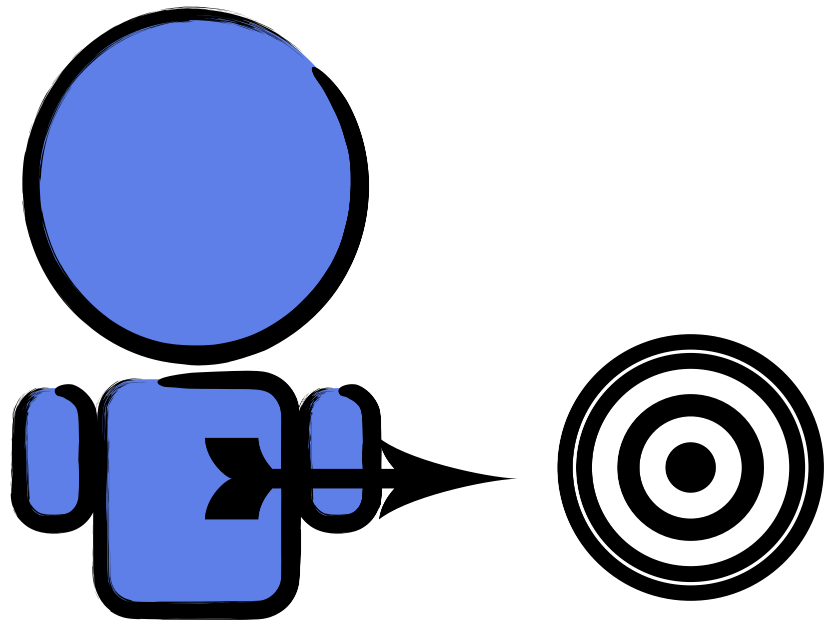
Evolution naturally produces self-recognition.
Self-recognition reinforces cooperation.

Winning Strategies (2024)

Characterised the properties of successful tournament strategies.
Strategies trained in environments with errors were more adaptable

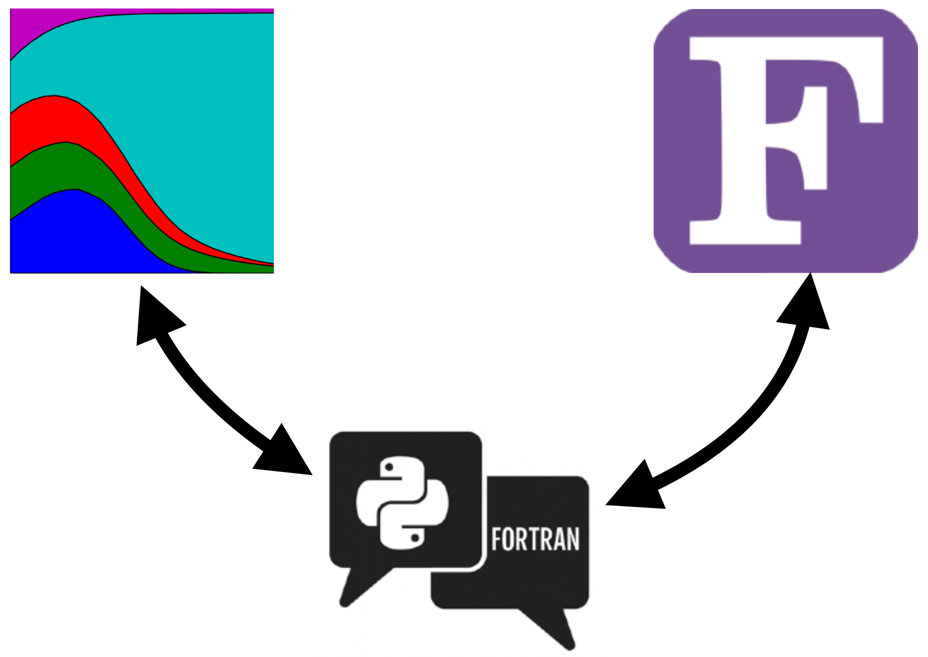
- Even simple games reveal rich social behaviours such as cooperation, exploitation, trust, and retaliation.
- Machine learning can discover highly successful strategies that are difficult for humans to interpret.
- **Future work:** bridge game theory, reinforcement learning, and interpretable AI to build autonomous agents whose social behaviour is both effective and understandable.

1. Training Strategies



2. Successful Properties

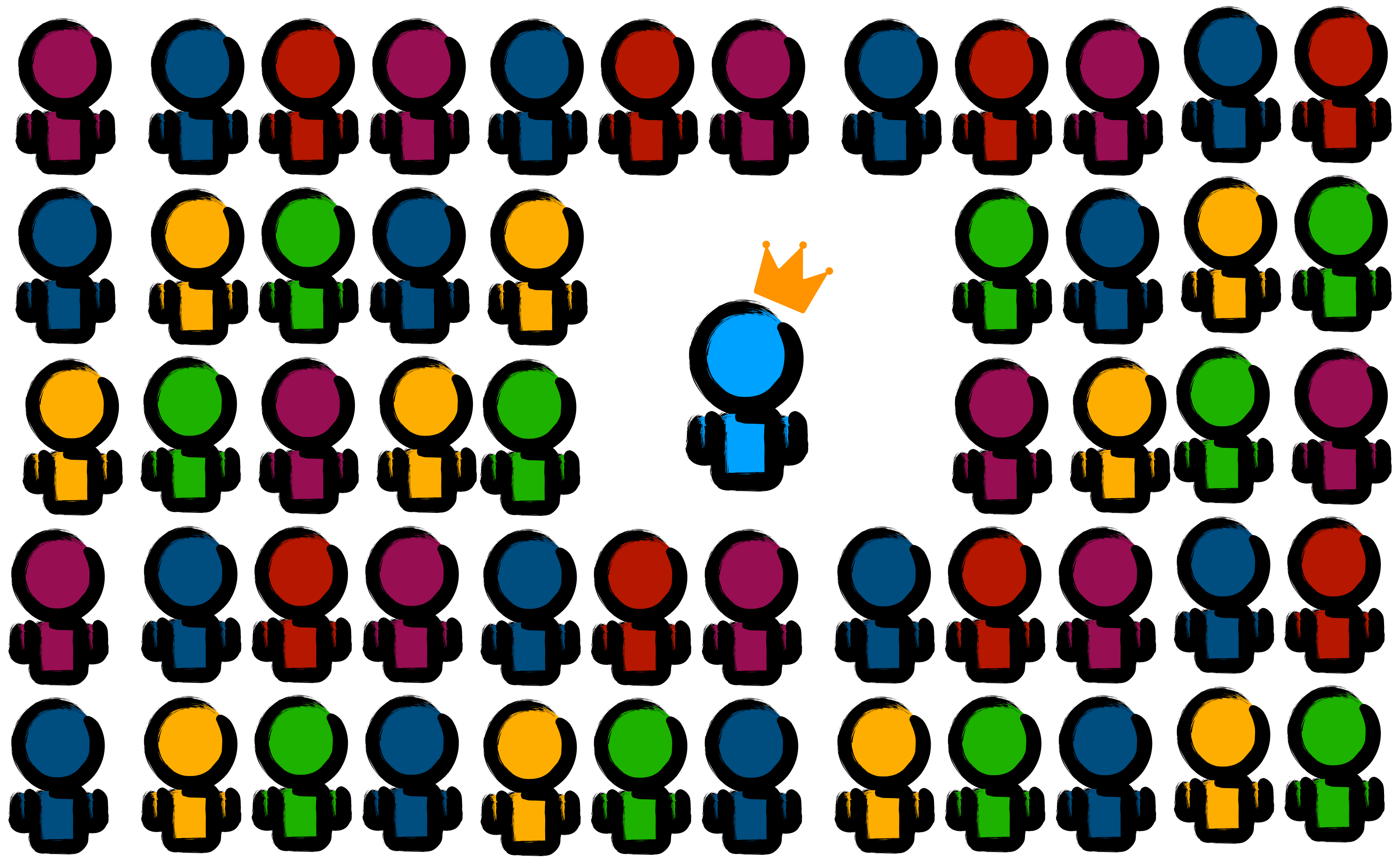
3. Reproducing Axelrod's second



3. Reproducing Axelrod’s second

14

63

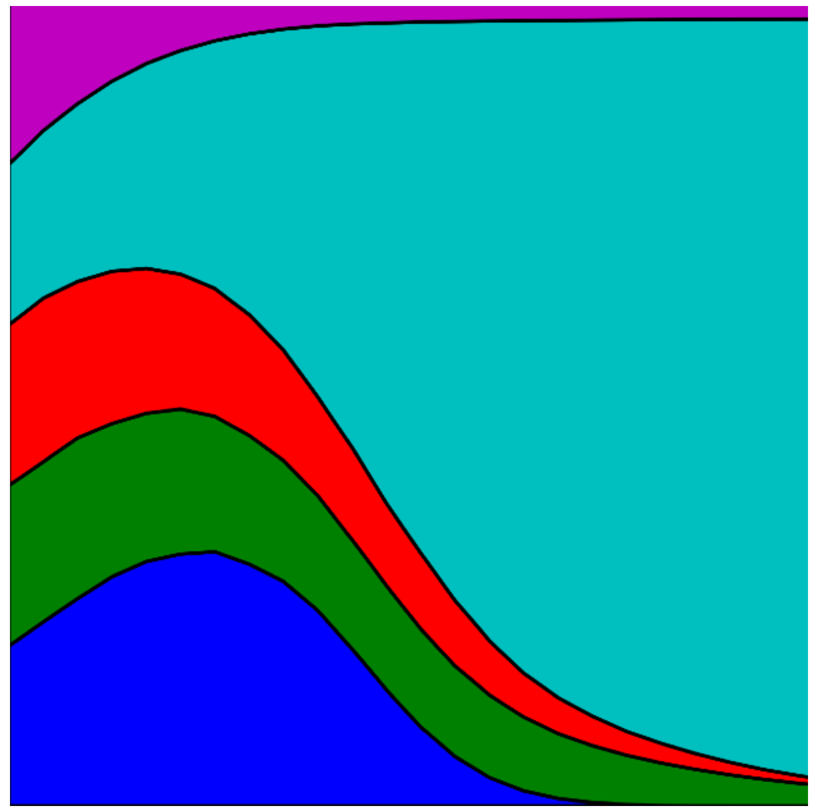


3. Reproducing Axelrod’s second

```
      FUNCTION K92R(J,M,K,L,R,JA)
C BY ANATOL RAPOPORT
C TYPED BY AX 3/27/79 (SAME AS ROUND ONE TIT FOR TAT)
c replaced by actual code, Ax 7/27/93
c  T=0
c    K92R=ITFTR(J,M,K,L,T,R)
      k92r=0
      k92r=j
c test 7/30
c    write(6,77) j, k92r
c77   format('␣test␣k92r.␣j,k92r:␣', 2i3)
      RETURN
      END
```

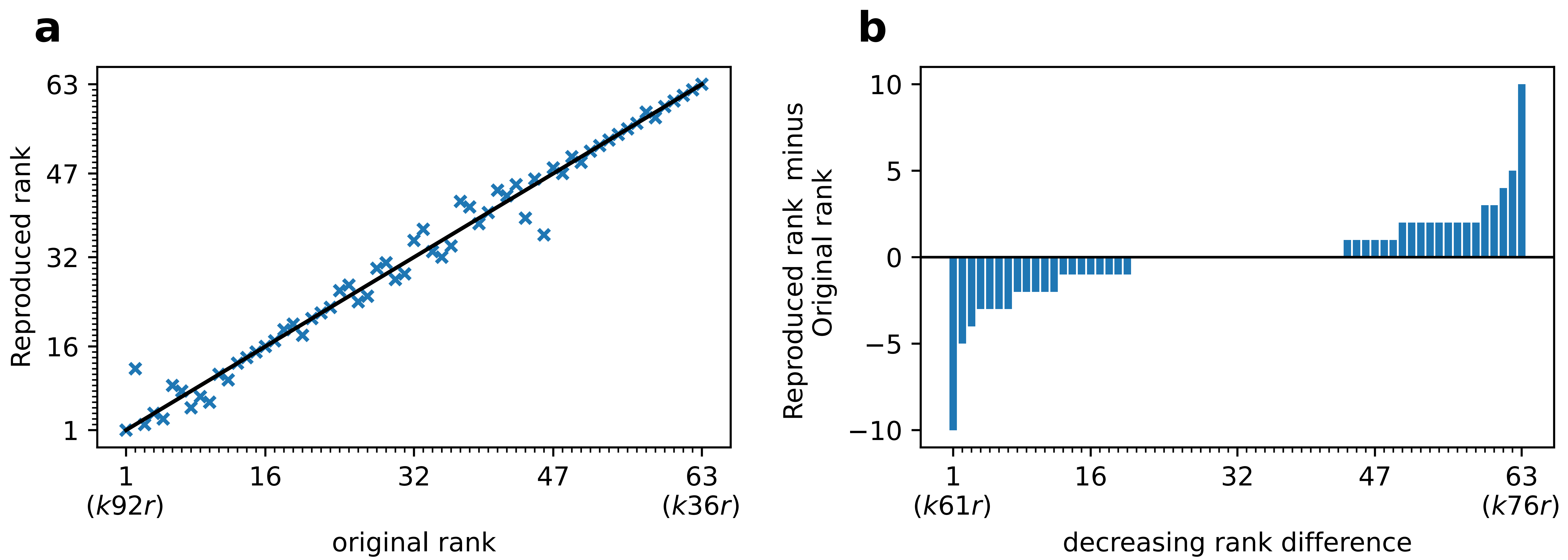
- J: Opponent’s previous move (0 = cooperate, 1 = defect),
- M: Current turn number (starting at 1),
- K: Player’s cumulative score,
- L: Opponent’s cumulative score,
- R: Random number between 0 and 1 (for stochastic strategies),
- JA: Player’s previous move.

3. Reproducing Axelrod's second

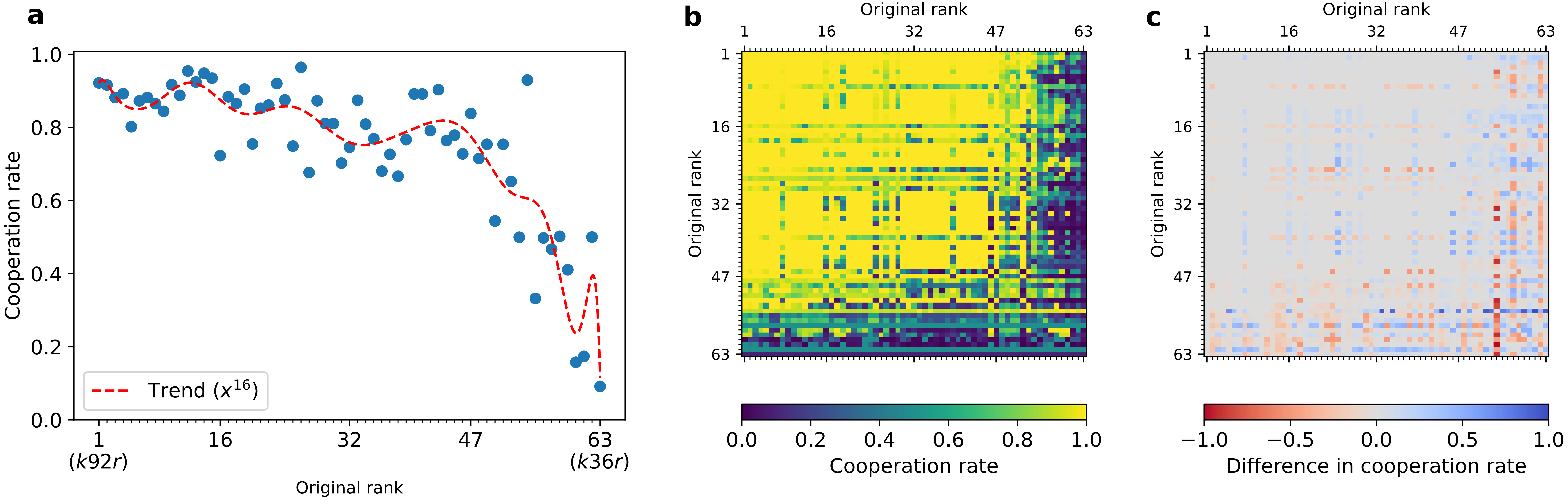


https://github.com/Axelrod-Python/axelrod-fortran

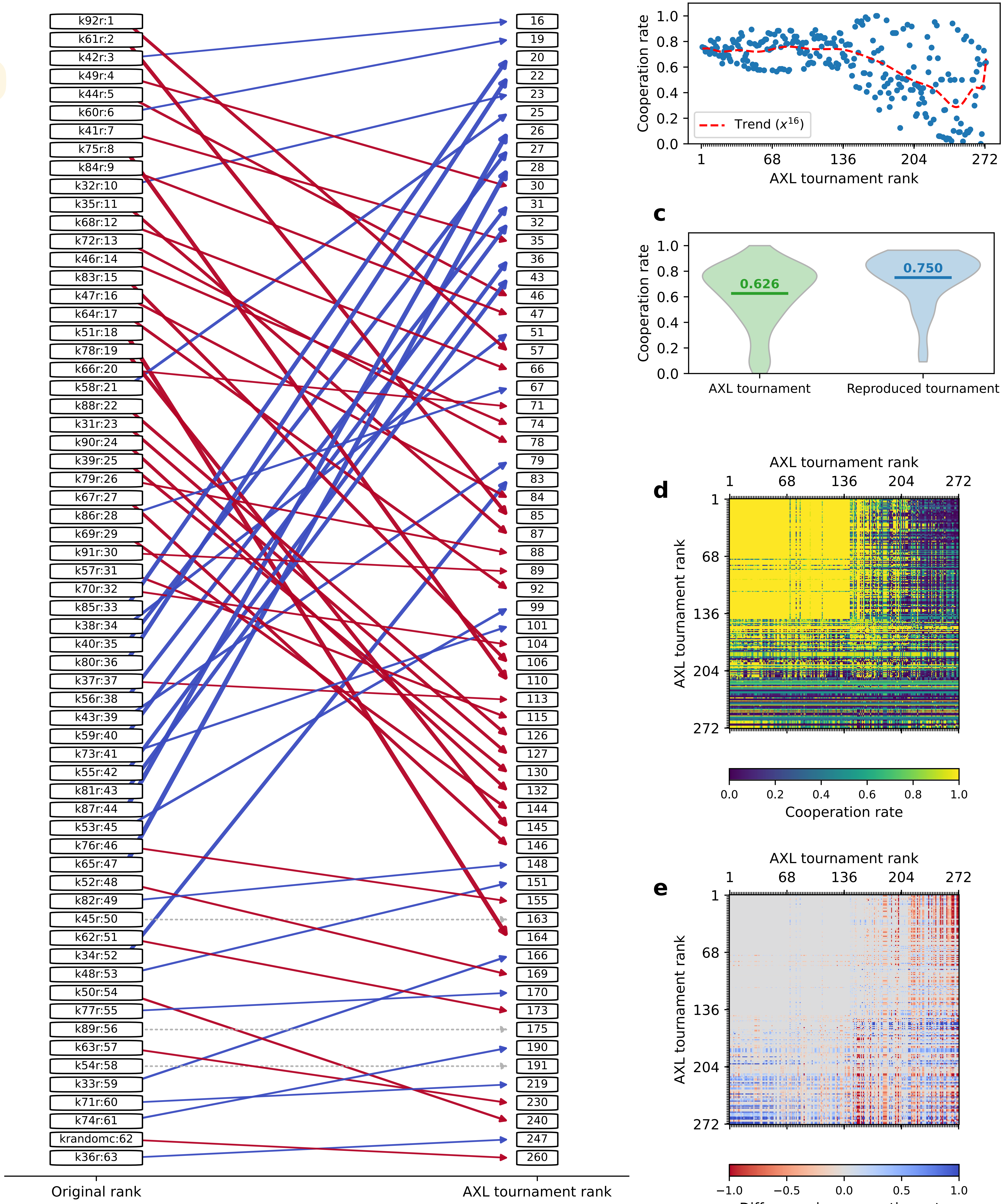
3. Reproducing Axelrod's second



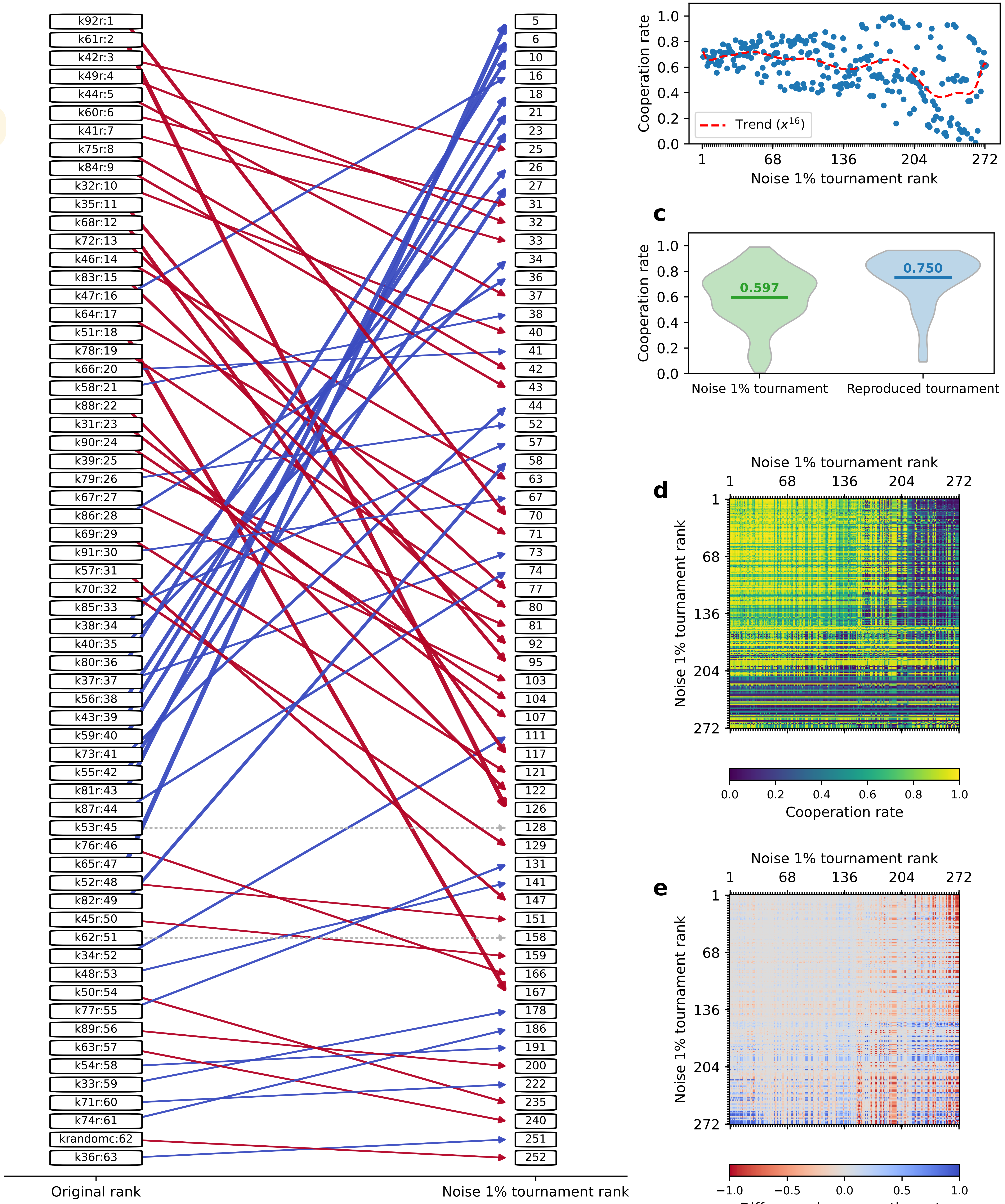
3. Reproducing Axelrod's second



1. Axelrod-Python tournament
2. Axelrod-Python tournament 1% Noise



1. Axelrod-Python tournament
2. Axelrod-Python tournament 1% Noise



Summary

Faithful replication of the original tournaments

No universal winner: rankings change across environments

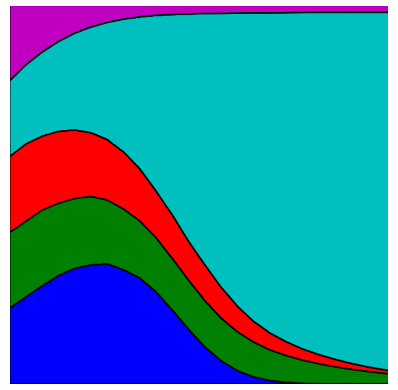
What makes a tournament representative?

Many strategies, or a balanced set of strategic behaviours?

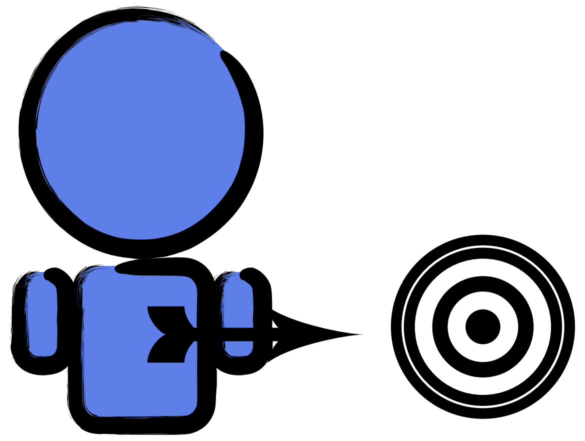
[5] Reviving, reproducing, and revisiting Axelrod's second tournament
arXiV: https://arxiv.org/abs/2510.15438

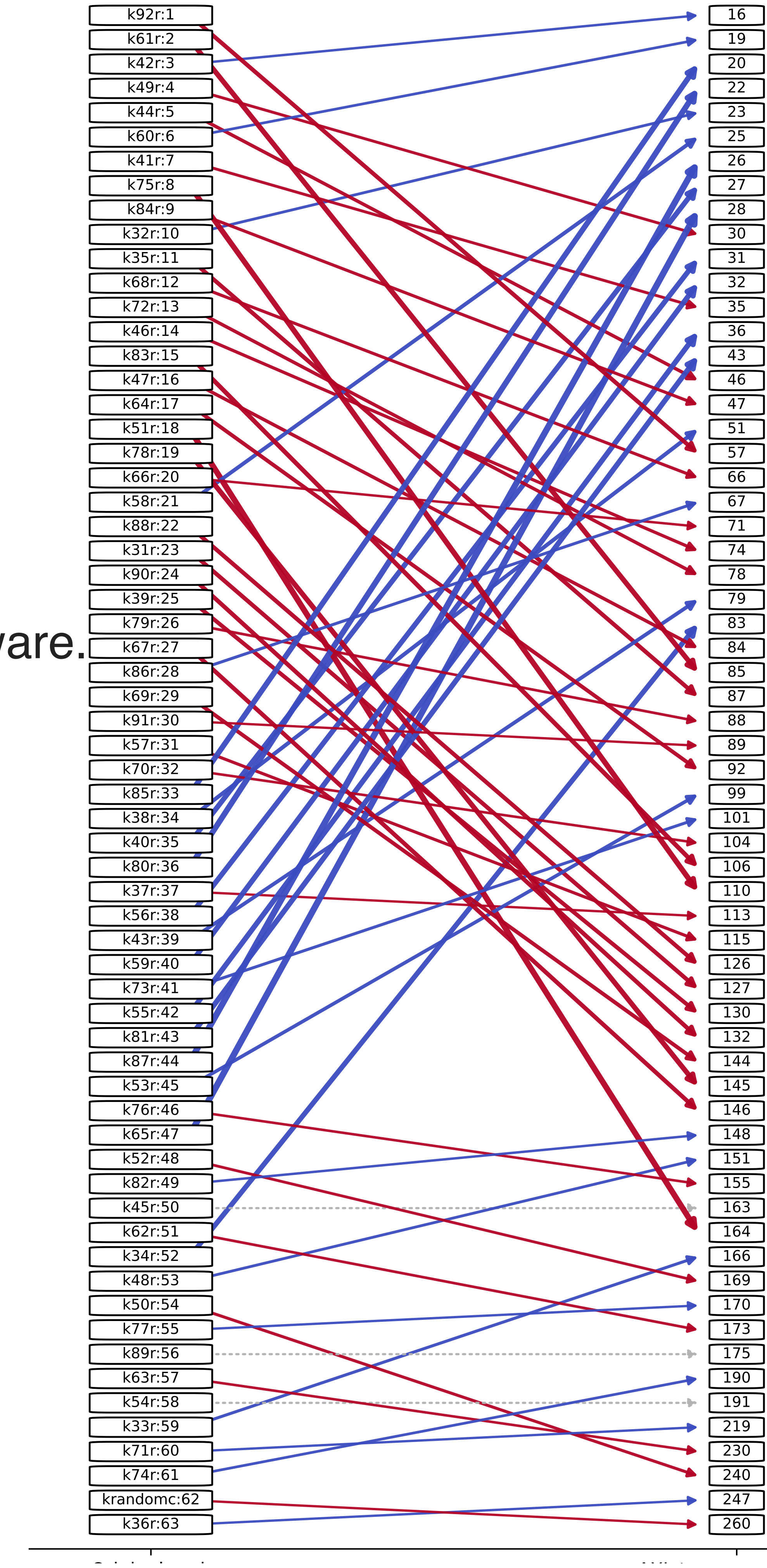
Soon to be published in *Communications AI & Computing*

Summary



Axelrod-Python





[1] An open framework for the reproducible study of the iterated prisoner's dilemma. Journal of Open Research Software.
https://doi.org/10.5334/jors.125

[2] Reinforcement learning produces dominant strategies for the iterated prisoner's dilemma. PloS one.
https://doi.org/10.1371/journal.pone.0188046

[3] Evolution reinforces cooperation with the emergence of self-recognition mechanisms. PloS one.
https://doi.org/10.1371/journal.pone.0204981

[4] Properties of winning iterated prisoner's dilemma strategies. PLOS Computational Biology.
https://doi.org/10.1371/journal.pcbi.1012644

[5] Reviving, reproducing, and revisiting Axelrod's second tournament
arXiV: https://arxiv.org/abs/2510.15438

Soon to be published in *Communications AI & Computing*

Nikoleta-v3

nikoleta.glynatsi@riken.jp

THANK YOU!