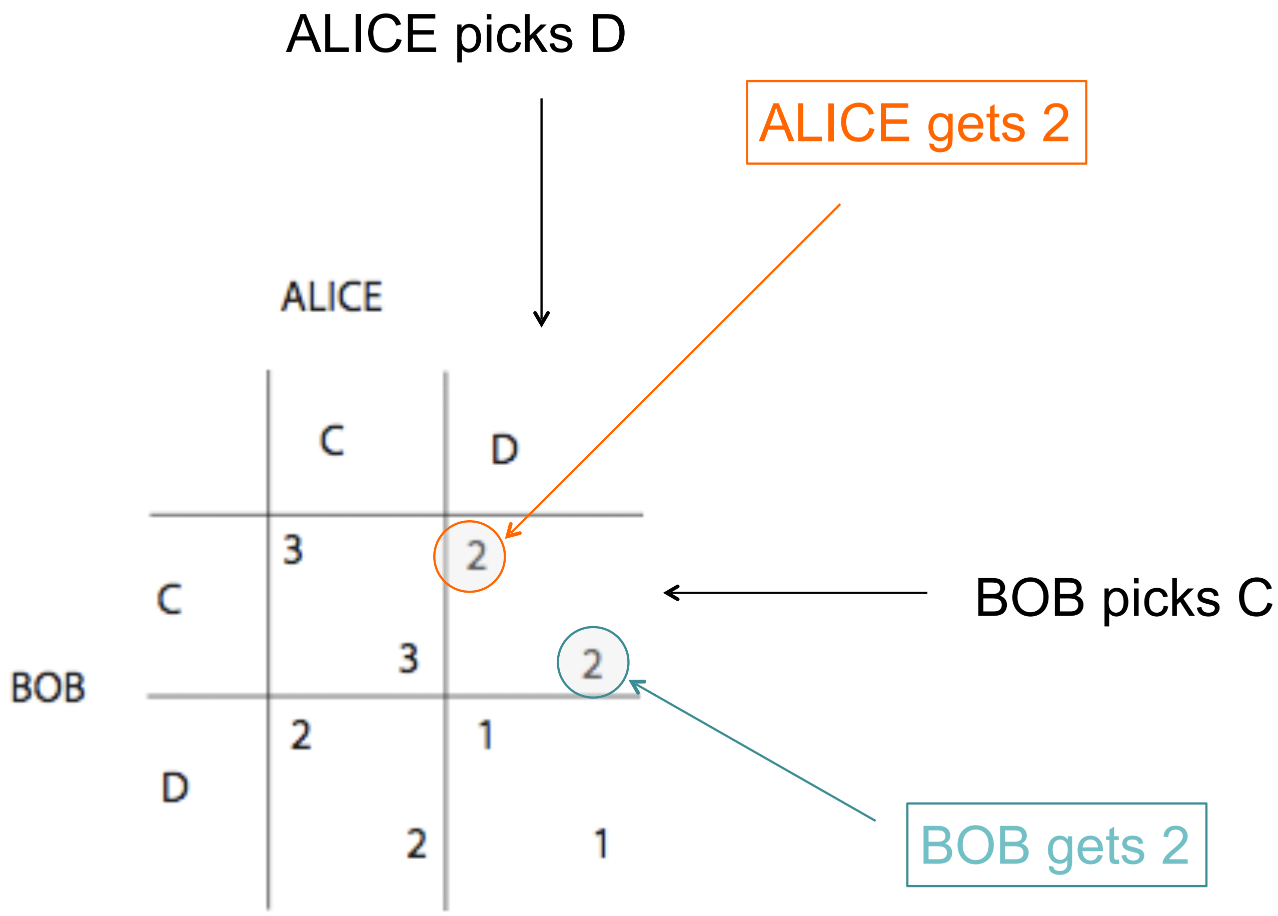
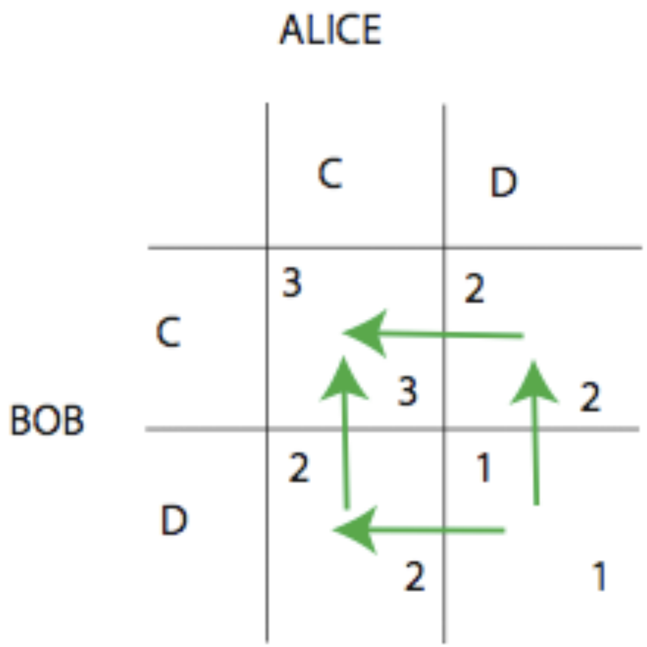
Game theory

		ALICE	
		C	D
BOB	C	3 / 3	2 / 2
	D	2 / 2	1 / 1


ALICE picks D
ALICE gets 2
ALICE
C
D
BOB
C
D
3
2
3
2
2
1
2
1
BOB picks C
BOB gets 2

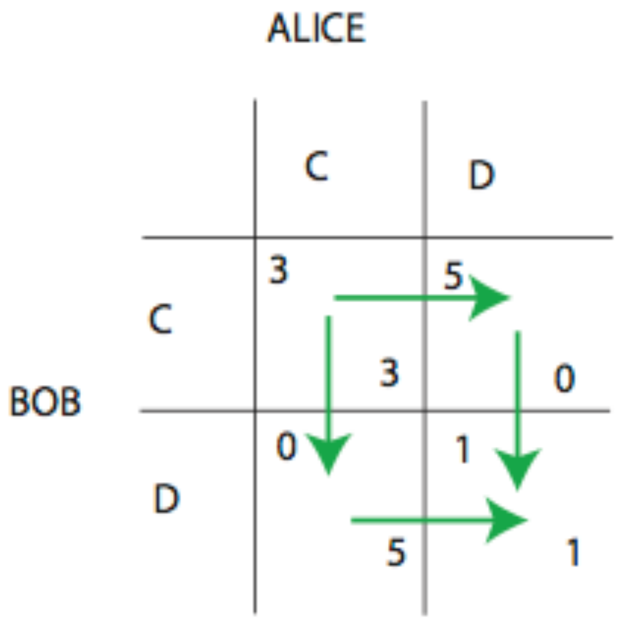
Nash equilibrium

Alice and Bob are in Nash equilibrium if Alice is making the best decision she can, taking into account Bob's decision, and Bob is making the best decision he can, taking into account Alice's decision.



Prisoner's dilemma

		Alice	
		C	D
	C	3, 3	5, 0
Bob	D	0, 5	1, 1


ALICE
C
D
BOB
C
D
3
5
3
0
0
1
5
1

Iterated Prisoner's dilemma (IPD)

The cooperator will get

$$3 + 3\gamma + 3\gamma^2 + 3\gamma^3 + \cdots = \frac{3}{1-\gamma}$$

And the defector will get

$$5 + \gamma + \gamma^2 + \gamma^3 + \cdots = 5 + \frac{\gamma}{1-\gamma}$$

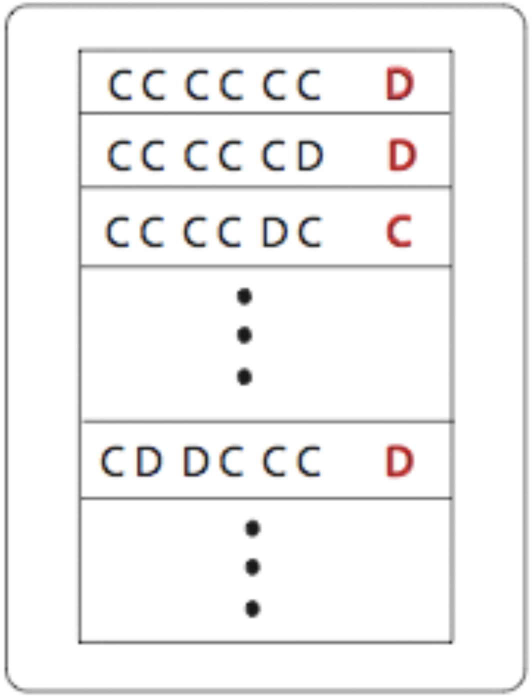
Axelrod, Evolution of Cooperation, 1984

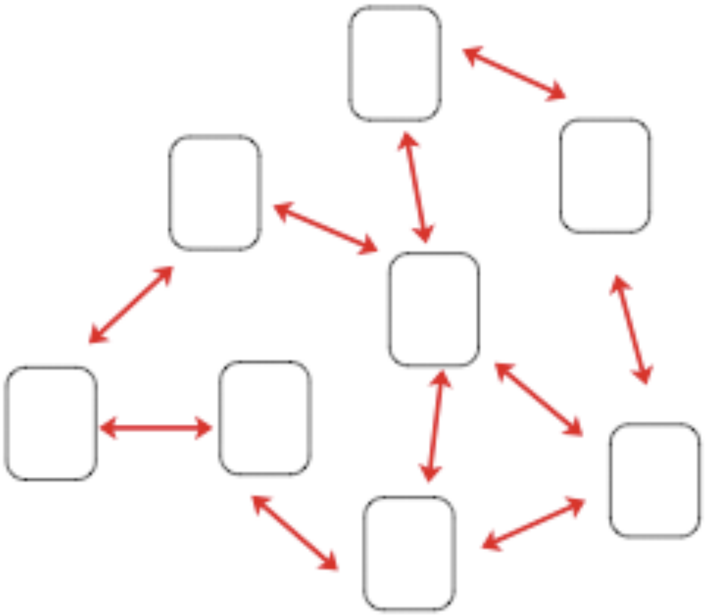
Groups play IPD

Winner’s in human v human tournament used Tit-For-Tat strategies

- **Be nice:** cooperate, never be the first to defect.
- **Be provocable:** return defection for defection, cooperation for cooperation.
- **Don't be envious:**: be fair with your *partner*.
- **Don't be too clever:** or, don't try to be tricky.

A genetic algorithm learns to beats the humans!



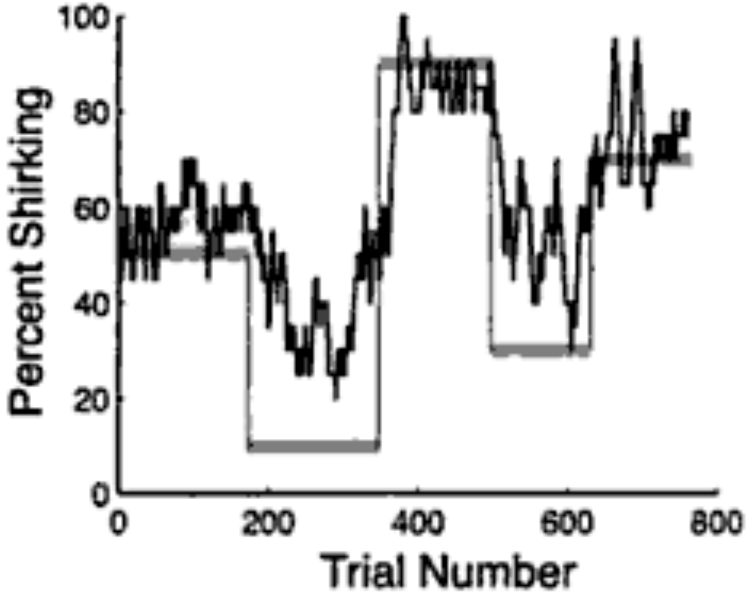


… because it punishes weak players

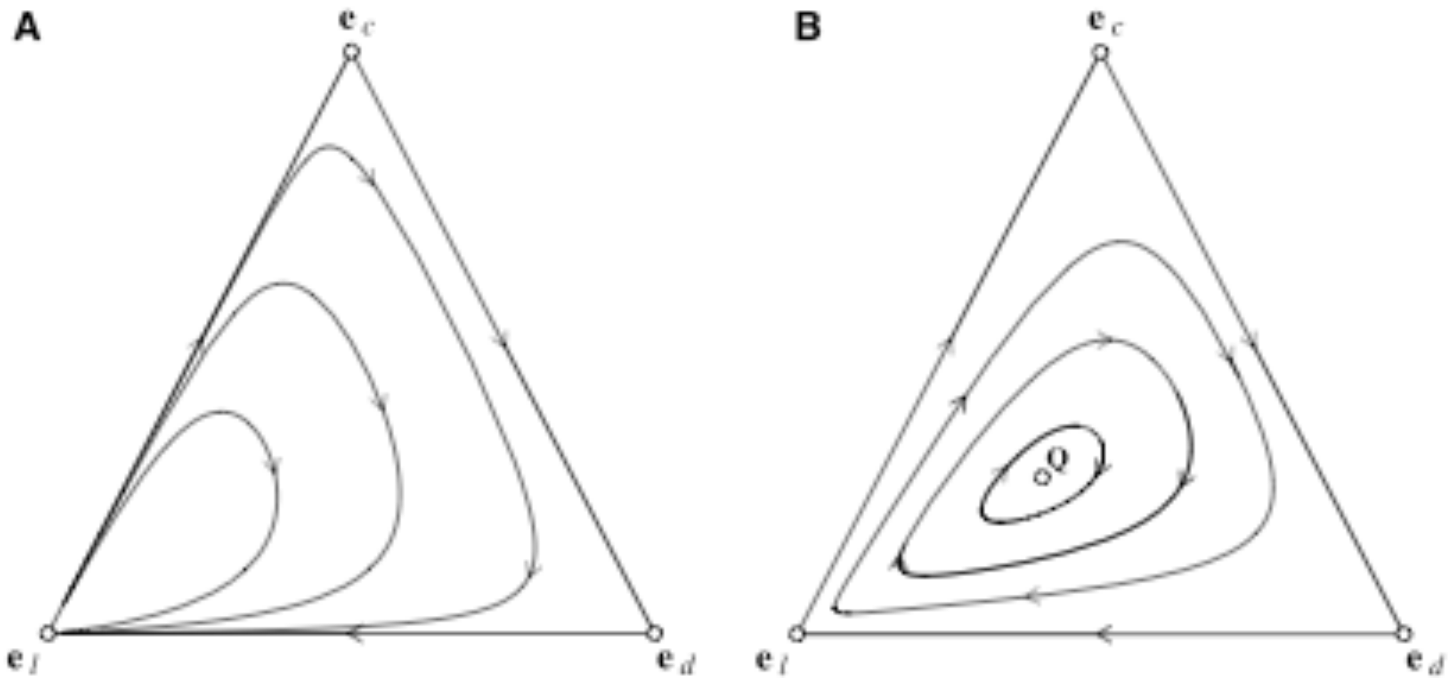
Probabilistic strategies e.g. 'Work or Shirk'

Alice

Bob		Insp.	~Insp.
	Shirk	-h 0	-w w
	Work	v-w-h w-g	v-w w-g



Monkeys can play work or shirk

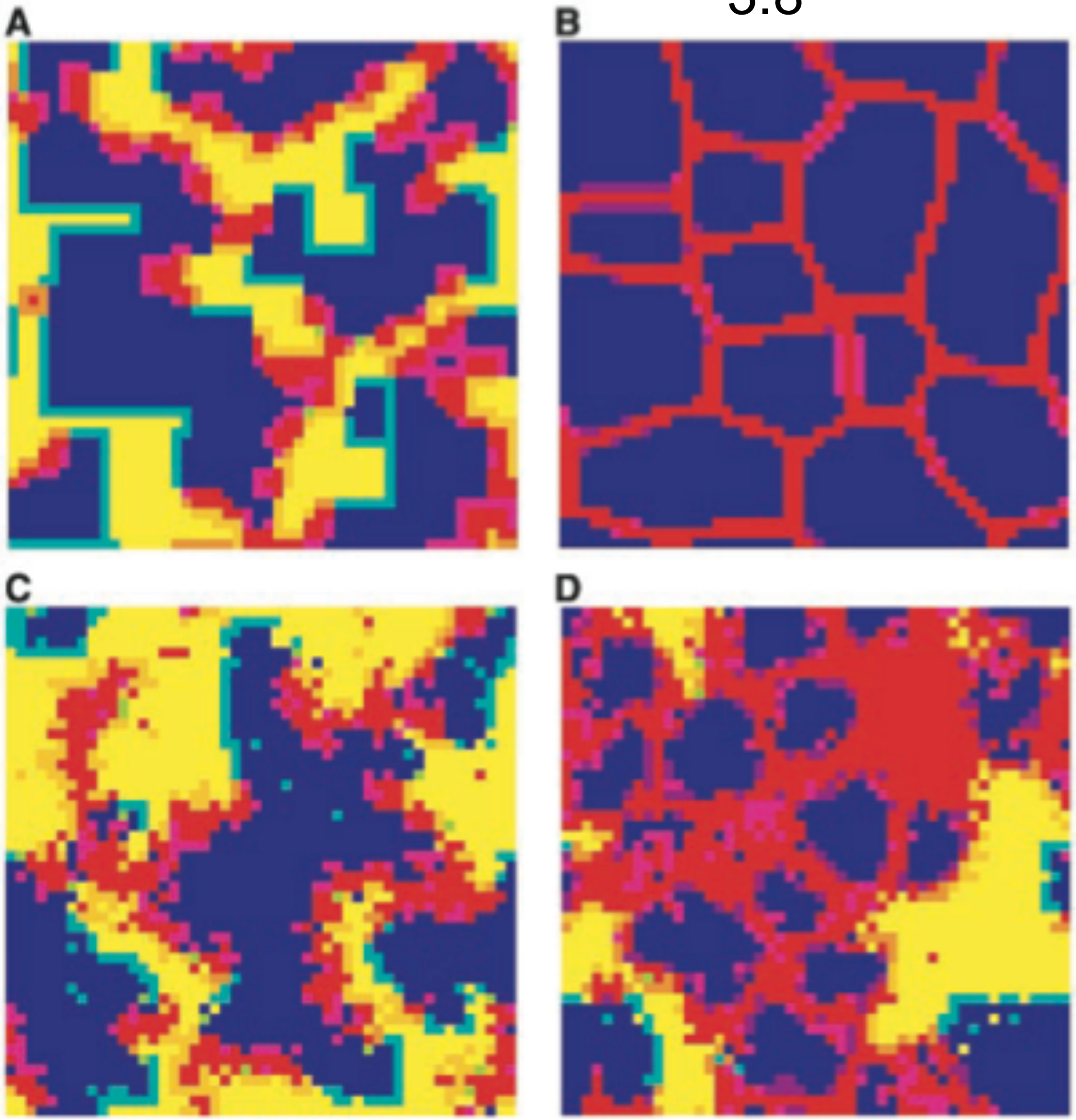

A
$\mathbf{e}_c$
$\mathbf{e}_l$
$\mathbf{e}_d$
B
$\mathbf{e}_c$
Q
$\mathbf{e}_l$
$\mathbf{e}_d$

r < 2 r > 2

Adopt the strategy of the best neighbor in 3x3

r=2.2 3.8

Sites adopt the strategy of a neighbor



80% of sites adopt the strategy of a neighbor w prob. ~ payoff diff.

Blue = cooperate; red = defect; yellow = sit out

1. Let α and δ be learning rates. Initialize,

$$Q(s,a) \leftarrow 0, \qquad \pi(s,a) \leftarrow \frac{1}{|\mathcal{A}_i|}.$$

2. Repeat,

   (a) From state s select action a with probability $\pi(s,a)$ with some exploration.

   (b) Observing reward r and next state s',

   $$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha\left(r + \gamma \max_{a'} Q(s',a')\right).$$

   (c) Update $\pi(s,a)$ and constrain it to a legal probability distribution,

   $$\pi(s,a) \leftarrow \pi(s,a) + \begin{cases} \delta & \text{if } a = \operatorname{argmax}_{a'} Q(s,a') \\ \frac{-\delta}{|A_i|-1} & \text{otherwise} \end{cases}.$$

$$R = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \quad C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

$$\begin{aligned} V_r(\alpha,\beta) &= \alpha\beta r_{11} + \alpha(1-\beta)r_{12} + \\ &\quad (1-\alpha)\beta r_{21} + (1-\alpha)(1-\beta)r_{22} \\ &= u\alpha\beta + \alpha(r_{12} - r_{22}) + \\ &\quad \beta(r_{21} - r_{22}) + r_{22} \end{aligned} \tag{1}$$

$$\begin{aligned} V_c(\alpha,\beta) &= \alpha\beta c_{11} + \alpha(1-\beta)c_{12} + \\ &\quad (1-\alpha)\beta c_{21} + (1-\alpha)(1-\beta)c_{22} \\ &= u'\alpha\beta + \alpha(c_{12} - c_{22}) + \\ &\quad \beta(c_{21} - c_{22}) + c_{22} \end{aligned} \tag{2}$$

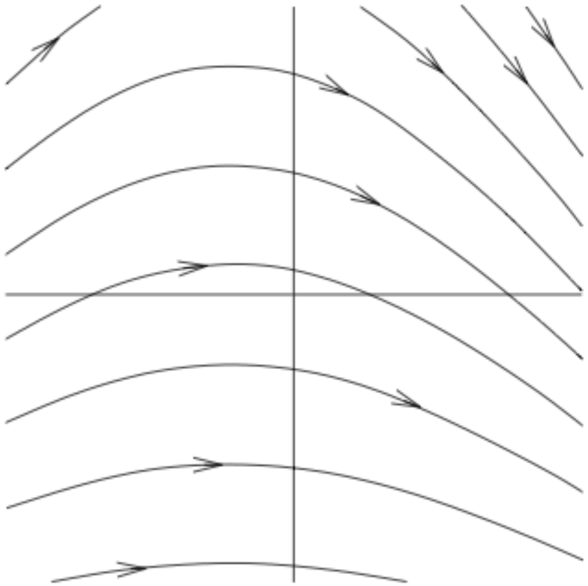
where,

$$\begin{aligned} u &= r_{11} - r_{12} - r_{21} + r_{22} \\ u' &= c_{11} - c_{12} - c_{21} + c_{22}. \end{aligned}$$

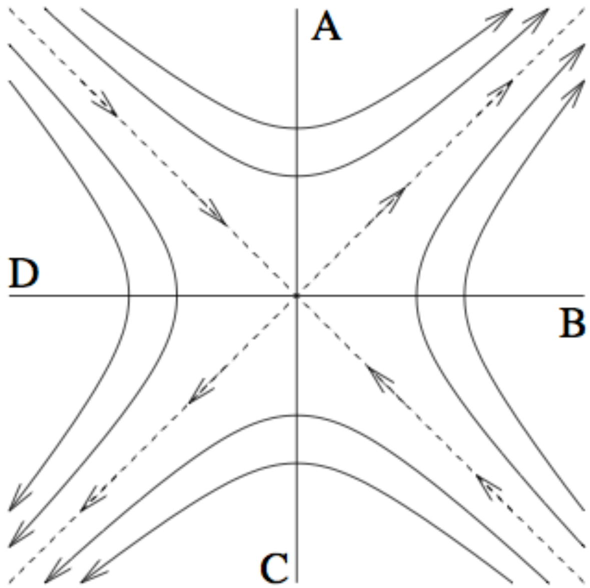
$$
\begin{aligned}
\alpha_{k+1} &= \alpha_k + \eta \frac{\partial V_r(\alpha_k, \beta_k)}{\partial \alpha_k} \\
\beta_{k+1} &= \beta_k + \eta \frac{\partial V_r(\alpha_k, \beta_k)}{\partial \beta_k}.
\end{aligned}
$$

$$\begin{bmatrix} \frac{\partial \alpha}{\partial t} \\ \frac{\partial \beta}{\partial t} \end{bmatrix} = \begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} (r_{12} - r_{22}) \\ (c_{21} - c_{22}) \end{bmatrix}$$

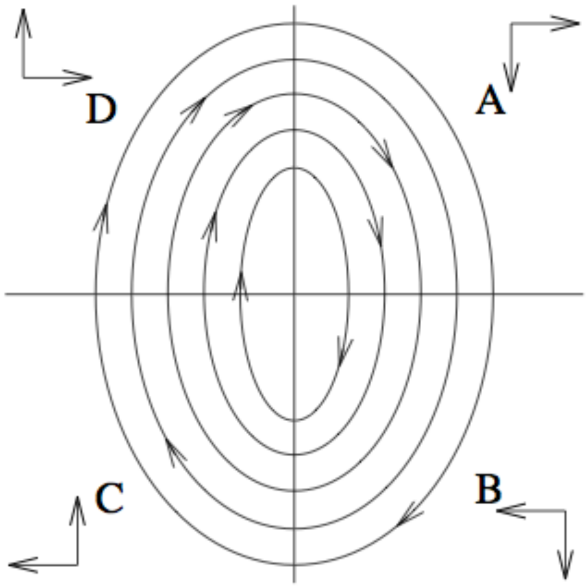
$$(\alpha^*, \beta^*) = \left(\frac{(c_{22} - c_{21})}{u'}, \frac{(r_{22} - r_{12})}{u} \right)$$



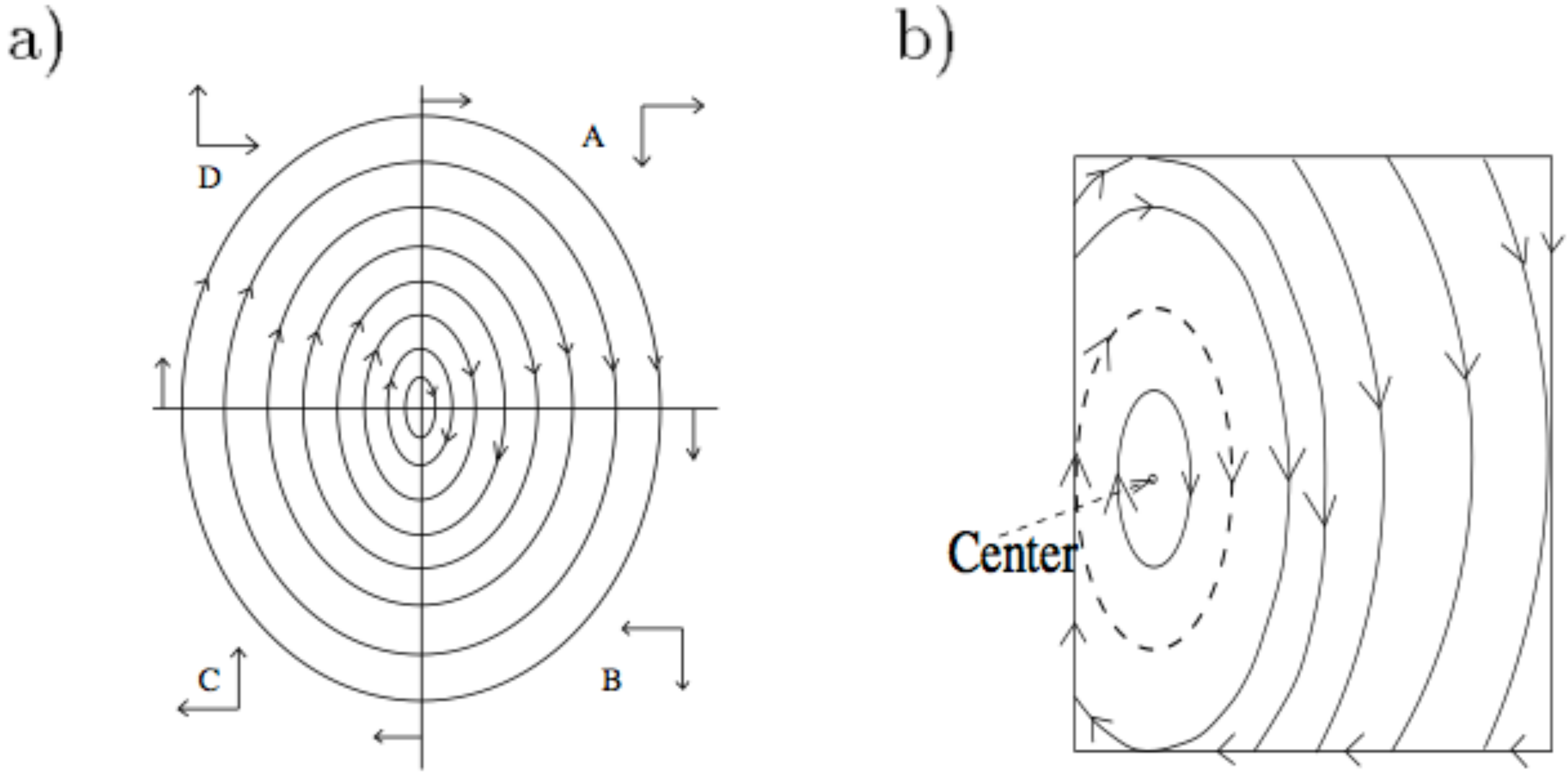
(a) U is not invertible



(b) U has real eigenvalues



(c) U has imaginary eigenvalues


a)
D
A
C
B
b)
Center

$$\ell_k^r = \begin{cases} \ell_{\min} & \text{if } V_r(\alpha_k, \beta_k) > V_r(\alpha^e, \beta_k) \quad \text{WINNING} \\ \ell_{\max} & \text{otherwise} \quad \text{LOSING} \end{cases}$$

$$\ell_k^c = \begin{cases} \ell_{\min} & \text{if } V_c(\alpha_k, \beta_k) > V_c(\alpha_k, \beta^e) \quad \text{WINNING} \\ \ell_{\max} & \text{otherwise} \quad \text{LOSING} \end{cases}$$

1. Let $\alpha, \delta_l > \delta_w$ be learning rates. Initialize,

$$Q(s,a) \leftarrow 0, \qquad \pi(s,a) \leftarrow \frac{1}{|\mathcal{A}_i|}, \qquad C(s) \leftarrow 0.$$

2. Repeat,

(a,b) Same as PHC in Table 1

(c) Update estimate of average policy, $\bar{\pi}$,

$$\begin{array}{rrcl} & C(s) & \leftarrow & C(s) + 1 \\ \forall a' \in \mathcal{A}_i & \bar{\pi}(s,a') & \leftarrow & \bar{\pi}(s,a') + \\ & & & \frac{1}{C(s)} \left(\pi(s,a') - \bar{\pi}(s,a') \right). \end{array}$$

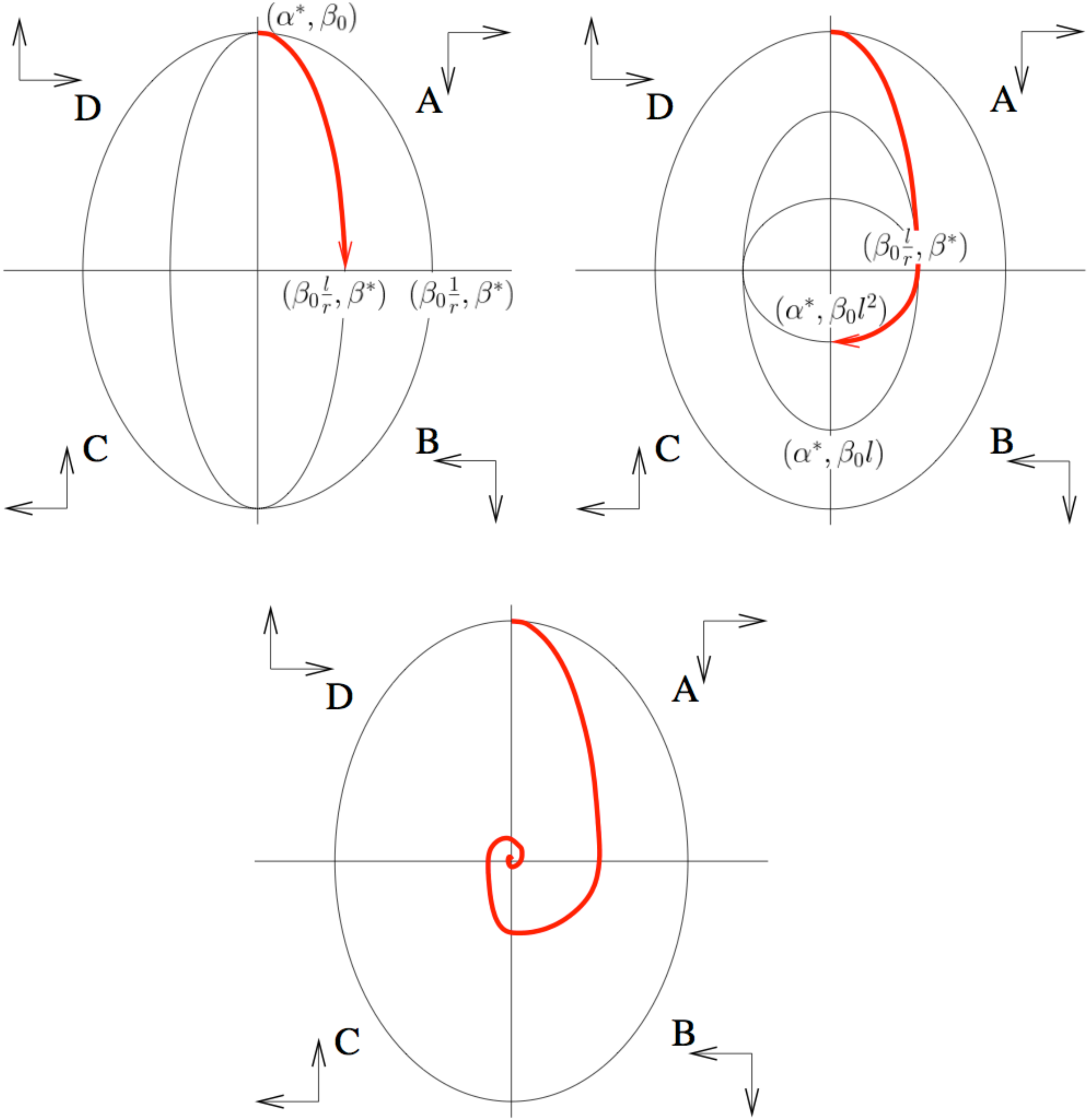
(d) Update $\pi(s,a)$ and constrain it to a legal probability distribution,

$$\pi(s,a) \leftarrow \pi(s,a) + \begin{cases} \delta & \text{if } a = \text{argmax}_{a'} Q(s,a') \\ \frac{-\delta}{|A_i|-1} & \text{otherwise} \end{cases},$$

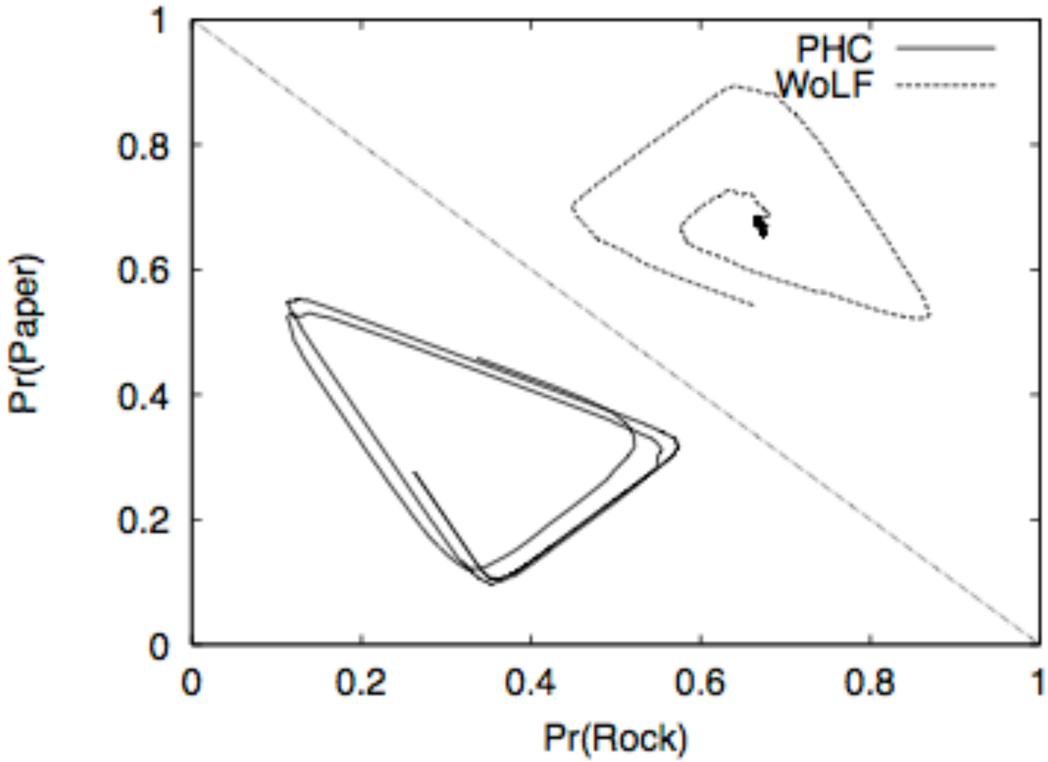
where,

$$\delta = \begin{cases} \delta_w & \text{if } \sum_a \pi(s,a) Q(s,a) > \sum_a \bar{\pi}(s,a) Q(s,a) \\ \delta_l & \text{otherwise} \end{cases}.$$

$$\begin{bmatrix} \frac{\partial \alpha}{\partial t} \\ \frac{\partial \beta}{\partial t} \end{bmatrix} = \begin{bmatrix} 0 & \ell^r(t)u \\ \ell^c(t)u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \ell^r(t)(r_{12} - r_{22}) \\ \ell^c(t)(c_{21} - c_{22}) \end{bmatrix}$$



(α^*, β_0)
D
A
$(\beta_0\frac{l}{r}, \beta^*)$
$(\beta_0\frac{1}{r}, \beta^*)$
C
B
D
A
$(\beta_0\frac{l}{r}, \beta^*)$
$(\alpha^*, \beta_0 l^2)$
C
$(\alpha^*, \beta_0 l)$
B
D
A
C
B



(b) Rock-Paper-Scissors Game

\ to play	C	D
opponent played C	$1-\lambda\delta$	$\lambda\delta$
opponent played D	δ	$1-\delta$

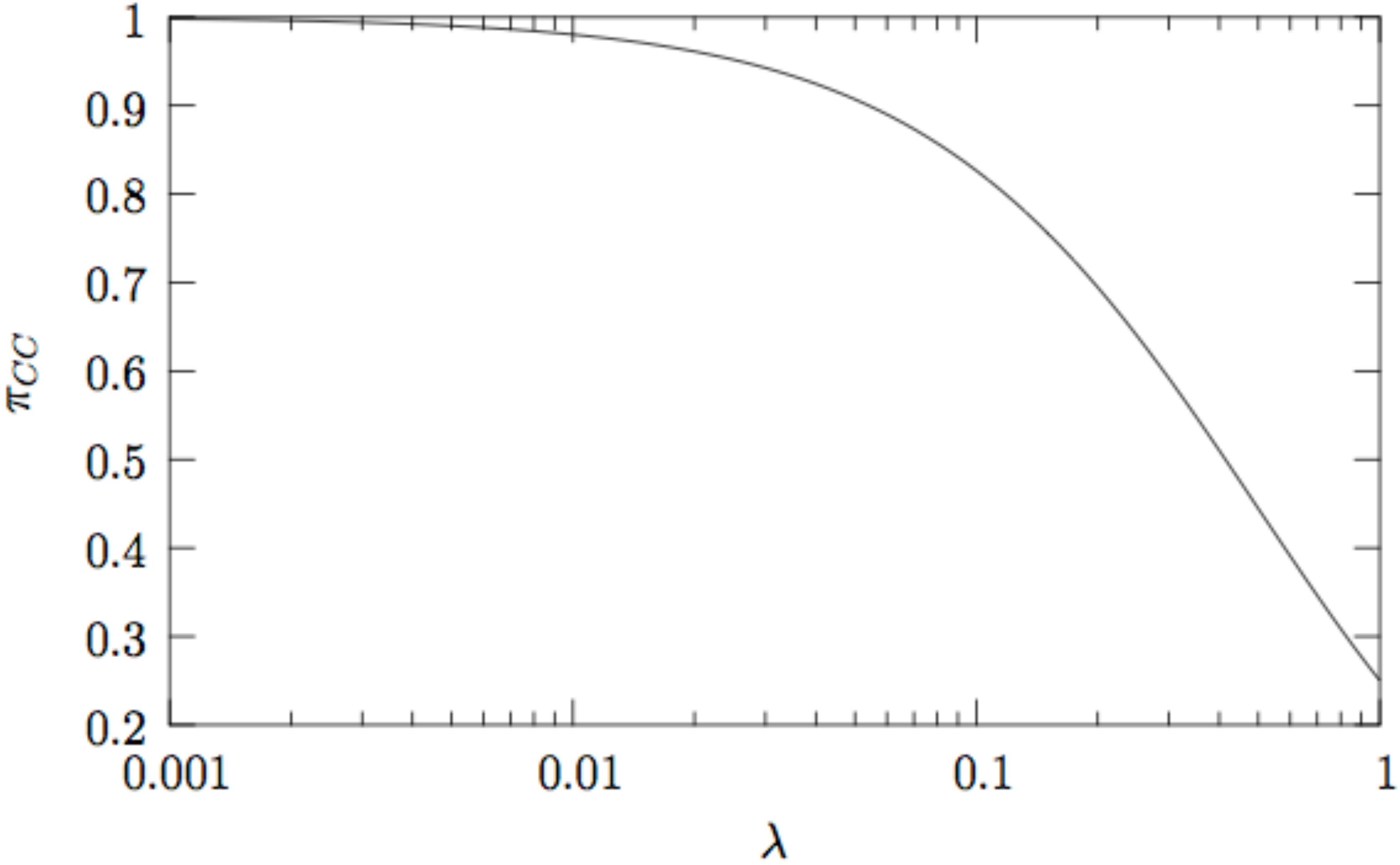
$$\delta S + (1-\delta)P = P - \delta(P - S).$$

$$\delta < \frac{\epsilon}{P - S}.$$

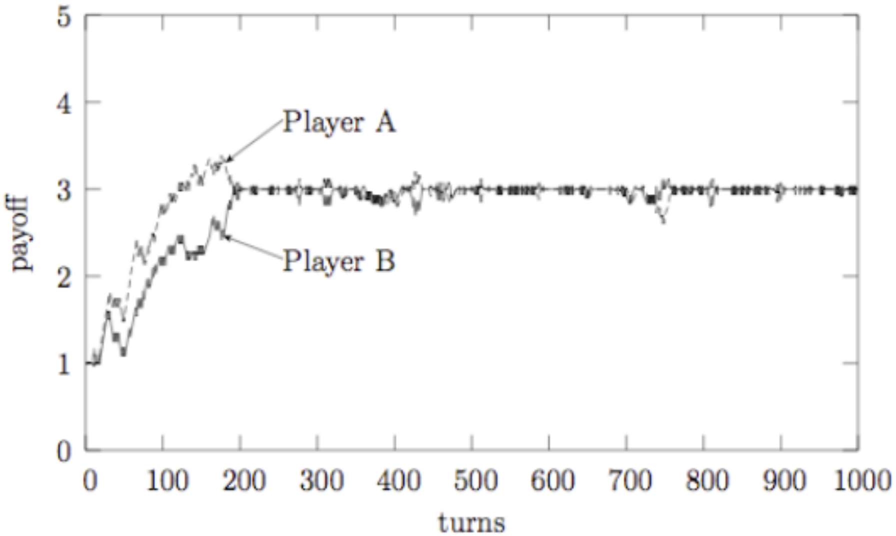
current \ next	CC	CD	DC	DD
CC	$(1-\lambda\delta)^2$	$\lambda\delta(1-\lambda\delta)$	$\lambda\delta(1-\lambda\delta)$	$(\lambda\delta)^2$
CD	$\delta(1-\lambda\delta)$	$\lambda\delta^2$	$(1-\lambda\delta)(1-\delta)$	$\lambda\delta(1-\delta)$
DC	$\delta(1-\lambda\delta)$	$(1-\lambda\delta)(1-\delta)$	$\lambda\delta^2$	$\lambda\delta(1-\delta)$
DD	δ^2	$\delta(1-\delta)$	$\delta(1-\delta)$	$(1-\delta)^2$

$$\pi = \pi T$$

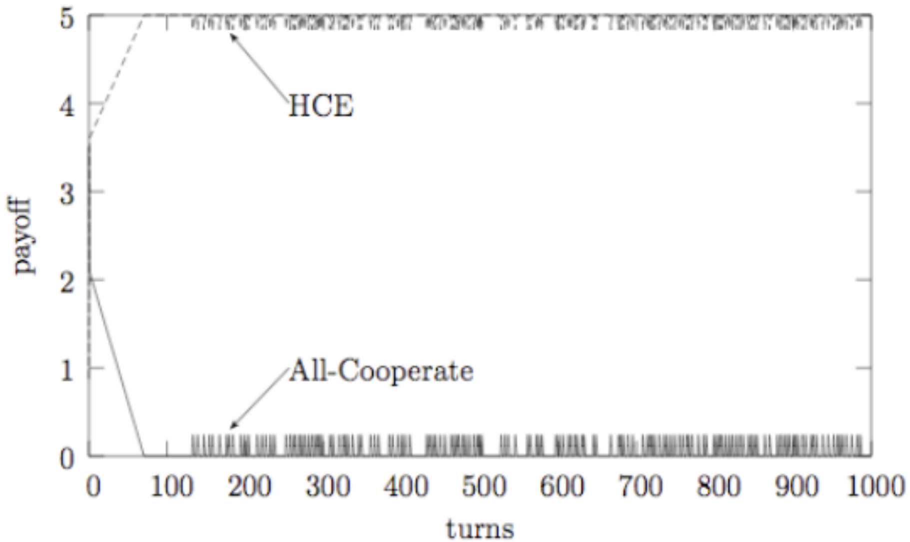
$$\pi = \left[\frac{1}{(1+\lambda)^2}, \frac{\lambda}{(1+\lambda)^2}, \frac{\lambda}{(1+\lambda)^2}, \frac{\lambda^2}{(1+\lambda)^2}\right]$$



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
π_{CC}
0.001
0.01
0.1
1
λ



5
4
3
2
1
0
payoff
Player A
Player B
0
100
200
300
400
500
600
700
800
900
1000
turns



5
4
3
2
1
0
payoff
HCE
All-Cooperate
0
100
200
300
400
500
600
700
800
900
1000
turns