电子科技大学
University of Electronic Science and Technology of China

Can people cooperate in prisoners dilemma

Xiaolin Yang

Data Mining Lab, Big Data Research Center, UESTC
School of Computer Science and Engineering
Email: xiaolinyn@gmail.com

2017.03.08

01 Prisoners Dilemma

Prisoners dilemma

Prisoners' dilemma:

	C	D
C	R, R	S, T
D	T, S	P, P

$T > R > P > S$

C: *cooperate*

D: *defect*

Dominant Strategy : It is always better than any other strategy, for any profile of other players' actions.

So (D,D) is the nash equilibrium.

but it's inefficient!!

Prisoners dilemma

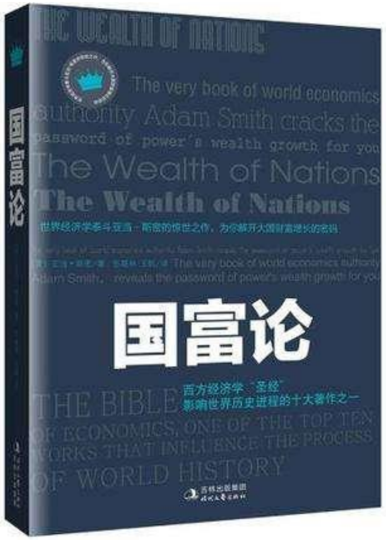
数据挖掘实验室
Data Mining Lab

Why nash equilibrium is so shocking?

Every individual . . . neither intends to promote the public interest, nor knows how much he is promoting it. . . . He intends only his own gain, and he is in this, as in many other cases, led by an ***invisible hand*** *to promote an end which was no part of his intention. By pursuing his own interest he frequently promotes that of the society more* ***effectually*** *than when he really intends to promote it.*

—— *《the Nature and Causes of the Wealth of Nations》*

Adam Smith



So that competition is really death?

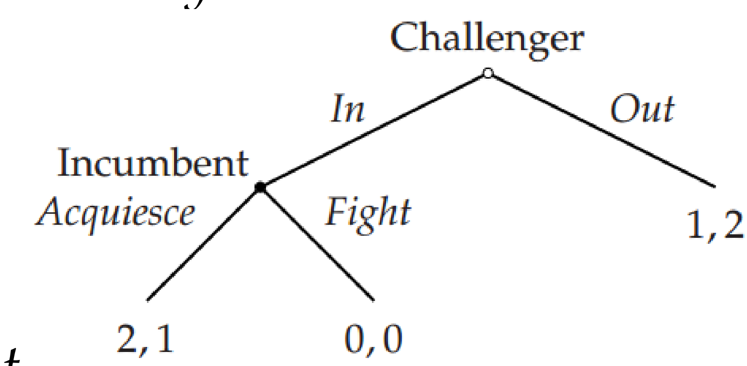
But how cooperation exists in reality? deal with it!!

02 Repeated Games

Repeated games

Repeated games:

- *a set of* ***players***
- ***histories****: actions sequences*
- *the player function: assigns a player to one history*
- *preference*



Entry game

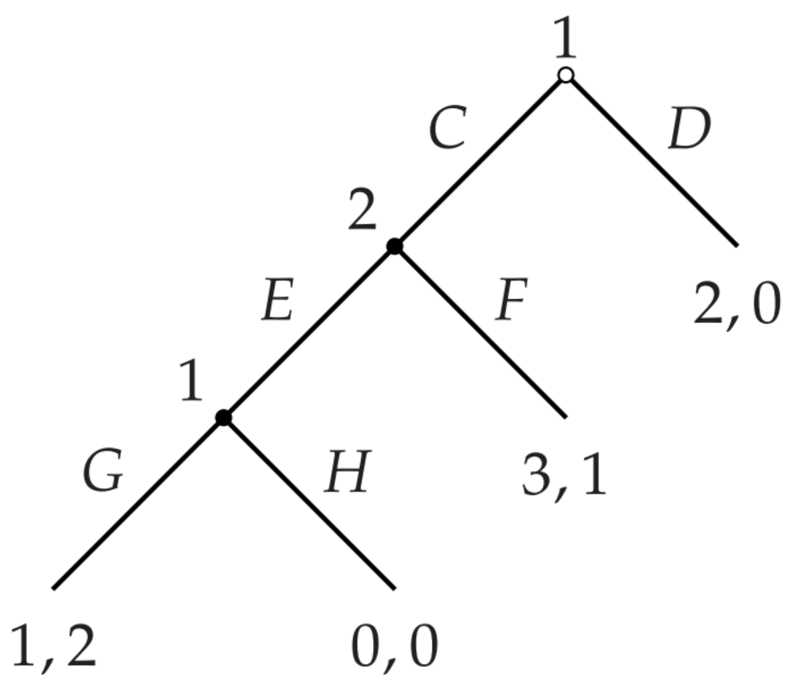
- ***players:*** *the challenger and the incumbent*
- ***histories****: (In, Acquiesce), (In, Fight), and Out*
- ***the player function****: P(Ø) = Challenger and P(In) = Incumbent*(在位者)
- ***preference****:*

$$u_1\,(In,\,Acquiesce) = 2,\ u_1(In,\,Fight) = 0 \text{ and } u_1(Out) = 1$$
$$u_2(In,\,Acquiesce) = 1,\ u_2(In,\,Fight) = 0 \text{ and } u_2\,(Out) = 2$$

Repeated games

The strategy in repeated games:

> *a function that assigns to each history h after which it is player i's turn to move (i.e. P(h) = i, where P is the player function) an action in A(h) (the set of actions available after h).*



Example:

The strategies of player 1 are

CG, CH, DG, DH

Repeated games

If we make actions contingent on the history, how many strategies are there?

All possible strategies in the second stage

The strategy set is too huge, it's should be simplified!!

$$A_1: (s_1, s_2) \quad A_2 : (t_1, t_2) \quad \text{->} \quad h_{ij}=(s_i, t_j) \quad \text{->} \quad f\left(h_{ij}\right): H \rightarrow A_1$$

$f_{1111} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_1	s_1	s_1

$f_{1112} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_1	s_1	s_2

$f_{1121} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_1	s_2	s_1

$f_{1122} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_1	s_2	s_2

$f_{1211} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_2	s_1	s_1

$f_{1212} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_2	s_1	s_2

$f_{1221} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_2	s_2	s_1

$f_{1222} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_2	s_2	s_2

$f_{2111} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_2	s_1	s_1	s_1

$f_{2112} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_2	s_1	s_1	s_2

$f_{1121} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_1	s_1	s_2	s_1

$f_{2122} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_2	s_1	s_2	s_2

$f_{2211} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_2	s_2	s_1	s_1

$f_{2212} : H \rightarrow S$

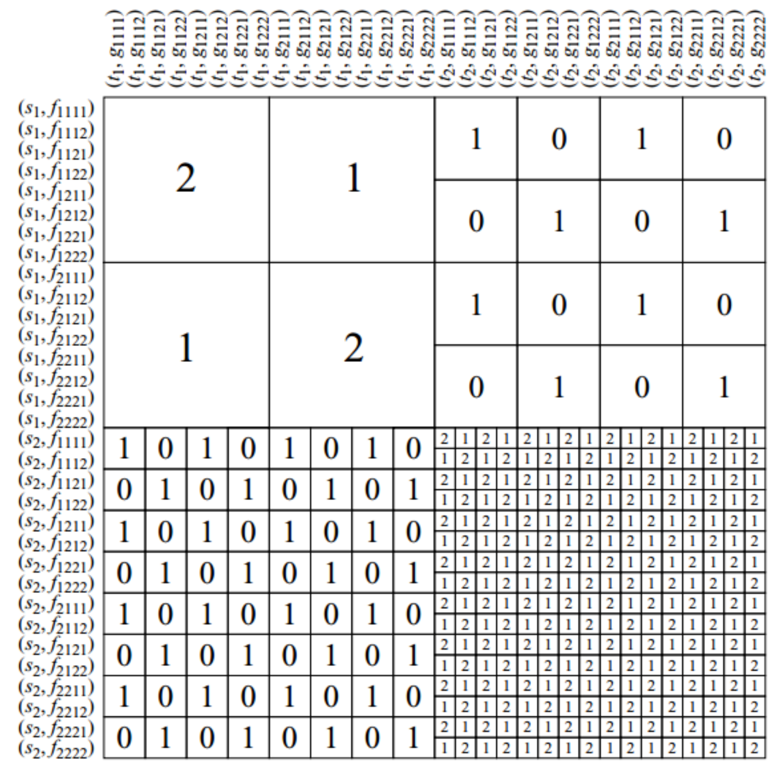
h_{11}	h_{12}	h_{21}	h_{22}
s_2	s_2	s_1	s_2

$f_{2221} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_2	s_2	s_2	s_1

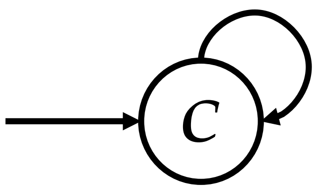
$f_{2222} : H \rightarrow S$

h_{11}	h_{12}	h_{21}	h_{22}
s_2	s_2	s_2	s_2

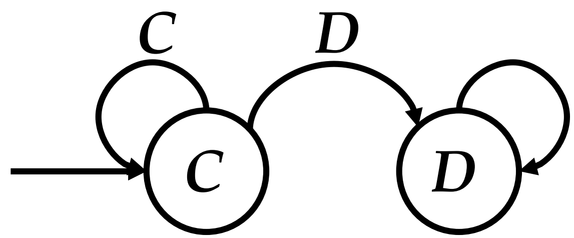


Repeated games

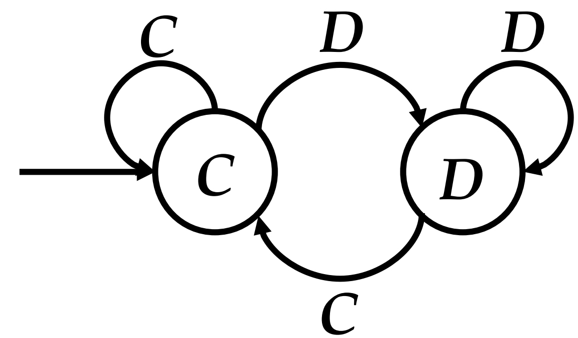
Representative strategies:



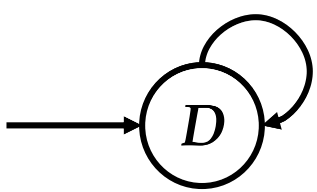
ALLC



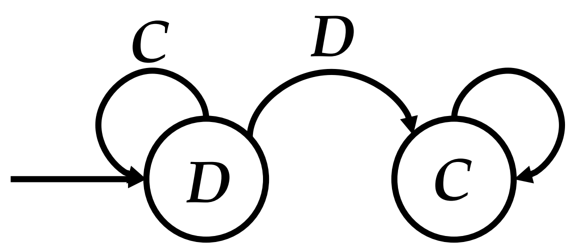
GRIM



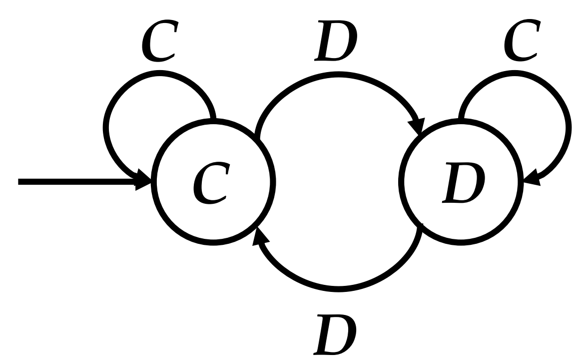
Tit-For-Tat



ALLD



CANARY(金丝雀)



WSLS

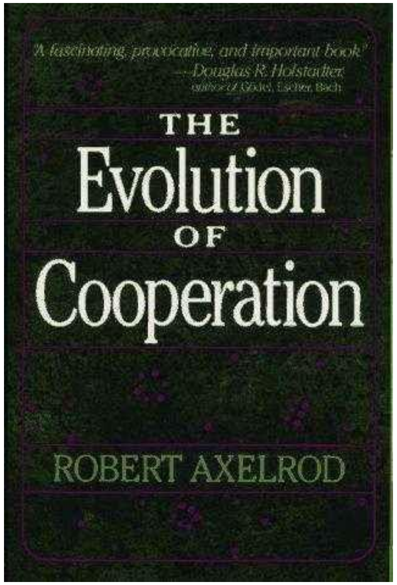
Repeated games

Axelrod's tournament :

In 1978, Robert Axelrod, a political scientist, decided to conduct a Prisoners' dilemma championship. He invited people form all over the world to submit strategies for the repeated game. The strategies had to be formulated in terms of computer programs. Every strategy played every other strategy. The payoff from each encounter were added up. In the end, Axelrod analyzed which strategy had the highest total payoff.

The final winner is Tit-For-Tat(TFT), even though there are many more "clever" strategies.

Axelrod summarizes the reason and published 《The evolution of cooperation》



Repeated games

Axelrod's tournament :

TFT: CCCCCCCCCC…
TFT: CCCCCCCCCC…

TFT: CCCCCCCCCC…
GRIM : CCCCCCCCCC…

TFT: CDDDDDDDD…
ALLD : DDDDDDDDD…

TFT: CDCCCCCCCC…
CANARY : DDCCCCCCCC…

TFT

- *It's never the first to defect;*
- *Never try to get more than the opponent in a direct pairwise comparison;*
- *However, the payoff sum over all matches is higher than any other strategy.*

It's apparently very successful at inducing cooperation from other strategies

Repeated games

数据挖掘实验室
Data Mining Lab

Achilles' heel:

Axelrod's original tournaments were conducted in an error-free digital universe, but real-world situations are permeated by mistakes. A "trembling hand" can lead to a misimplementation of one's own action. A "fuzzy mind" can cause the misinterpretation of the opponent's move.

- *TFT cannot correct mistakes*

*TFT: CCC**D**CDCDDD…*

*TFT: CCCCDCD**D**DD…*

- *When play with ALLC, TFT is not a Evolutionarily Stable Strategy, random drift(随机漂变) can lead from TFT to ALLC.*

03 Evolutionary Equilibrium

Symmetric games:

A two-player strategic game with ordinal preferences is symmetric if the players' sets of actions are the same and the players' preferences are represented by payoff functions u1 and u2 for which u1(a1, a2) = u2(a2, a1) for every action pair (a1, a2).

Player1:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{A} & \mathrm{B} \end{array} \\ \begin{array}{c} \mathrm{A} \\ \mathrm{B} \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

Player2:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{A} & \mathrm{B} \end{array} \\ \begin{array}{c} \mathrm{A} \\ \mathrm{B} \end{array} & \begin{pmatrix} a & c \\ b & d \end{pmatrix} \end{array}$$

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{A} & \mathrm{B} \end{array} \\ \begin{array}{c} \mathrm{A} \\ \mathrm{B} \end{array} & \begin{pmatrix} a, a & b, c \\ c, b & d, d \end{pmatrix} \end{array}$$

Prisoners' dilemma:

	C	*D*
C	*R, R*	*S, T*
D	*T, S*	*P, P*

Evolutionary Equilibrium

Strict nash equilibrium:

$$\begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

- *A is a Strict Nash Equilibrium if* $a > c$
- *A is a Nash Equilibrium if* $a \geq c$
- *B is a Strict Nash Equilibrium if* $d > b$
- *B is a Nash Equilibrium if* $d \geq b$

There may be no dominant strategies, so we should use the primitive definition of Nash Equilibrium:

> *A Nash equilibrium is an action profile* a^* *with the property that no player i can do better by choosing an action different from* a_i^**, given that every other player j adheres to* a_j^*.

Player1:

$$\begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ \\ B \end{array} & \begin{pmatrix} a & b \\ \vee & \\ c & d \end{pmatrix} \end{array}$$

Player2:

$$\begin{array}{c c} & \begin{array}{ccc} A & & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a & > & c \\ b & & d \end{pmatrix} \end{array}$$

(A, A)

Evolutionary Equilibrium

Evolutionarily stable strategy(ESS):

*In an intense competitive environment, imagine there is a large population of **A** players. A single mutant of type **B** is introduced. What is the condition for selection to oppose the invasion of **B** into **A**?*

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{A} & \mathrm{B} \end{array} \\ \begin{array}{c} \mathrm{A} \\ \mathrm{B} \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

Let us assume there is an infinitesimally small quantity of B invaders. The frequency of B is ε, and the frequency of A is 1-ε.

The fitness of A is greater than B if:

$$a(1-\varepsilon) + b\varepsilon > c(1-\varepsilon) + d\varepsilon$$

Evolutionary Equilibrium

数据挖掘实验室
Data Mining Lab

Evolutionarily stable strategy(ESS):

canceling the ε terms, this inequality leads to:

$$a(1-\varepsilon) + b\varepsilon > c(1-\varepsilon) + d\varepsilon$$

$$a > c$$

If, however, it happens that $a = c$, the inequality leads to:

$$b > d$$

Summary:

Strategy A is ESS if either

(i) $a > c$ or (ii) $a = c$ and $b > d$

Strict nash equilibrium -> ESS -> nash equilibrium

Evolutionary Equilibrium

Evolutionarily stable strategy(ESS):

$$
\begin{array}{rcc}
 & \text{TFT} & \text{ALLD} \\
\begin{array}{r}\text{TFT}\\ \text{ALLD}\end{array} & \multicolumn{2}{c}{\begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix}}
\end{array}
$$

TFT: CDDDDDDDDD…

ALLD : DDDDDDDDDD…

If m > (T - P)/(R - P), TFT can't be invaded by ALLD. However, ALLD can always oppose TFT's invasion.

$$
\begin{array}{rcc}
 & \text{ALLC} & \text{ALLD} \\
\begin{array}{r}\text{ALLC}\\ \text{ALLD}\end{array} & \multicolumn{2}{c}{\begin{pmatrix} mR & mS \\ mT & mP \end{pmatrix}}
\end{array}
$$

ALLC: CCCCCCCCCC…

ALLD : DDDDDDDDDD…

ALLC is exploited by ALLD, and can't defense the latter's intrusion.

Evolutionary Equilibrium

数据挖掘实验室
Data Mining Lab

Adaptive dynamics:

Replication equation:

The frequency of i

Fitness of i: $f_i(x) = \sum_{j=1}^{n} r_{ij} x_j$

$$x_i' = x_i[f_i(x) - \varphi(x)]$$

Payoff matrix: $R = [r_{ij}]_{n*n}$

Time derivative of the frequency of i

Average fitness: $\varphi(x) = \sum_i f_i(x) x_i$

Evolutionary Equilibrium

Adaptive dynamics:

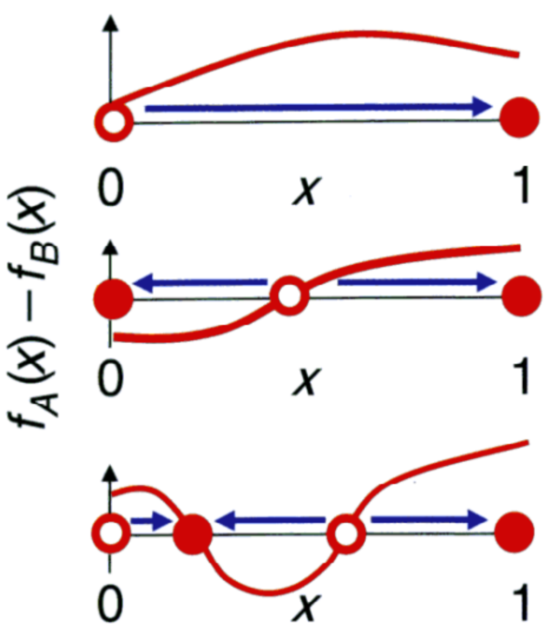
Two strategies:

$$x_A' = x_A[f_A(x) - \varphi(x)]$$
$$x_B' = x_B[f_B(x) - \varphi(x)]$$

$$\begin{cases} f_A(x) = r_{AA}x_A + r_{AB}x_B \\ \varphi(x) = f_A(x)x_A + f_B(x)x_B \end{cases}$$

$$x_B = 1 - x_A$$

$$x_A' = x_A(1-x_A)[f_A(x) - f_B(x)]$$



x: abundance of A

1-x: abundance of B

fitness difference between A and B

selection dynamics

stable equilibrium

unstable equilibrium

Evolutionary Equilibrium

Adaptive dynamics:

Two strategies:

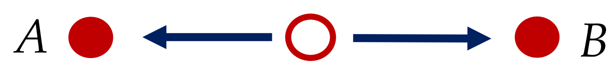
$$x_A' = x_A(1\text{-}x_A)[(a-b-c+d)x+b-d]$$

$$\begin{array}{ccc} & A & B \\ A & \multirow{2}{*}{$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$} & \\ B & & \end{array}$$

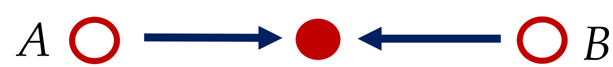
A dominates B, if a > c and b > d

B dominates A, if a < c and b < d

A and B are bistable, if a > c and b < d



A and B coexist, if a < c and b > d



A and B are neutral, if a = c and b = d

→ *Selection dynamics*

● *Stable equilibrium*

○ *Unstable equilibrium*

Evolutionary Equilibrium

Adaptive dynamics:

TFT: CCCCCCCCCC…
ALLC: CCCCCCCCCC…

TFT and ALLC are neutral

Random drift(Genetic drift): A finite population of randomly reproducing organisms would experience changes from generation to generation in the frequencies of the different genotypes.[Wiki]

ALLC: CCCCCCCCCC…
ALLD: DDDDDDDDD…

ALLD dominates ALLC

TFT: CDDDDDDDD…
ALLD : DDDDDDDDD…

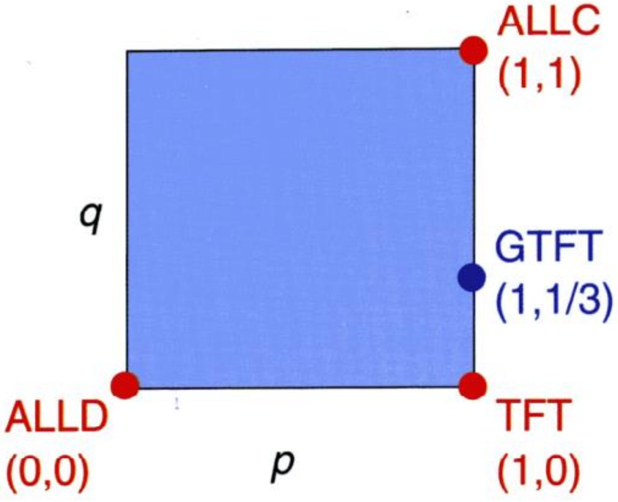
TFT and ALLD are bistable

"In silico" evolution:

Reactive strategies:

These strategies are given by two parameters:

- *p denotes the probability of cooperating if the opponent has cooperated in the previous move;*
- *q denotes the probability of cooperating if the opponent has defected in the previous move.*



Evolutionary Equilibrium

数据挖掘实验室
Data Mining Lab

"In silico" evolution:

Repeating prisoners' dilemma:

CC

CC

Markov Matrix: $S_1 = (p_1, q_1)$ *vs.* $S_2 = (p_2, q_2)$

$$\boldsymbol{M} = \begin{array}{c} \\ CC \\ CD \\ DC \\ DD \end{array} \begin{array}{c} \begin{array}{cccc} CC & CD & DC & DD \end{array} \\ \begin{pmatrix} p_1 p_2 & p_1(1-p_2) & (1-p_1)p_2 & (1-p_1)(1-p_2) \\ q_1 p_2 & q_1(1-p_2) & (1-q_1)p_2 & (1-q_1)(1-p_2) \\ p_1 q_2 & p_1(1-q_2) & (1-p_1)q_2 & (1-p_1)(1-q_2) \\ q_1 q_2 & q_1(1-q_2) & (1-q_1)q_2 & (1-q_1)(1-q_2) \end{pmatrix} \end{array}$$

If the entries of M or any power of M are all positive, markov matrix is irreducible, and it has a stationary vector x: $x^T M = x^T$.

Evolutionary Equilibrium

数据挖掘实验室
Data Mining Lab

"In silico" evolution:

Repeating prisoners' dilemma:

$$\begin{array}{cccc} \text{CC} & \text{CD} & \text{DC} & \text{DD} \end{array}$$

$$x = [s_1 s_2, s_1(1 - s_2), (1 - s_1)s_2, (1 - s_1)(1 - s_2)]$$

$$\begin{array}{c|cc} & \text{C} & \text{D} \\ \hline \text{C} & R & S \\ \text{D} & T & P \end{array} \qquad \begin{pmatrix} R & S \\ T & P \end{pmatrix}$$

$$s_1 = \frac{q_2 r_1 + q_1}{1 - r_1 r_2} \qquad r_1 = p_1 - q_1$$

$$s_2 = \frac{q_1 r_2 + q_2}{1 - r_1 r_2} \qquad r_2 = p_2 - q_2$$

Thus

$$\boxed{E_1(S_1, S_2) = R * s_1 s_2 + S * s_1(1 - s_2) + T * (1 - s_1)s_2 + P * (1 - s_1)(1 - s_2)}$$

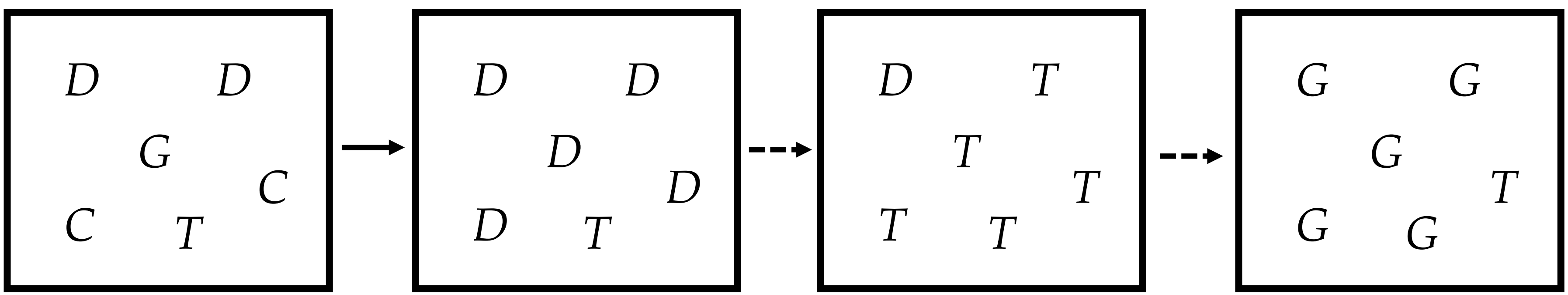
*Then use this to calculate the n*n payoff matrix, insert the matrix into the replicator equation, observe the evolutionary trajectory.*

"In silico" evolution:

Repeating prisoners' dilemma:

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

D: ALLD
C: ALLC
T: TFT
G: GTFT — *Generous TFT*

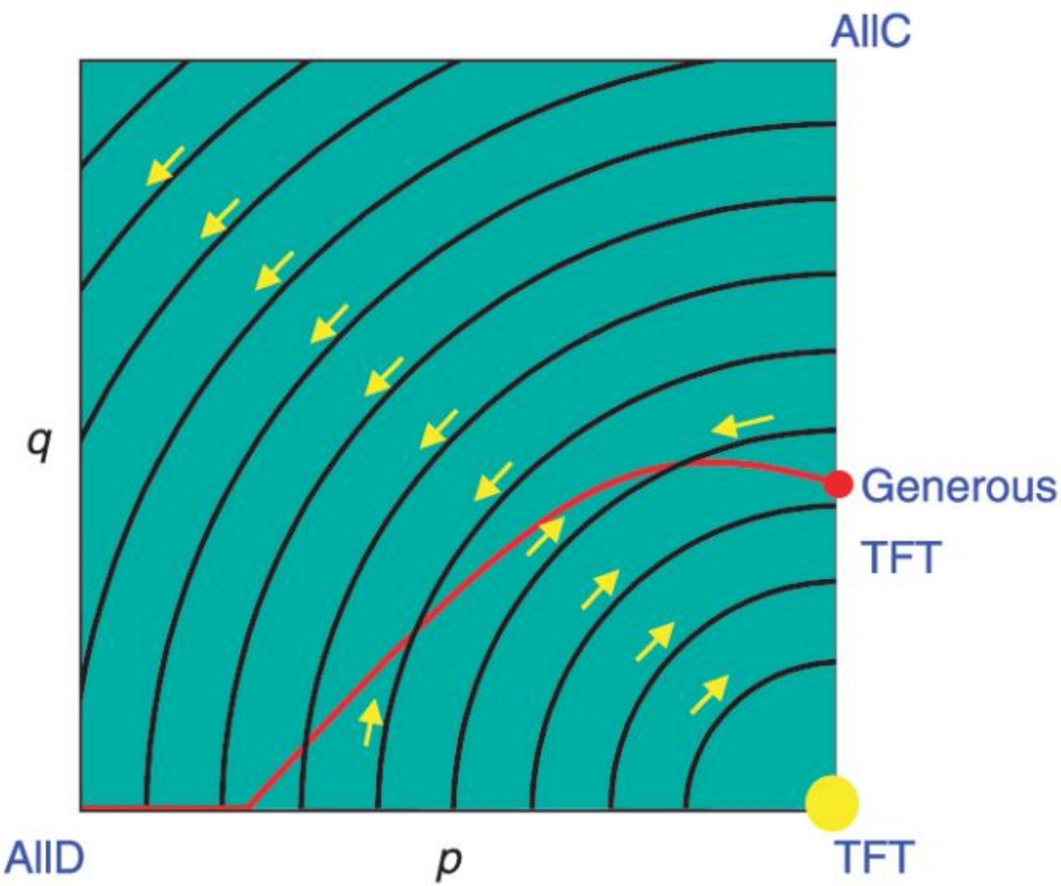


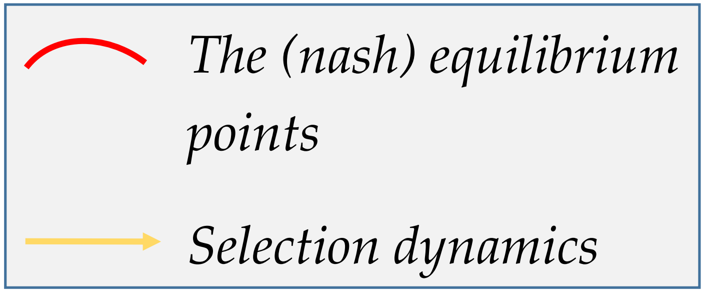
Evolutionary Equilibrium

数据挖掘实验室
Data Mining Lab

"In silico" evolution:

Repeating prisoners' dilemma:





Evolutionary Equilibrium

数据挖掘实验室
Data Mining Lab

"In silico" evolution:

Repeating prisoners' dilemma:

GTFT(Generous TFT):

- *Can make up for mistakes. With a certain probability, a sequence of cooperation and defection leads back to mutual cooperation.*

  *GTFT: CCC**D**CDCCCC…*

  *GTFT: CCCCDC**C**CCC…*

- *When GTFT plays ALLD, it has a slightly lower payoff than when TFT plays ALLD. Hence we need TFT to initiate cooperation. But once a cooperation population has been established, TFT will be replaced by GTFT.*

Evolutionary Equilibrium

"In silico" evolution:

Return to reactive strategies:

Previous strategies base their decision only on the last move of the opponent, but now we consider those also base their decision on their own last action.

Last Round	CC	CD	DC	DD
S	p_1	p_2	p_3	p_4

The conditional probability to cooperate

$ALLC = S(1, 1, 1, 1)$

$ALLD = S(0, 0, 0, 0)$

$TFT = S(1, 0, 1, 0)$

$GTFT = S(1, 1/3, 1, 1/3)$

Evolutionary Equilibrium

"In silico" evolution:

Return to reactive strategies:

Last Round	CC	CD	DC	DD
S	p_1	p_2	p_3	p_4

Markov Matrix: $S_1=(p_1, p_2, p_3, p_4)$ *vs.* $S_2=(p_1', p_2', p_3', p_4')$

CC

CC

$$
\boldsymbol{M} = \begin{array}{c} \\ \text{CC} \\ \text{CD} \\ \text{DC} \\ \text{DD} \end{array}
\begin{array}{c}
\begin{array}{cccc} \text{CC} & \text{CD} & \text{DC} & \text{DD} \end{array} \\
\begin{pmatrix}
p_1p_1' & p_1(1-p_1') & (1-p_1)p_1' & (1-p_1)(1-p_1') \\
p_2p_3' & p_2(1-p_3') & (1-p_2)p_3' & (1-p_2)(1-p_3') \\
p_3p_2' & p_3(1-p_2') & (1-p_3)p_2' & (1-p_3)(1-p_2') \\
p_4p_4' & p_4(1-p_4') & (1-p_4)p_4' & (1-p_4)(1-p_4')
\end{pmatrix}
\end{array}
$$

To compute the stationary vector x: $x^T M = x^T$

Evolutionary Equilibrium

"In silico" evolution:

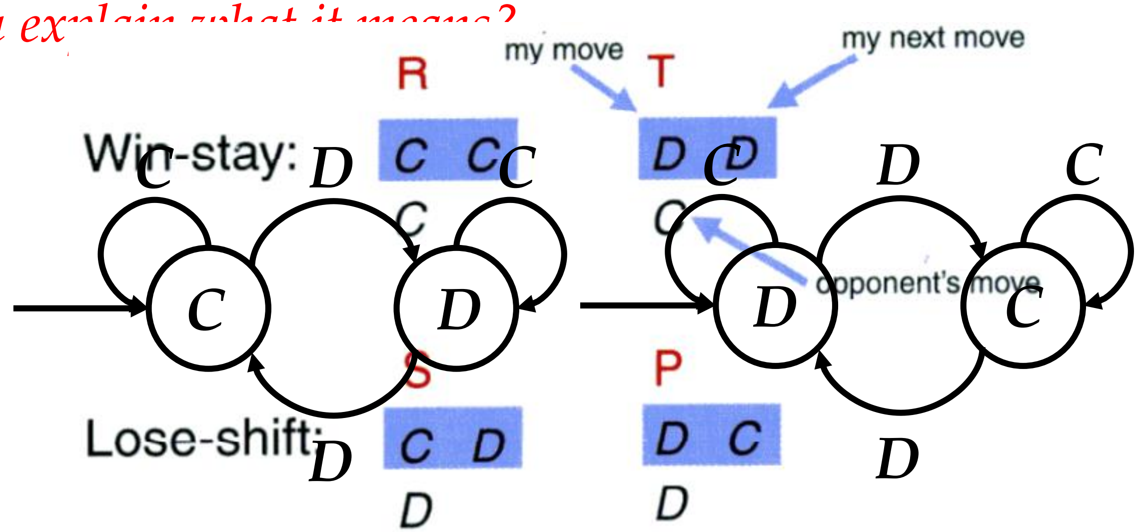
Return to reactive strategies:

Originally this experiment was performed to confirm the success of GTFT. Indeed for some instantiations of this evolutionary process, GTFT is the ultimate winner. But more often another strategy is discovered unexpectedly. This strategy is **S(1, 0, 0, 1)**.

Can you explain what it means?

Win-stay, Lose-shift(WSLS)

Last Round	CC	CD	DC	DD
S	1	0	0	1
	C	D	D	C



Evolutionary Equilibrium

数据挖掘实验室
Data Mining Lab

"In silico" evolution:

Return to reactive strategies:

Win-stay, Lose-shift(WSLS):

- *Can correct occasional mistakes, and it's a deterministic corrector, while GTFT is a stochastic one.*

WSLS: CCC**D**DCCCCC…

WSLS: CCCCDCCCCC…

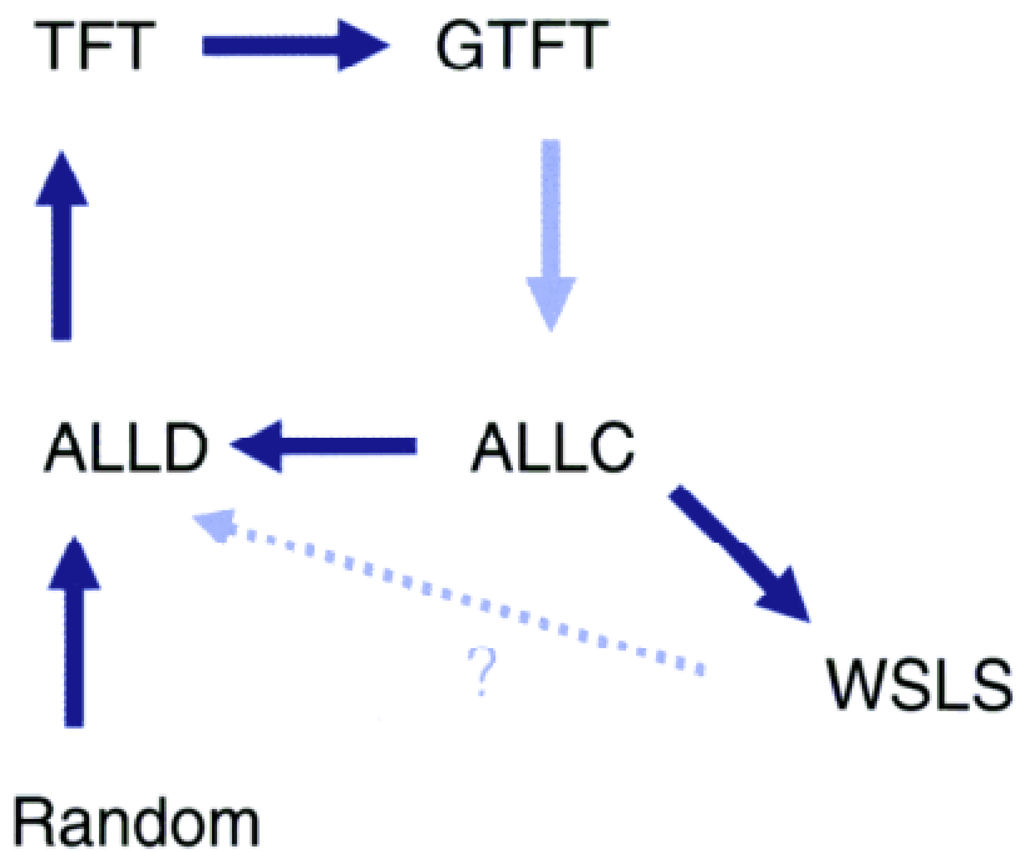
- *WSLS also has another advantage over GTFT when both plays ALLC. Briefly speaking, WSLS can exploite ALLC, but GTFT can not.*

WSLS: CCC**D**DDDDD…

ALLC: CCCCCCCCCC…

Sadly, this behavior is also compatible with human psychology: unconditional(defenseless) cooperators tend to be exploited!

"In silico" evolution:



Oscillations of cooperation and defection

04 Conclusion

Conclusion

*Review **Prisoners' Dilemma***

✓ *Why nash equilibrium is so shocking*

*Review **Repeated Games***

✓ *Strategy*

✓ *Axelrod's tournament*

✓ *TFT(Tit-For-Tat)*

*Introduce **Evolutionary Equilibrium***

✓ *Evolutionarily stable strategy(ESS)*

✓ *Adaptive dynamics*

✓ *"In silico" evolution*

✓ *GTFT*

✓ *WSLS*

Reference

- ✓ *Osborne M J. An introduction to game theory[M]. New York: Oxford university press, 2004.*
- ✓ *Nowak M A. Evolutionary Dynamics: Exploring the Equations of Life. 2006[J]. Massachusetts: Belknap Press Google Scholar.*
- ✓ *Osborne M J, Rubinstein A. A course in game theory[M]. MIT press, 1994.*
- ✓ *Binmore K. Playing for real: a text on game theory[M]. Oxford university press, 2007.*
- ✓ *Nowak M A, Sigmund K. Evolutionary dynamics of biological games[J]. science, 2004, 303(5659): 793-799.*

Thanks!