Evolution of Diversity and Cooperation

1 / 3

Jorge M. Pacheco

atpgroup

Departamento de Matemática & Aplicações
Universidade do Minho
Portugal

UIMP Universidad Internacional Menéndez Pelayo

RSEM

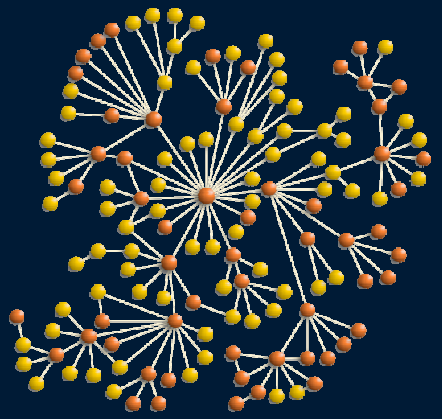
FisyMat

Luis Santaló School, 15th of July, 2013

synopsis

1st part

- What do I mean by “cooperation”?
- Game theory ⊗ Evolutionary game theory (EGT)
- Classic mechanisms of cooperation
- Other mechanisms of cooperation



how to quantify cooperation ?

Donor
Pays a cost c

→

Receiver
Receives a benefit b

b>c

Why should we cooperate?

If natural selection is based on competition, how can it lead to cooperation ?

how to quantify cooperation ?

Donor
Pays a cost c

Receiver
Receives a benefit b

b>c

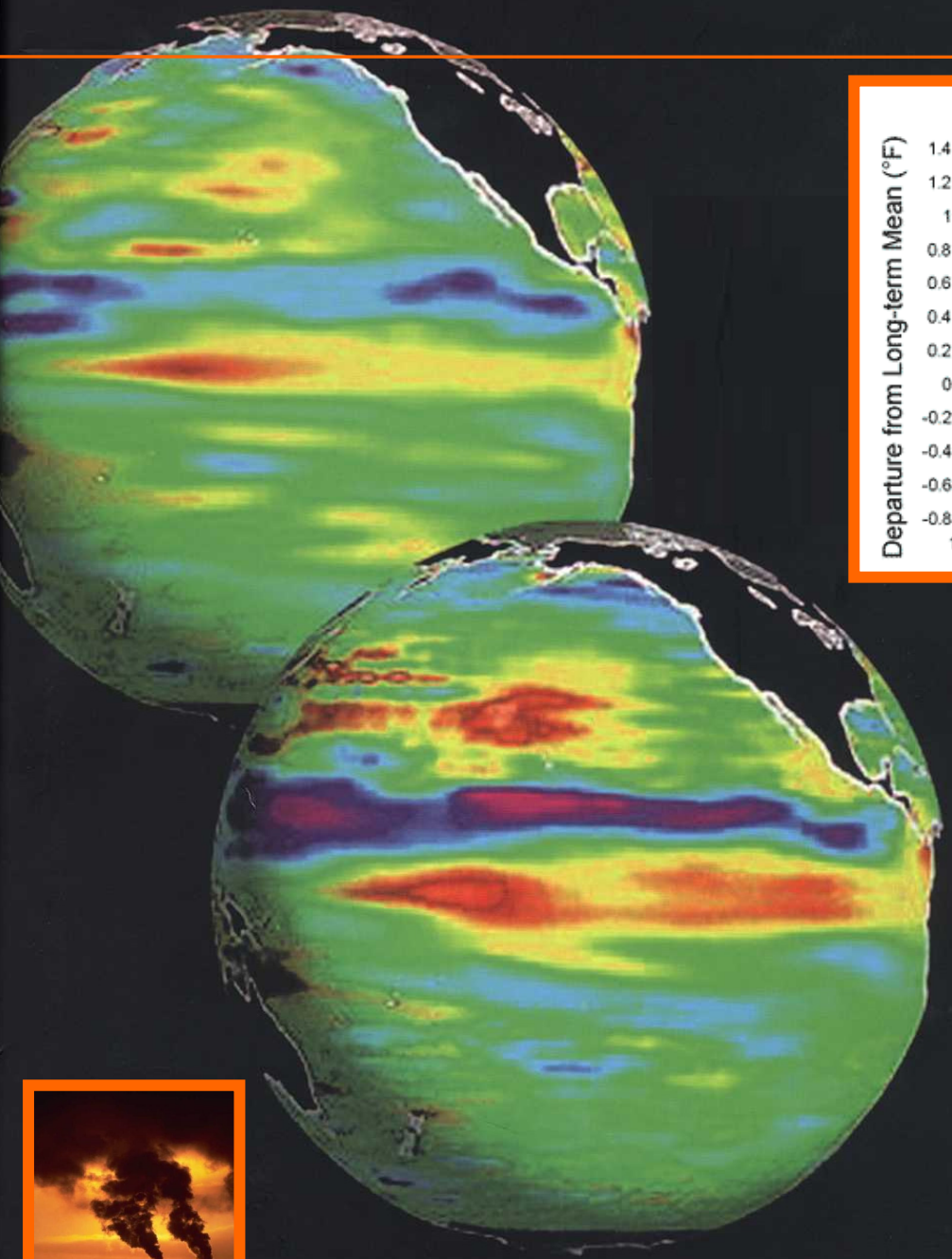
Social insects, etc.

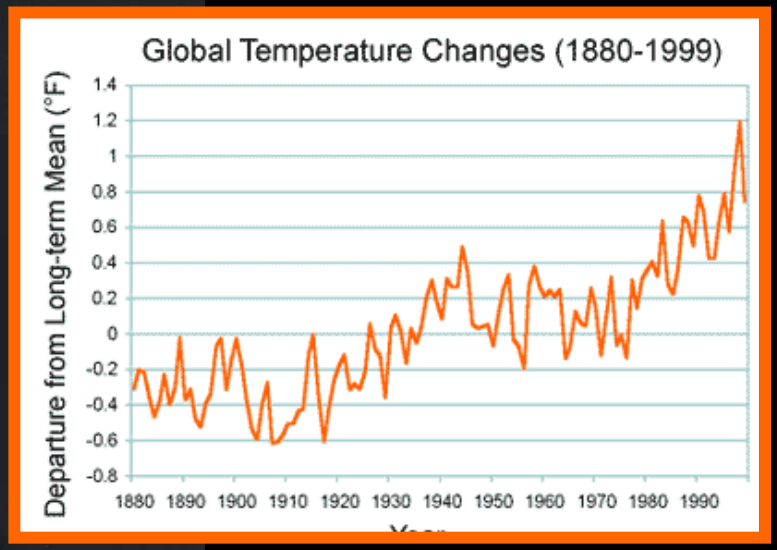
by not reproducing, workers reduce their own reproductive success to help others (the Queen)

cooperation among humans

- collective action to protect, hunt, nourish, etc.
- water sharing
- tax paying and social welfare
- open source projects…

global warming





taming the climate requires …

cooperation !

[Santos, Pacheco, *PNAS* 108 (2011) 10421-10425]

evolution of complexity

cooperation is on the basis of some of the *major transitions in evolution*
emergence of multicellularity

cooperation is essential for the evolution of reproductive entities

genes *cooperate* to form cells
cells *cooperate* to form multicellular organisms
individuals *cooperate* to form groups and societies
human culture is a *cooperative* process.

[Maynard-Smith & Szathmáry, *The major transitions in evolution*, OUP95]

Interdisplinary subject

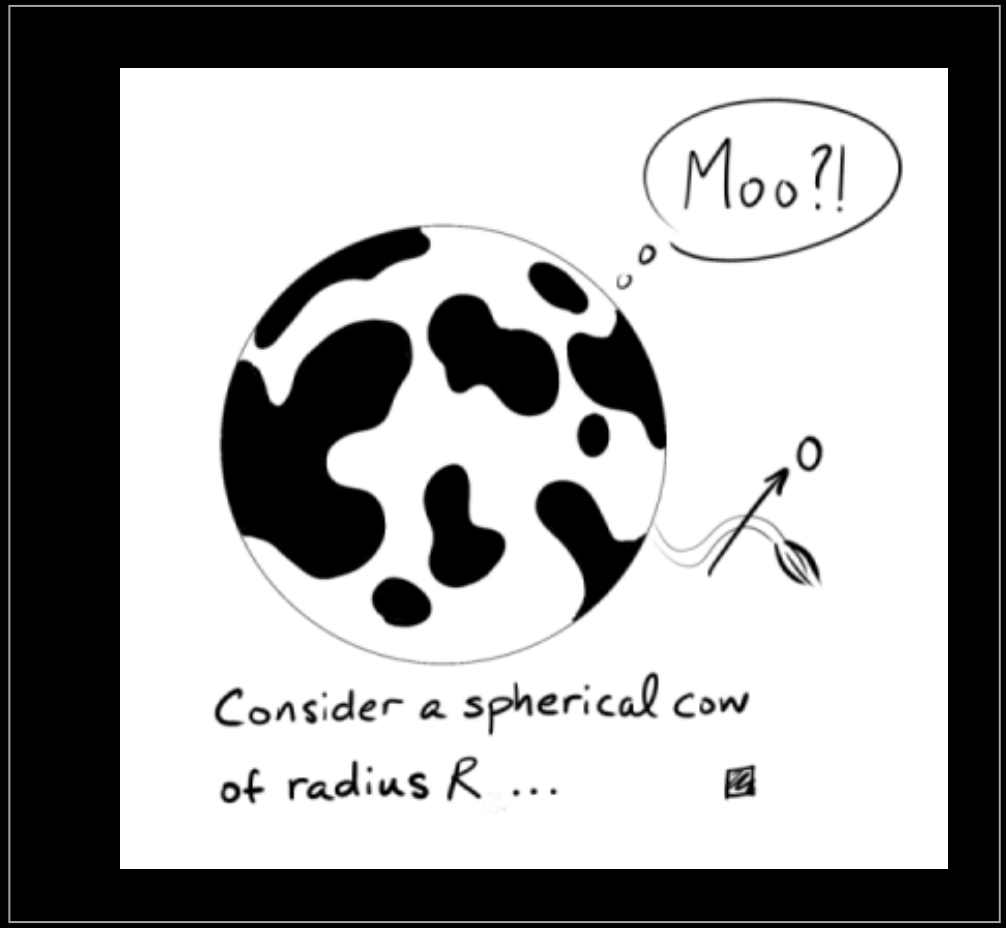
understanding *cooperation* remains a *major challenge to* scientists from fields as diverse as *biology*, *political science*, *economics, anthropology, history, mathematics, physics, computer science, law,* etc.

many disciplines… one common language... same math!!

mathematical framework : (evolutionary) Game Theory
metaphors : prisoner's dilemma, ultimatum game, etc.

classical game theory &
2-person dilemmas of cooperation

models & simple models



models should be as simple as possible, but not simpler . . .

Albert Einstein

prisoner's dilemma &
the *cost-benefit* dilemma

Donor → **Receiver**

C : *"pays" a cost c* — *receives a benefit b*

D : *"pays" nothing* — *receives nothing*

when 2 players interact simultaneously and decide independently

"rational" goal :
maximize your own payoff !

if your opponent plays **C** :
you better play **D**.
if your opponent plays **D** :
you better play **D**.

BUT:
CC is better than **DD**

your opponent

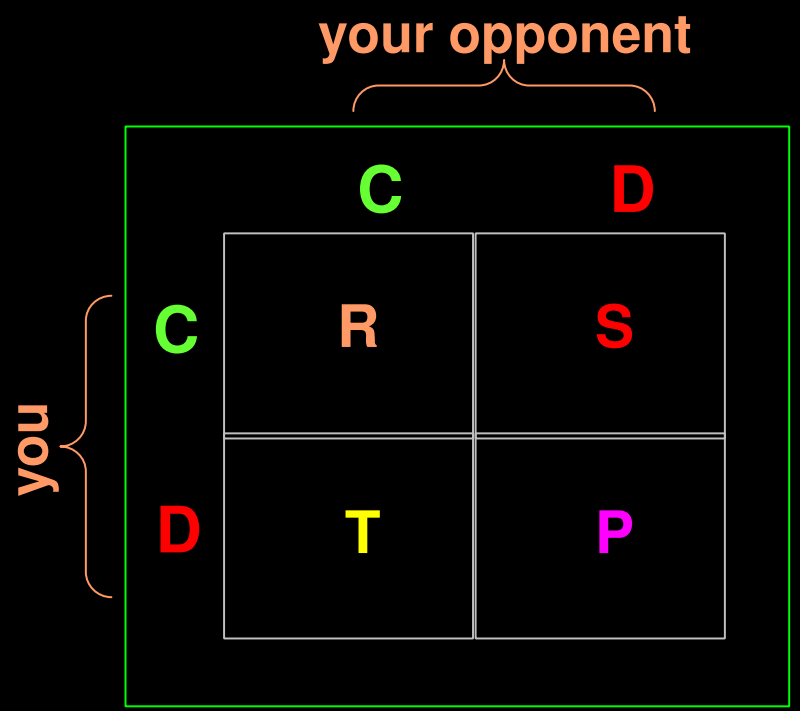
you \ opponent	C	D
C	$b - c$	$-c$
D	b	0

dilemma : although **mutual cooperation (CC)** is better than **mutual defection (DD)**, individual "rational choice" leads to **DD**

symmetric 2-person dilemmas of cooperation

- symmetric 2-player games :
- 2 individuals meet
- each player uses 1 of 2 strategies (Cooperate or Defect)
- each possible outcome has an associated payoff
  (tabulated in the *payoff-matrix*)

R : mutual cooperation
P : mutual defection
S : sucker's payoff
T : temptation to defect



fear & greed

P > S → (DD is better than CD)
one may associate S with fear (of being cheated)

T > R → (DC is better than CC)
one may associate T with greed (temptation to cheat)

---

we can fix R=1 and P=0, and vary the intensities of greed and fear at will.
as a result, we obtain the most popular social dilemmas of cooperation:

	C	D
C	R	S
D	T	P

	greed	fear
SG: snowdrift game	T > R	> S > P
SH: stag-hunt game	R > T >	P > S
PD: prisoner's dilemma	T > R >	P > S

snowdrift game

T > R > S > P

they may get out and start shovelling (**cooperate**) or remain in the car (**defect**).

if they both **cooperate**, each gets a benefit *b* and shares the cost *c* of shovelling; they end up benefitting both **R** = *b* - ½*c*.

if only one does the work, then he gets **S** = *b* - *c* , whereas the defector gets **T** = *b* . If they both remain in the car, thet both get **P** = 0. *As a result . . .*

snowdrift game

$$T > R > S > P$$

you \ your opponent	C	D
C	b-c/2	b-c
D	b	0

cooperate if the other defects

&

defect if the other cooperates

stag-hunt dilemma (*ex of* coordination games)

R > T > P > S

individual **hunting: does not depend on others… but offers lower benefits.**

defect (go for hare)

collective **hunting: highest benefits.**

cooperate (go for stag = collective hunting)

cooperate if the other cooperates

&

defect if the other defects

2D space of 2-person dilemmas

SG: snowdrift game

$T > R > S > P$

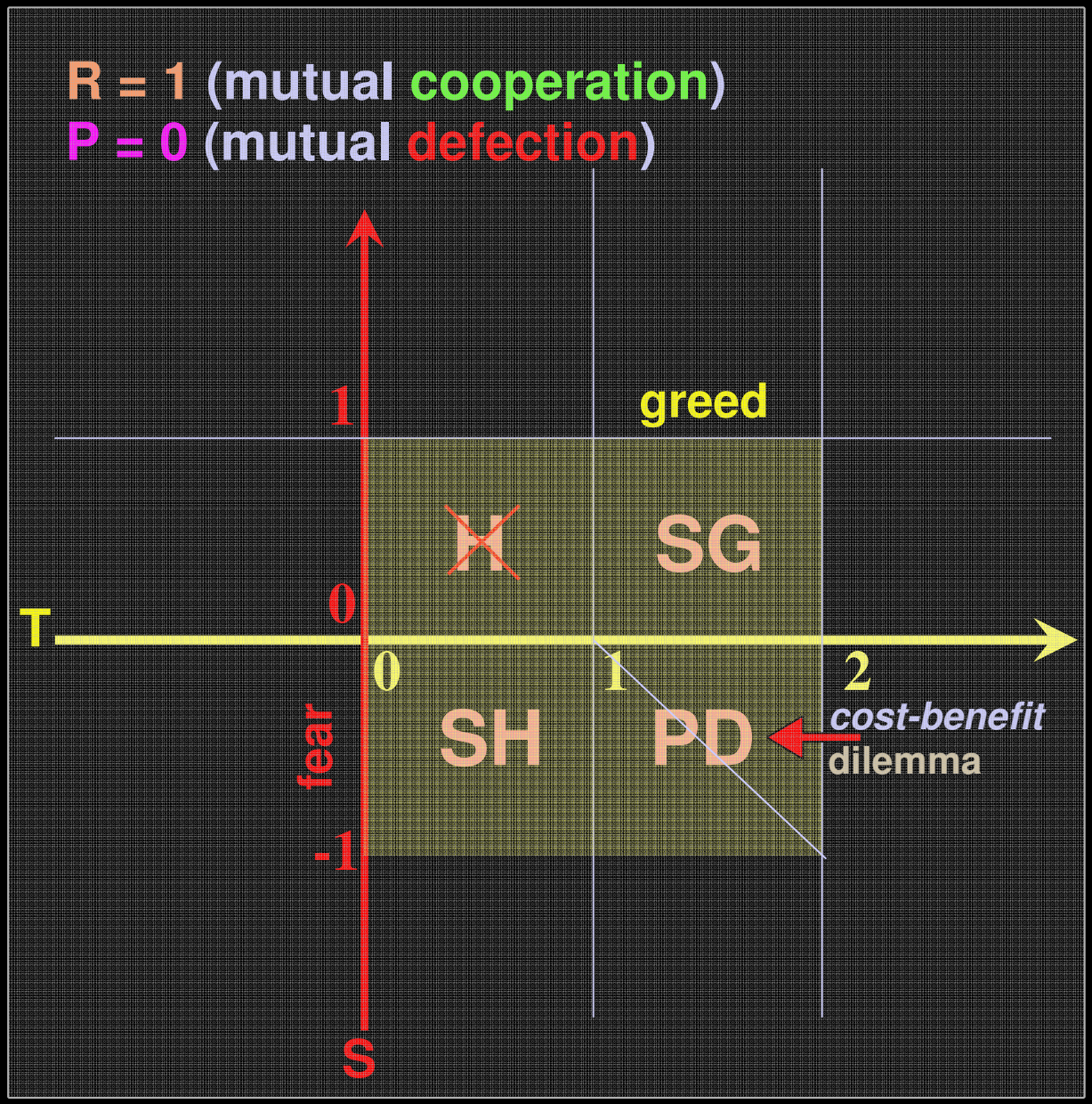
SH: stag-hunt game

$R > T > P > S$

PD: prisoner's dilemma

$T > R > P > S$

	C	D
C	R	S
D	T	P



[Santos, Pacheco, Lenaerts, *PNAS* 103 (2006) 3490]

evolutionary game theory

payoff ⇨ fitness ⇨ reproductive or social success
2 individuals ⇨ a population of individuals
dynamics : composition of the pop will change in time

genetic evolution

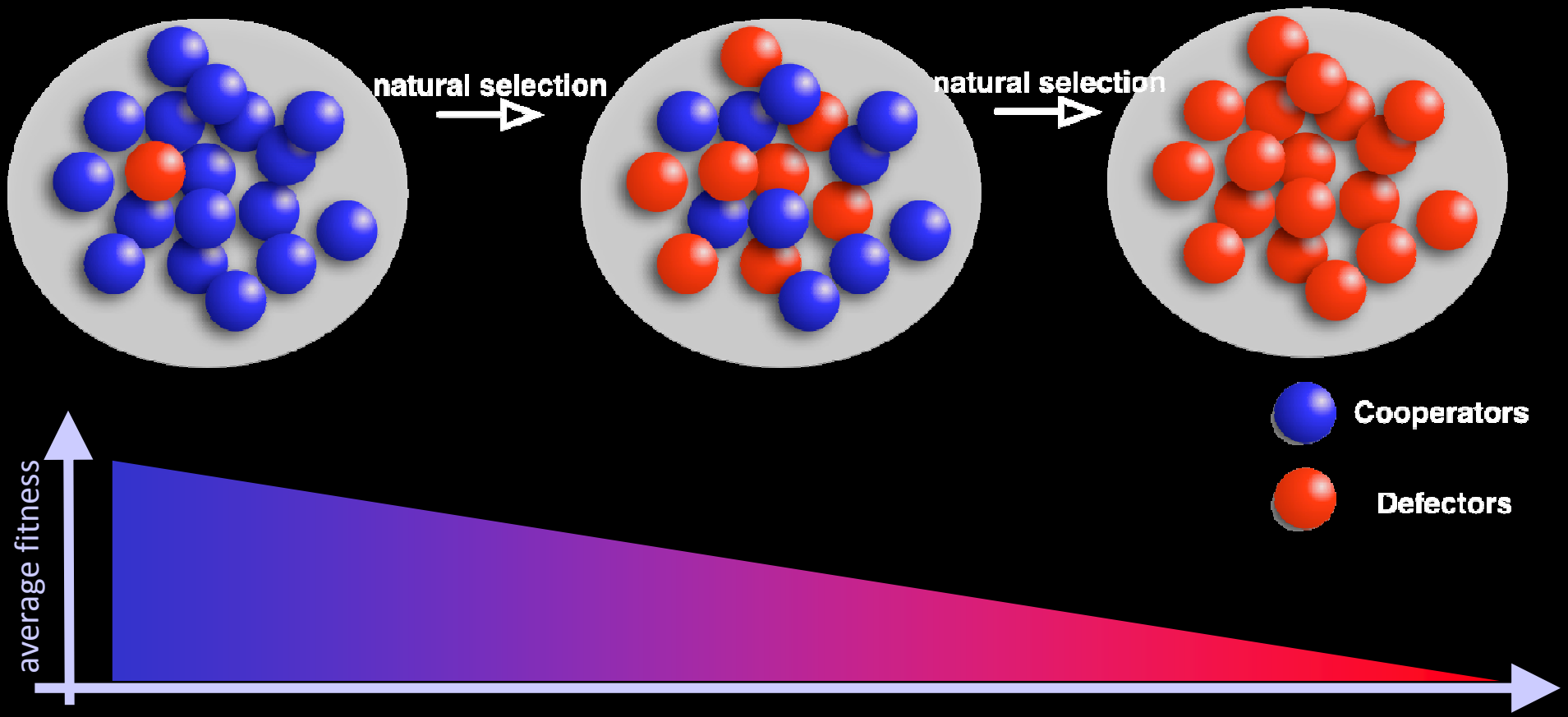
individuals with *higher fitness* will have *more offspring*

social/cultural evolution

***individual behaviors* with *higher fitness* will be *imitated more often* (social learning)**

evolutionary game theory

prisoner's dilemma



natural selection leads to the *demise of cooperators*

general stability concepts

game theory

Nash equilibrium

If a strategy is a Nash equilibrium, and if both players play that strategy, then neither person can deviate from that strategy and increase her payoff

evolutionary game theory

Evolutionarily Stable Strategy

If a strategy is an ESS, then an infinitesimally small amount of players of the other strategy will never be able to invade (spread over the entire — *infinite* — population)

	C	D
C	R	S
D	T	P

C is a NASH equilibrium *if* $R \geq T$
D is a NASH equilibrium *if* $P \geq S$

C is an ESS *if* $R > T \quad \vee$
$R = T \wedge S > P$

C is a **NASH** equilibrium in **H** & **SH**
D is a **NASH** equilibrium in **SH** & **PD**

C is an **ESS** in **H** & **SH**
D is an **ESS** in **SH** & **PD**

the *math* of (standard) evolutionary game theory

populations are *infinite*; there is a fraction *x* of Cs & (*1-x*) of Ds

populations are *well-mixed* ; everybody is equally likely to interact with everybody else (*mean field*); the frequency with which each C interacts with a D is given by (*1-x*) & vice versa; hence ALL Cs have the same fitness & also ALL Ds have the same fitness

evolution → *replicator dynamics* :
strategies' evolution follow the gradient of (natural) selection determined by relative fitness

replicator equation

remember: populations are *infinite*

$$\begin{cases} \dot{x}_C = x_C(f_C(\vec{x}) - \phi) \\ \dot{x}_D = x_D(f_D(\vec{x}) - \phi) \end{cases}$$

those strategies whose fitness (reproductive success) exceeds the average fitness ϕ of the population will increase in frequency; those that don't will decline.

$$\vec{x} = (x_C, x_D)$$

$$\phi = x_C f_C + x_D f_D$$

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{bmatrix} R & S \\ T & P \end{bmatrix} \end{array}$$

$$f_C(\vec{x}) = x_C R + x_D S$$

$$f_D(\vec{x}) = x_C T + x_D P$$

replicator equation

$$x_C + x_D = 1 \rightarrow x \equiv x_C \Rightarrow x_D = 1 - x \rightarrow$$ **1 equation !**

$$\dot{x} = x(1-x)\underbrace{[f_C(x) - f_D(x)]}_{\Delta(x)}$$

(gradient of selection)

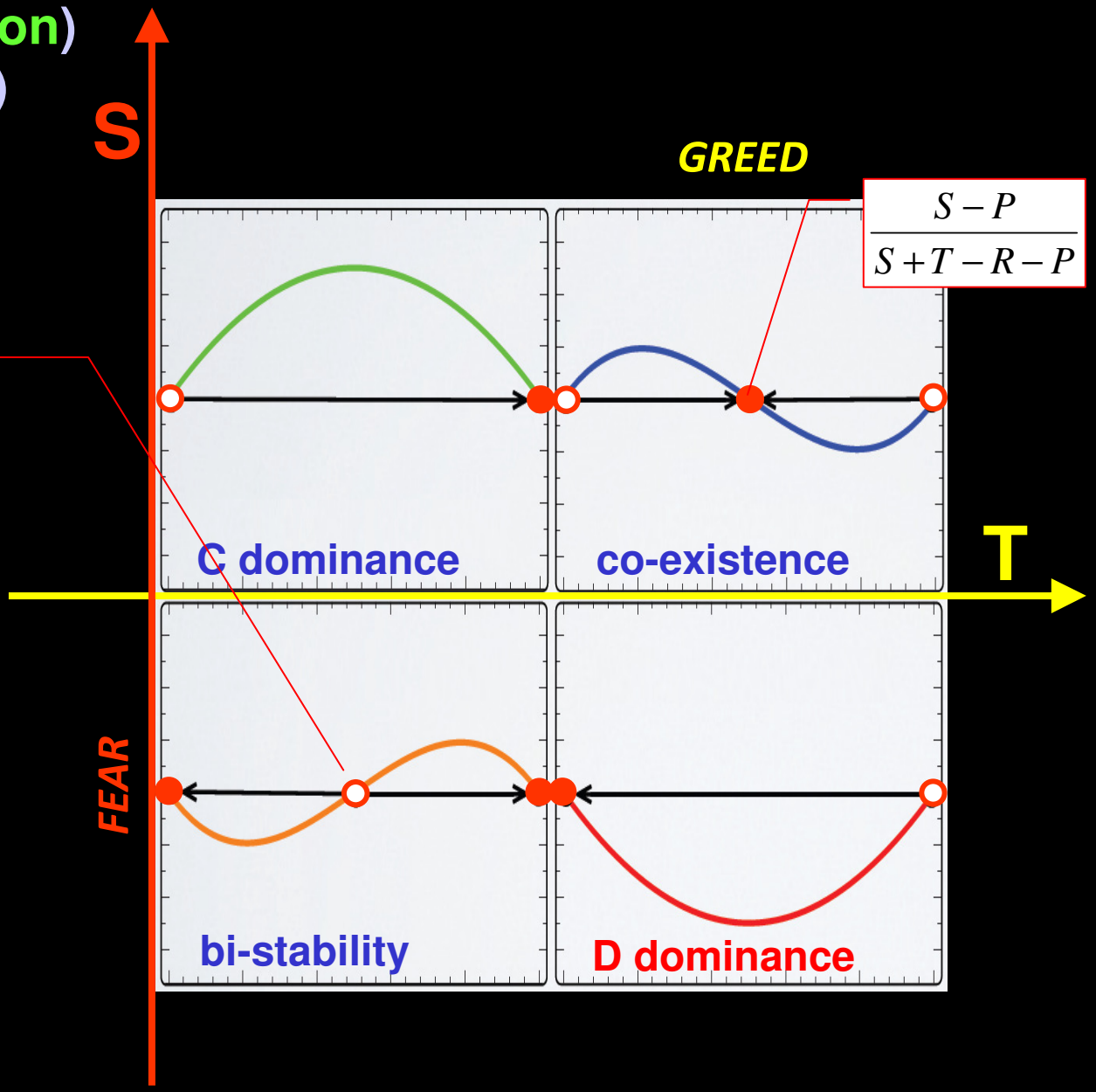
equilibria of the replicator equation

$$x = 0 \vee x = 1 \vee \Delta(x) = 0$$

$$\Delta(x) = (R - T - S + P)x + (S - P)$$

social dilemmas and the equilibria of the replicator dynamics

R = 1 (mutual cooperation)
P = 0 (mutual defection)



$\frac{S-P}{S+T-R-P}$

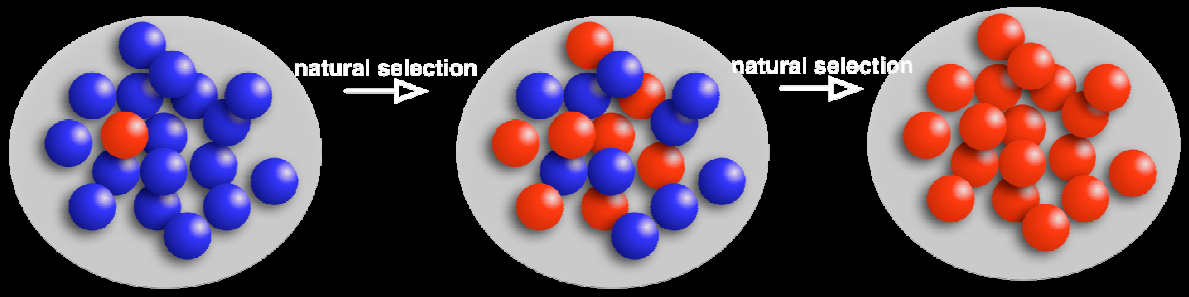
C is an ESS *if*

$$R > T \quad \vee$$
$$R = T \wedge S > P$$

C is an **ESS** in **H** & **SH**
D is an **ESS** in **SH** & **PD**

the paradox of cooperation

the prisoner's dilemma is the most famous metaphor of cooperation, but



natural selection leads to the extinction of cooperation ! However,

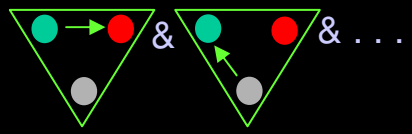
cooperation surrounds us !!!

what are we missing here ?!

Escaping the paradox of cooperation

- Kin selection — *all in the family . . .*

- Direct reciprocity — *I scratch your back & you scratch mine . . .*



- Indirect reciprocity — *I scratch your back & someone else scratches mine . . .*

  & evolution of moral systems — (*reputation & the evolution of the concepts of "good" and "evil"*)

Kin selection

all in the family . . .

the more individuals *are related*, the more cooperation is feasible. ***How ?***

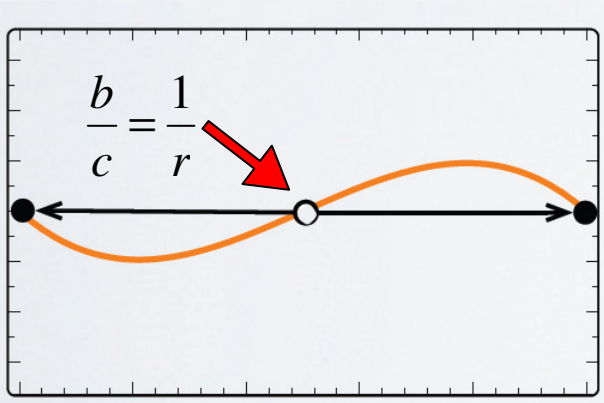
r : (genetic) relatedness between individuals → *your action means r to me; hence, I also get r of what you get;* then

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{bmatrix} (b-c)(1+r) & -c+br \\ b-cr & 0 \end{bmatrix} \end{array}$$

→ ***ESS*** condition : $\frac{b}{c} > \frac{1}{r}$

this is the famous ***Hamilton's rule of kin-selection***

how does the *gradient of selection* look like ?

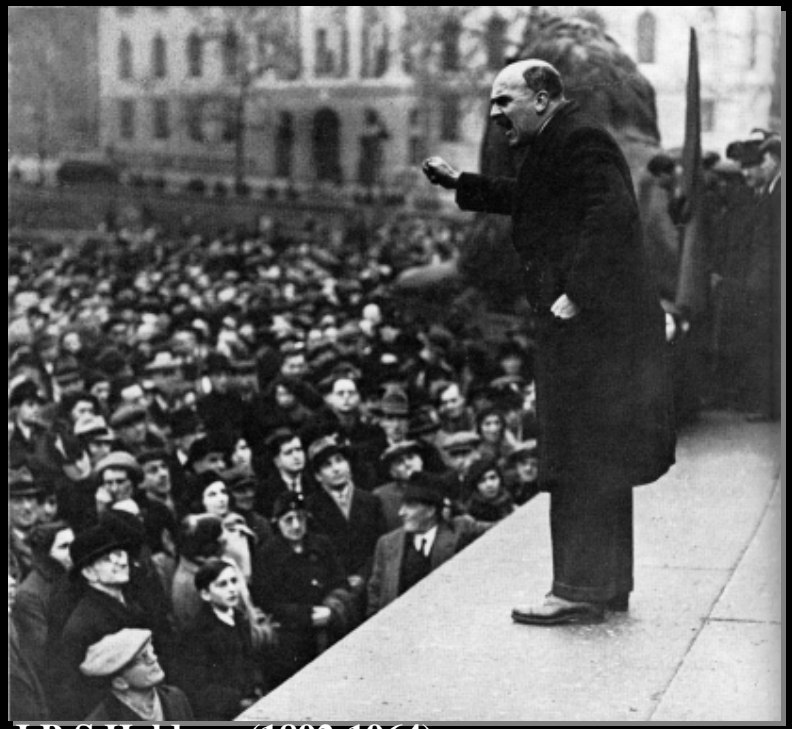


kin selection transforms a ***PD*** into a ***coordination game***

[W. D. Hamilton, J. Theor. Biol. 7, 1 (1964)]

Kin selection

all in the family . . .

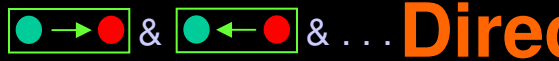


J.B.S Haldane (1892-1964)

"I will jump into the river to save two brothers or eight cousins."

$$\frac{b}{c} > \frac{1}{r}$$

for instance, *kin selection* explains the behaviour of eusocial insects



Direct reciprocity

I scratch your back & you scratch mine . . .

the same players interact more than once → *repeated games*

→ *complex strategies*

strategy : specifies ***next action*** given ***previous history***

simple strategies

ALLC → always cooperate (no memory)

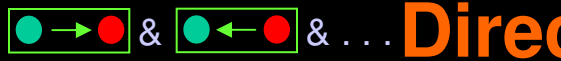
ALLD → always defect (no memory)

GRIM → (1 step memory:
start cooperating and change into permanent defection when your opponent defects for the first time)

T F T → tit-for-tat (1 step memory:
start cooperating and repeat the action of your opponent in previous move)

can direct reciprocity solve the paradox of cooperation ?

R. Trivers, Q. Rev. Biol. 46, 35 (1971).



Direct reciprocity

I scratch your back & you scratch mine . . .

folk theorem : if fully-rational players engage in a finite round repeated **PD**,

ALLD is an ***ESS***

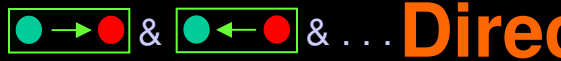
proof : backwards induction

experiments show that humans are **NOT** fully rational players

workaround : ***tomorrow never dies . . .***

w : probability for the occurrence of another round; then

results remain valid while we circumvent the ***folk-theorem***



Direct reciprocity

I scratch your back & you scratch mine . . .

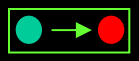
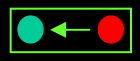
are there ideal strategies for the repeated PD ? → **Axelrod's tournaments !**

1st tournament : 14 players ; winner : *tit-for-tat*

Anatol RAPOPORT:

Start by cooperating, and repeat the strategy of your opponent in the previous move

2nd tournament : 63 players ; many of the new strategies would have won the 1st tournament

 &  & . . . **Direct reciprocity**

I scratch your back & you scratch mine . . .

are there ideal strategies for the repeated PD ? → **Axelrod's tournaments !**

1st tournament : 14 players ; winner : *tit-for-tat*

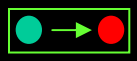
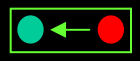
Anatol RAPOPORT:

Start by cooperating, and repeat the strategy of your opponent in the previous move

2nd tournament : 63 players ; many of the new strategies would have won the 1st tournament

winner : *tit-for-tat*

It's very hard to predict the outcome against an unknown set of strategies!

 &  & . . . Direct reciprocity

I scratch your back & you scratch mine . . .

public enemy = **ALLD**

let's play **TFT** ⊗ **ALLD**

assume a probability w of a next round (of the ***benefit-cost*** game) :

1 round :

$$\begin{array}{c} \\ TFT \\ ALLD \end{array}\begin{array}{c} \begin{array}{cc} TFT & ALLD \end{array} \\ \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array}$$

1/(1-w) rounds :

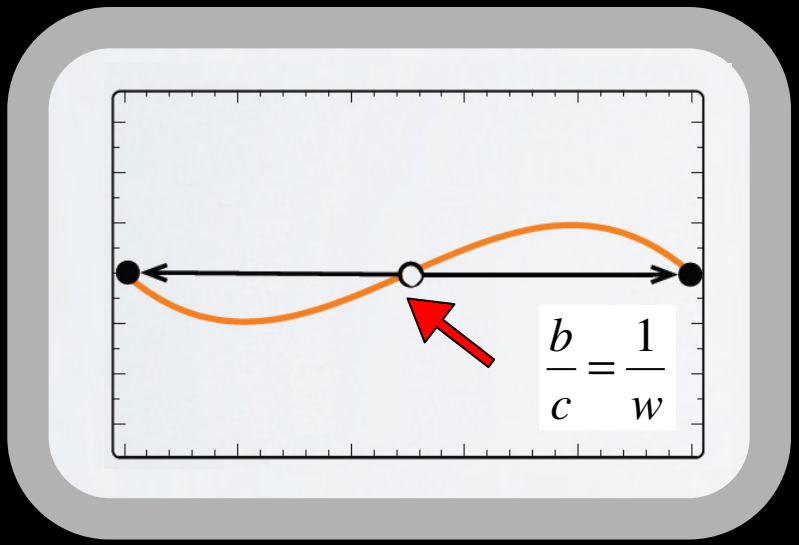
$$\begin{array}{c} \\ TFT \\ ALLD \end{array}\begin{array}{c} \begin{array}{cc} TFF & ALLD \end{array} \\ \begin{pmatrix} (b-c)/(1-w) & -c \\ b & 0 \end{pmatrix} \end{array}$$

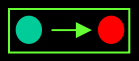
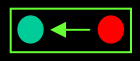
ESS condition : $\frac{b}{c} > \frac{1}{w}$ for a sufficiently large number of rounds, TFT gives cooperation a chance to invade.

how does the ***gradient of selection*** look like ?

reciprocity & TFT transform a repeated ***PD***

into a ***coordination game***




& 
& . . . **Direct reciprocity**

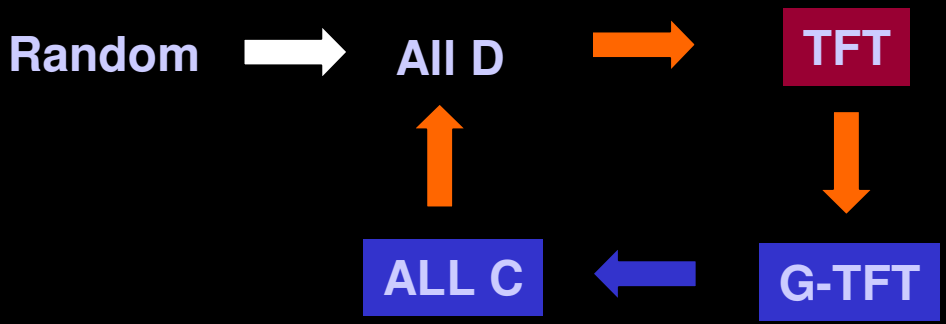
I scratch your back & you scratch mine . . .

*is **tit-for-tat** the final word ? **NO !***

tit-for-tat **is BAD at correcting errors (** ***trembling hand effect . . .*** **)**

Generous TFT: ***Never forget a good move, sometimes forget a bad one***

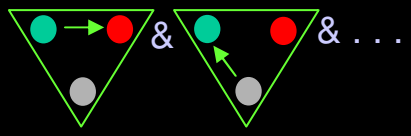
Let evolution decide:



Cycles of war and peace

win-stay, lose-shift **: if in the previous move I got** $(b\text{-}c)$ **or** (b)**, I stick to the "winning strategy"; if not, I change to the other strategy.**

[Nowak & Sigmund, Nature 364 (1993) 56–58]



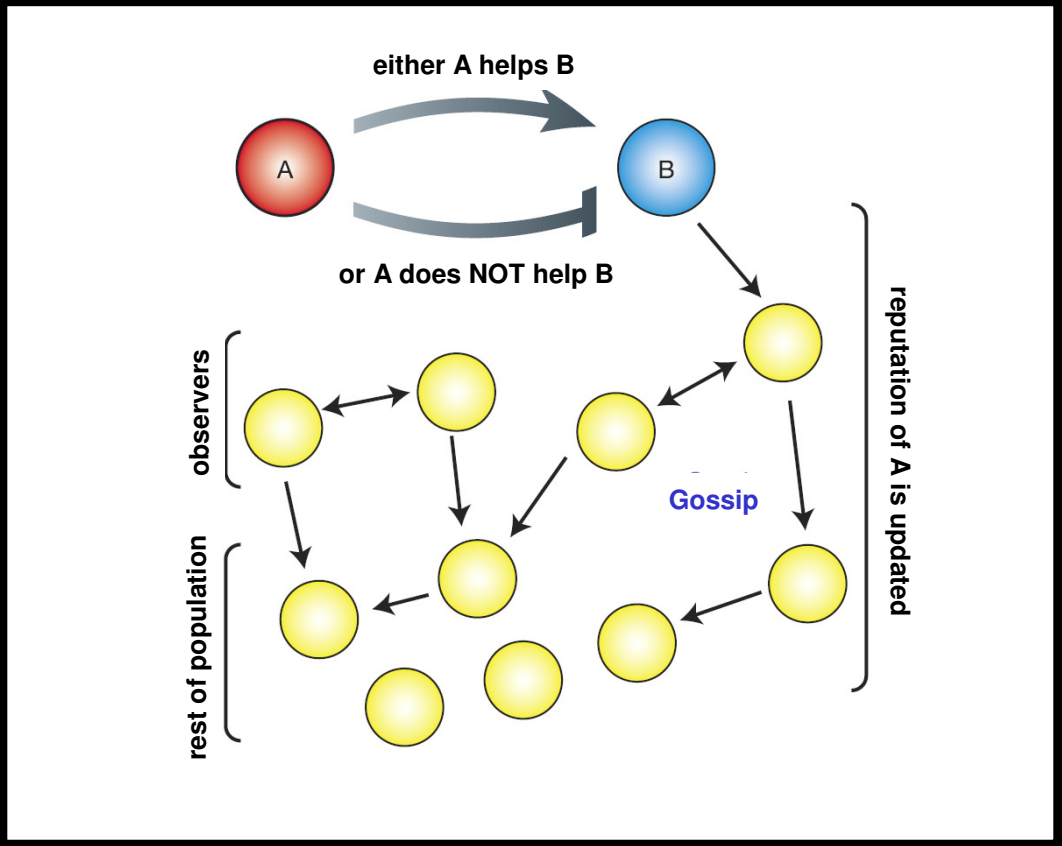
Indirect reciprocity

I scratch your back & someone else scratches mine . . .

I help you and somebody else will help me

indirect reciprocity → *reputation building* → *moral judgements* → . . .

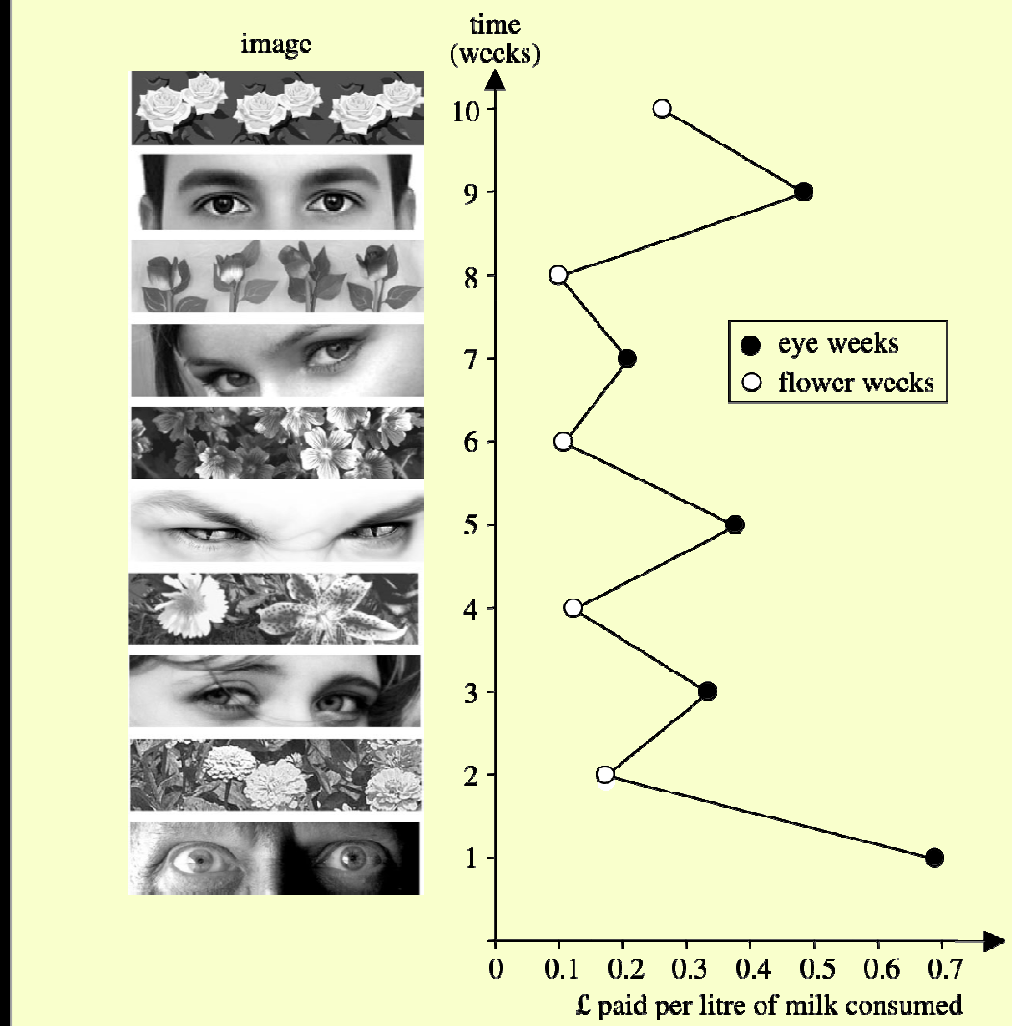
building a reputation



By helping (or not) another individual, a given player may modify its reputation, which may change the pre-disposition of others to help in the future

- *trust, gossip, assessment of actions (social norms)*

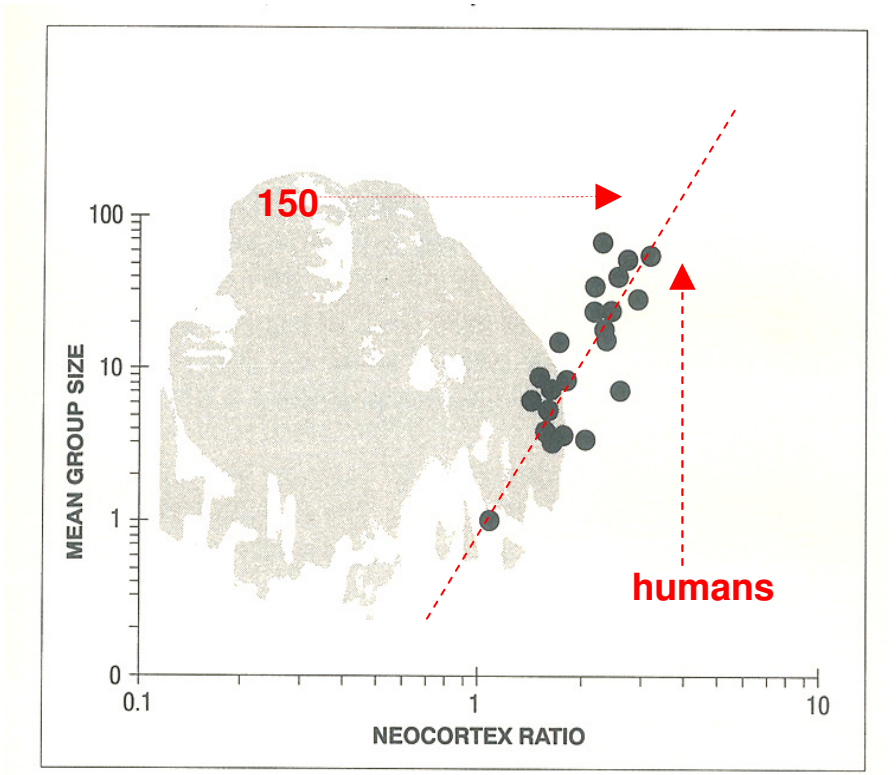
the strength of third parties



[Bateson et al., Biology Letters 2, 412–414 2006]

managing our image is not easy…

indirect reciprocity is the mechanism of cooperation most dependent on **cognition and social control**



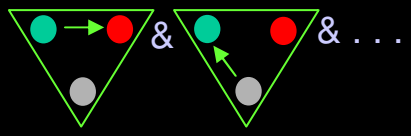
[R. Dunbar, "*Grooming, giossip & the evolution of human language*", HUP'02]

the large neocortex ratio reflects the larger cognitive capacities of humans compared to other primates;

the estimated group size for humans suggests that **humans would spend all their time grooming each other**; **language & gossip** arise as a **natural replacement** for the role **grooming** plays in other primates ;

***indirect reciprocity** presumably provides the mechanism which distinguishes us, humans, from all other living species on Earth.*

[R. D. Alexander, *The Biology of Moral Systems,* Aldine de Gruyter- 87]



Indirect reciprocity

I scratch your back & someone else scratches mine . . .

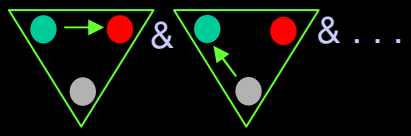
simplest possible setting :

1. world in ***black & white*** :
   a. people's reputation can be GOOD (1) or BAD (0)
   b. people can HELP (1) or NOT-HELP (0) (cooperate or defect as usual) depending on the reputation of the receiver and their own reputation.

each player needs a 4-bit string to take a decision

and 16 (2^4) possible strategies

Donor's reputation	Receiver's reputation	Donor's action
GOOD	GOOD	GIVE/NOT GIVE
GOOD	BAD	GIVE/NOT GIVE
BAD	GOOD	GIVE/NOT GIVE
BAD	BAD	GIVE/NOT GIVE



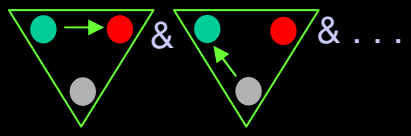
Indirect reciprocity

I scratch your back & someone else scratches mine . . .

simplest possible setting :

1. world in ***black & white*** :
   a. people's reputation can be GOOD (1) or BAD (0)
   b. people can HELP (1) or NOT-HELP (0) (cooperate or defect as usual) depending on the reputation of the receiver and their own reputation.

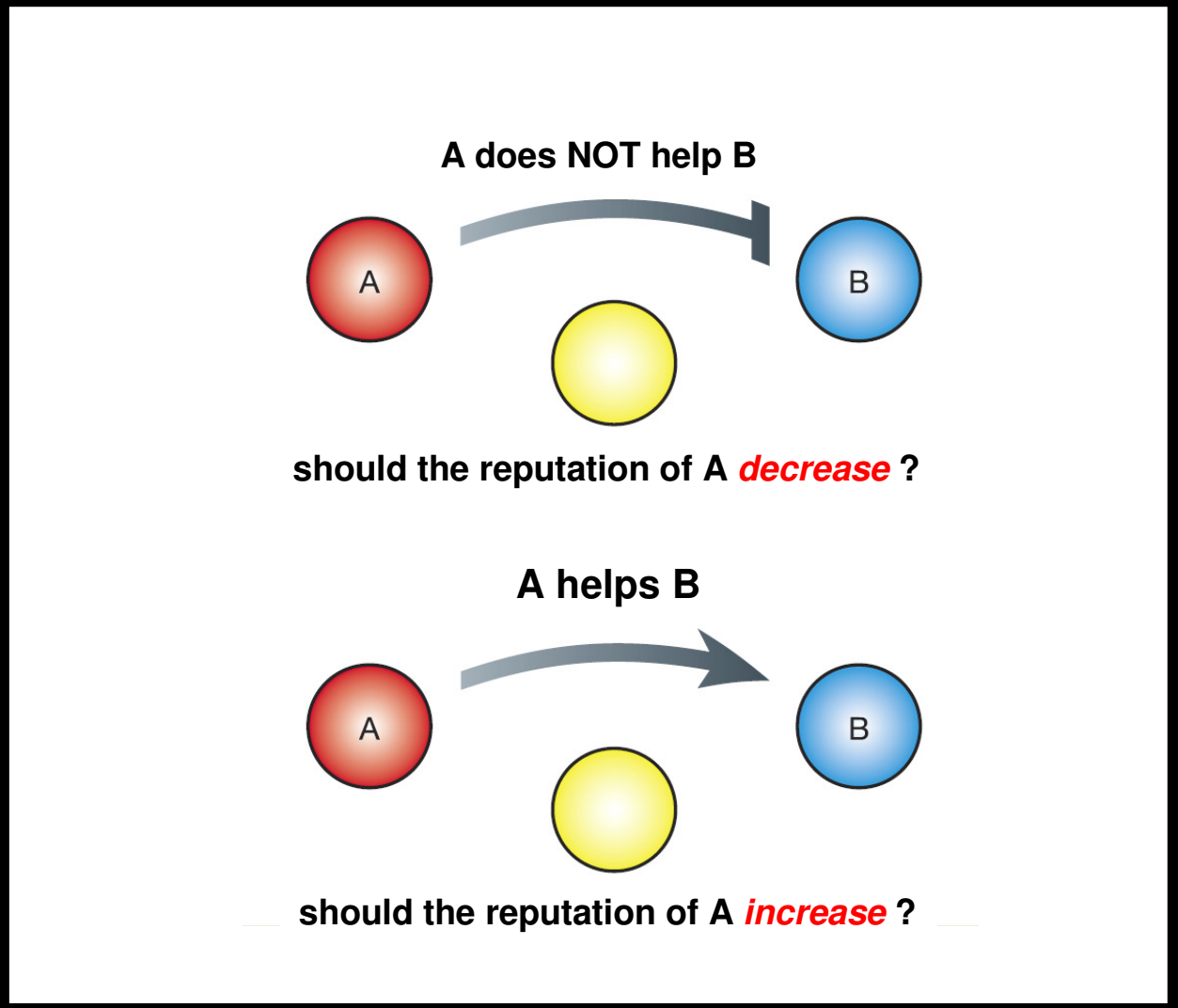
   c. gossip is error-free ; each action is witnessed by one observer & all others acquire the same information
   d. interactions are not repeated (eliminate direct reciprocity effects)
   e. population lives under one ***social norm*** which defines the reputation of one acting individual

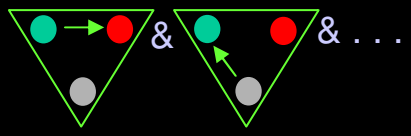


Indirect reciprocity

I scratch your back & someone else scratches mine . . .

How to define a *social norm*?

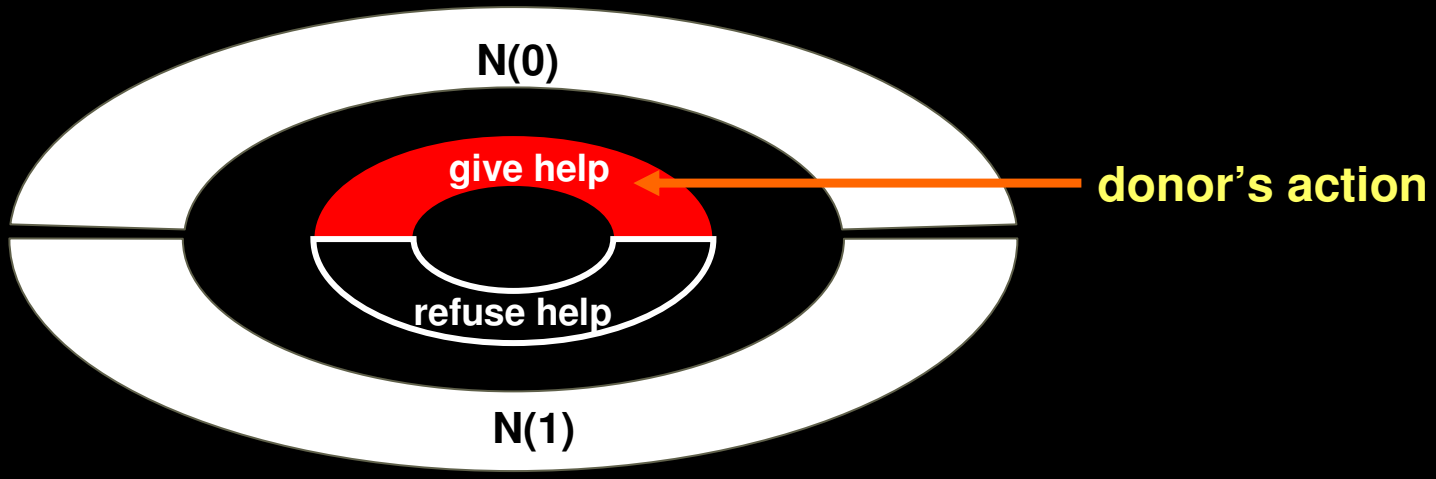




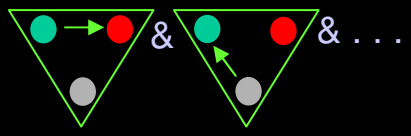
Indirect reciprocity

I scratch your back & someone else scratches mine . . .

The simplest assessment rule… 1st order norm



2^2 possible norms



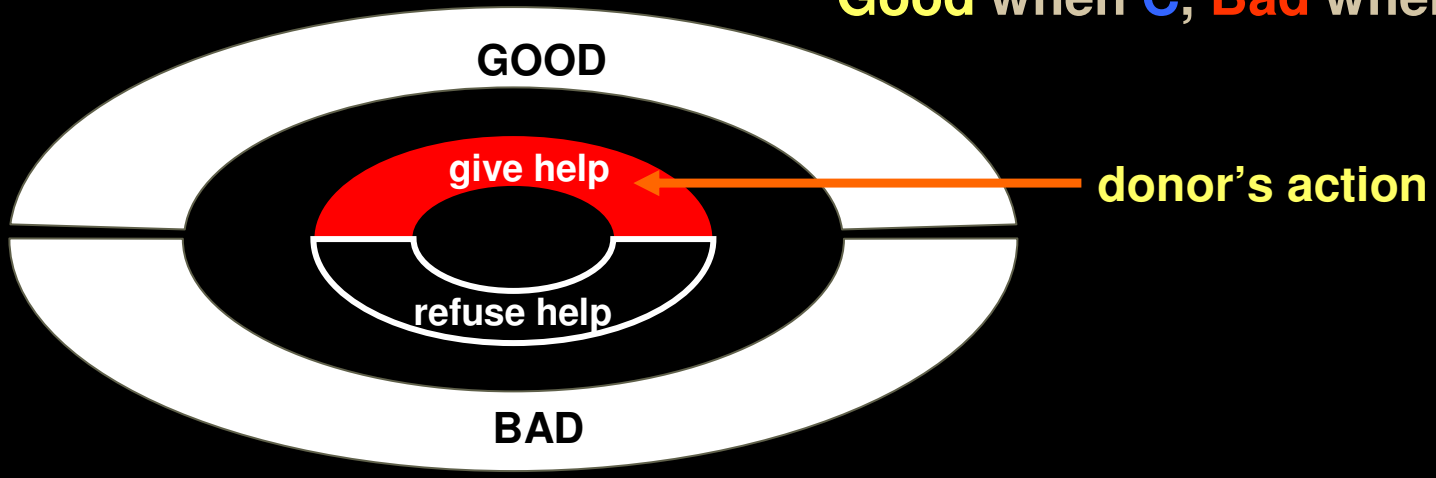
Indirect reciprocity

I scratch your back & someone else scratches mine . . .

the simplest assessment rule… 1st order norm

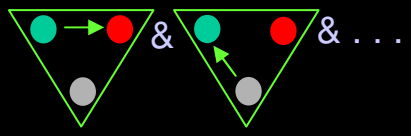
only one makes sense (*image scoring*):

Good when C, Bad when D



Does it promote cooperation? *not really* !! *in particular*

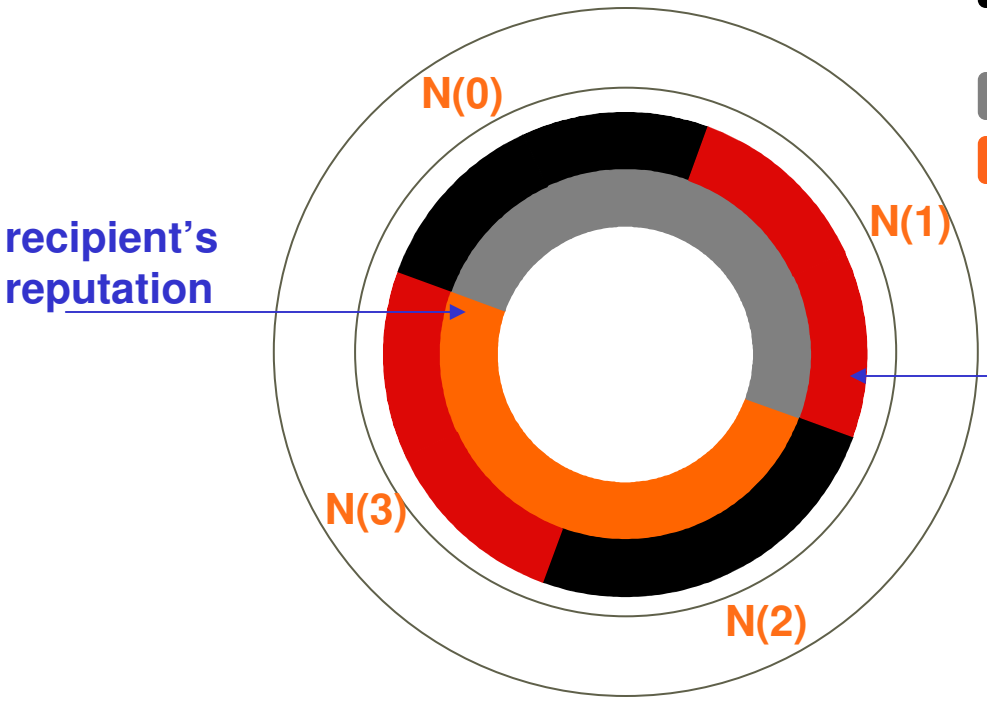
If defection brings a BAD reputation, why should you defect against bad players… Defection is never justifiable !!



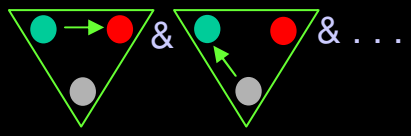
Indirect reciprocity

I scratch your back & someone else scratches mine . . .

2nd order norms: adding the recipient's reputation



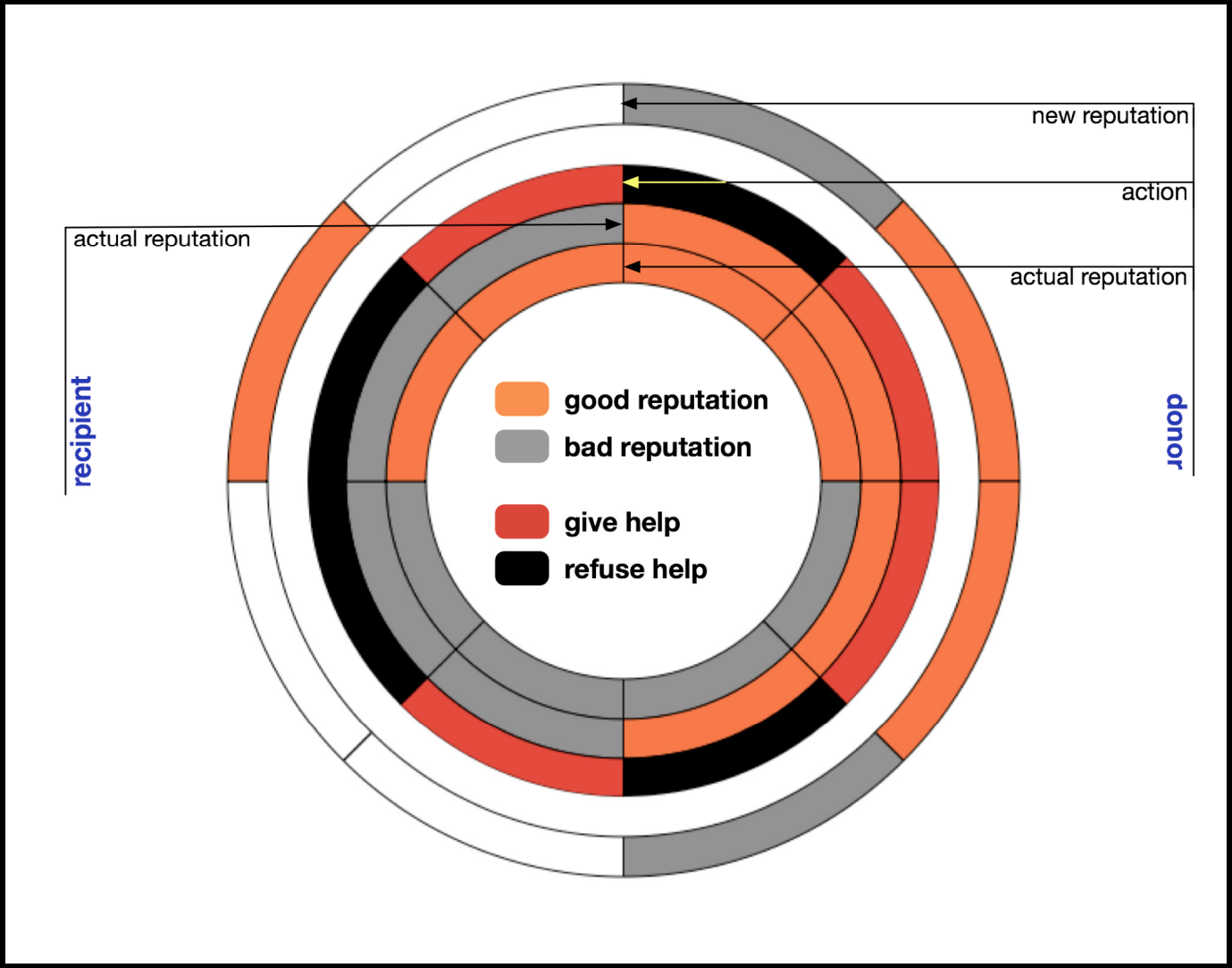
2^4 possible norms

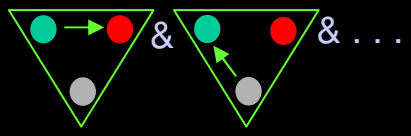


Indirect reciprocity

I scratch your back & someone else scratches mine . . .

3rd order norms: adding the past -> reputation of the donor





Indirect reciprocity

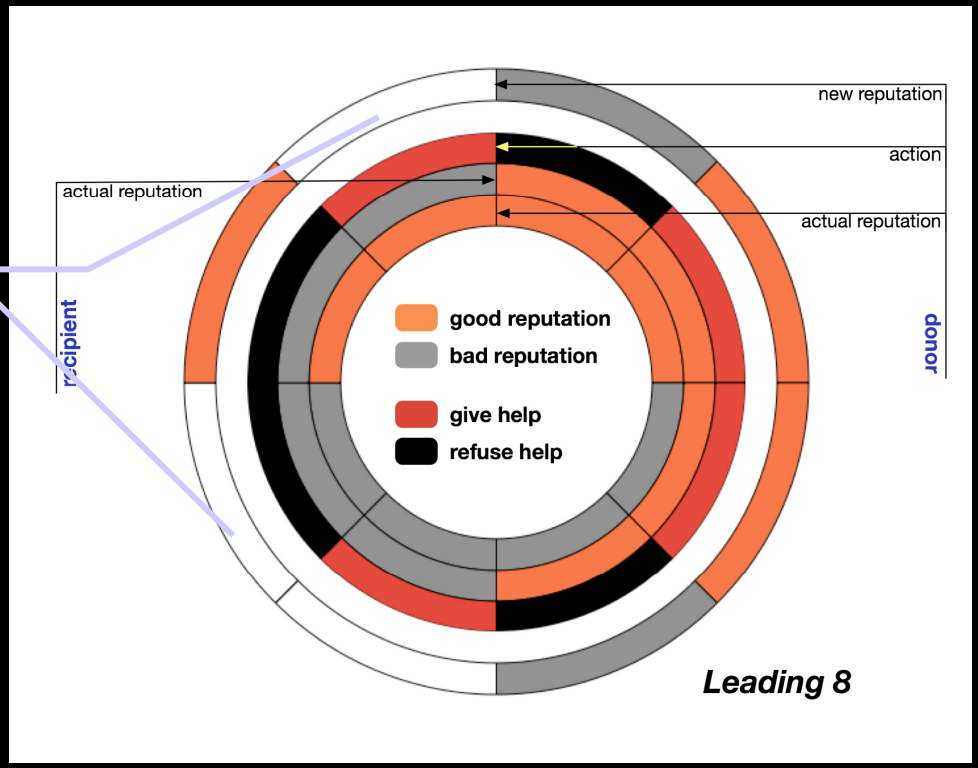
I scratch your back & someone else scratches mine . . .

approximate analytic study :

There are only 8 norms which turn cooperation into a ESS – the leading eight

[H. Ohtsuki & Y. Iwasa, *JTB* 231 *(2004) 107*]

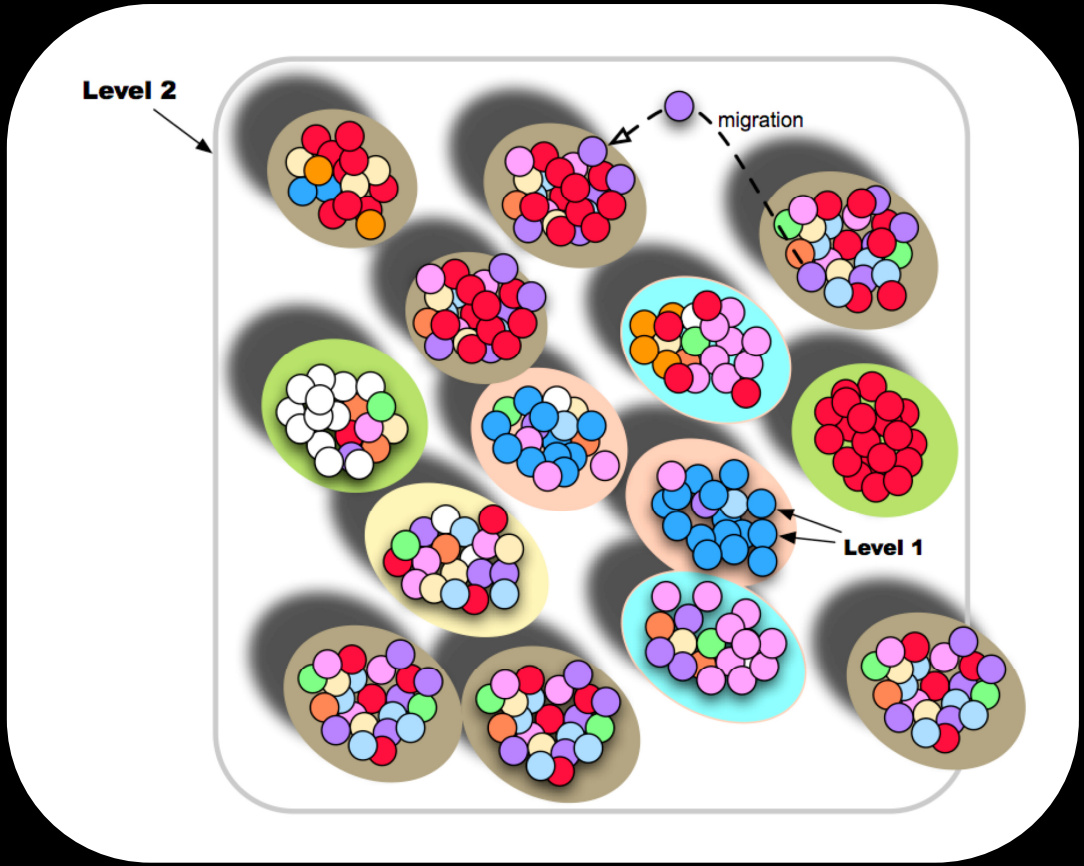
wild cards



one norm at a time; leading eight are ESS only when coupled to a specific strategy

what happens if norms compete, individuals have different strategies and errors occur ?

evolution of social norms

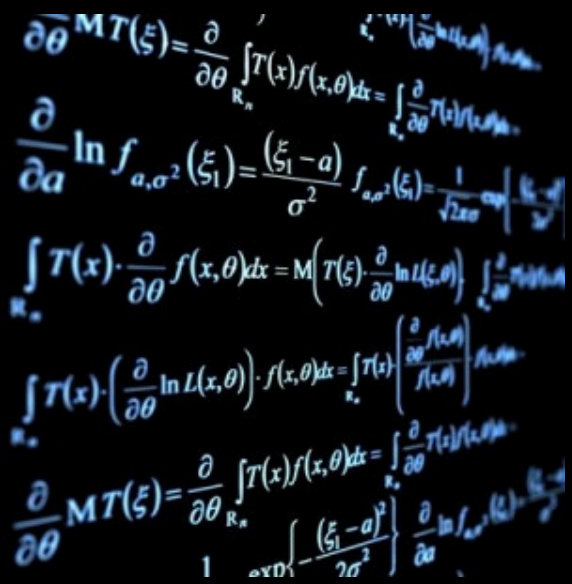


one tribe, one norm ;
each individual, a different strategy;
strategies evolve in a tribe under a single norm
individual fitness is the payoff of a cooperation game in each tribe;
tribes compete with each other; competition modeled by different games;
the norm of the loosing tribe changes towards the norm of the winning tribe;

[Chalub, Santos & Pacheco, J Theoretical Biology 271 *(2006) 233*]
[Pacheco, Santos & Chalub, *PloS Computational Biology (2006)*]

evolution of social norms

no analytical solution → numerical simulations

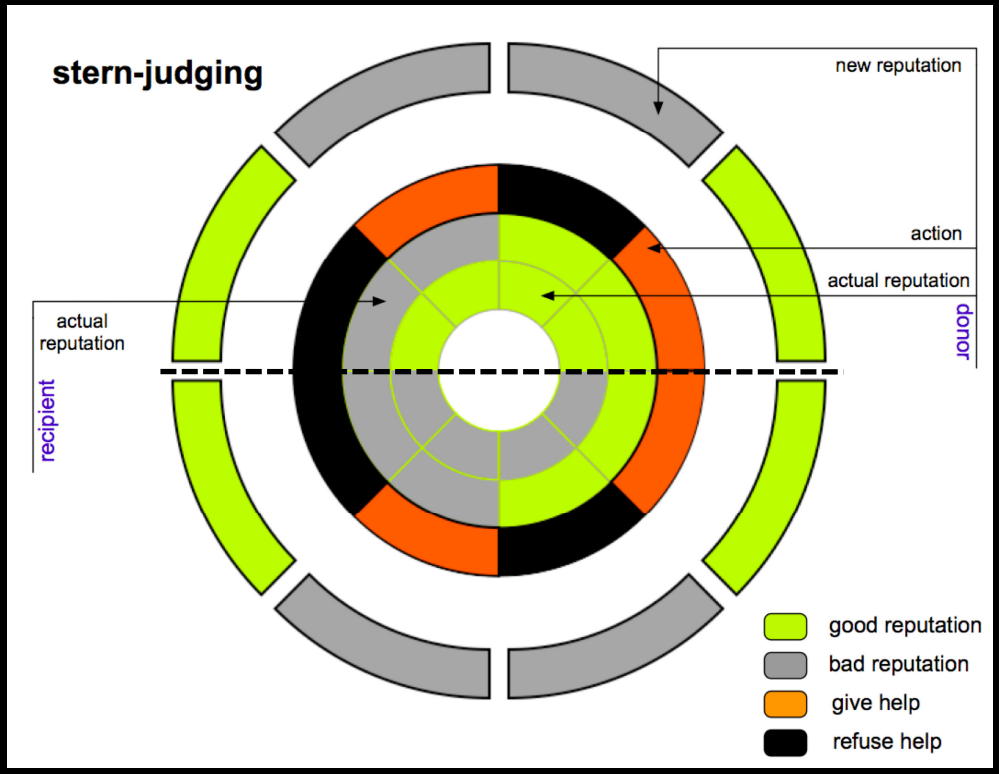


we include mutations at the level of strategies and when norms are adopted by the tribes that loose the “war”; we also include a small amount of migration among tribes.

evolution of social norms

Stern-Judging

a single norm emerges as the most evolutionary successful



stern-judging :

emerges **independently of the type of conflict** between tribes

is part of the **leading-eight** norms

is the ***simplest*** of the **leading-eight** norms : ***2nd order norm*** :

all that matters is the ***action of the donor*** & the ***reputation of the recipient***

evolution of social norms

Stern-Judging

a simple and successful norm → *Stern-Judging*

Help the good and refuse help to the bad ; otherwise you will be punished

1) **Maintenance of cooperation.** a GOOD guy who gives to a GOOD is seen as GOOD. Any refusal to help a GOOD player is seen as BAD.
2) **Justified & implacable Punishment.** A GOOD guy HAS to defect against BAD guys, otherwise he will be seen as BAD.
3) **Apologies.** *single move is enough to reach the worst but also the best standard*
4) **never being morally dubious...** In each case there's a single move that leads to GOOD

Pacheco, Santos, Chalub, PLoS Comp Biol 2 (12) e178 (2006)

evolution of social norms

Stern-Judging forgets the past…

The simplicity of stern-judging (2nd order) correlates nicely with real reputation-based systems…

e-trade :

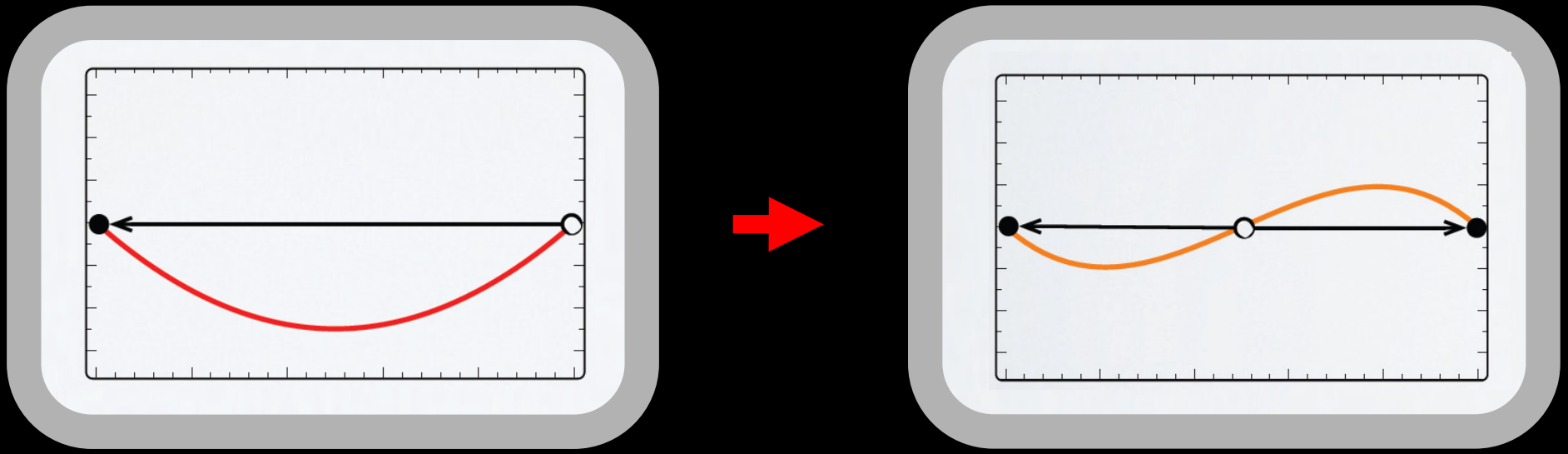
- **relies on reputation-based mechanisms of cooperation;**
- **exhibits *high levels* of cooperation;**
- **dominated by anonymous one-shot interactions between individuals loosely connected and geographically dispersed;**
- **reputation in e-trade is introduced via a feedback mechanism which**
- **announces rating of sellers;**
- **it has been found that publicizing a detailed account of the seller's (the donor) feedback history does not improve cooperation, as compared to publicizing only the seller's most recent rating;**

kin, reciprocity, memory, reputations, moral systems, punishments, gossip, etc.

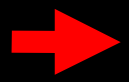
help solving the paradox of cooperation

how ?

so far, in all mechanisms studied



are we missing other mechanisms ?

yes 

Evolution of Diversity and Cooperation

2 / 3

Jorge M. Pacheco

atpgroup

Departamento de Matemática & Aplicações
Universidade do Minho
Portugal

UIMP Universidad Internacional Menéndez Pelayo

RSEM

FisyMat

Luis Santaló School, 18th of July, 2013