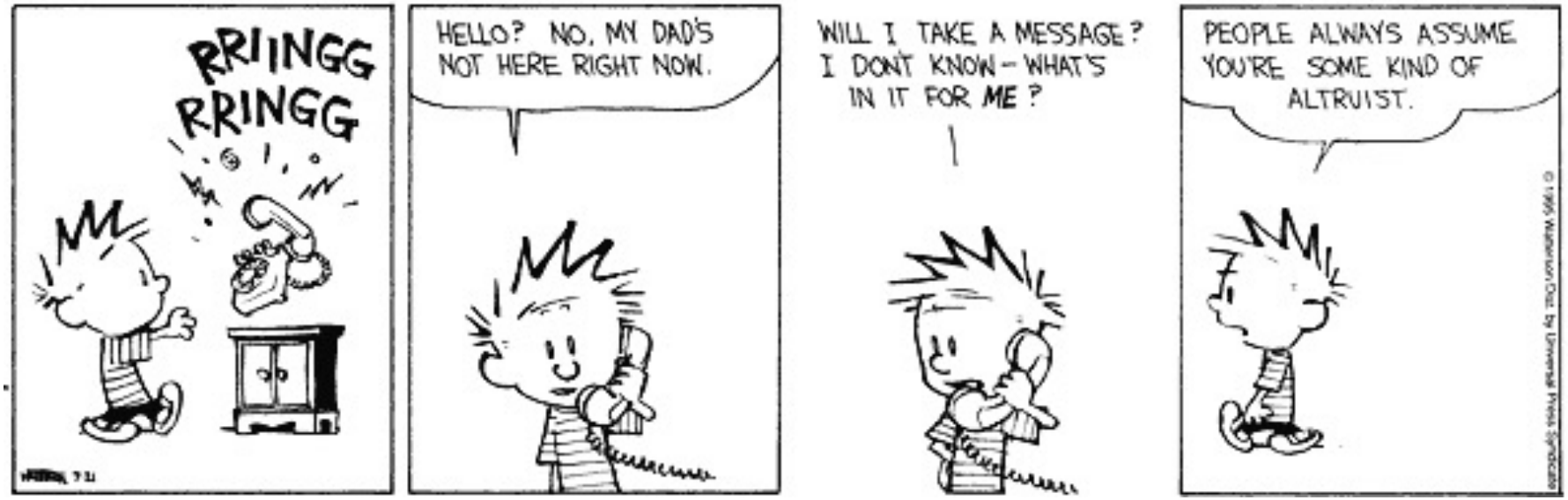
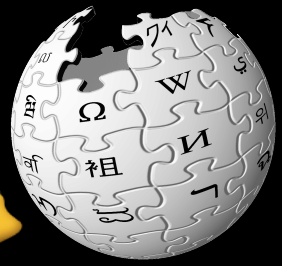
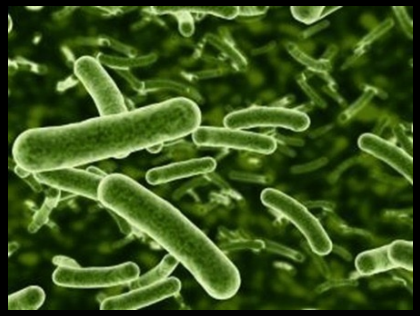
DEI
DEPARTAMENTO
DE ENGENHARIA INFORMÁTICA
TÉCNICO LISBOA

The Calculus of Selfishness (cont.)



cooperation at all scales





Prisoner's dilemma or
the *cost-benefit* dilemma

Donor
Pays a cost c

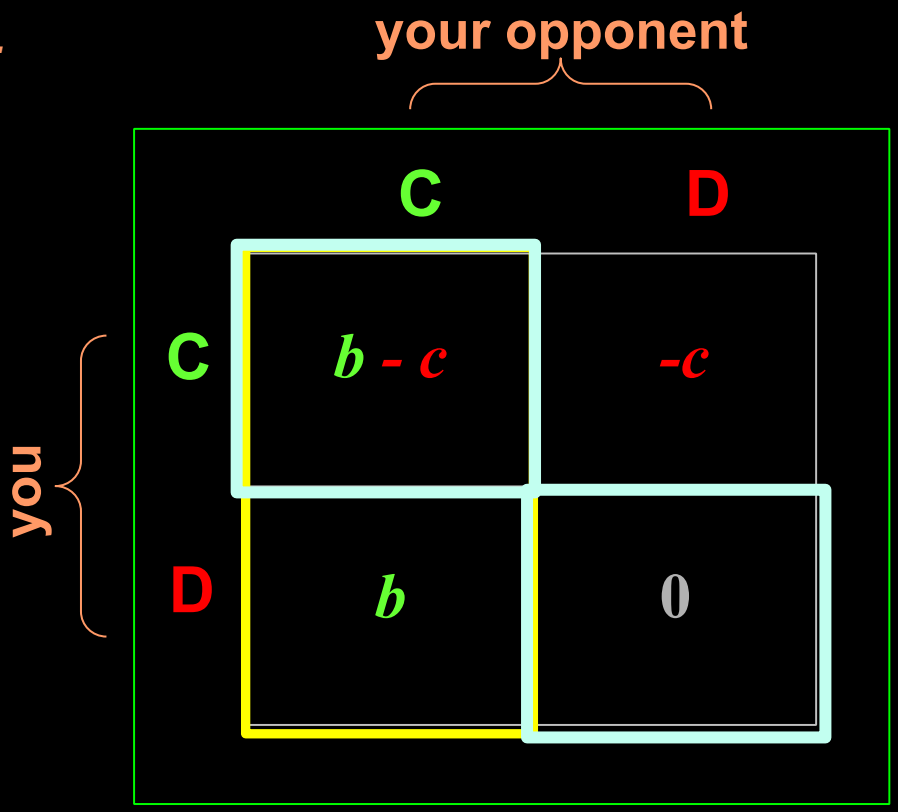
Receiver
Receives a benefit b>c

If both play as a donor and as a receiver...

"RATIONAL" GOAL :
maximize your own payoff !

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

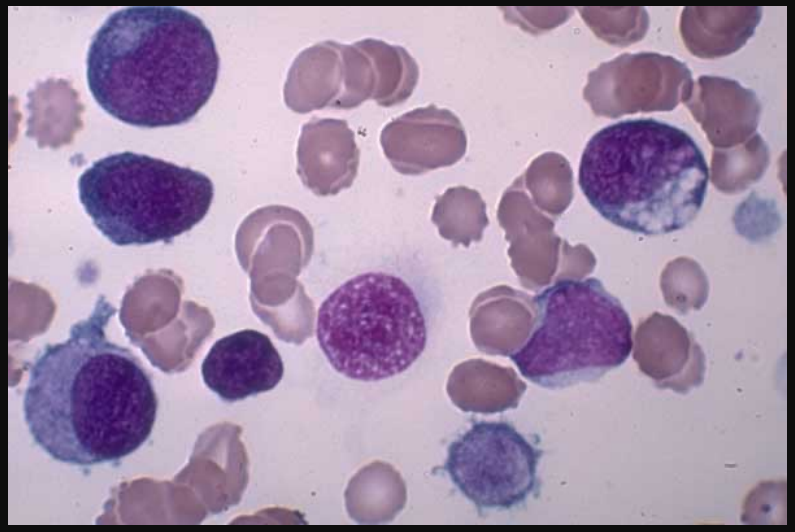
BUT:
CC is better than DD.



Dilemma : despite **mutual cooperation** (**CC**) being better than **mutual defection** (**DD**), individual "rational choice" leads to **DD**

evolutionary game theory

payoff ➡ fitness ➡ social success



genetic evolution

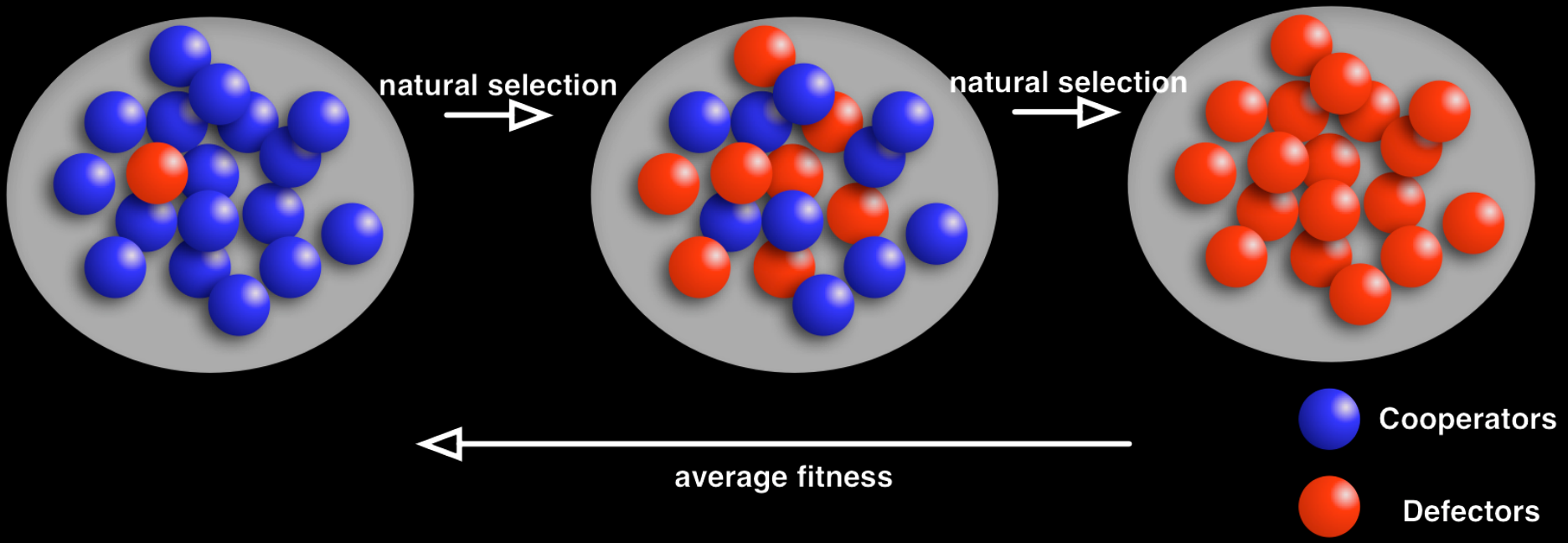
social/cultural evolution

Individuals with higher fitness will *reproduce more*

or their behavior will be *imitated* more often
(social learning)

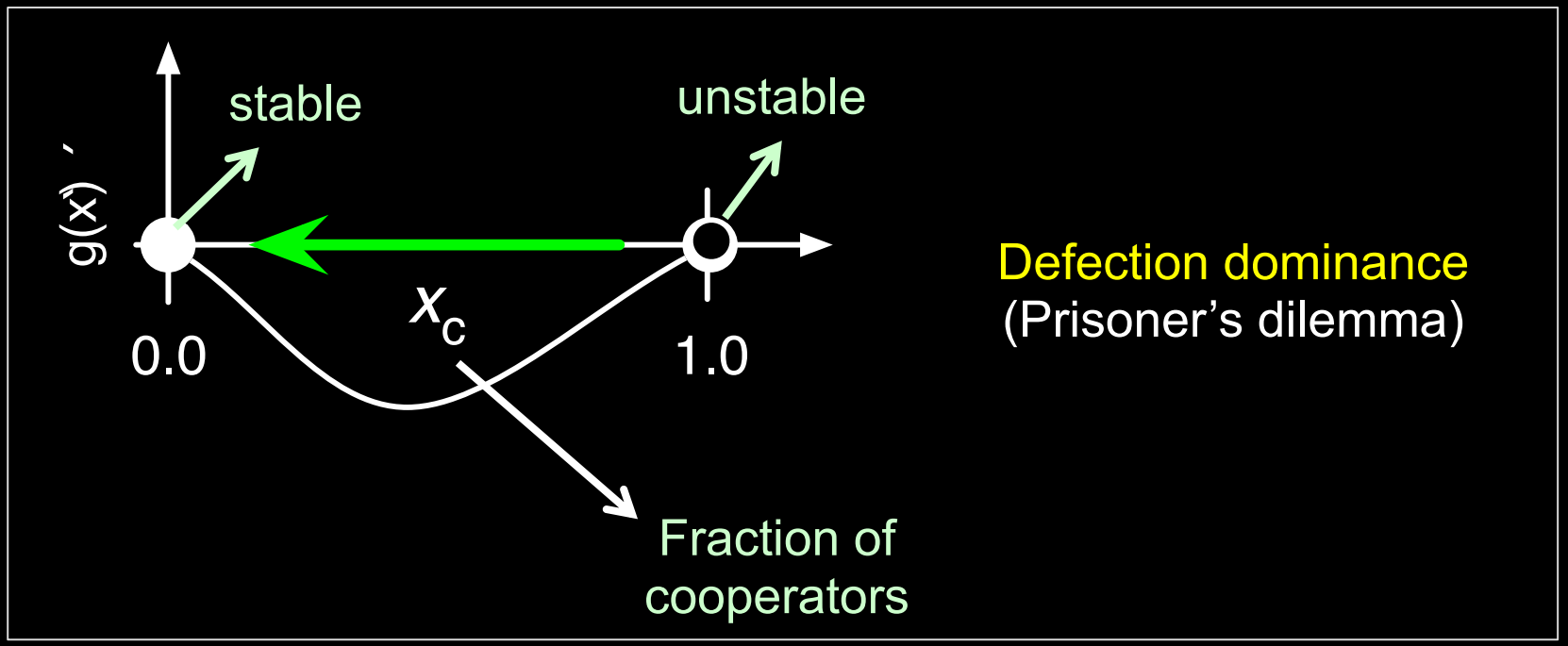
evolutionary game theory

cost-benefit dilemma (prisoner's dilemma)



Natural selection & self-organization
leads to the tragedy of the commons

The calculus of selfishness



Defection dominance
(Prisoner's dilemma)

Gradient of selection

$$\dot{x} = x(1-x)\left[f_C(x) - f_D(x)\right] = g(x)$$

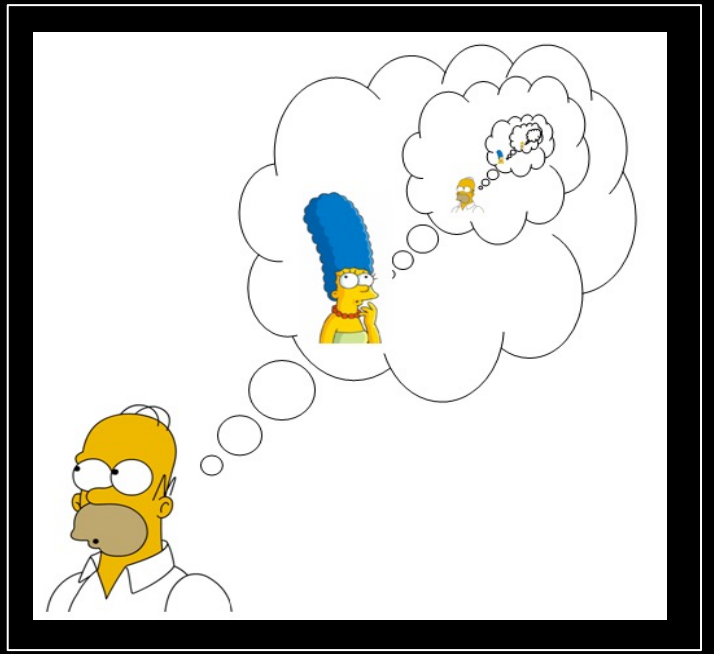
Replicator equation

General stability concepts

Game Theory

Nash equilibrium

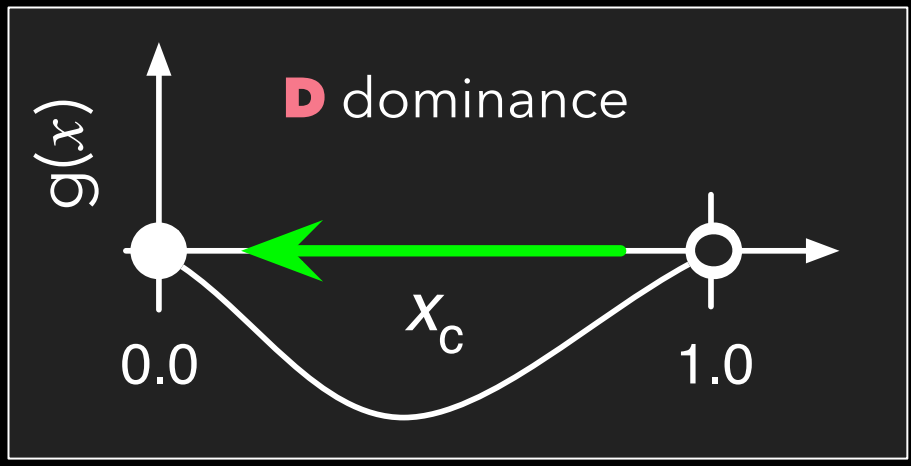
If a strategy is a Nash equilibrium, and if both players play that strategy, then neither person can deviate from that strategy and increase her payoff



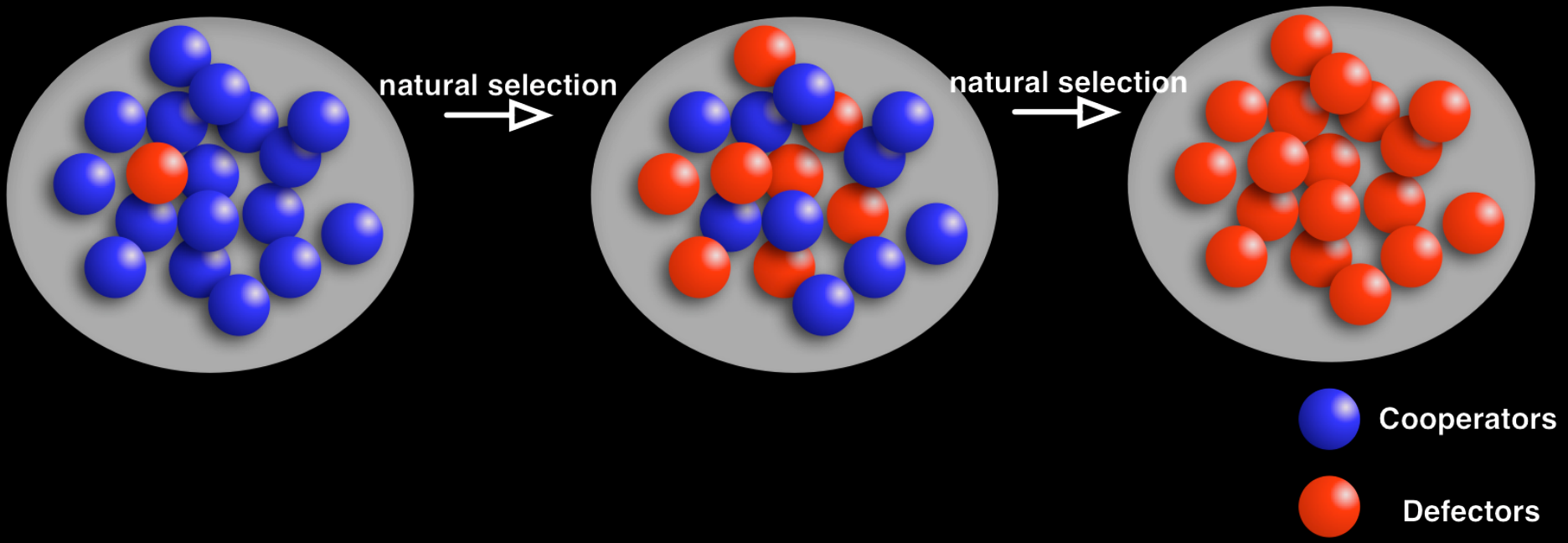
Evolutionary Game Theory

Evolutionarily Stable Strategy (ESS)

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 population)



Dealing with an apparent paradox



but… cooperation surrounds us!!

What are we missing here ?

mechanisms of cooperation

kin selection (Hamilton's rule)

multi-level selection (a.k.a., group selection)

direct reciprocity

indirect reciprocity

signaling & quorum sensing

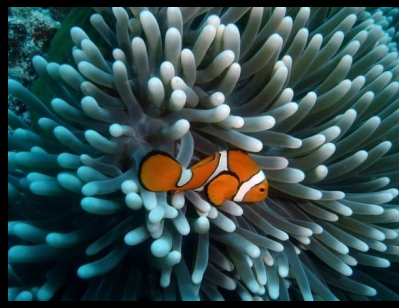
networked populations

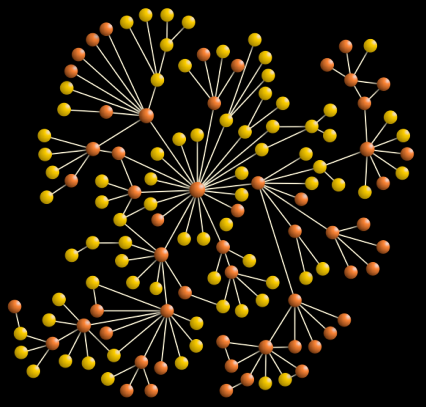
commitments

institutions, sanctions and rewards

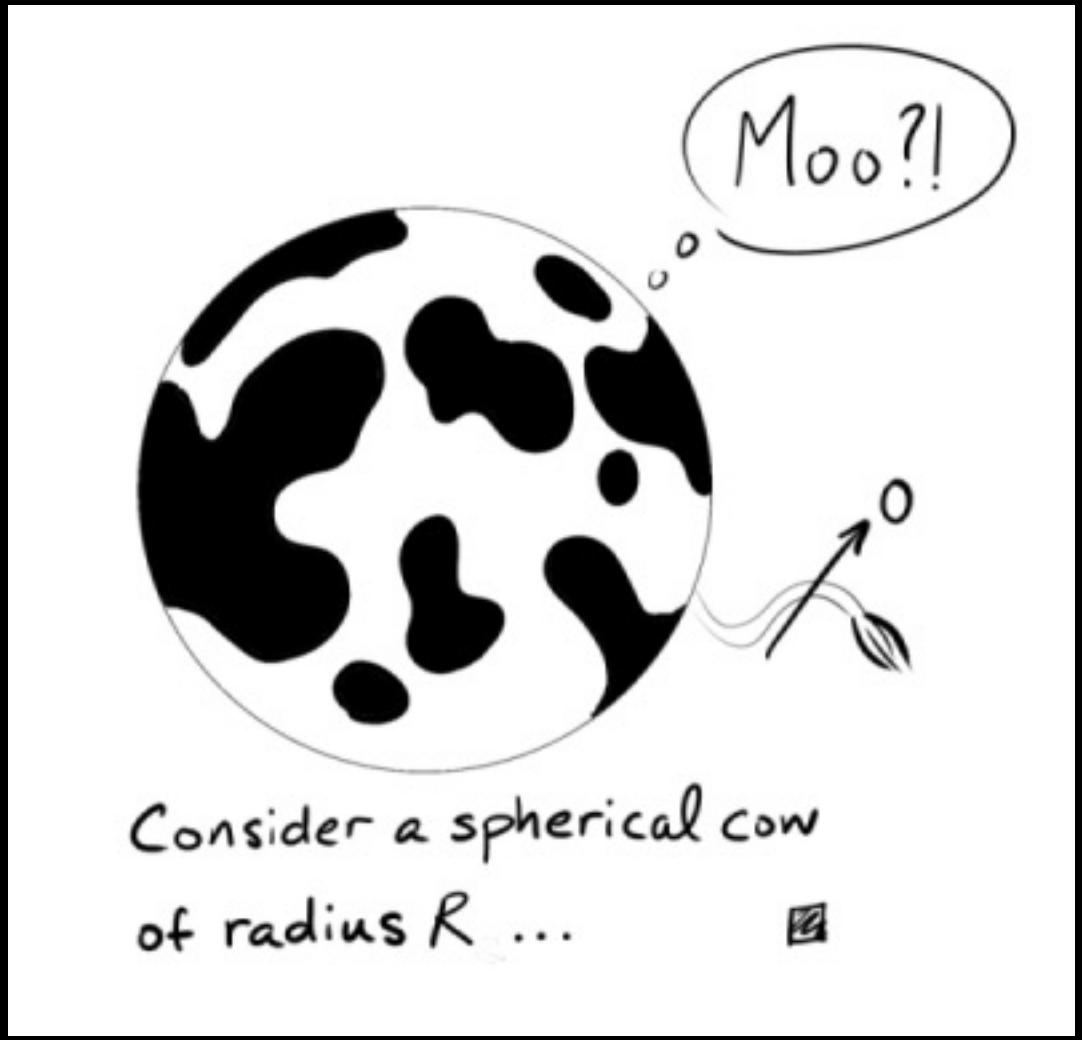
social diversity

(...)





The beauty of modelling core principles

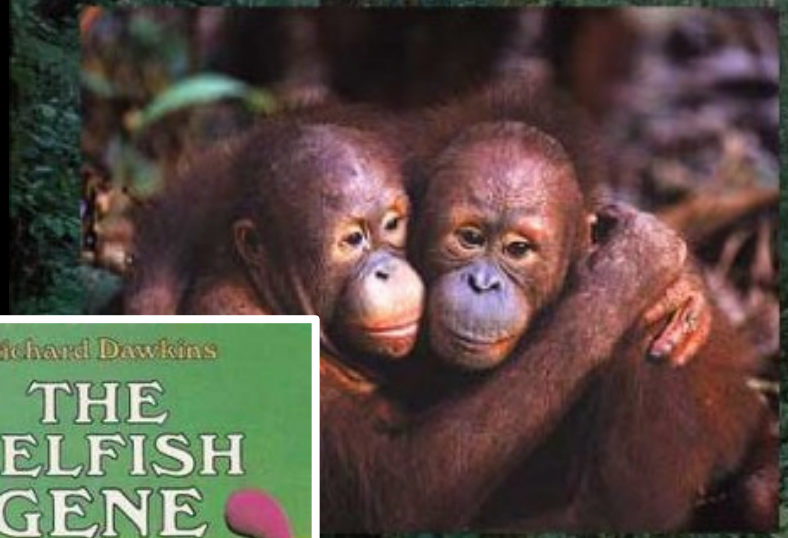


Ways out of the dilemma

all in the family . . .

WD Hamilton, JTB (1964)
G. Price, J Maynard-Smith, etc.

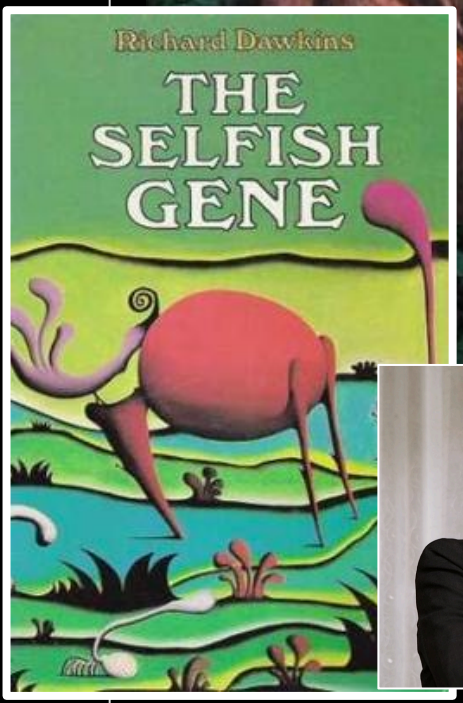
the more individuals are related, the more cooperation is feasible

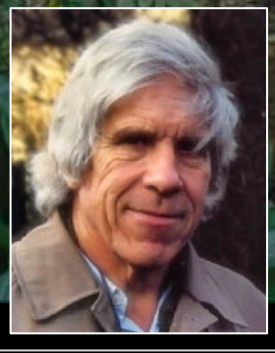


$$\frac{b}{c} > \frac{1}{r}$$

genetic relatedness

kin selection





The Genetical Evolution of Social Behaviour.

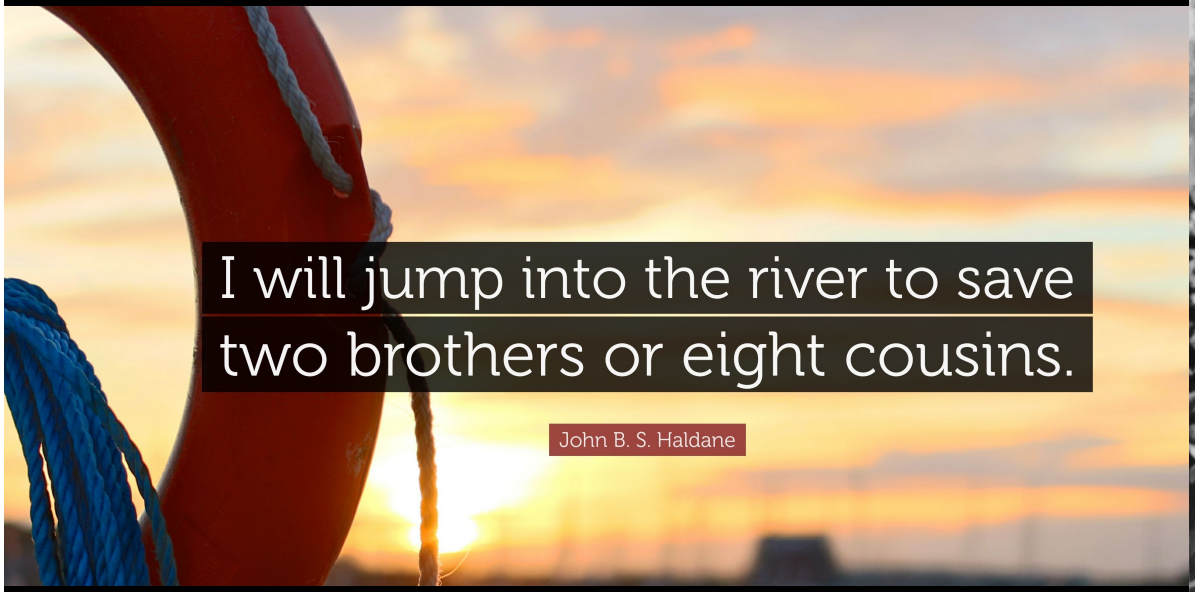
W. D. HAMILTON

The Galton Laboratory, University College, London, W.C.2

J. Theoret. Biol. (1964)

Ways out of the dilemma

all in the family . . .



John B. S. Haldane (1892-1964)
Evolutionary / Mathematical Biologist

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

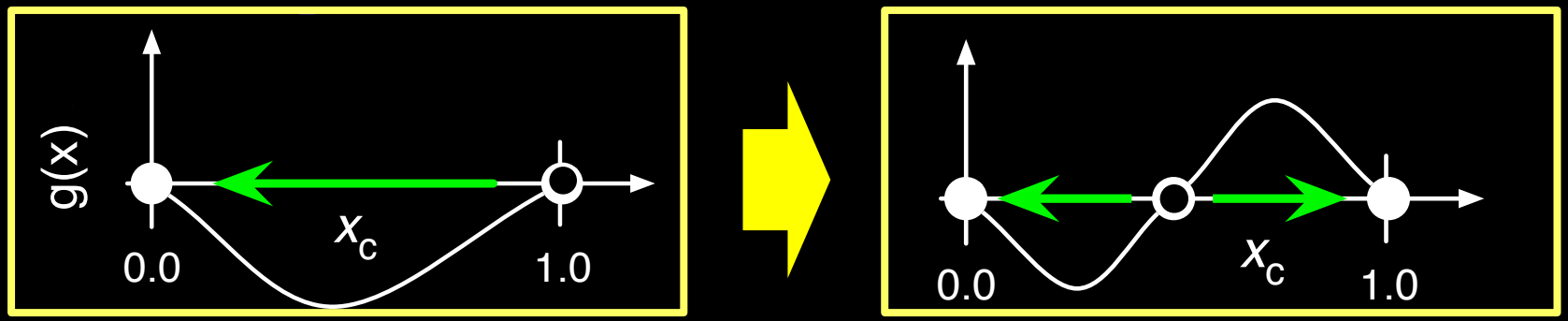
$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{bmatrix} (b-c)(1+r) & -c+br \\ b-cr & 0 \end{bmatrix} \end{array}$$

ESS condition :

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

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

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



Ways out of the dilemma

pick the right team…

Wynne-Edwards (1962),
DS Wilson & EO Wilson, Am Sci (2008)
Traulsen & Nowak, PNAS (2006), etc.

Cooperative groups tend to be more successful than groups of defectors

Multi-level selection

Multi-level selection

pick the right team…

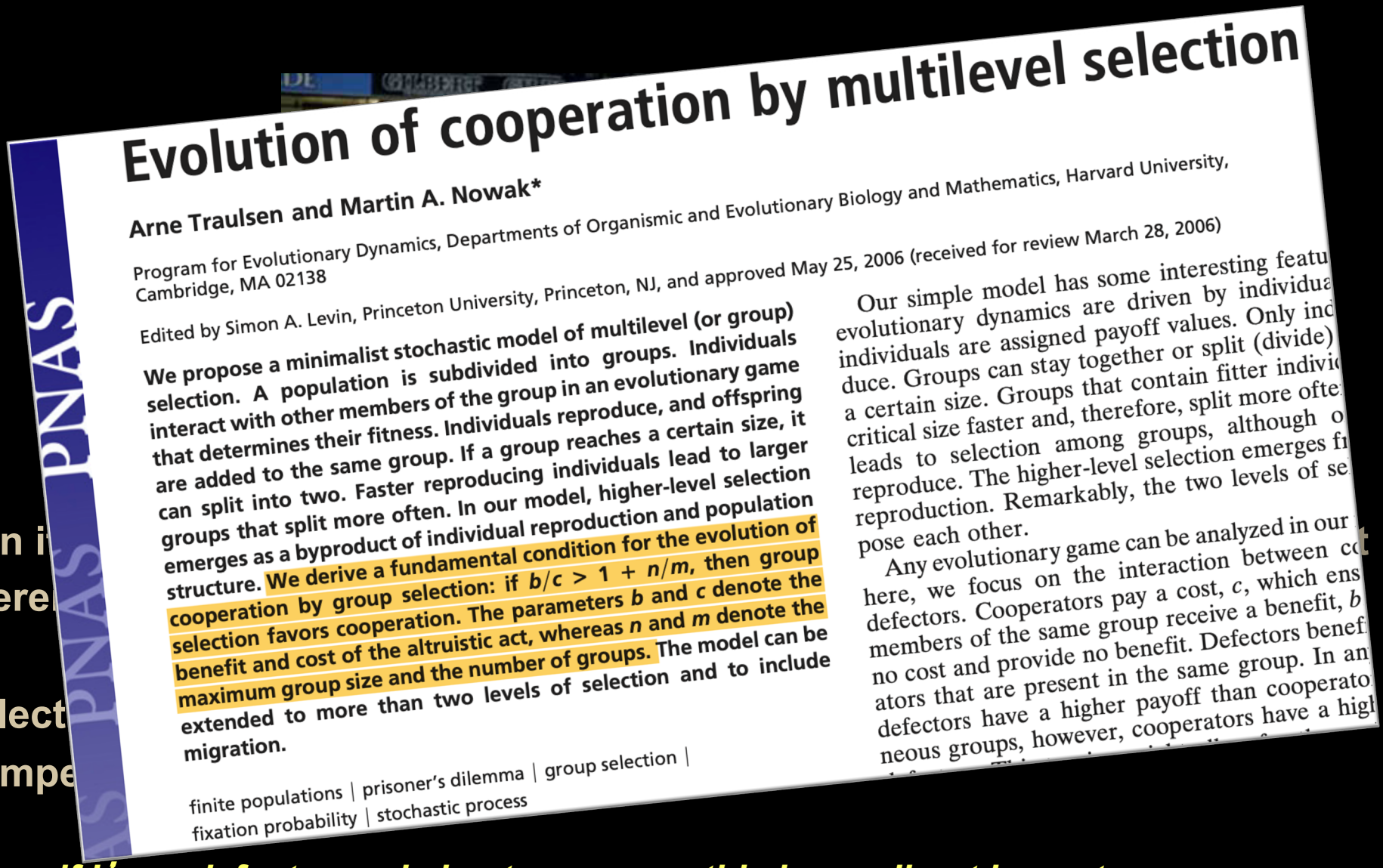
Even if the unity of selection is solely the individual, selection may occur at different levels

- selection inside groups
- competition between groups

Hence, if I'm a defector and cheat someone, this has a direct impact on my own success whenever conflicts raise between groups (creating a kind of "kin" relation).

Multi-level selection

pick the right team…



Evolution of cooperation by multilevel selection

Arne Traulsen and Martin A. Nowak*

Program for Evolutionary Dynamics, Departments of Organismic and Evolutionary Biology and Mathematics, Harvard University, Cambridge, MA 02138

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved May 25, 2006 (received for review March 28, 2006)

We propose a minimalist stochastic model of multilevel (or group) selection. A population is subdivided into groups. Individuals interact with other members of the group in an evolutionary game that determines their fitness. Individuals reproduce, and offspring are added to the same group. If a group reaches a certain size, it can split into two. Faster reproducing individuals lead to larger groups that split more often. In our model, higher-level selection emerges as a byproduct of individual reproduction and population structure. We derive a fundamental condition for the evolution of cooperation by group selection: if $b/c > 1 + n/m$, then group selection favors cooperation. The parameters b and c denote the benefit and cost of the altruistic act, whereas n and m denote the maximum group size and the number of groups. The model can be extended to more than two levels of selection and to include migration.

finite populations | prisoner's dilemma | group selection | fixation probability | stochastic process

Our simple model has some interesting featu
evolutionary dynamics are driven by individua
individuals are assigned payoff values. Only ind
duce. Groups can stay together or split (divide)
a certain size. Groups that contain fitter indivi
critical size faster and, therefore, split more ofte
leads to selection among groups, although o
reproduce. The higher-level selection emerges fr
reproduction. Remarkably, the two levels of se
pose each other.

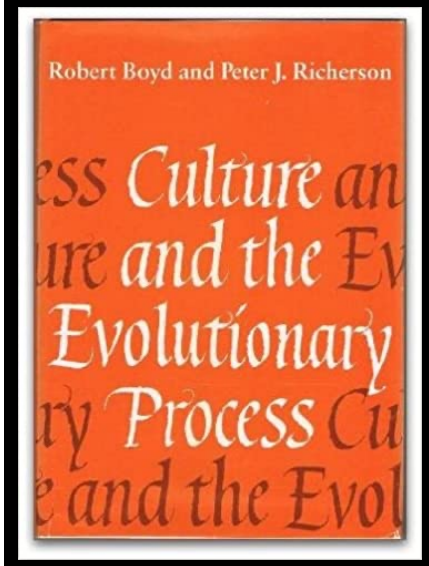
Any evolutionary game can be analyzed in our
here, we focus on the interaction between co
defectors. Cooperators pay a cost, c, which ens
members of the same group receive a benefit, b
no cost and provide no benefit. Defectors benef
ators that are present in the same group. In an
defectors have a higher payoff than cooperato
neous groups, however, cooperators have a high

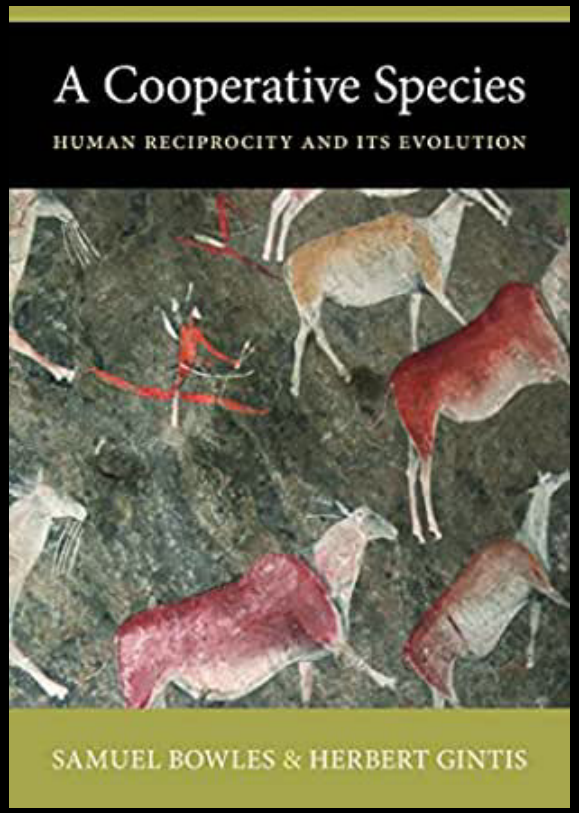
Even i
differe

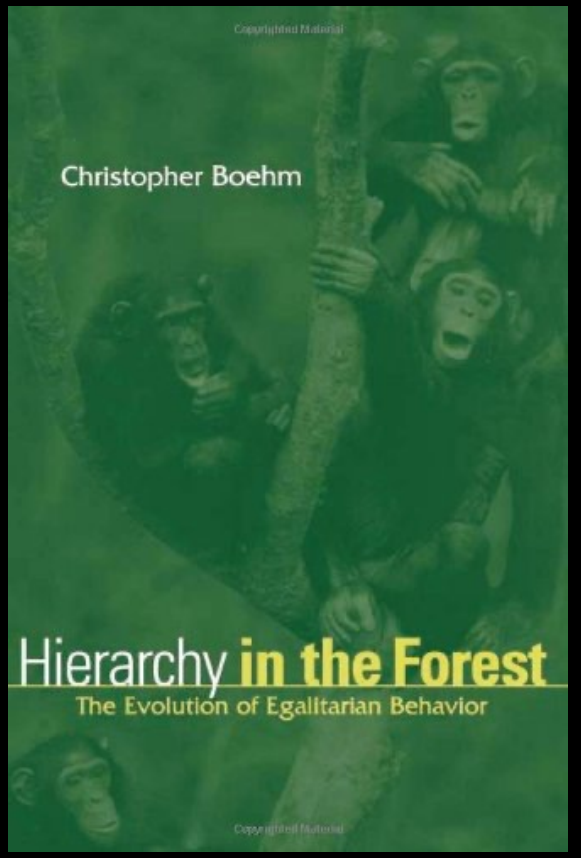
- select
- compe

Hence, if I'm a defector and cheat someone, this has a direct impact on my own success whenever conflicts raise between groups (creating a kind of "kin" relation).

Gene & culture evolution





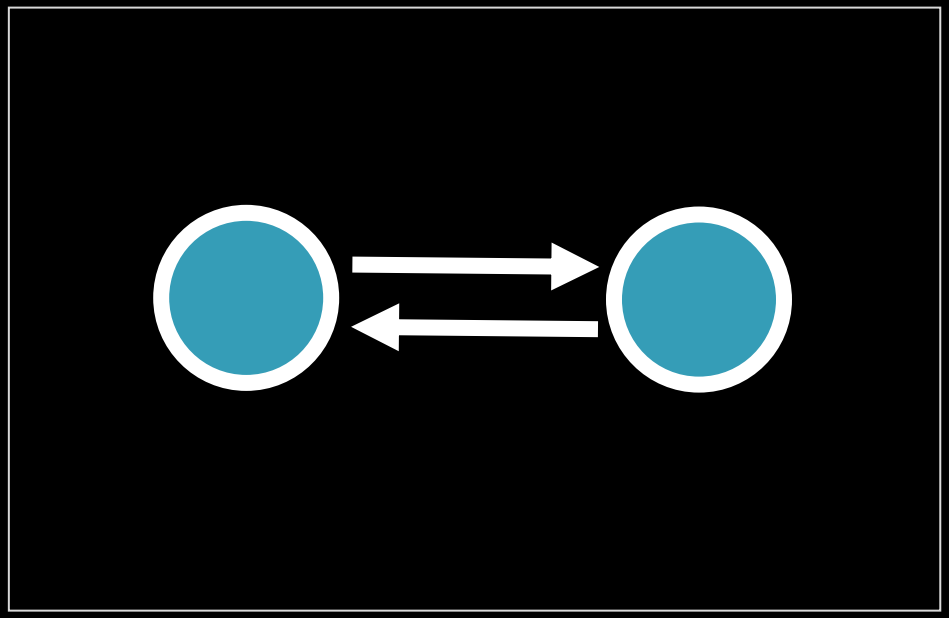


Boyd, Richerson, Bowles, Gintis and others show how the act of social learning and cooperation within groups, have been essential for survival, and may allow human populations to accrue information over many generations.

Direct reciprocity

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

repeated interactions, memory, complex strategies, …



I scratch your back & you scratch mine . . .

direct reciprocity
relies on individuals' experience & memory

Direct reciprocity

I scratch your back & you scratch mine . . .

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 : *RAPOPORT & TIT for TAT*

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

Ecology of direct reciprocity

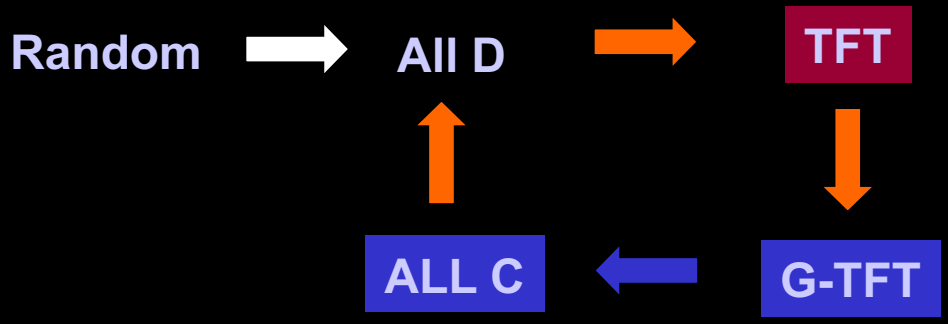
I scratch your back & you scratch mine . . .

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

Challenge: implement a agent-based model addressing this dynamics

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]

Ecology of direct reciprocity

I scratch your back & you scratch mine . . .

PD: T>R>P>S

What's the trick?

WSLS vs WSLS: takes a couple of moves to correct an error

WSLS : C C C C C C D C C C …
WSLS : C C C C C D D C C C …
error

…creating a positive feedback

Public enemy: dumb *AllCs*

ALLC : C C C C C C C C C C C …
WSLS : C C C C C D D D D D D …

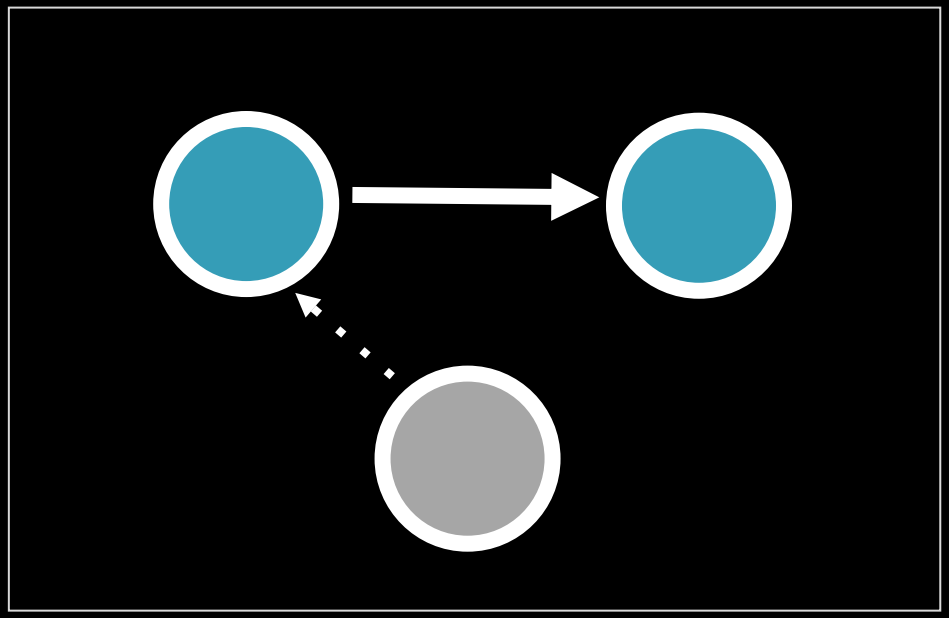
ALLC : C C C C C C C C C C C …
TFT : C C C C C D C C C C C …

… contrary to TFTs, WSLS will shamelessly exploit AllCs

Indirect reciprocity

RD Alexander, The Biology of Moral Systems, 1987

One shot interactions with reputation based strategies

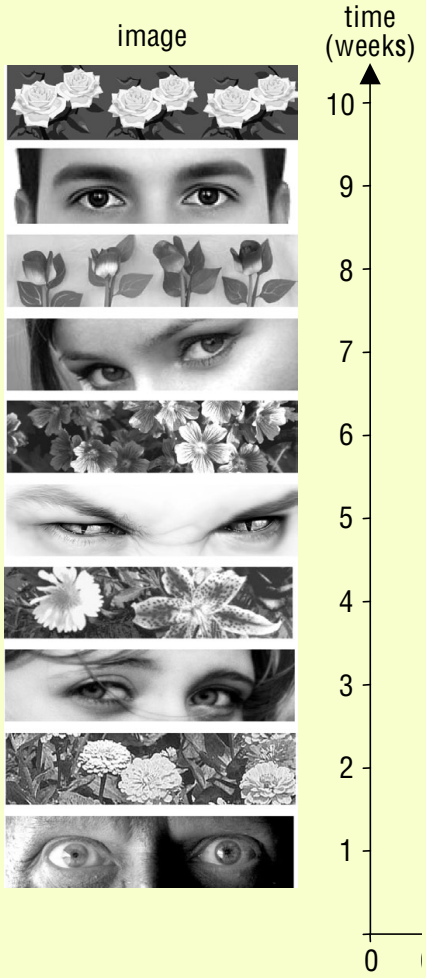


I scratch your back & someone else scratches mine . . .

indirect reciprocity

relies on reputations, gossip, ***social control***, moral systems, etc.

The strength of third parties & social control



Prices:

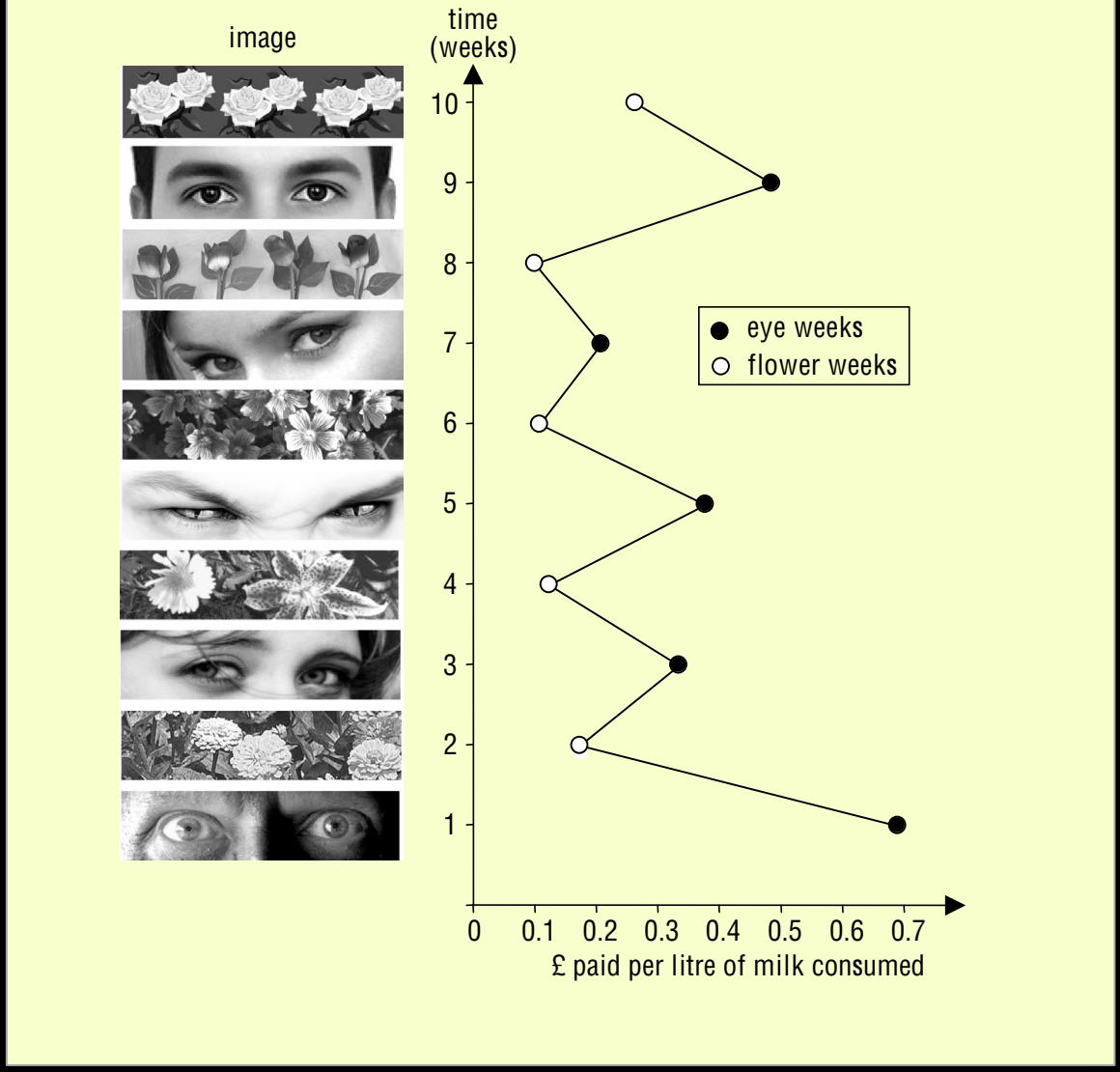
Coffee (with or without milk): 50p
Tea (with or without milk): 30p
Milk only (in your own coffee or tea): 10p
Full cup of milk: 30p

Please put your money in the blue tin.

Thanks, Melissa.

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

The strength of third parties & social control



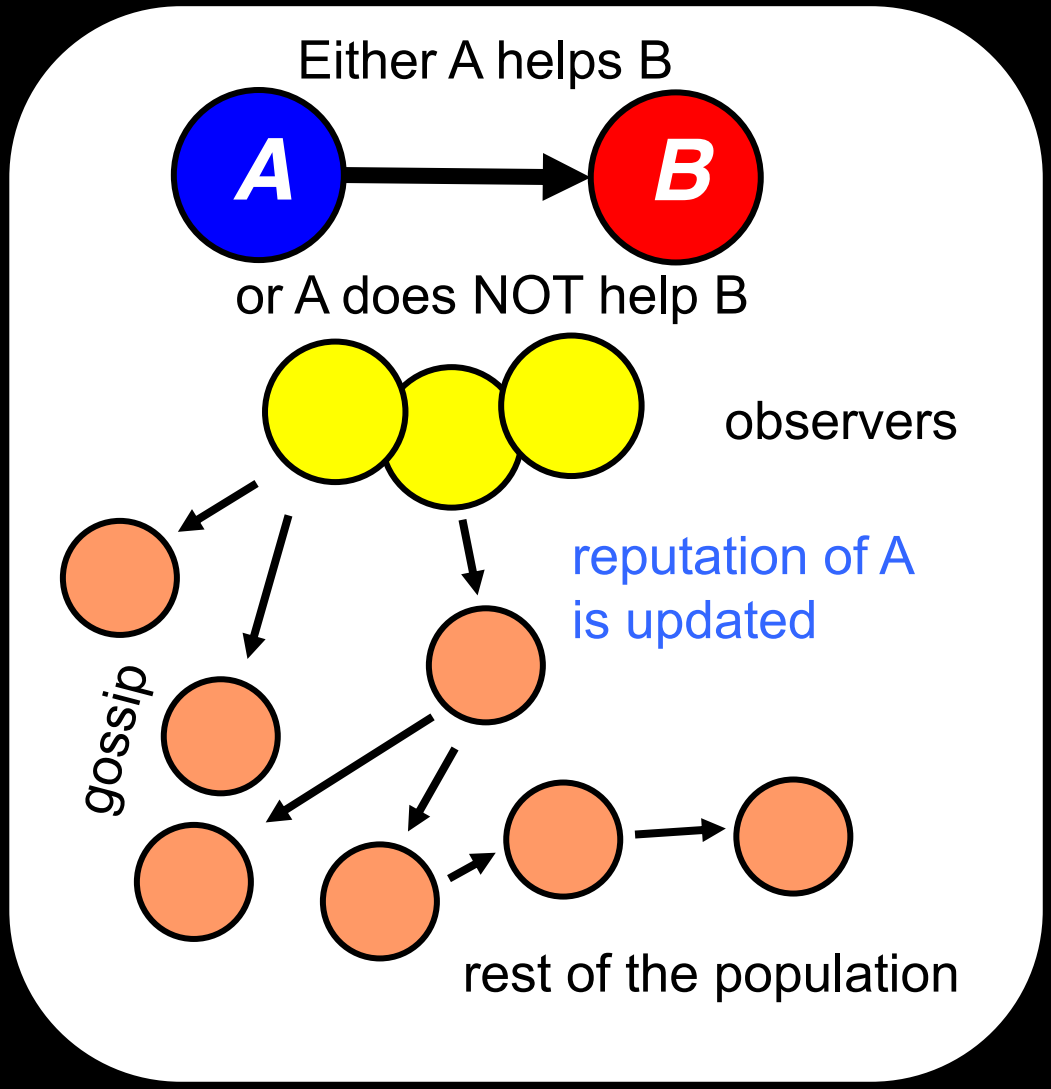
[Melissa Bateson et al., Biology Letters 2, 412–414 2006]

Indirect reciprocity

RD Alexander, The Biology of Moral Systems, 1987
Nowak & Sigmund, Nature 1998, Nature 2005
Ohtsuki & Iwasa, J Theor Biol 2004, 2006, etc.

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

I help you and somebody else will help me



By helping someone, individuals may increase their reputation, which may change the predisposition of others to help them in the future

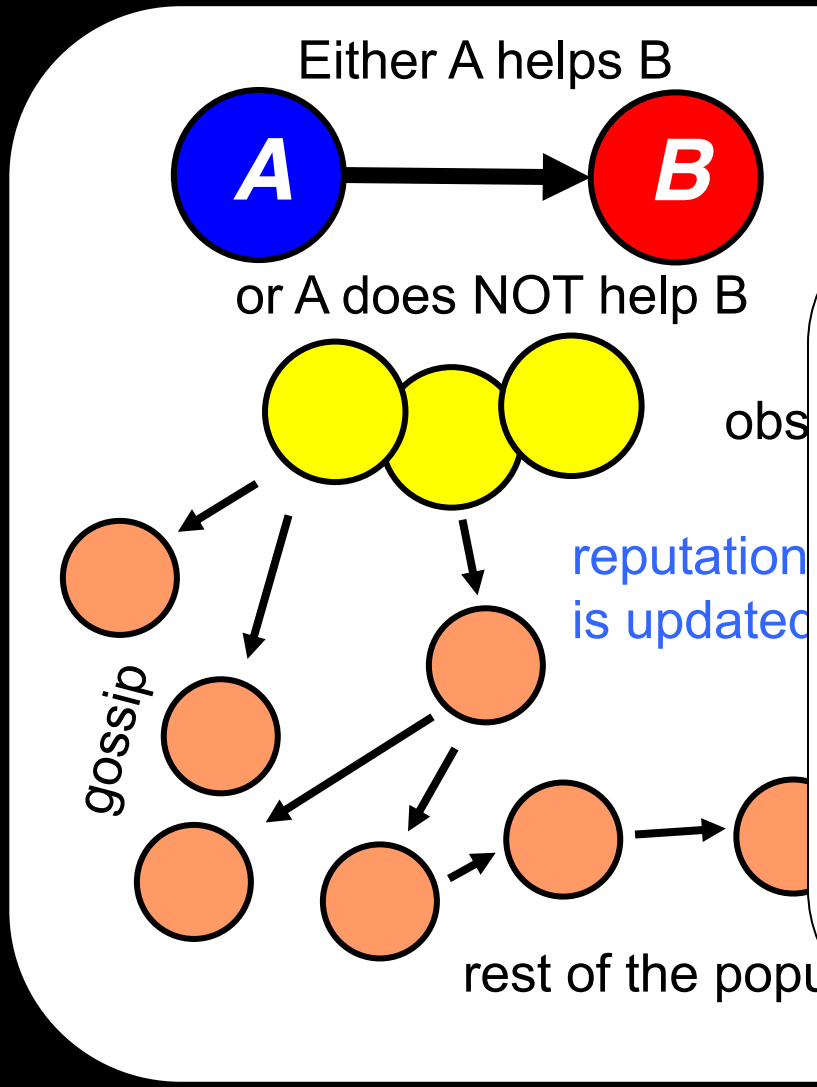
-> *social control, gossip, assessment of actions (social norms)*

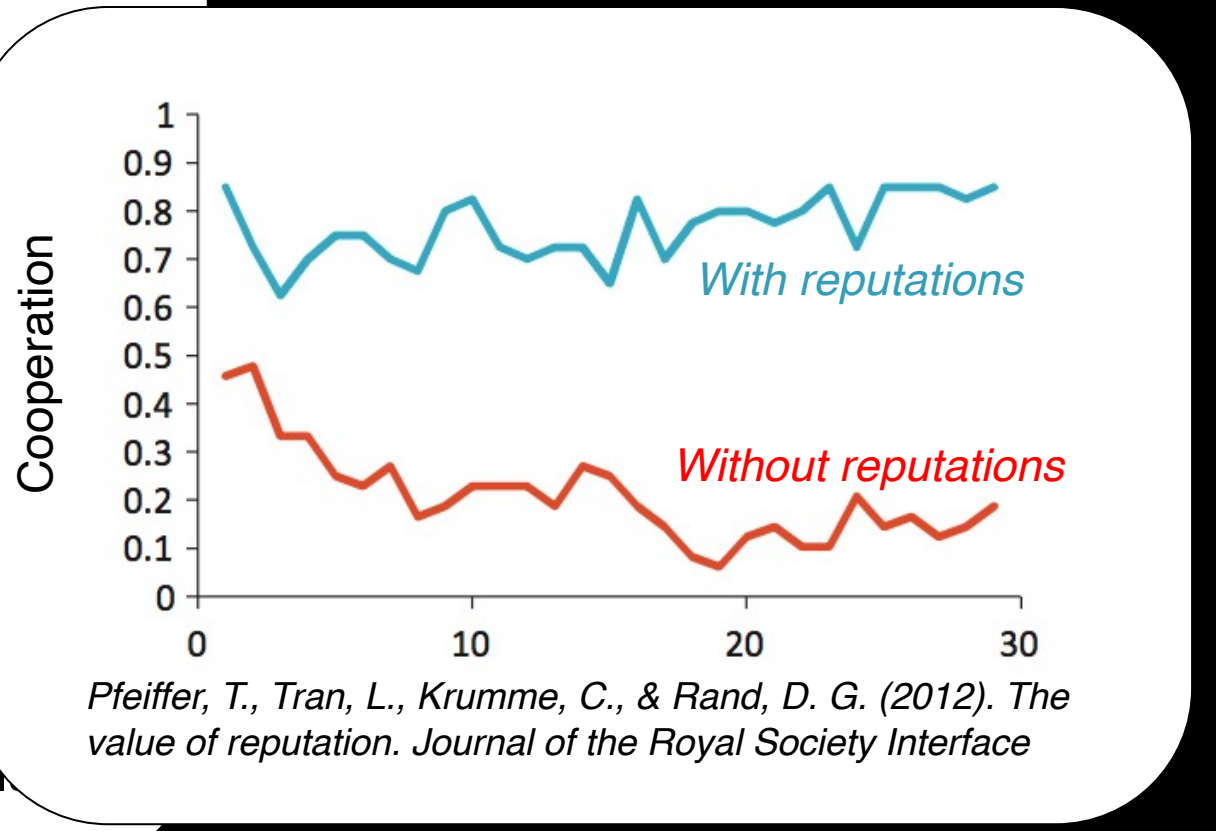
Indirect reciprocity

RD Alexander, The Biology of Moral Systems, 1987
Nowak & Sigmund, Nature 1998, Nature 2005
Ohtsuki & Iwasa, J Theor Biol 2004, 2006, etc.

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

I help you and somebody else will help me





Pfeiffer, T., Tran, L., Krumme, C., & Rand, D. G. (2012). The value of reputation. Journal of the Royal Society Interface

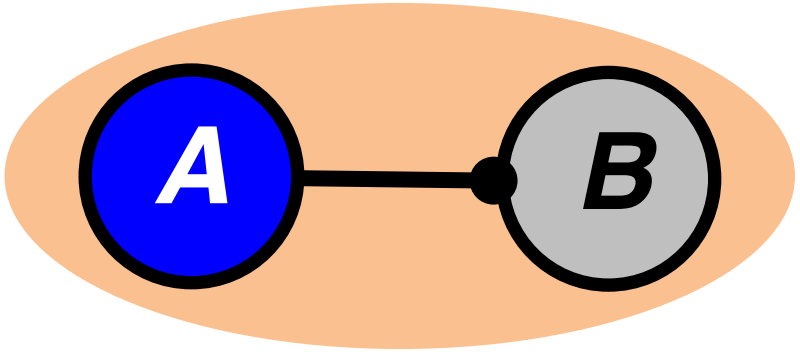
Indirect reciprocity and the evolution of moral systems

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

After each interaction a new reputation is assigned…

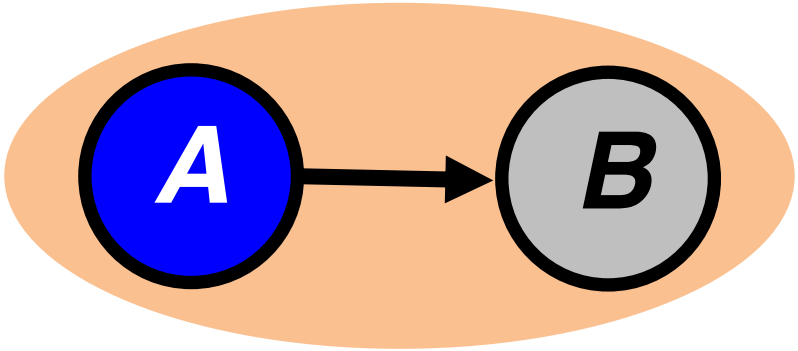
How do we define a *social norm*?

A does not help B



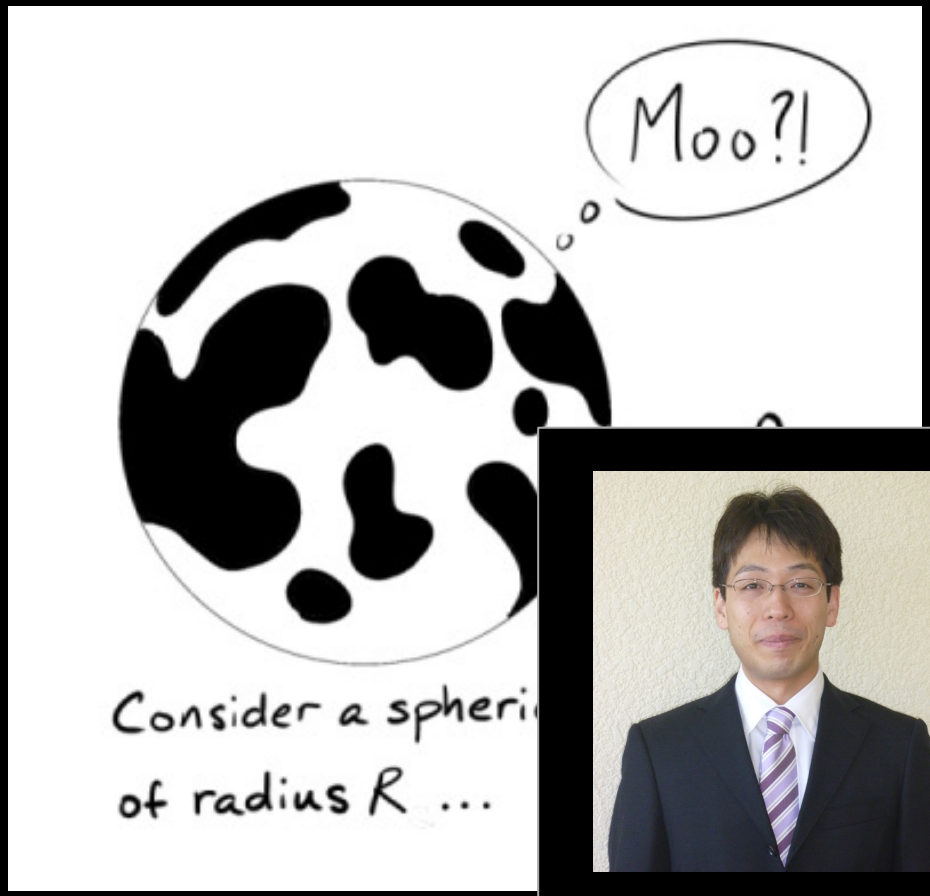
Should the reputation of A decrease?

A helps B



Should the reputation of A increase?

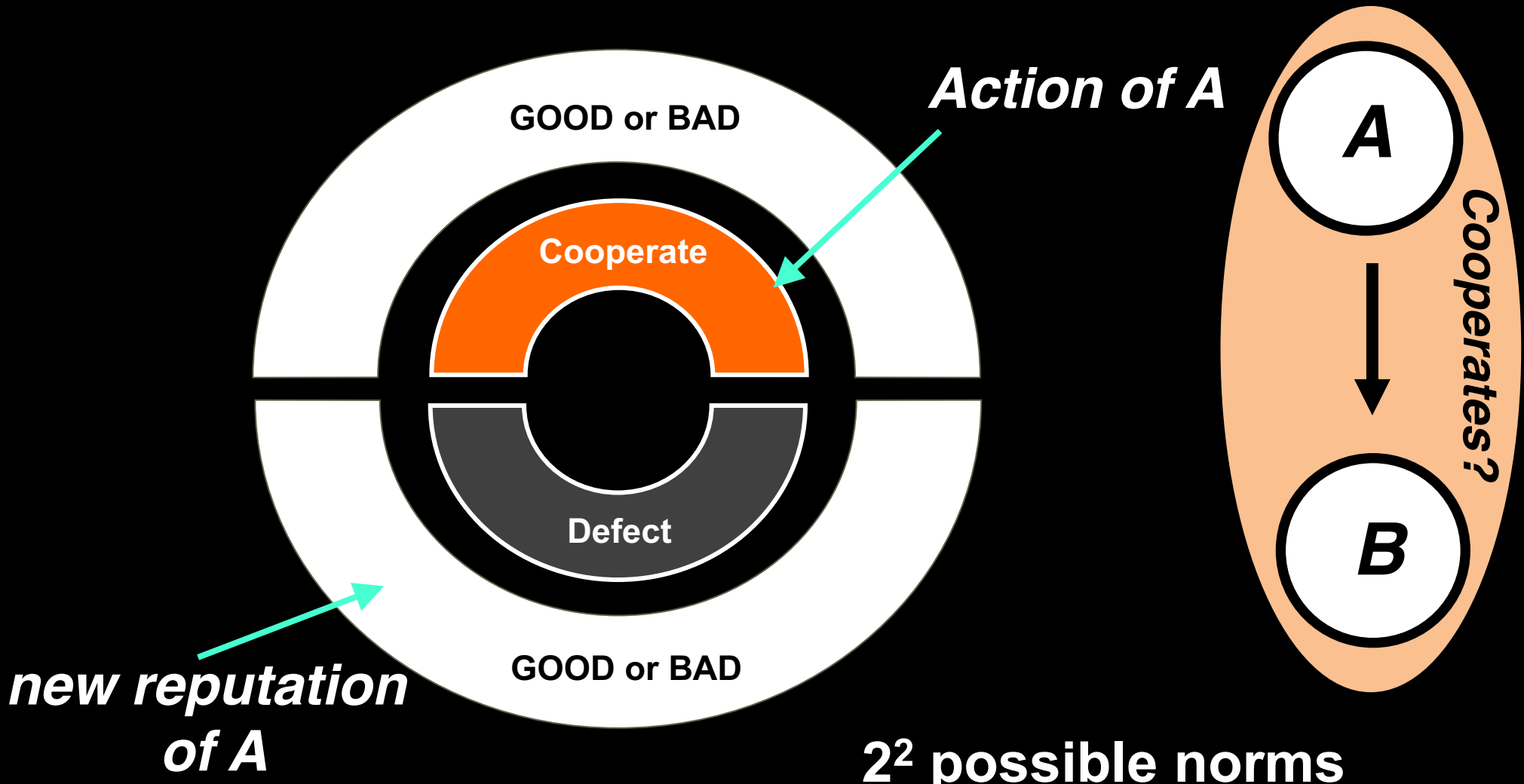
The beauty of modelling core principles



Ohtsuki, H., & Iwasa, Y.,
J Theor. Biol. (2004), J Theor Biol. (2006),
J Theor Biol. (2007)

1st order norms in a black & white world

The simplest assessment rule (2 bits)… 1st order norm *('98)*



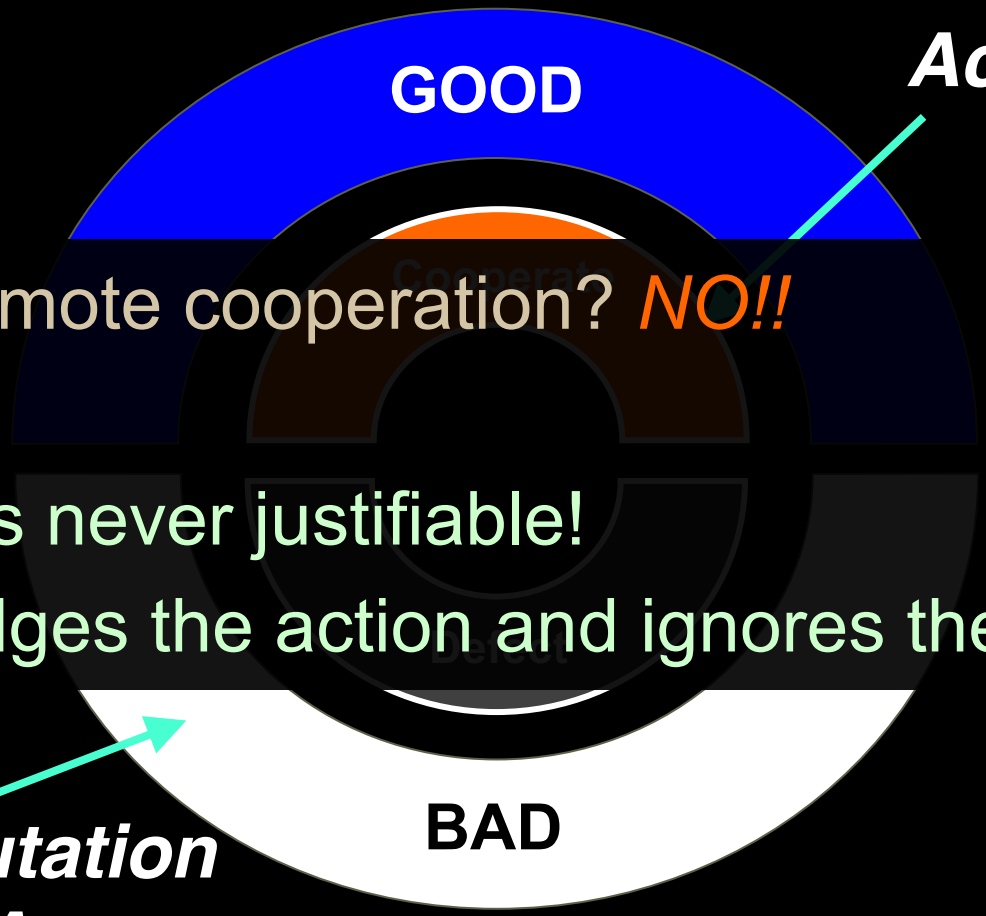
2^2 possible norms

1st order norms: Image scoring ('98)

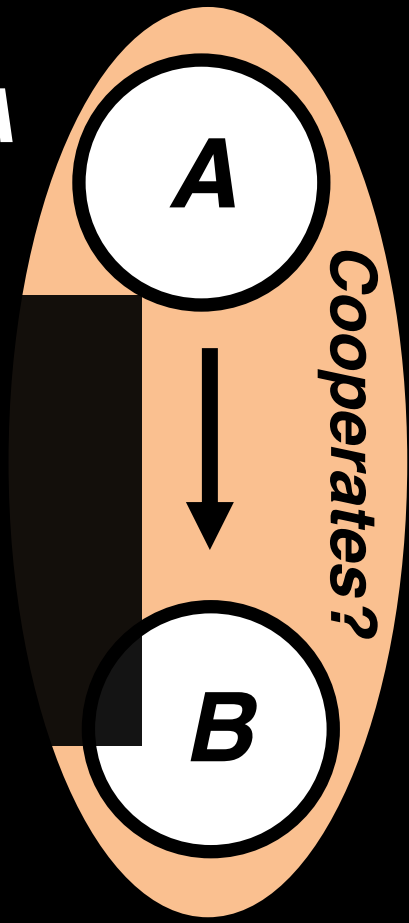
The simplest assessment rule (2 bits)… 1st order norm *('98)*

Only one makes sense (*Image scoring*):

Good when C, Bad when D



Action of A



Does it promote cooperation? *NO!!*

Defection is never justifiable!

Scoring judges the action and ignores the stories.

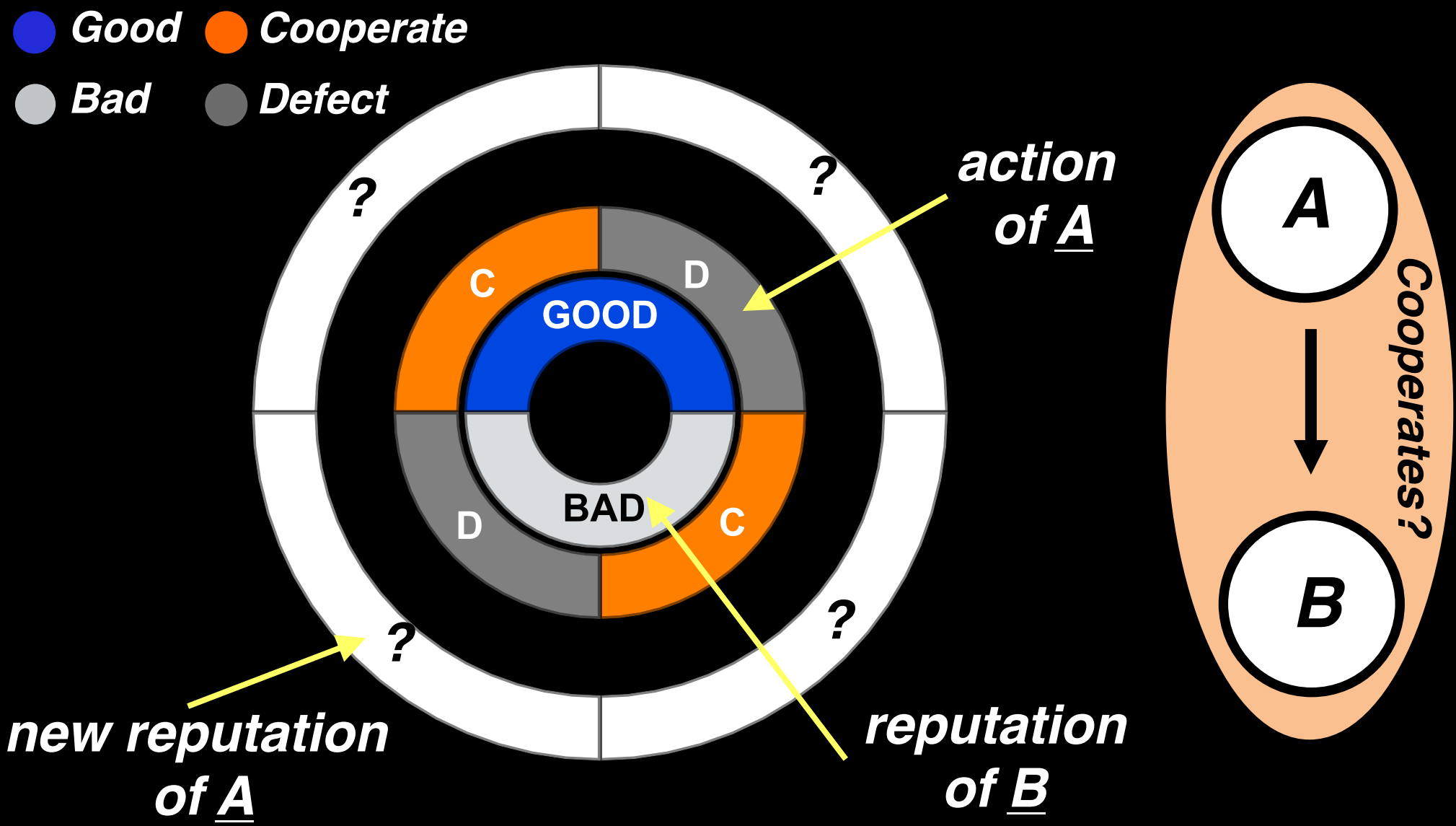
new reputation of A

2^2 possible norms

2nd order norms (4 bits): adding the recipient's reputation

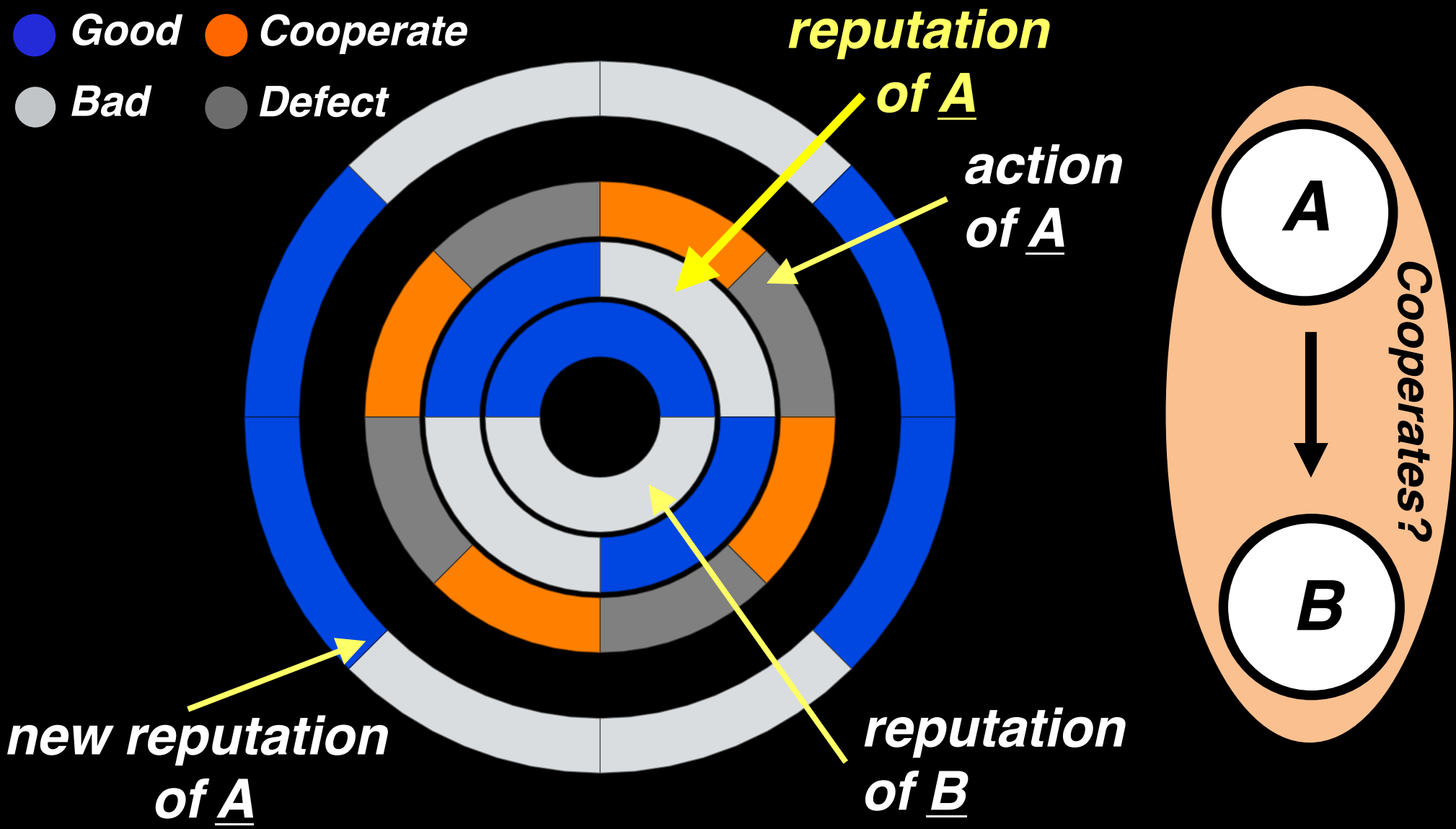
Leimar, Hammerstein, PRSB 2001
Ohtsuki, Iwasa, JTB, 2004, JTB 2006

2⁴ norms



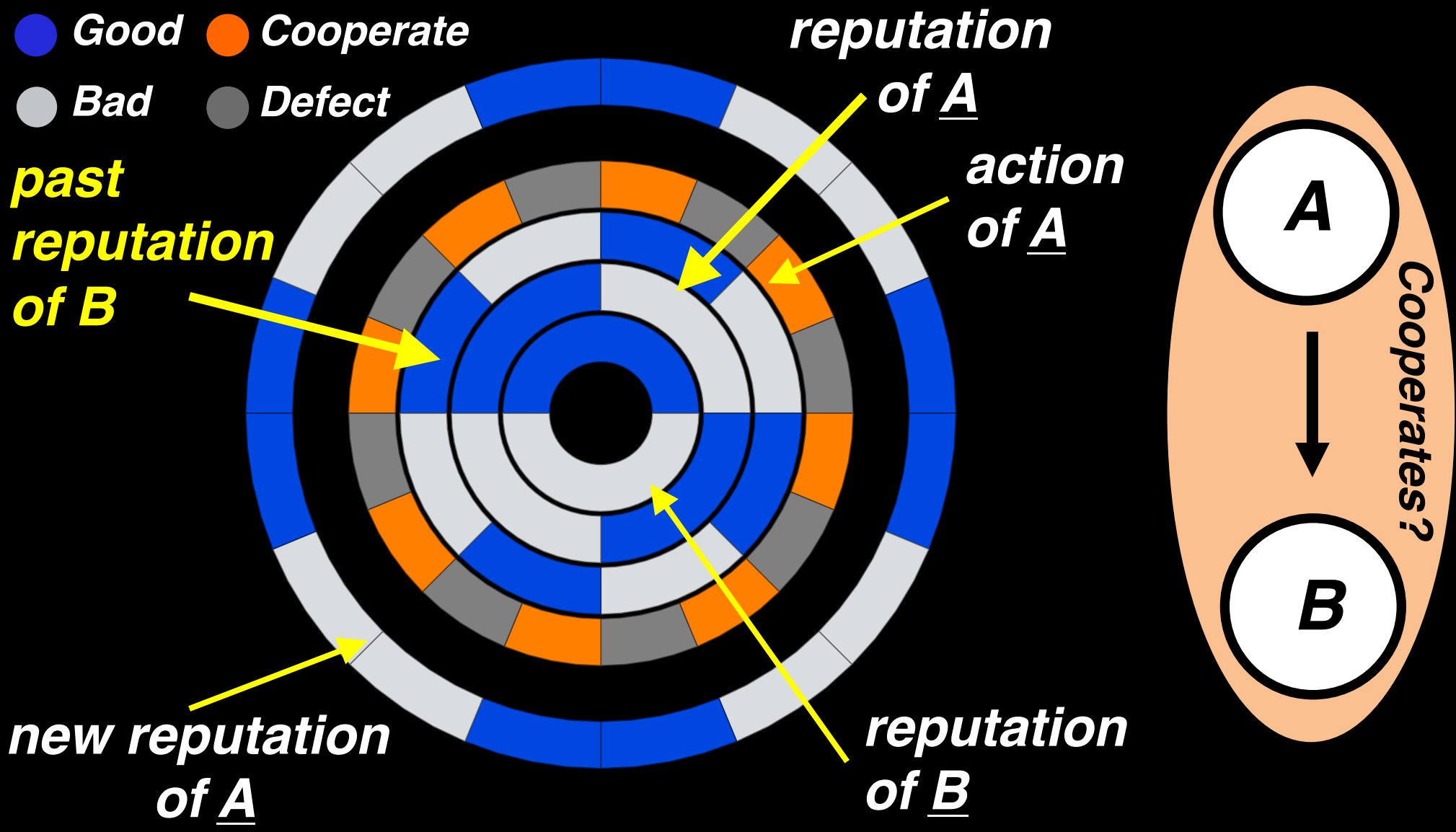
3rd order norms (8 bits): adding the donor's reputation

2^8 norms



4th order norms (16 bits): adding the previous reputation of the receiver

2^{16} norms



Who should be considered good for cooperation to thrive?

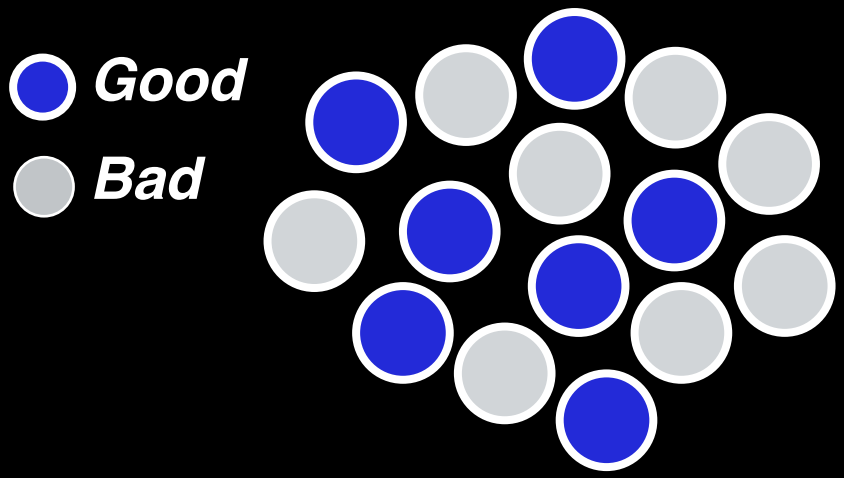
What's the most "efficient" social norm?

Let evolution decide!

Ohtsuki, H., & Iwasa, Y.,
J Theor. Biol. (2004, 2006 & 2007)

FP Santos, FC Santos, JM Pacheco,
Nature 2018, Plos Comput. Biol. 2016, 2007, etc.

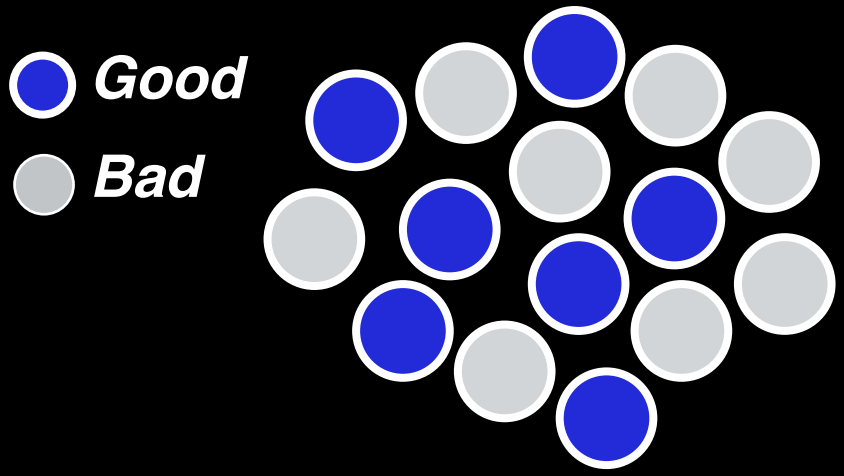
Each individual has a reputation...



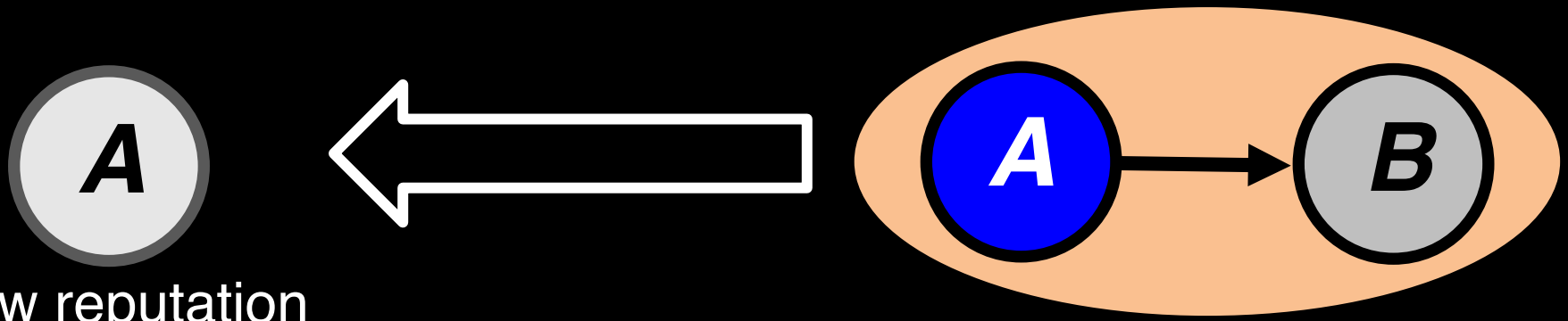
Each individual has a reputation...
and a strategy (or action rule)

2nd order

Rep.	AllC	AllD	Disc	*p*Disc
G	C	D	C	D
B	C	D	D	C



Action rules that do well spread in the population



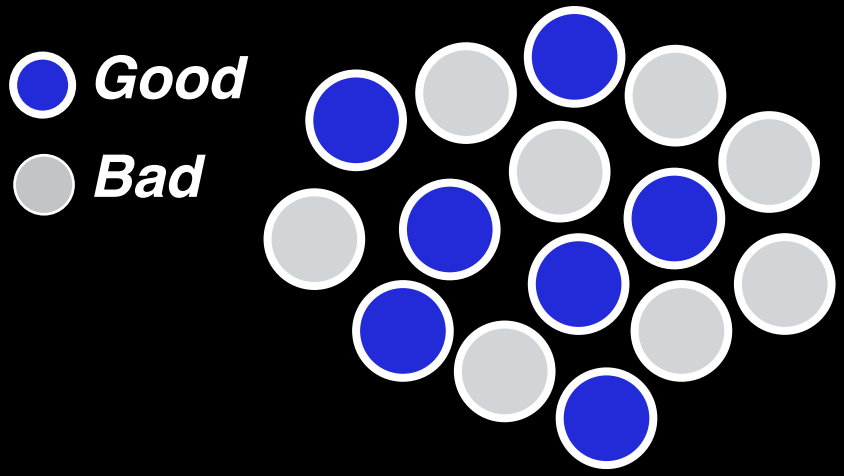
New reputation based on the social norm

Choose C/D based on the adopted action rule

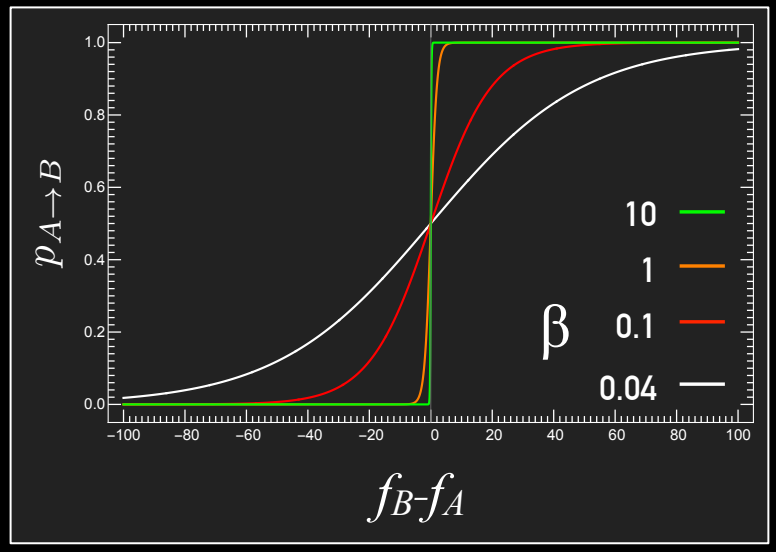
Each individual has a reputation...
and a strategy (or action rule)

2nd order

Rep.	AllC	AllD	Disc	*p*Disc
G	C	D	C	D
B	C	D	D	C



Action rules that do well spread in the population

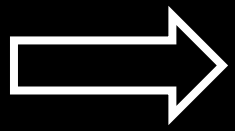


A imitates a random individual B with a probability that increases with the fitness difference.

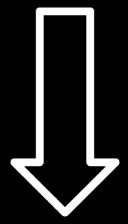
$$p = \left[1 + e^{-\beta(f_B - f_A)}\right]^{-1}$$

The model

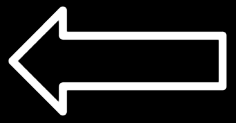
social norms are fixed in the population



different norms lead to different reputations



Reputations influence the success and frequency of strategies

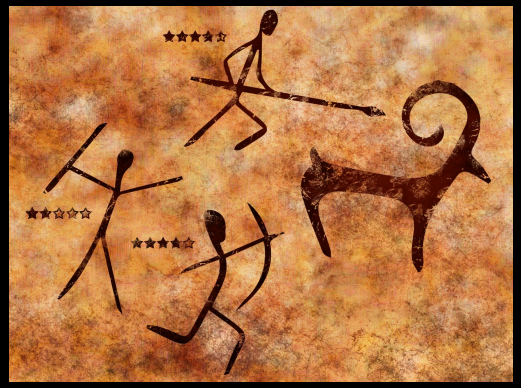


different levels of cooperation

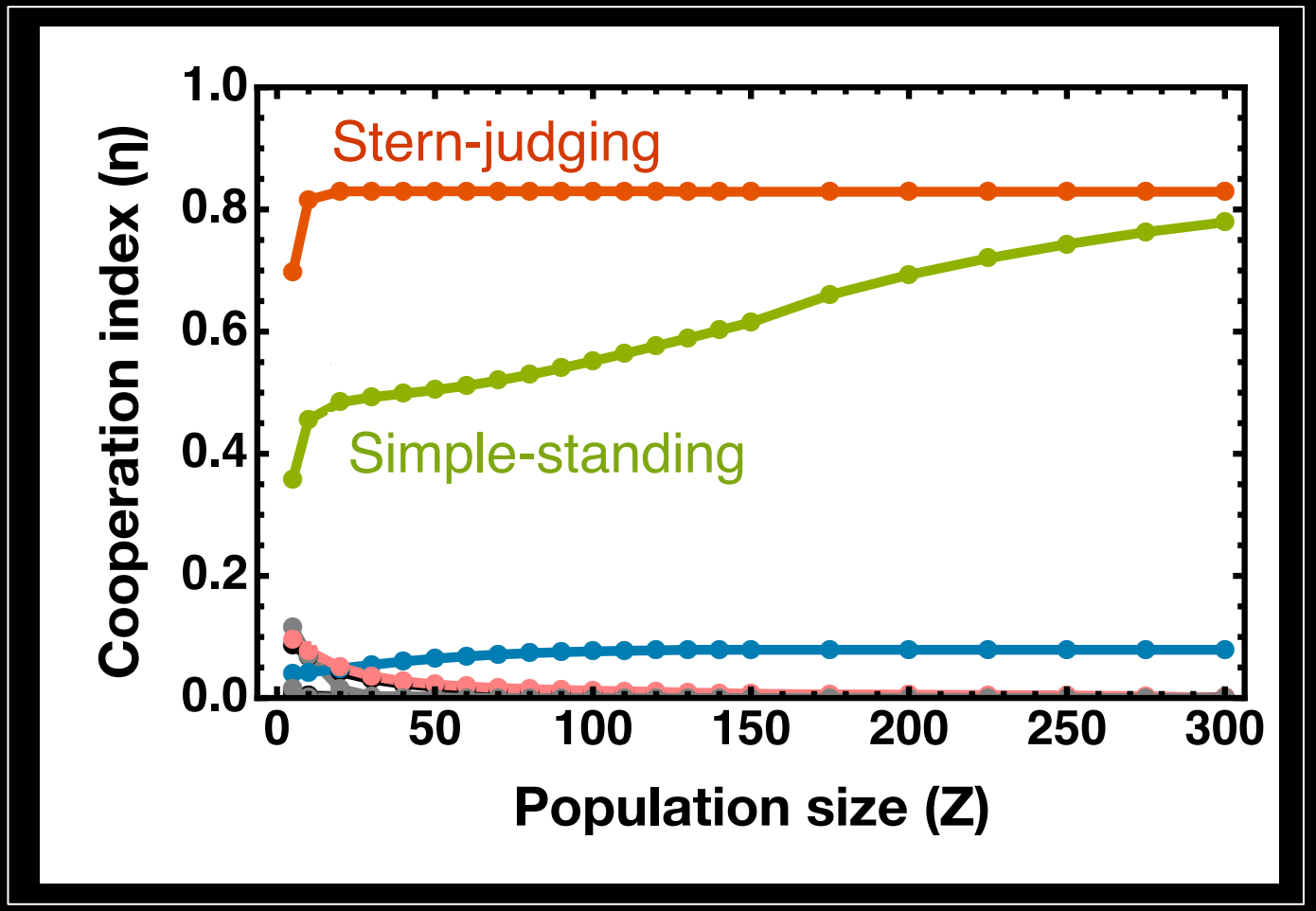
Who should be considered good for cooperation to thrive?

A special norm: Stern-judging (in small-scale societies)

Help the good and refuse help to the bad ; otherwise you will become *bad*.



Fernando P. Santos
Amsterdam & Princeton
Recruting PhD students:
https://fp-santos.github.io/



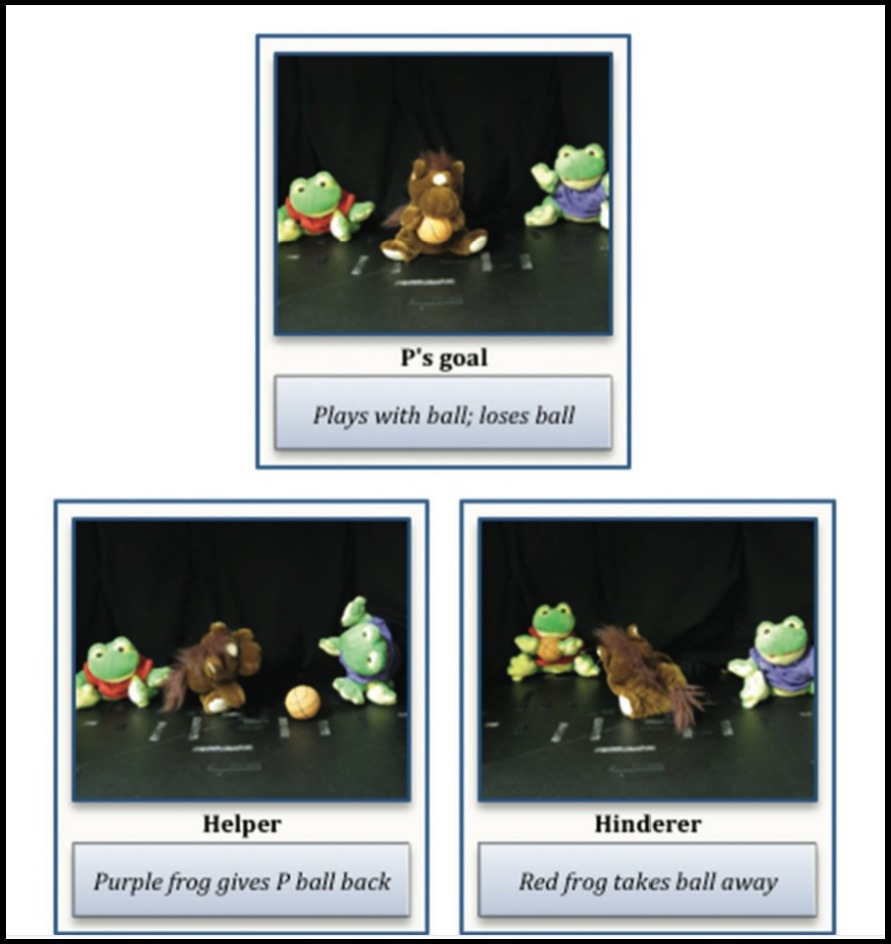
FP Santos, JM Pacheco & FC Santos, Sci. Rep., 2016
FP Santos, FC Santos & JM Pacheco, PLoS Comput Biol, 2016

A special norm: Stern-judging

Help the good and refuse help to the bad ; otherwise you will become *bad*.

Experiments: infants prefer puppets that help others and puppets that defect against hinderers

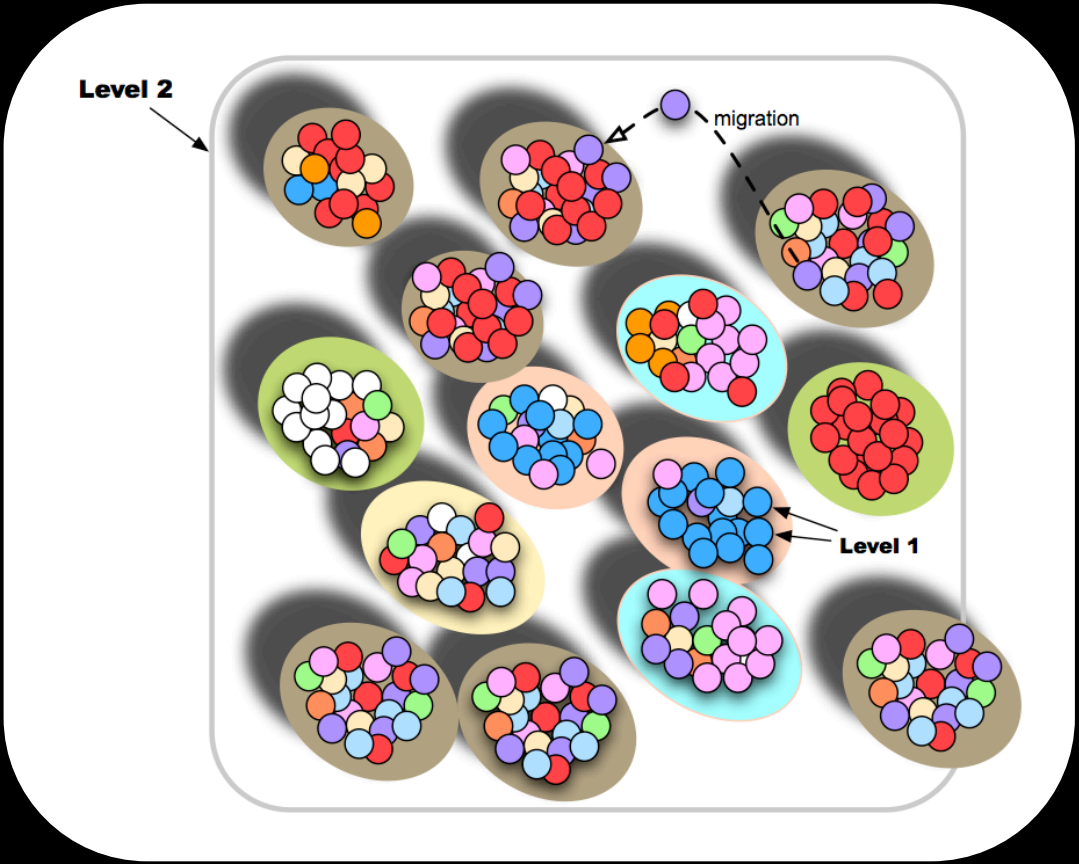
Help the good and refuse help to the bad



Hamlin, J. K. (2013). Moral Judgment and Action in Preverbal Infants and Toddlers Evidence for an Innate Moral Core. Current Directions in Psychological Science

Hamlin, J. K. (2011) How infants and toddlers react to antisocial others. PNAS

Other types of models



- *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;*
- *the loosing tribe changes towards the norm of the winning tribe;*

Pacheco, J. M., Santos, F. C., & Chalub, F. A. C. (2006). PLoS computational biology, 2(12), e178.

Conclusion: bits of good and evil!

- A single norm maximizes cooperation, while minimizing its cognitive complexity & order (particularly, in small societies).
- Cooperate with good and defect against bad; all else is *bad*… Even toddlers know it! ☺
- Cooperative societies can be sustained by very simple moral judgments that neglect past reputations.

Conclusion: bits of good and evil!

*Indirect reciprocity allows the characterization of the **evolution** of complex cultural concepts, such as:*

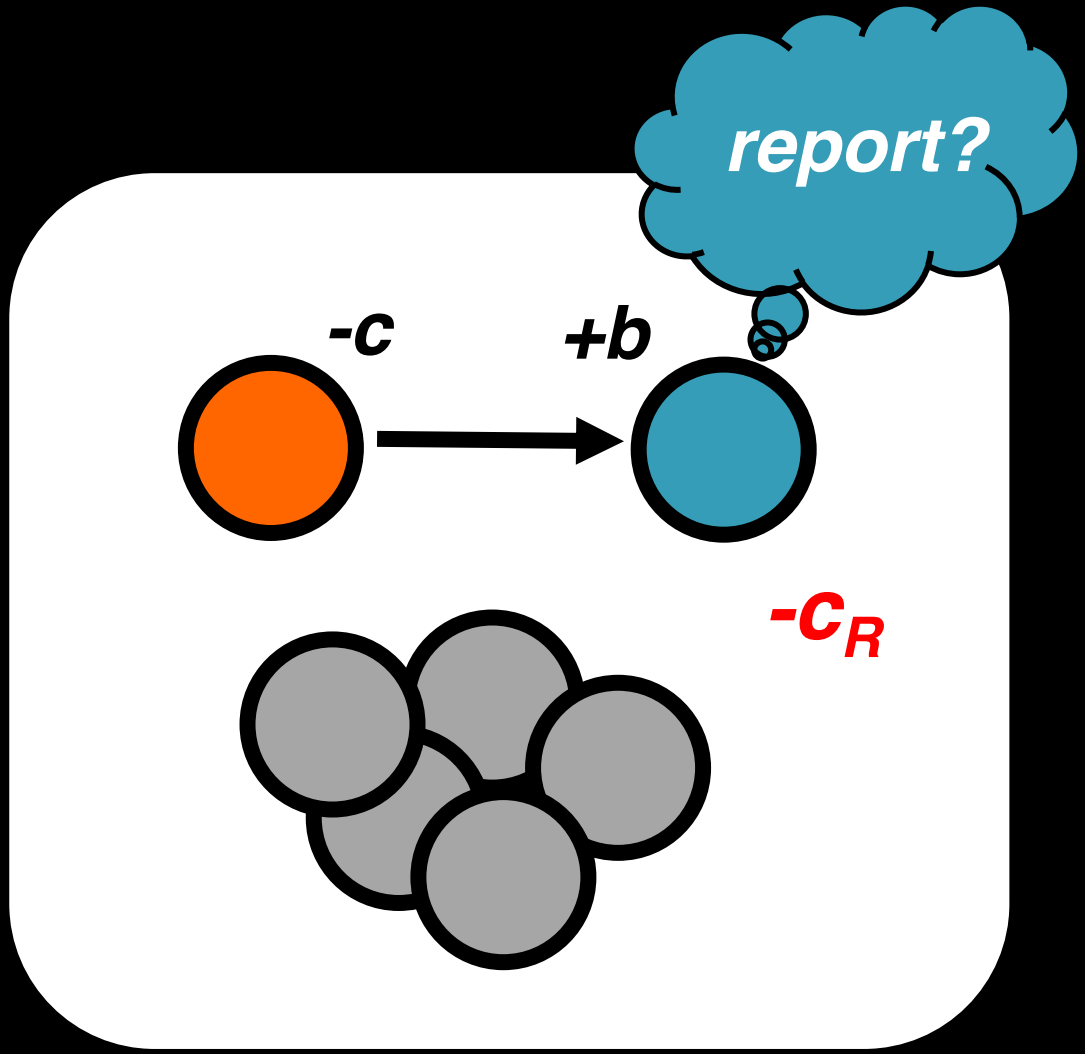
- Punishment
- Forgiveness
- Evolution of laws and moral systems

- Appealing definitions of GOOD and EVIL.

...but also has several limitations

The Achilles heel of indirect reciprocity:

Inaccurate information about reputations

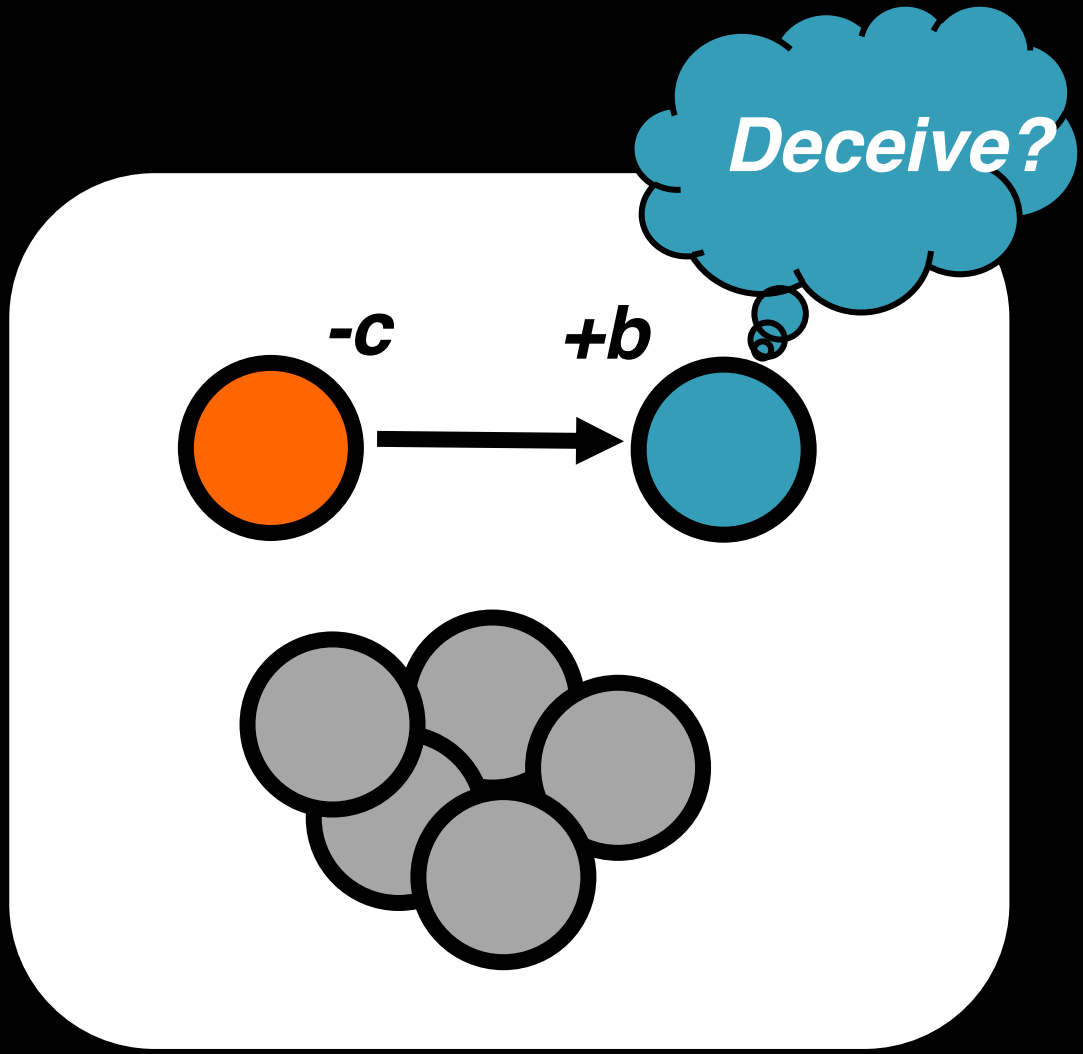


Fernando P. Santos
Amsterdam & Princeton

Santos, F. P., Pacheco, J. M. & Santos, F. C. (2018). Social Norms of Cooperation with Costly Reputation Building. AAAI'2018

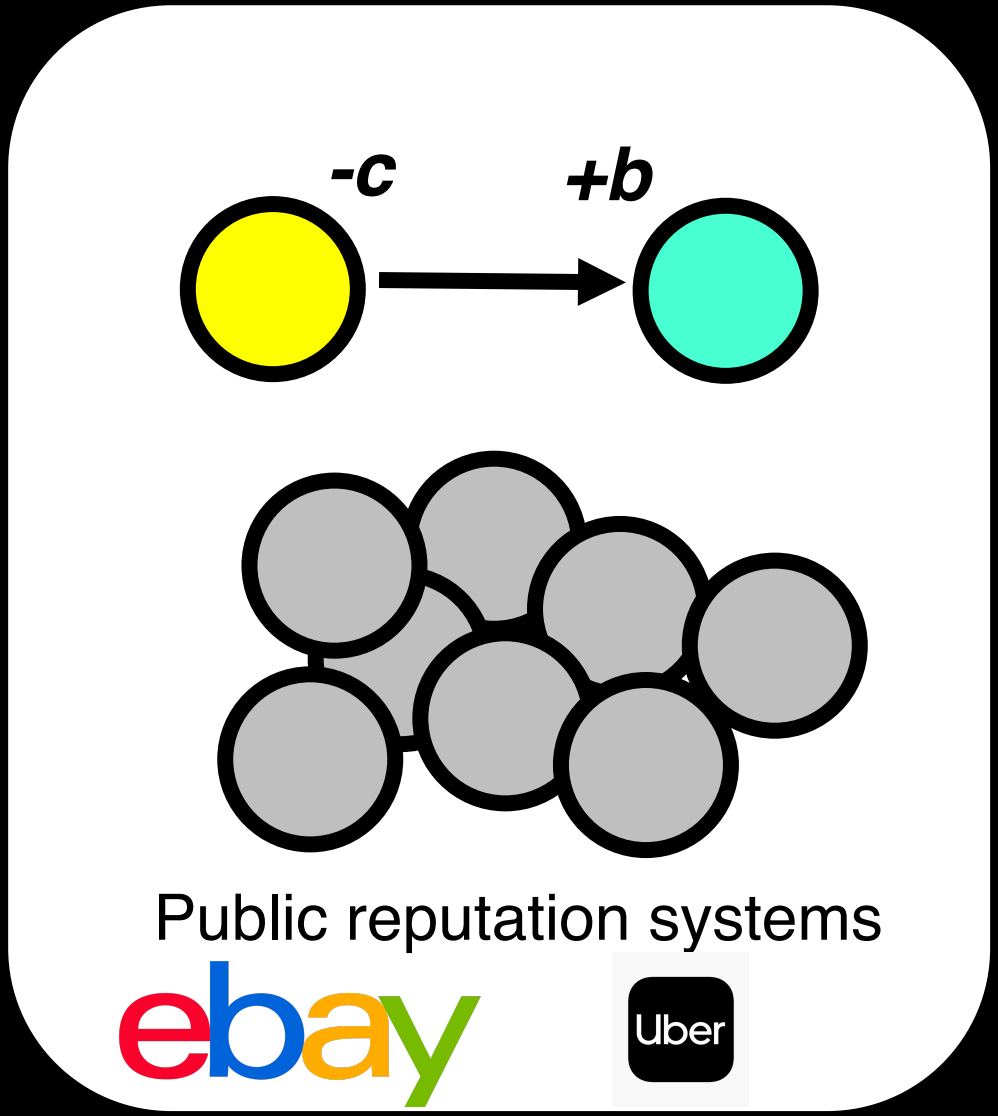
The Achilles heel of indirect reciprocity:

Inaccurate information about reputations



Számadó et al. Deception undermines the stability of cooperation in games of indirect reciprocity PloS one (2016)

From public to private information systems



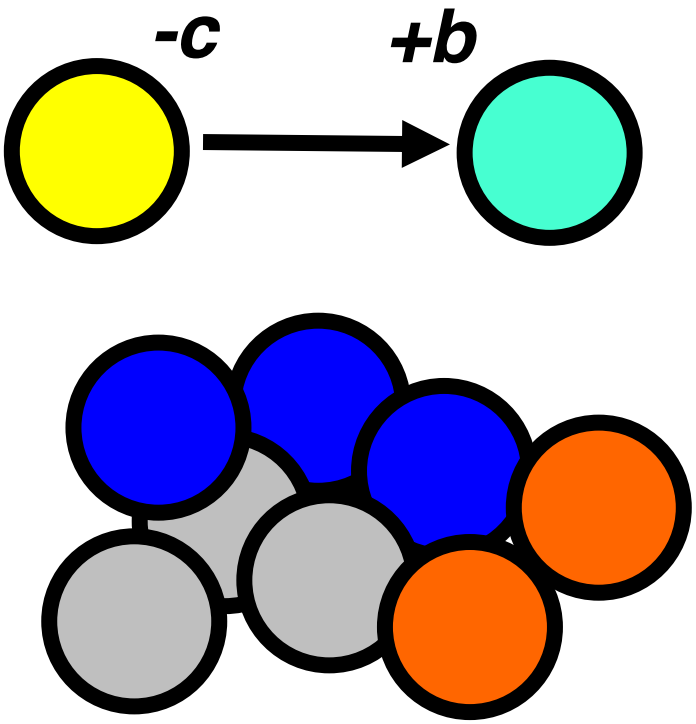
Binary reputations

G ***GOOD***

B ***BAD***

From public to private information systems

H. Fonseca, et al. 2020



Partial observability?

Ternary reputations

G *GOOD*

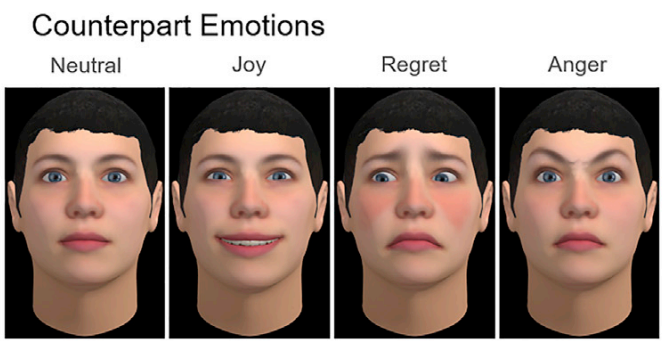
B *BAD*

U *UNKNOWN*

Emotion expressions & indirect reciprocity

Emotion-based reputations and social norms (3x2x3 experiments, n=711)

Emotions are used to forgive those that defect but also to punish those that cooperate.



Competitive

Counterpart

Participant	Cooperation	Defection
Cooperation	Regret	Joy
Defection	Anger	Neutral

Cooperative

Counterpart

Participant	Cooperation	Defection
Cooperation	Joy	Regret
Defection	Anger	Neutral

Initial Reputation	Strategy	Emotion	Final Reputation
Negative	Competitive	Competitive	Negative
		Neutral	Negative
		Cooperative	Negative
	Cooperative	Competitive	Neutral
		Neutral	Neutral
		Cooperative	Positive
Unknown	Competitive	Competitive	Negative
		Neutral	Negative
		Cooperative	Neutral
	Cooperative	Competitive	Positive
		Neutral	Positive
		Cooperative	Positive
Positive	Competitive	Competitive	Negative
		Neutral	Neutral
		Cooperative	Positive
	Cooperative	Competitive	Positive
		Neutral	Positive
		Cooperative	Positive

Emotion expressions shape human social norms and reputations
C. de Melo, K.Terada, and F.C. Santos
iScience, 24, 102141, Cell Press, Feb 2021.

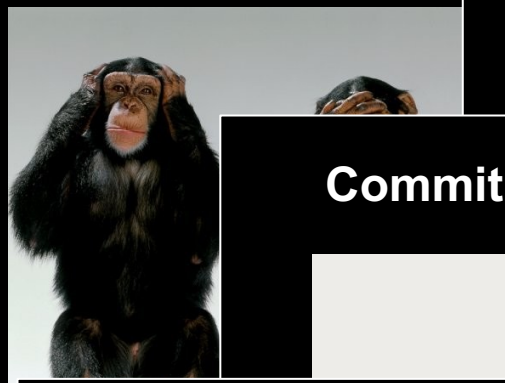
Celso de Melo
U.S. ARL, Los Angeles

Kazunori Terada
Gifu Univ., Japan

Other ways out of the dilemma

Tag-based cooperation

Signaling for cooperation



Commitments

Co-evolution of Cognitive/Social Skills & Cooperation

Santos, Pacheco, Skyrms, Journal of Theoretical Biology (2011)
Han, Pereira, Santos, Adaptive Behavior (2011), IJCAI'2011, IJCAI'2013, Sci. Rep. 2015
Pacheco, Vasconcelos, Santos, Skyrms, PLoS Computational Biology (2015).

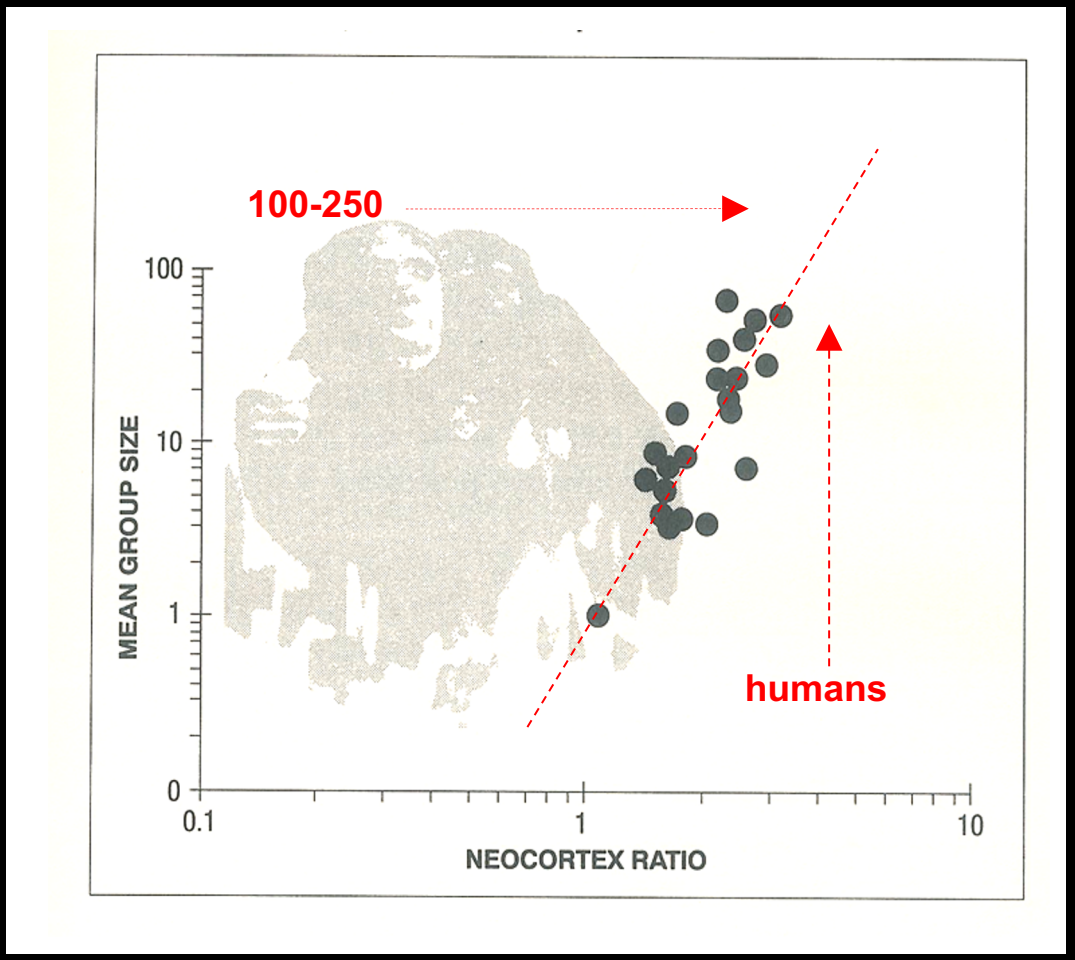
Evolution of social-cognition & cooperation

Co-evolution of different features

i) Social cognition

ii) Group size

iii) Cooperation and social cohesion



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

A new evolutionary paradigm

Evolution in networked populations