Gossip, Reputation, Norms, and Cooperation

Károly Takács

**Institute for Analytical Sociology,
Linköping University, Sweden**

co-authors:

Károly Takács
karoly.takacs@liu.se
Linköping University, Sweden

Shirsendu Podder
s.podder@live.co.uk
University College London

Simone Righi
simone.righi@unive.it
Ca'Foscari University of Venice

International online seminar
Artificial Societies and
Information Technology
October 15, 2021

*Acronym: EVILTONGUE
Project number 648693*

THE PROBLEM OF COOPERATION: WHO IS A FREE RIDER?

Kyoto protocol
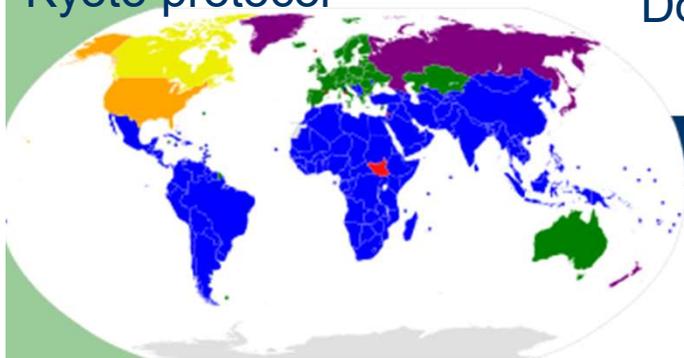
Doha amendment

CO₂ emissions per capita, 2016
Average carbon dioxide (CO₂) emissions per capita measured in tonnes per year.

Kyoto successes (blue) and failures (red)

Status of the Paris Agreement (February 2021)

By Our World in Data - https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=95471080

LiU LINKÖPINGS UNIVERSITET
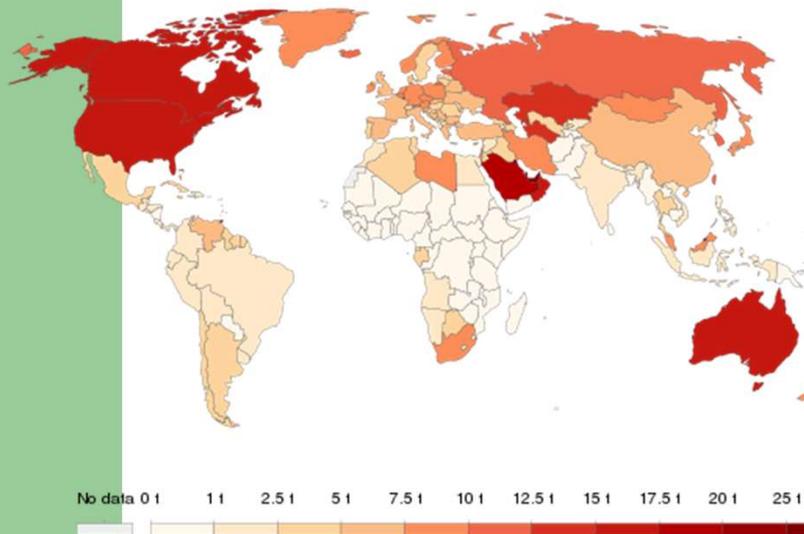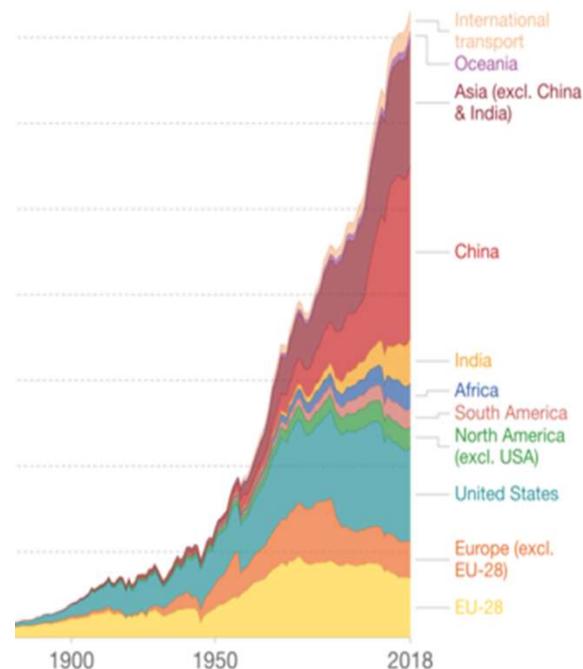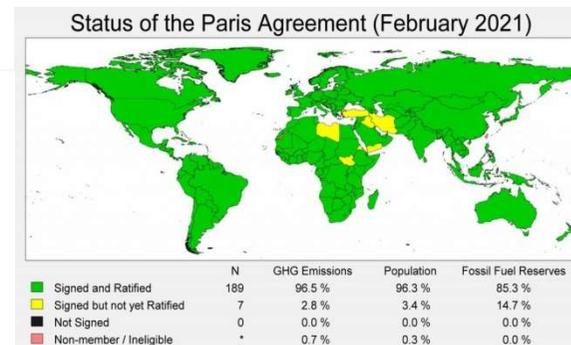
PROBLEM: HUMAN COOPERATION

Unprecented high level of large-scale cooperation

The Prisoner's Dilemma

	cooperate	defect
cooperate	R, R	S, T
defect	T, S	P, P

T > R > P > S

social optimum ≠ equilibrium

3

li.u LINKÖPINGS UNIVERSITET

SOLUTION OF THE PROBLEM: DIFFICULT

The Prisoner's Dilemma

	cooperate	defect
cooperate	R, R	S, T
defect	T, S	P, P

?

T > R > P > S

4

li.u LINKÖPINGS UNIVERSITET

DIRECT RECIPROCITY

The Prisoner's Dilemma

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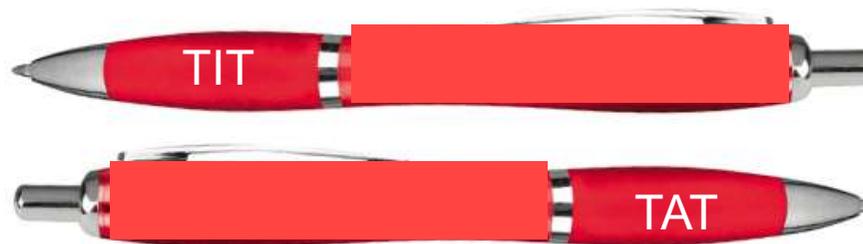
T > R > P > S

- Trivers 1971 *reciprocal altruism*
- Axelrod and Hamilton 1981
- Axelrod 1984

- *Repeated interaction*
- Shadow of the future
- Direct reciprocity is not feasible in large populations

?

TIT

TAT

THE EVOLUTION OF COOPERATION

Robert Axelrod
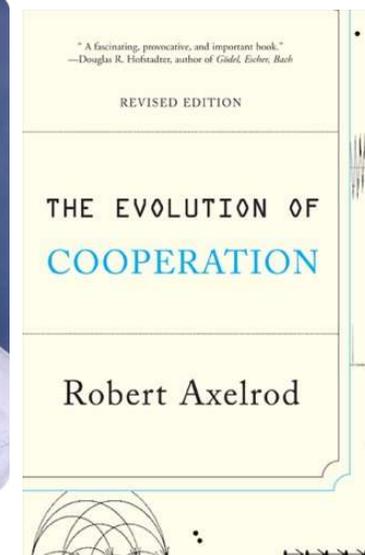
KINSHIP SELECTION / POSITIVE ASSORTMENT

The Prisoner's Dilemma

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

?

T > R > P > S

- Hamilton's rule: Hamilton, 1964 $c < r * b$
- Németh, A., Takács, K. 2010. *J. Theor. Biol.* 264: 301-311.
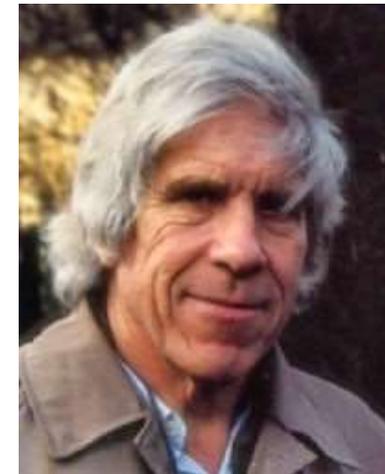
GROUP SELECTION / TEAM COMPETITION

The Prisoner's Dilemma

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

?

$T > R > P > S$

- Multilevel selection
- Sober and Wilson 1998 *Unto Others*
- Boyd and Richerson
- Team games
- Social identity and group categorization theories

7

LI.U LINKÖPINGS UNIVERSITET

EMBEDDED PLAY

The Prisoner's Dilemma

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

?

T > R > P > S

- Spatial games
- Nowak and May, 1992; 1993
- Németh, A., Takács, K. 2007. *JASS*, 10(3): 4.
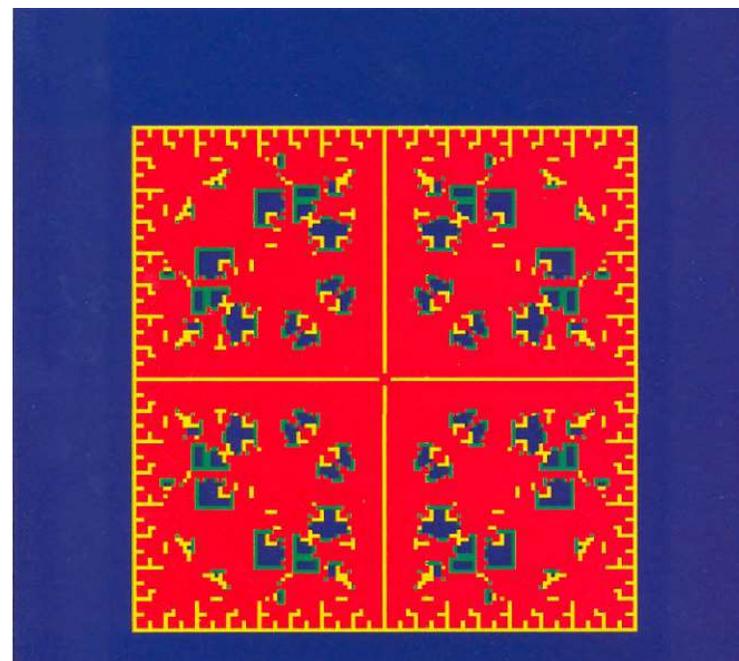
COOPERATION IN NETWORKS

The Prisoner's Dilemma

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

T > R > P > S

?

- Games in networks/graphs: Hauert, 2004; Lieberman, Hauert, Nowak, 2005; Santos et al., 2006; Szabó and Fáth, 2007; Takács et al., 2021 *Phil.Trans.Roy.Soc.B.*
- Ohtsuki et al., 2006: $b/c > k$ (density)
  - Note: death/birth updating, local

9
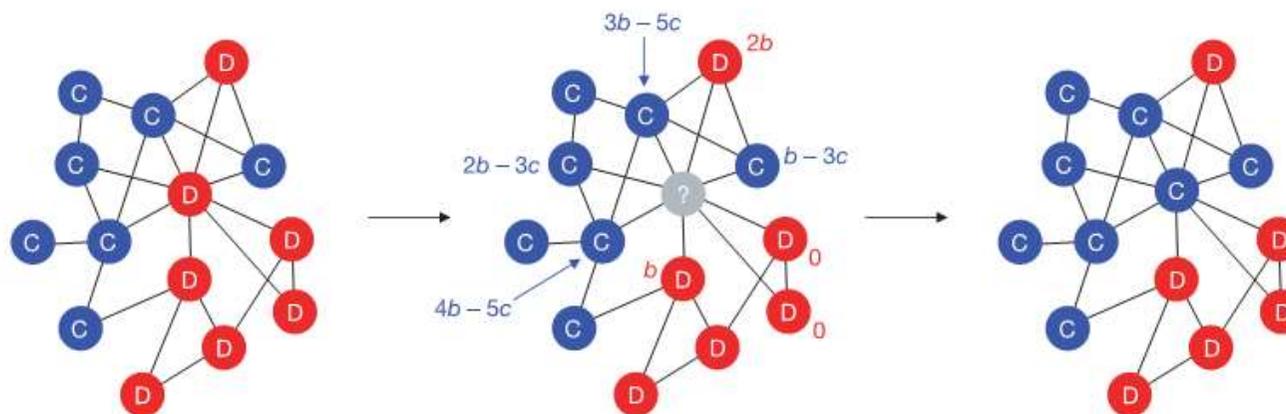
FOCUS: INDIRECT RECIPROCITY AND REPUTATION

The Prisoner's Dilemma

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

T > R > P > S
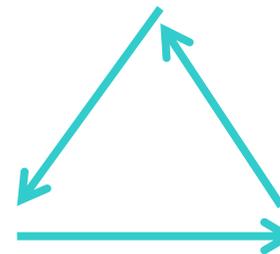
Indirect reciprocity: Help (or retaliation) does not come from the interaction partner who was helped (cheated) by the individual, but from somebody else (Nowak and Sigmund, 2005)

Wedekind and Milinski (2000): *image scoring*
        - eBay and other reputation-based markets
Nowak (2006): reputation increases cooperation if information is efficiently transmitted (publicly available and objective)

10

THE PROBLEM OF COOPERATION

- ❏ **Who are the cooperators?**
- ❏ **Not always known / not always public knowledge**
  - ❏ **Especially if cooperation problems are local interactions**
- ❏ **Free riders might not be condemned by social judgment**
- ❏ **Who are perceived to be good?**

REPUTATION

HOW SHOULD WE DEFINE GOODNESS?

❑ **Ohtsuki Hisashi & Iwasa Yoh (2004)** *J. Theor. Biol.*

	B	
	C	D
A C	(b-c, b-c)	(-c, b)
D	(b, -c)	(0, 0)

COOPERATORS	DEFECTORS
GOOD ☺	BAD ☹

12

LINKÖPINGS UNIVERSITET

SOCIAL NORMS

❑ **But what to do when somebody meets a defector?**

❑ **Conditional cooperation on reputation of opponent**

Cooperation against cooperators	Defection against cooperators
GOOD ☺	BAD ☹

Cooperation against defectors	Defection (punishment) against defectors
BAD?	GOOD?

SOCIAL NORMS THAT CAN SUPPORT LARGE-SCALE COOPERATION

Many different social norms are possible

What are those norms that - if generally followed - **sustain** large-scale cooperation among strangers?
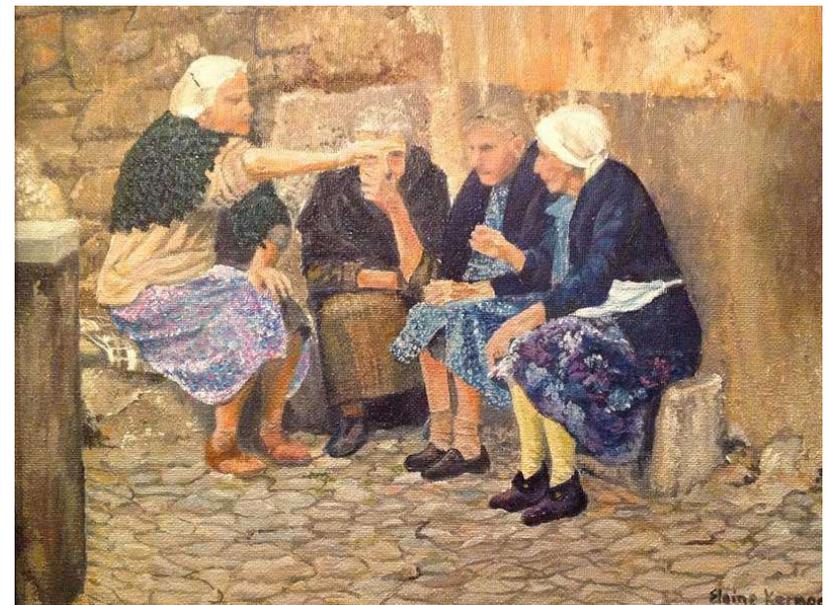
THE LEADING EIGHT NORMS

Norms s define:

1: good reputation, 0: bad reputation

- reputational update (left 8 columns)
  - e.g., third column: assigned reputation to individuals with good reputation who cooperated with an individual of bad reputation
- conditional action (right 4 columns)
  - e.g., last column: prescribed action for an individual with bad reputation against an individual with bad reputation

Group	d_{11C}	d_{11D}	d_{10C}	d_{10D}	d_{01C}	d_{01D}	d_{00C}	d_{00D}	P_{11}	P_{10}	P_{01}	P_{00}	
I	1	0	1	1	1	0	1	0	C	D	C	C	s_1
	1	0	0	1	1	0	1	0	C	D	C	C	s_2
II	1	0	1	1	1	0	1	1	C	D	C	D	s_3
	1	0	1	1	1	0	0	1	C	D	C	D	s_4
	1	0	0	1	1	0	1	1	C	D	C	D	s_5
	1	0	0	1	1	0	0	1	C	D	C	D	s_6
III	1	0	1	1	1	0	0	0	C	D	C	D	s_7
	1	0	0	1	1	0	0	0	C	D	C	D	s_8

THE LEADING EIGHT NORMS

Coloured sections are defining characteristics

1: good reputation, 0: bad reputation

Group	d_{11C}	d_{11D}	d_{10C}	d_{10D}	d_{01C}	d_{01D}	d_{00C}	d_{00D}	p_{11}	p_{10}	p_{01}	p_{00}	
I	1	0	1	1	1	0	1	0	C	D	C	C	s_1
	1	0	0	1	1	0	1	0	C	D	C	C	s_2
II	1	0	1	1	1	0	1	1	C	D	C	D	s_3
	1	0	1	1	1	0	0	1	C	D	C	D	s_4
	1	0	0	1	1	0	1	1	C	D	C	D	s_5
	1	0	0	1	1	0	0	1	C	D	C	D	s_6
III	1	0	1	1	1	0	0	0	C	D	C	D	s_7
	1	0	0	1	1	0	0	0	C	D	C	D	s_8

X

i

j

Focal Agent

Opponent

Reputational update
d_{ijX}

Behavioural Strategy
p_{ij}
(Conditional action)

16

CHARACTERISTICS OF SUCCESSFUL NORMS

- Maintenance of cooperation (1st column)
- Identification of defectors (2nd and 6th columns)
  - Assigning bad reputation for those who refuse to cooperate with an opponent with good reputation
- Justified punishment (4th column)
  - Assigning good reputation for justified punishment by good players
- Forgiveness (5th column)
- Key differences between groups of norms are indicated with color

Group	d_{11C}	d_{11D}	d_{10C}	d_{10D}	d_{01C}	d_{01D}	d_{00C}	d_{00D}	P_{11}	P_{10}	P_{01}	P_{00}	
I	1	0	1	1	1	0	1	0	C	D	C	C	s_1
	1	0	0	1	1	0	1	0	C	D	C	C	s_2
II	1	0	1	1	1	0	1	1	C	D	C	D	s_3
	1	0	1	1	1	0	0	1	C	D	C	D	s_4
	1	0	0	1	1	0	1	1	C	D	C	D	s_5
	1	0	0	1	1	0	0	1	C	D	C	D	s_6
III	1	0	1	1	1	0	0	0	C	D	C	D	s_7
	1	0	0	1	1	0	0	0	C	D	C	D	s_8

SOCIAL NORMS SOLVING THE PROBLEM OF COOPERATION

Well-mixed populations

Globally available reputations

Global strategy update

We relax these assumptions
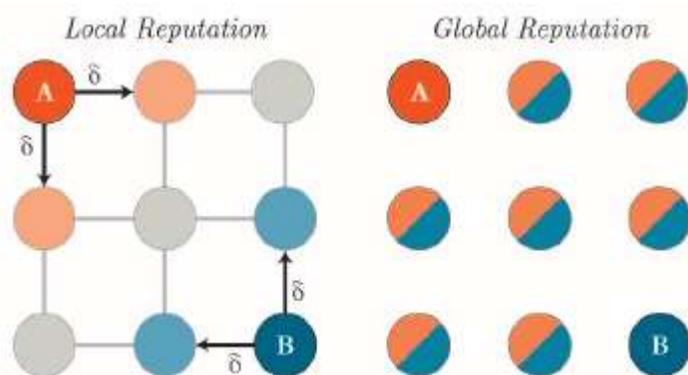
Global Reputation
An individual's reputation is well known by <u>anyone</u>.

Local Reputation
An individual's reputation is only known by their <u>neighbour</u>.
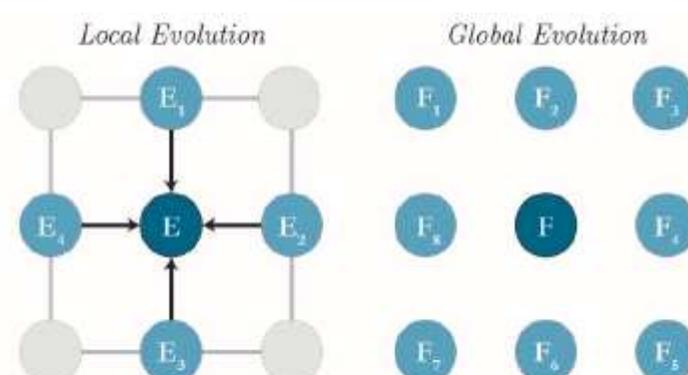
Global Evolution
Individuals adopt the population's best <u>average</u> strategy.

Local Evolution
Individuals adopt the best strategy in their <u>neighbourhood.</u>

METHOD: ABM

Setup

Individuals in a static network (single component, with min. degree of 2)

- Erdős-Rényi random graph; lattice; scale-free network; small world network
- Role: local reputation, local evolution (norm update)

Leading 8 norms + ALL D assigned

Dynamics

- Random matching for PD (min. 1 for all)
- Reputations assigned
- Agent update to a better social norm with probability α
- Run till convergence or till maximum time

19

RESULTS: TYPE II NORMS UNDERPERFORM

The Leading 8 norms are all successful in sustaining cooperation also with local reputation and local evolution
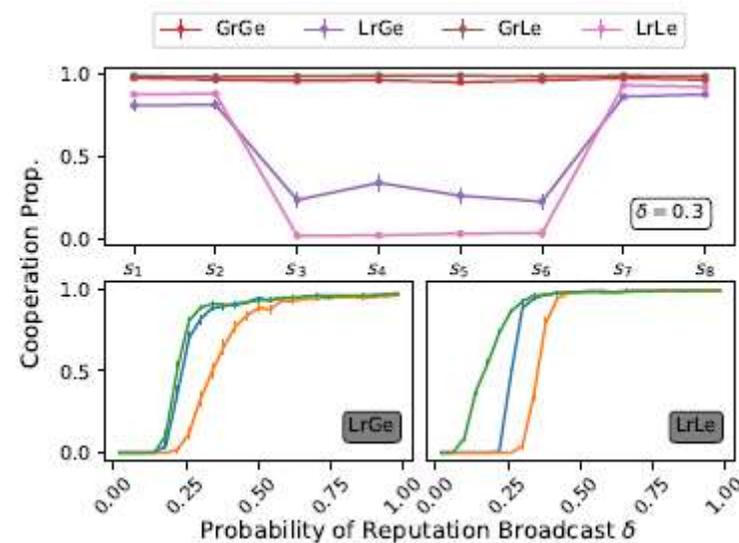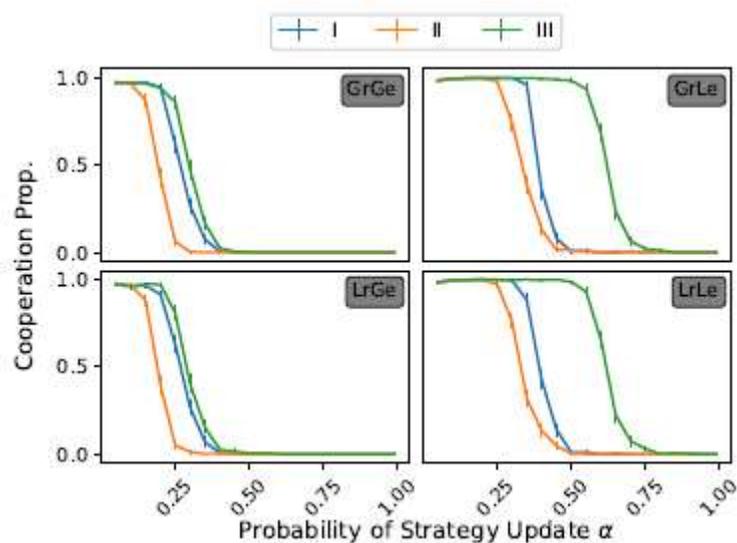
RESULTS: TYPE III NORMS PERFORM BEST

The Leading 8 norms are all successful in sustaining cooperation also with local reputation and local evolution
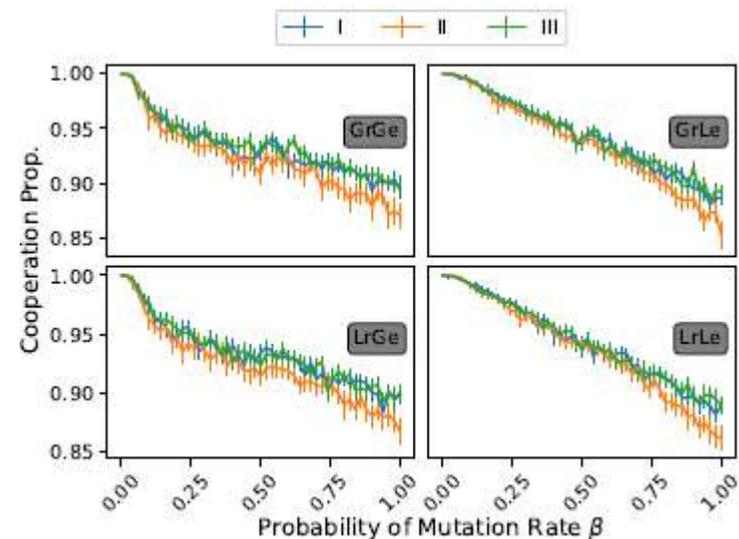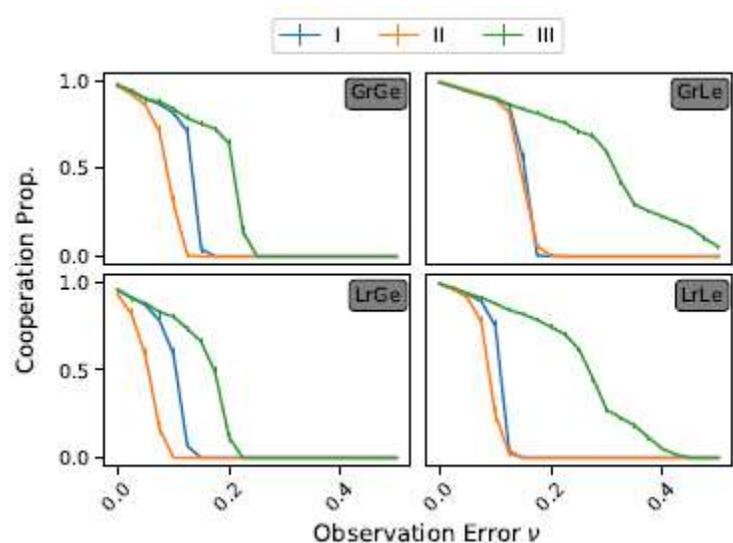
RESULTS: TYPE III NORMS PERFORM BEST

The Leading 8 norms are all successful in sustaining cooperation also with local reputation and local evolution
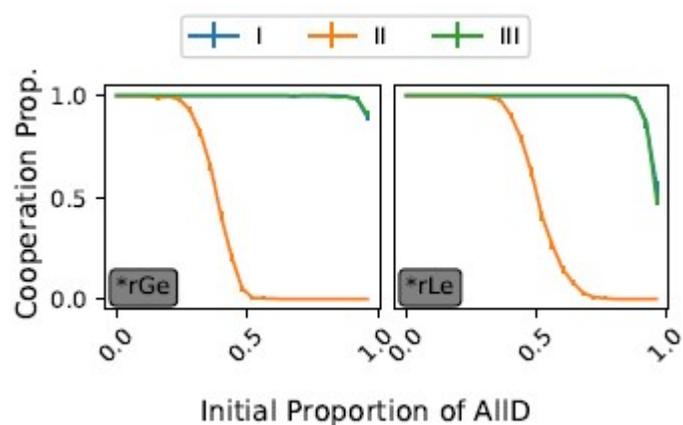
ALL D proportion present at the outset
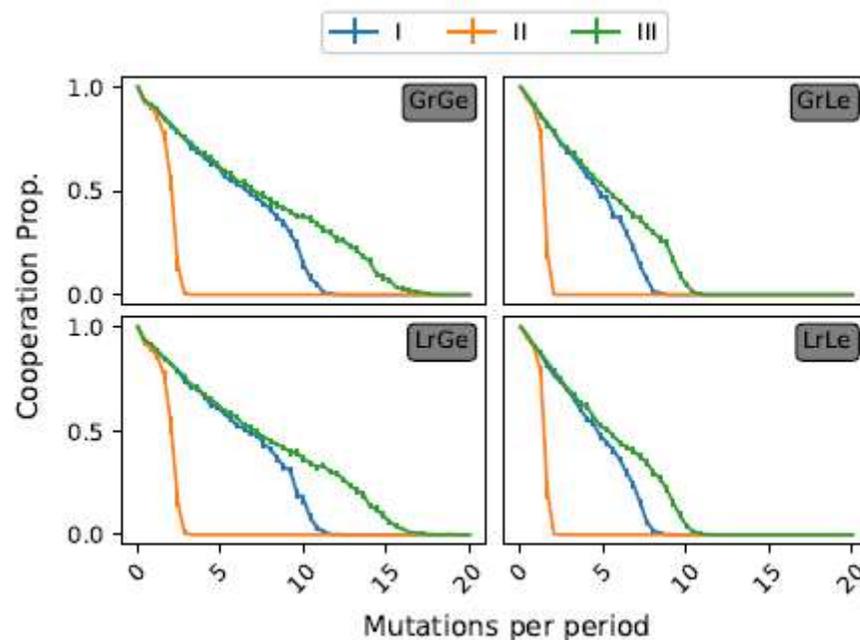
Multiple ALL D mutations per period

THE BEST OF SUCCESSFUL NORMS

For local reputation, local evolution

Robust result (network type, density, speed of evolution, reputation broadcast, mutations, observation error, proportion of ALL D strategies)

- No repair of reputations (7th and 8th columns)
  - No action is rewarded with good reputation after the meeting of individuals both with bad reputation
  - Type III are the most critical from the Leading 8

Group	d_{11C}	d_{11D}	d_{10C}	d_{10D}	d_{01C}	d_{01D}	d_{00C}	d_{00D}	P_{11}	P_{10}	P_{01}	P_{00}	
I	1	0	1	1	1	0	1	0	C	D	C	C	s_1
	1	0	0	1	1	0	1	0	C	D	C	C	s_2
II	1	0	1	1	1	0	1	1	C	D	C	D	s_3
	1	0	1	1	1	0	0	1	C	D	C	D	s_4
	1	0	0	1	1	0	1	1	C	D	C	D	s_5
	1	0	0	1	1	0	0	1	C	D	C	D	s_6
III	1	0	1	1	1	0	0	0	C	D	C	D	s_7
	1	0	0	1	1	0	0	0	C	D	C	D	s_8

KÁROLY TAKÁCS

Related experimental evidence

Samu F., Számadó Sz., Takács K. 2020. *Sci. Rep.*

This study is published as:
Podder, S., Righi, S., and Takács, K. 2021. Local Reputation, Local Selection, and the Leading Eight Norms. *Scientific Reports*, 11, 16560.

LINKÖPINGS UNIVERSITET

SOCIAL NORMS OF COOPERATION

People are characterized as having good or bad reputation

Here: privately assigned reputations

Clear effect of first order norms (reputation), but uncertainties about higher order norms

25

		GG		GB		BG		BB	
Panel A	C	G maintenance of cooperation		*		G forgiveness		*	
	D	B identification of defectors		G justification of punishment		B identification of defectors		*	
Panel B	C	64.1	75.5	61.5	69.9	31.6	40	30.6	38.3
		52.3	55.3	68.7	62.8	35.3	44.9	33.8	32.1
	D	59	75.9	57.5	69.1	22.3	34.4	18.6	30.2
		46.6	48.4	54.2	55.2	27	28.8	23.0	28.6
Panel C	C	-1.38	1.18	0.08	2.75	13.60	18.78	14.35	27.90
		-1.78	0.27	-3.14	-3.08	12.94	16.22	17.15	7.46
	D	-4.36	-3.66	-4.04	-5.00	3.95	19.08	4.76	16.48
		-8.42	-4.98	-8.08	-10.08	2.27	2.57	2.84	2.17

Table 1. Means of predicted reputation scores (Panel B) and mean of predicted changes in reputation scores (Panel C) of the focal player after the observation of a play between the focal player and his opponent in our experiment. Rows show the action of the focal player (C: cooperate, D: defect). Columns show the potential combinations of reputation scores for the focal player (first letter) and the opponent (second letter) (G: good, B: bad). Results can be compared to the table of social norms[19]. Expected common properties of the leading eight norms are indicated with red (good) and purple (bad) font. Each cell contains predicted reputation scores divided by treatment condition (upper-left: A-NP, upper-right: A-PW bottom-left: S-NP, bottom-right: S-PW).

REPUTATION-BASED COOPERATION

□ **Constructing reputations:**

    □ **Direct experience (interactions, observations)**

    □ **Exchange of information -> gossip**

Gossip is an informal device to maintain cooperation in larger groups

- objected towards norm violators and shirkers
- sanctioning potential: punishment / damage to reputation of the target (Feinberg et al., 2014; Hess & Hagen, 2006; Nowak & Sigmund, 2005)
- information on others (potential partners) without direct observation

26

COOPERATION
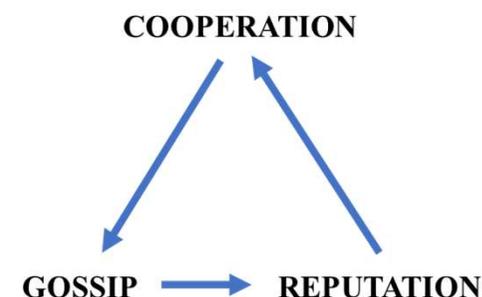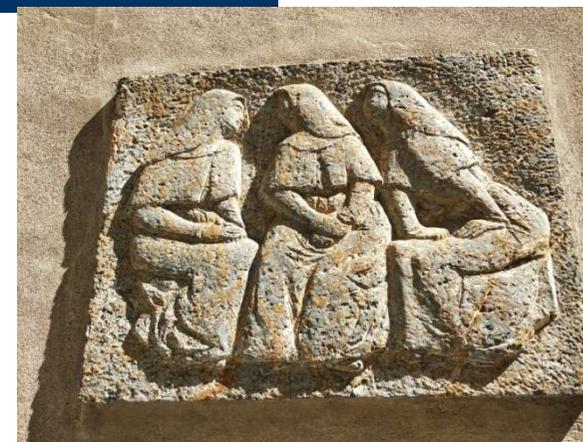
GOSSIP → REPUTATION

THE PUZZLE:
HONESTY AND JUSTIFIED PUNISHMENT

- **Reputation is not binary**
- **But how can reputation work, if it is just a *social construct***
  - **and built on potentially *dishonest* communication such as gossip?**
  - **and *justified punishment* can easily be misinterpreted**

27

Gossip must be more than the exchange of subjective evaluations of others!

Hypotheses:

Reputation-based conditional cooperation can work if gossip is also

A SOCIAL MIND READER

- *Perspective taking* to calculate expected actions of others

A SOCIAL MIRROR

- A tool to acquire information about the reputation of the self

28

LI.U LINKÖPINGS UNIVERSITET

METHOD: ABM

SocNetV: 00.png
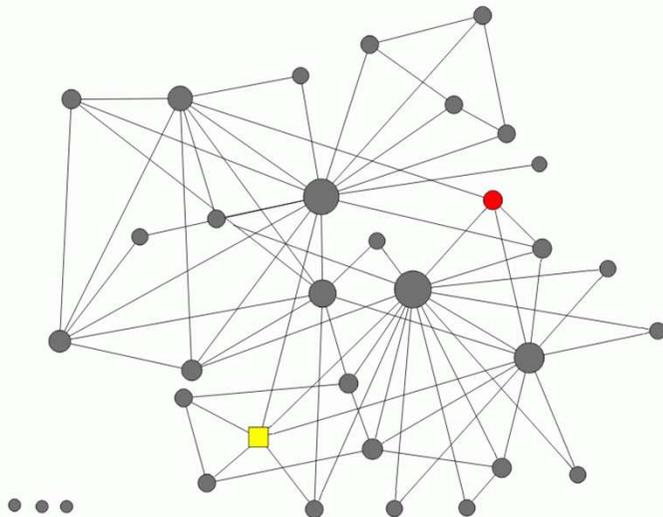
gossip with no limits

Setup
Random initialization of reputations [0, 100] and of conditional strategies [c_i]

Dynamics
- Random matching for 2-person PDs
- Reputations are updated after play
- Gossip
- Reputations are updated
Update conditional strategies

Manipulations
- What is passed on in gossip [reputations only, thresholds, self-image]
- Frequency of gossip
- Memory
- Gossip partner and target selection

LIU LINKÖPINGS UNIVERSITET

MAIN RESULTS

Average cooperation level for different contents of gossip

Legend:
- no gossip
- Gossip updates R_{iz}
- Pass only R_{zi}
- Pass R_{zi} and c_z
- Pass only C_z
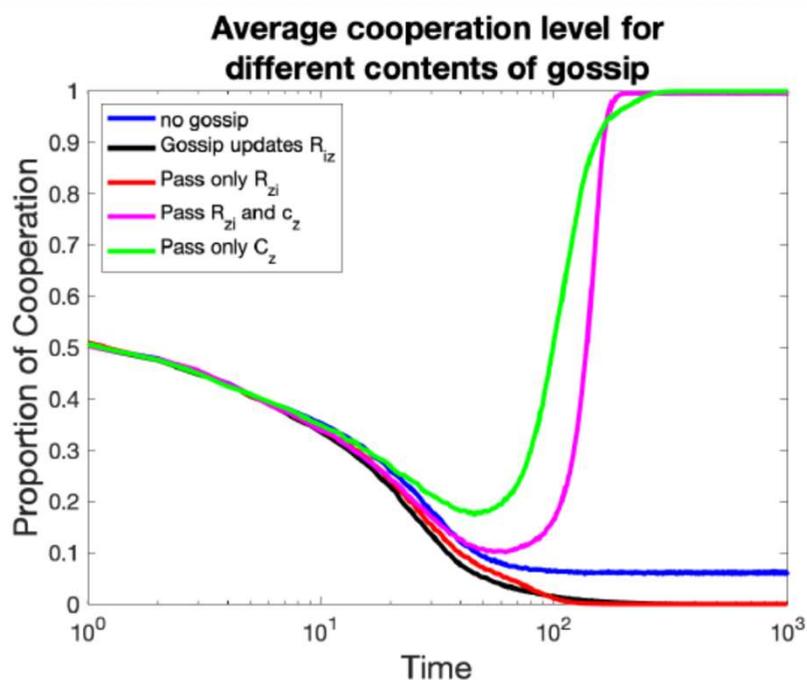
Figure 1: Time evolution of cooperation under different assumptions concerning the content of gossip. Baseline model with perfect - up to date memory. Averages computed at each time step over 100 simulations for each type of gossip. The distribution of outcomes is reported in the SI (Section S1.2).

- Passing on reputations only does not establish large-scale cooperation
- Gossip including a social mirror function only is insufficient to produce cooperation
- Gossip must contain perspective taking (social mind reading) on conditional thresholds in order to choose the proper behavior about an opponent
- It matters not just how GOOD the opponent is, but also how STRICT the opponent is with others

30

MORE GOSSIP AND LONGER MEMORY HELPS COOPERATION

Average cooperation level for different contents of gossip and information characteristics

Average cooperation level for different memory lenghts and amounts of gossip
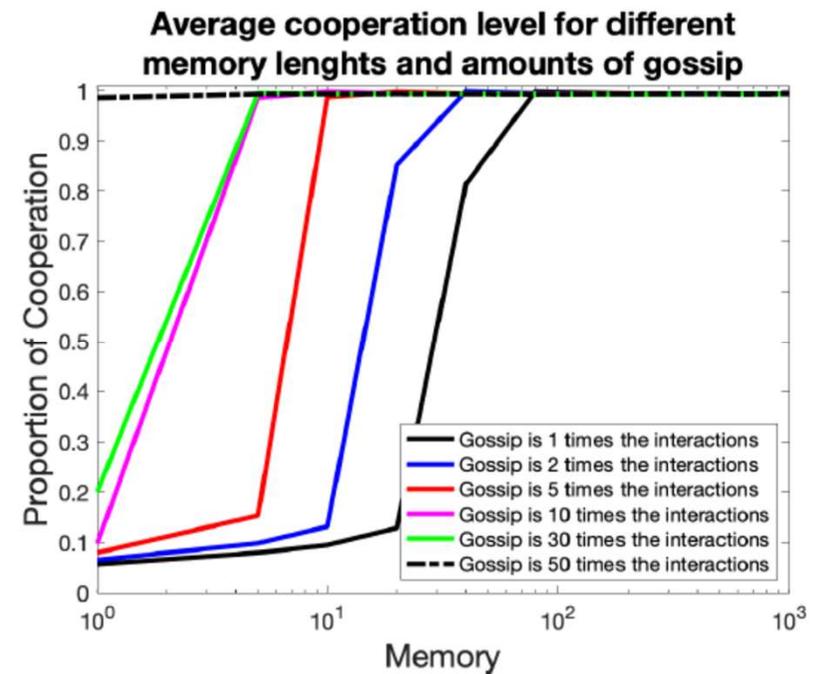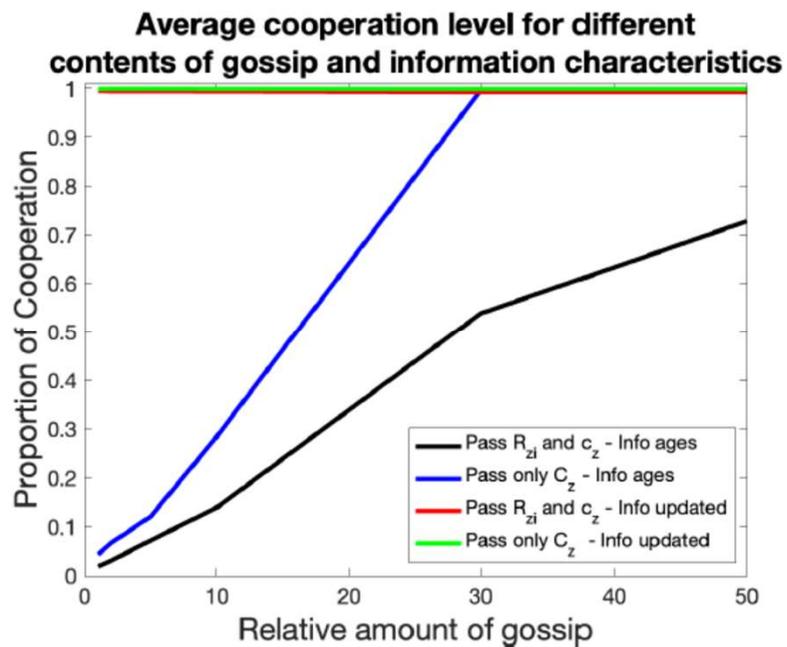
CONCLUSIONS

- Public reputation systems can contribute to large-scale cooperation, it is more difficult if reputations are privately assigned
- *Perspective taking* is the key element of gossip that *in addition to* evaluation of others can lead to large-scale cooperation
  - by separating cooperators
  - and leaving defectors with lower payoffs
- Not just goodness matters, but the intentions as well
- Results are robust to:
  - different gossip partner selection mechanisms,
  - target selection,
  - population size
    (few-hundreds range)