Towards feasible social norms in divided societies: Evolutionary game theory in indirect reciprocity under private assessment

Sam Okada

Soka University, Tokyo

and

Hannelore De Silva

Vienna University of Economics, Austria

Social Norms Workshop at Ascona Monte Verita
July 10, 15:00-15:30 (30min), 2024

A fragmented society

Several aspects of divided worldviews
- – Divided relationships
- – Divided information
- – Divided opinions

Social division is accelerating. Specifically, not only is there a "division of relationships" where communication is dysfunctional due to geographical and hierarchical divisions, but there is also a "division of information" caused by closed information environments such as filtering bubbles and echo chambers, and a "division of opinion" where opinions on issues are conflicting and social consensus building is difficult, and these are interconnected and becoming more serious. We live in a society that is different from the past, and it is necessary to reexamine our previous way of thinking. This is because it is not self-evident what norms will function in a divided society.

Normative approach

Normative research on whether division is right or wrong is generally difficult to ensure objectivity. Furthermore, in order to avoid division, strong policies such as political leadership and the promotion of national consciousness are necessary, but this is generally difficult to achieve in a democracy. In addition, social infrastructure for social activities such as schools, and neighborhood associations do not anticipate a divided society, and the acceleration of division is a serious concern. If we continue to search for ways to avoid social division, we may not be able to address the current urgent issues.

Two approaches

Positive approach

In this talk, we take an empirical and elucidating standpoint that assumes social division and considers that "diversified worldviews," in which dialogue cannot take place because people see the world differently, are the cause of the various aspects of division in terms of relationships, information, and opinions. We then take into account of what kind of norms and principles of behavior can build a desirable cooperative regime in such a divided society. Although a standpoint that assumes division is normatively debatable, it is one realistic approach to solving problems and is expected to present realistic guidelines that lead to scientific social system design. By carrying out this research, we will be able to answer the academic "questions" of what norms and principles of behavior are necessary for social systems and institutions to function even in a divided society, and how incentives to promote them should be designed.
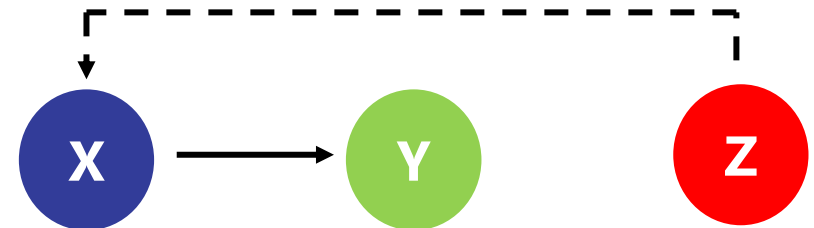
Three types of reciprocity

SOKA University

Direct reciprocity

First help

Second help

X helps Y because Y helped X

X → Y
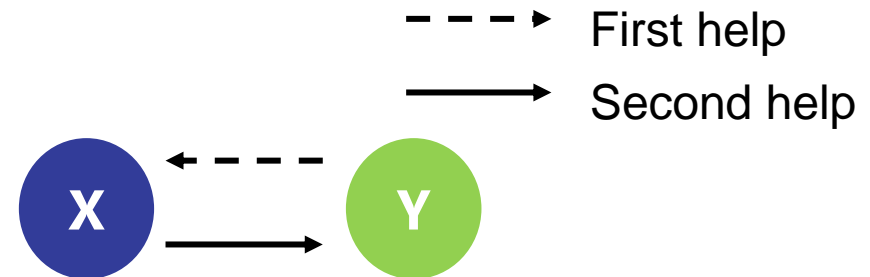
Upstream indirect reciprocity

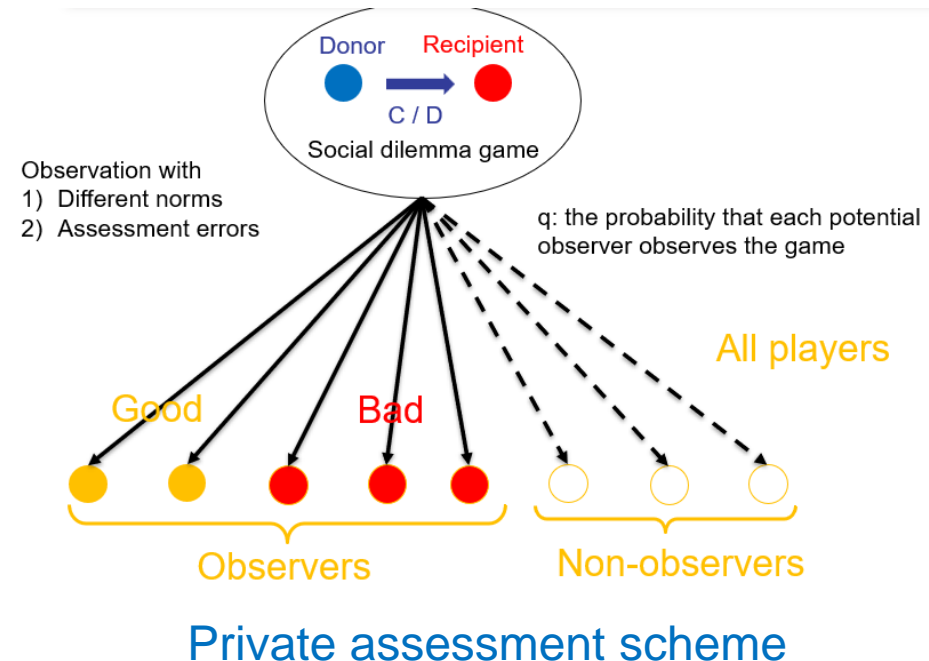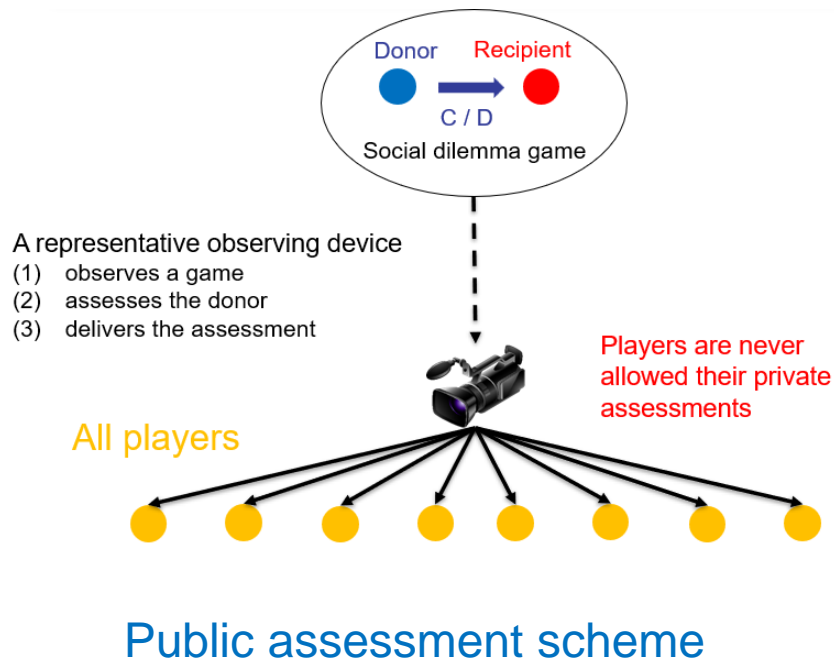X helps Y because Z helped X

X → Y   Z

Downstream indirect reciprocity

X helps Y because Y helped Z

X → Y → Z

- A fragmented society (A divided society)
  - evaluations by others do not match
- A public assessment scheme
  - reputation of a player is publicly and uniquely determined, that do not allow for personal impressions

Donor   Recipient
C / D
Social dilemma game

A representative observing device
(1) observes a game
(2) assesses the donor
(3) delivers the assessment

Players are never allowed their private assessments

All players

Public assessment scheme

Donor   Recipient
C / D
Social dilemma game

Observation with
1) Different norms
2) Assessment errors

q: the probability that each potential observer observes the game

All players

Good   Bad

Observers   Non-observers

Private assessment scheme

The dilemma of punishment

The impressions of Bob are different between Alice and Chris.

(1) Alice thinks that Bob is a bad person, so Alice defected against Bob.

(2) But, for Chris, Alice's action does not seem to be justified in any way because Bob is a good person from Chris's point of view.

-> With private assessment, a justified defection is not necessarily justified because two discriminators may not correspond a focal target
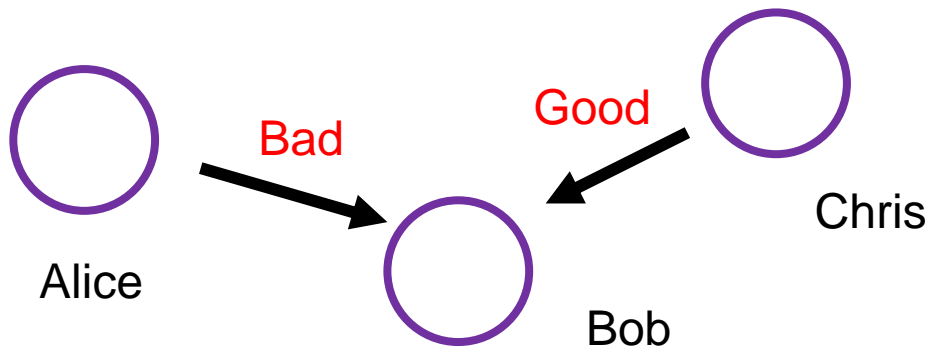
The third dilemma (Okada 2021 SciRep)

Different individuals disagree on who deserves punishment

(3) However, Chris can make the following inference:

"Until now, I thought Bob was a good person, but the fact that good Alice's defection against Bob means that maybe Bob is a bad person?"
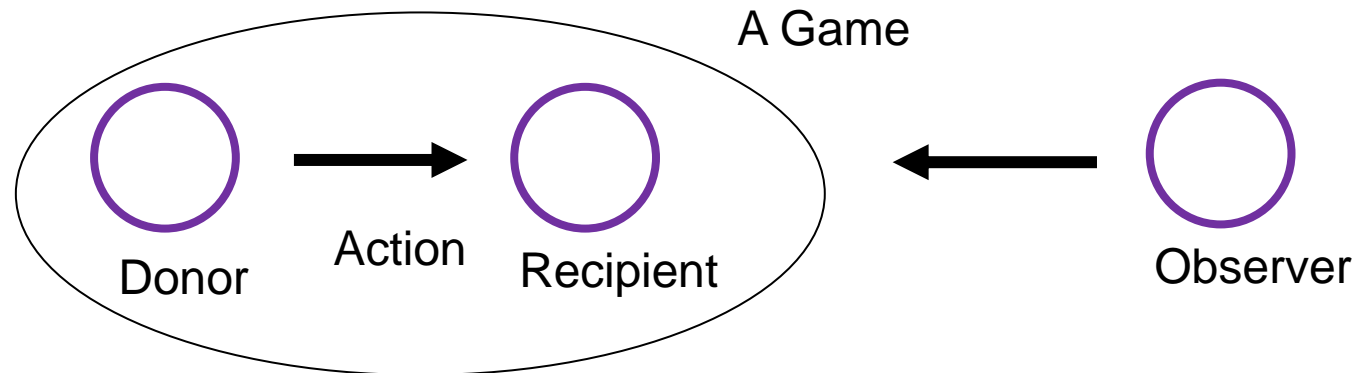
It is quite conceivable that Chris will update his impression of Bob as a result of such inference.

Assessment rules, so far

In previous studies …
An OLD observer's assessment function:

Donor's image{G,B} × Recipient's image {G,B} × Donor's action {C,D}
→ New donor's image {G,B}

A Game

Donor   Action   Recipient   Observer

We adopt
a NEW observer's assessment function (in a private assessment scheme):

D's image{G,B} × R's image {G,B} × D's action {C,D}
→ New Donor's image {G,B} × New Recipient's image {G,B}

D's image{G,B} × R's image {G,B} × D's action {C,D}
→ New Donor's image {G,B} × New Recipient's image {G,B}

If an observer feels that a game is not consistent with one's own norm, the observer have to change an image of the players playing the game, but which one?

(1) A consistency function number
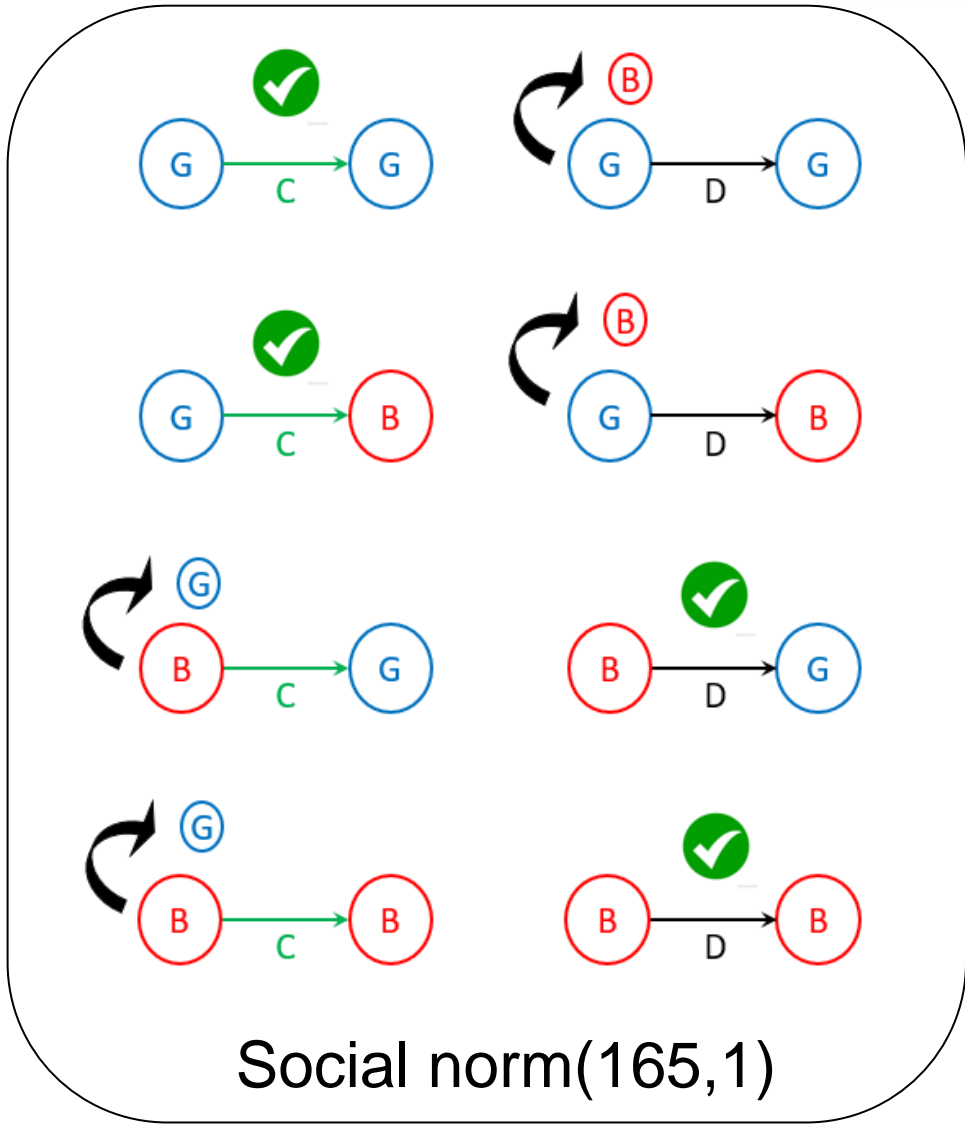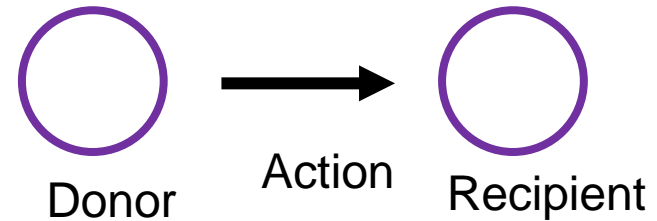(2) A rule updating images

Ex) Image Scoring

Social norm(165,1)

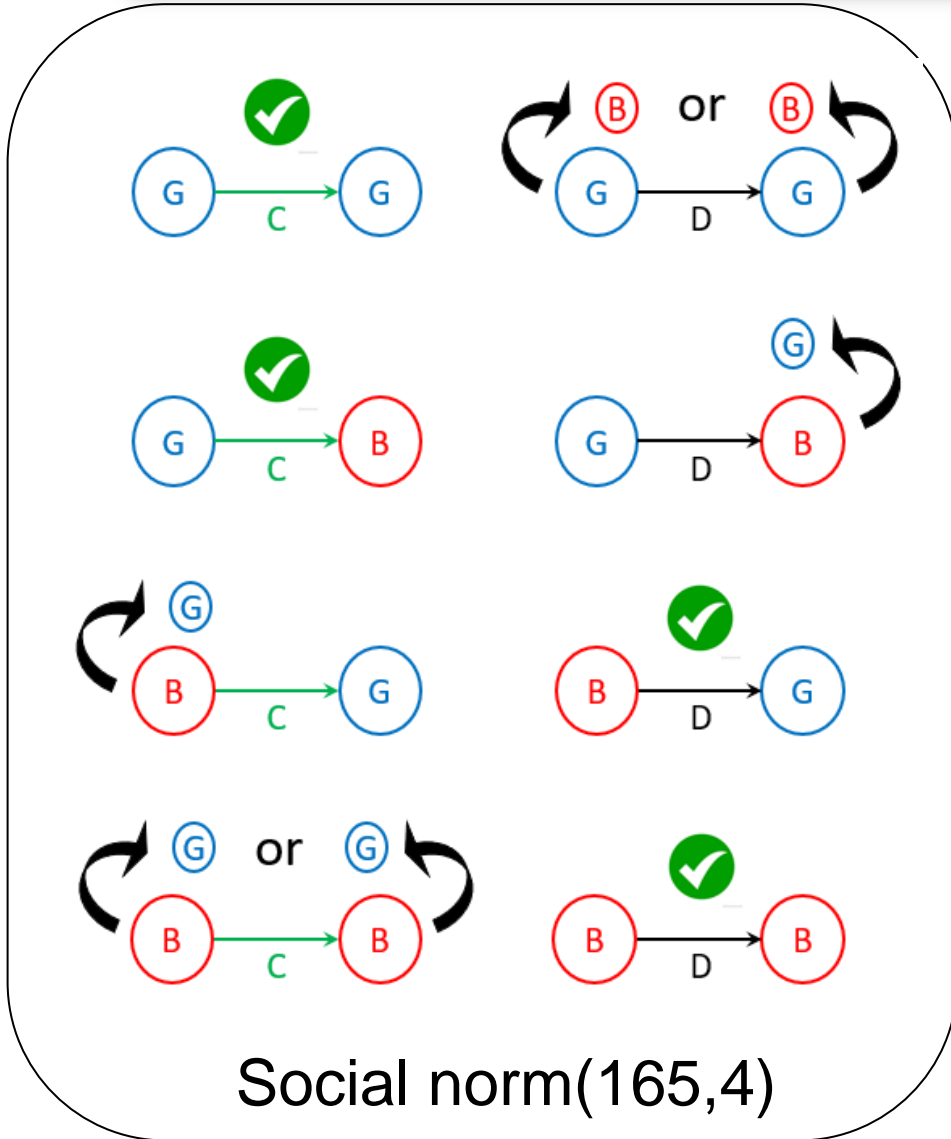Image Scoring rule is the simplest
- Cooperative donors are Good
- Defective donors are Bad

Donor   Action   Recipient

(1) Consistency function number
[1010 0101] = 165

(2) Rule updating images
[Rule 1] = Donor's image is revised.

Social norm(165,4)

The consistency function number is the same of Image Scoring.

[Rule 4: Prioritizing Good image] Despite of donor or recipient, the bad image should be changed to good image.
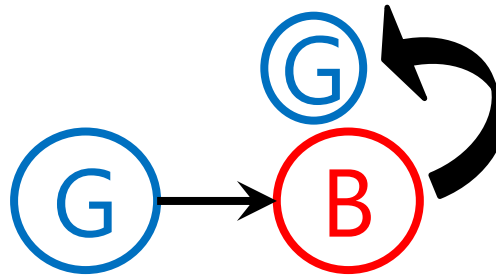
Note that the donor and recipient images may have the same priority; thus, one image (either donor or recipient) is selected at random for updating.
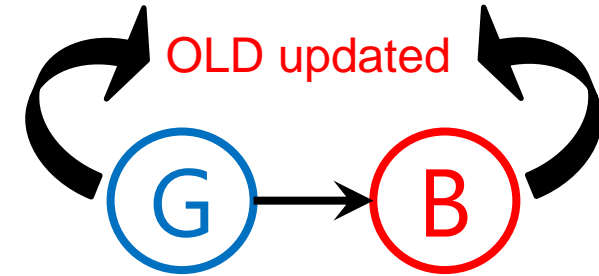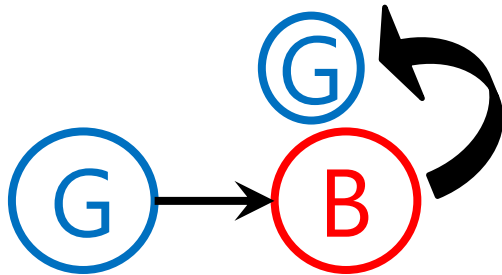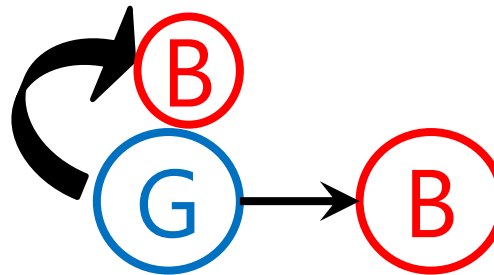
Analyzing six rules

Rule 1
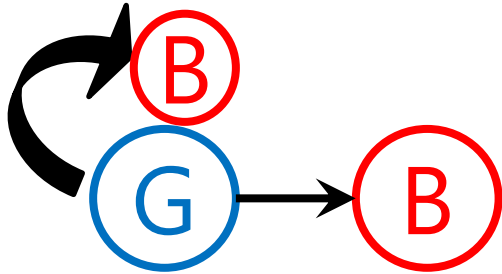
Rule 2

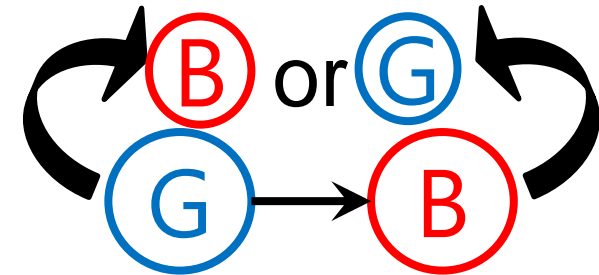### Rule 3
OLD updated

Rule 4

Rule 5

Rule 6
or

[Rule 1: Prioritizing Recipient's Image] Prioritize the recipient's image and update the donor's image

[Rule 2: Prioritizing Donor's Image] Prioritize the donor's image and update the recipient's image

[Rule 3: Prioritizing New Image] Prioritize the most recently updated image and update the older image

[Rule 4: Prioritizing Good Image] Prioritize good images and update bad images

[Rule 5: Prioritizing Bad Image] Prioritize bad images and update good images

[Rule 6: Random Updating] Randomly decide what image to update

Almost all previous studies adopt Rule 1

Model detailed

- Agent-based simulations
- Finite population: N agents
- Binary image {G,B} in private assessments
- Action rule: [C to G and D to B]
  - Note that, in private assessments, this image is used for the choice of C or D.
- In each round, a pair of two agents are randomly selected and play a donation game with $b>c>0$, with unilateral action error, e1
- Observation probability, q, with observation error, e2
- A generation consists of T1 rounds
- Data the last T2 rounds are used for the payoff calculation
- All initial images are G at the beginning of any generation
- N=100, (b,c)=(3,1), q=10%, e1=e2=1%, (T1,T2)=(1000,100)
- Three strategies (X=ALLC, Y=ALLD, Z=Norm adopter)(Init= 1:1:98)
- The fermi updating with mutation: $\beta=3$, $\mu=1\%$

Rate of maintaining cooperation in all possible social rules
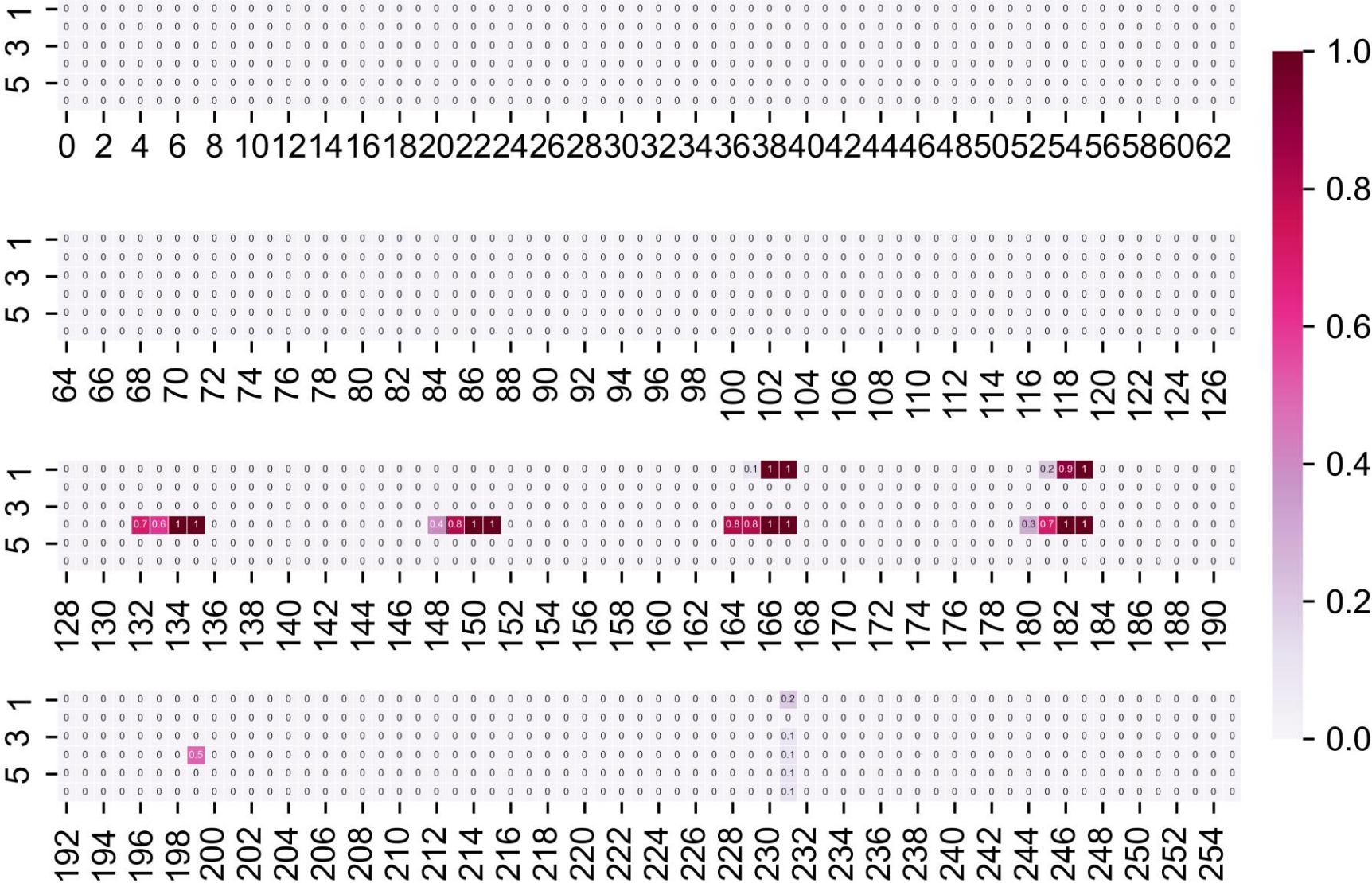
Stable Cooperative Norm (SCN) under Rule 4 (Good Image Prioritizing)

SCN under Rule 1 (Donor Updating)

Leading four for private scheme (Okada 2020)

Leading eight for public scheme (Ohtsuki and Iwasa 2004)

182(L4) 183(L7) 181(L1) 180(L3) 148(L5) 149(L2)

166 167

165(IS)

150(L6) 151(L8)

134 135

IS-like

Heider-like

Take-home messages

- For considering a fragmented society, a private assessment scheme in indirect reciprocity study should be analyzed
- We performed an exhaustive analysis using evolutionary game theory and agent-based (numerical) simulations
- We theoretically discovered the new norms for maintaining cooperative regimes
  - IS-like norms:
    - Image-scoring: Theory versus Empirical fact
  - Heider-like norms:
    - Heider's balance theory (1958)
- Hopefully, this paper may be published soon!

This talk is based on

Isamu Okada I and Hannelore De Silva.

Norms prioritizing positive assessments area likely to maintain cooperation in private indirect reciprocity. *Sci. Rep.* 2024